ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA

SCUOLA DI LETTERE E BENI CULTURALI

Corso di Laurea in Italianistica, Culture Letterarie Europee
e Scienze Linguistiche

(r) in Heritage Calabrese Italian: Cross-generational nativeness

Tesi di laurea in

Fonetica e Fonologia

Relatrice: Prof.ssa Gloria Gagliardi

Correlatrice: Prof.ssa Naomi Nagy

Presentata da: Angela Cristiano

Sessione
Terza

Anno accademico
2020-2021

A Flavio

Table of Contents

Introduction

The present work is an empirical study on how rhotics, or r-sounds, are realized in the spontaneous speech of Heritage and Homeland Calabrese Italian speakers. Within a Comparative Variationist paradigm, we will analyze whether the behavior of the Heritage speakers is influenced by language-internal tendencies or processes, or by contact with a majority language, i.e., English.

Our interest in the general topic of speech was sparked by the will to further understand and analyze spoken language in a systematic fashion. Since speech is our primary communicative modality, a linguistic model whose ambition is to provide an account of how language is produced, perceived, and used, cannot prescind from the phenomena occurring during communicative linguistic interactions. Hence, empirical data regarding said phenomena is needed. Furthermore, the aforementioned hypothetical model should be able to account for the linguistic behavior of different types of communities or individuals: thus, data reaching outside the scope of monolingual speakers or majority-language communities is also needed.

This thesis serves exactly these two goals: it enriches the literature about the patterns of production occurring in natural conversation, but going beyond the analysis of Monolingual, Young, Available, Literate (MYAL) speakers (Polinsky, 2018). The third purpose of the research is to help rethink the notion of "nativeness". Hence, different ideas linked to said notion will be explored, to evaluate whether their base assumptions hold up to empirical testing.

To reach the aforementioned goals, we focus our attention on the Heritage Calabrese Italian community of Toronto (ON), Canada, and on a phonetic feature of their language (often socially imbued among communities of the world): the realization of rhotics.

Chapter 1 is dedicated to a discussion about Heritage Language speakers, with the aim of providing a comprehensive summary of the frameworks they have been analyzed within, and of the notions and ideas that have emerged in the last decades.

Heritage speakers are described by Montrul (2010) as individuals, part of a minority language community, who are exposed to both the minority variety and the majority variety of the community in which they live. However, a unanimous definition of said group has not been reached: therefore, section 1.1. is a review of the debate surrounding the definition. Section 1.2. explores how ethnic and social factors intertwine with the linguistic production of Heritage speakers. Successively, we focus our attention on two of the most discussed aspects regarding

these speakers: bilingualism and acquisition (sections 1.3. and 1.4.). The most common idea of Heritage speakers is that of individuals who "incompletely" acquired the Heritage Language, or who lack proficiency in it, and who experience either attrition or transfer from the majority language, leading to a simplification of their grammar. The consequence is that they have rarely been considered native speakers. Results from approaches employing different assumptions on Heritage speakers, and different methodologies, tend to contradict the idea of inevitable grammar simplification: however, we think that it is important to evaluate the positions illustrated in 1.3. and 1.4., to have an idea of the state of the art, and to formulate research questions and hypotheses that take into account the current discussion.

Since our thesis analyzes phonetic detail, we specifically discuss the sound system of Heritage speakers in section 1.5., illustrating the advantages and disadvantages of the application of the Speech Learning Model by Flege (1995) to Heritage Languages. Section 1.6. draws conclusions on the matter of the grammar of Heritage Languages: are they incomplete, deficient, divergent, or not? Two answers are provided, coming from the two most common approaches to the study of Heritage Languages: the Experimental approach and the Comparative Variationist approach. Section 1.7. presents two very common sociolinguistic scenarios for Heritage Languages: maintenance and shift. Finally, section 1.8. argues for the attribution of the status of "native speakers" to Heritage speakers, by appealing to a re-shaping of the notion of nativeness itself.

Chapter 2 is a review of current literature regarding rhotics, of the variants in the linguistic varieties relevant for this thesis, and of the processes that may affect rhotics.

Rhotics are a class of phonemes not defined by a unifying articulatory or acoustic feature, whose origin lies in the derivation from the Greek character *rho*. Examples of rhotics are the Italian alveolar tap [ɾ] (as in *caro*, [ˈkaːɾo], "dear"), or the North American alveolar approximant [ɹ] (as in *more*, [ˈmɔɹ]). They are characterized, among many features, by a tendency to be socio-indexed. These and other aspects make them a particularly interesting category, whose boundaries, members, and "essence" are frequently discussed.

Section 2.1. provides a typological description of the types of rhotics. Section 2.2. illustrates the phonological behavior of rhotics, and their involvement in patterns of lenition. Section 2.3. lists, according to both traditional descriptions and empirical research, which rhotic variants are most diffused in Canadian English and in Italian, Calabrese Regional Italian, and Calabrese dialect, since these are the varieties possibly affecting the linguistic behavior of the speakers analyzed. Furthermore, this section serves the purpose of establishing which rhotic variants are already present in the Homeland grammar; since this thesis runs a parallel analysis of the

Heritage community and the Homeland comparator community, this is crucially important. Finally, section 2.4. summarizes the most recent proposals – from different frameworks - on how to define the nature of the class of rhotics, and tries to critically offer a temporary solution to this theoretical issue.

Chapter 3 illustrates the empirical study we conducted about rhotics in Heritage and Homeland Calabrese Italian. Section 3.1. provides a brief description of the theoretical framework we employ for the case study, i.e., the Comparative Variationism paradigm (while also considering the contribution third-wave sociolinguistics, socio-phonetics, and usage-based models gave to the shaping of the theoretical setting). Section 3.2. describes the features of the speech community we investigated in Toronto. Section 3.3. illustrates the project whose scope this research rests within: the HLVC project (Nagy, 2011). The HLVC project pushes variationist research beyond a monolingual approach. To do so, it built a multilingual corpus for inter-generational, cross-linguistic, and diatopic (Heritage vs. Homeland varieties) parallel comparison: the HerLD corpus, which contains speech data, coming from sociolinguistic interviews, for ten Heritage Languages of the Great Toronto Area, and comparator Homeland samples. Our case study utilized data from this corpus, and followed the tested methodology proposed by the HLVC project.

Section 3.4. describes the hypotheses and methods of the case study. Phonetic realizations of word-internal singleton /r/ are analyzed in a sample of 29 speakers, ranging from 19 to 75 years of age, and coming from three social groups (Homeland, Generation 1 Heritage speakers, and Generation 2 Heritage speakers). The total number of tokens analyzed is 1555. The dependent variable, (r), is considered to have three values: Tap&trill, Fricative, or Approximant. The respective prototypical realizations for the three values are: [r]/[ɾ], [ɾ̥], [ɹ]. A variety of social, ethnic, and linguistic independent variables has been considered, including measures of Ethnic Orientation for the Heritage speakers. The coding of the linguistic variables, such as suprasegmental features of the token, has been conducted based on phonetic realizations, and not on citation forms. Data has been analyzed using Mixed Effects models, a choice that has gained success among sociolinguists in the last decade.

The hypotheses of the case study all revolve around the lack of simplification in the grammar of Heritage speakers: they claim that the contact with the majority language, i.e., Canadian English, has not influenced the rhotic production of Heritage speakers, who are hypothesized to maintain the grammar of the Homeland speakers they are compared with. Moreover, the hypotheses support the idea of Approximant and Fricative rhotics in Italian as lenited variants

of Taps&trills, naturally emerging during relaxed speech, and being favored by the coda position, a context particularly prone to lenition.

Section 3.5. presents the distributional and statistical results; section 3.6. discusses them in detail, examining cross-generational and cross-linguistic comparisons.

1. Heritage Languages

What does "being a native speaker" mean in the 21st century?

Linguistics tradition has been focusing for years on monolingual, ideal, individuals and/or idealized languages: from Bloomfield to Italy's earliest dialectologists to Chomsky; from the search for a purer, archaic language to the one for the internal, perfect competence, the reality of linguistic knowledge, production, and usage has been often ignored. Both for convenience and for theoretical idealizations, studies have been analyzing what Polinsky (2018) refers to as MYALs: Monolingual, Young, Available, Literate individuals. There has been full consent over their status of "native speakers", and they have even become synonymous with the term.

Nowadays, we must rethink this concept to account for the reality: the world is mostly multilingual, multilingual people are not an exceptional and minority category, and monolinguals are not "*the benchmark of true nativeness*" (Rothman and Treffers-Daller, 2014, 1). Indeed, *more than half* of the global population is multilingual from the first years of their lives (Tucker, 1999). Yet not all multilinguals suffer the problem of not being labeled "native speakers" equally. While only a few people would hesitate to define a balanced, simultaneous bilingual as a native speaker of both the languages they acquire, or a sequential bilingual who keeps clear dominance over their L1 a native speaker of this L1, there is a group of bilingual speakers that researchers are more reluctant to consider natives: Heritage Language speakers.

We will address in section 1.1. what Heritage Languages are. In section 1.2. we will explore the social and ethnic dynamics related to Heritage Languages. In section 1.3. we will frame them in the context of studies around bilingualism and L2 proficiency. Section 1.4. will be dedicated to the acquisition of Heritage Languages. Section 1.5. will be a focus on the sound system of Heritage Languages, since this research investigates sound production in Heritage Language speakers. Section 1.6. will present a discussion on the grammar of Heritage Languages. In section 1.7., after having established a more or less comprehensive picture of what Heritage Languages are, we will describe patterns of maintenance and shift in Heritage Languages. Finally, we will turn back, in section 1.8., to the issue of describing Heritage Language speakers as native speakers or not of their Heritage Language.

1.1. An overview of what Heritage Languages are

Heritage Languages (henceforth HLs) are varieties subjected to multiple definitions. The first one comes from the mid-1970s, when the term was developed and used in Canada for the Ontario Heritage Languages government program (Aalberse, Backus and Muysken, 2019), but the term spread only during the 1990s in North America, and later in Europe (Montrul, 2010; Kupisch and Rothman, 2018).

The Canadian Heritage Languages Institute Act of 1991 uses the term HL to refer to every mother tongue language different from the two official languages of the country, English and French, including both indigenous languages and immigrant languages (from non-French and non-English immigrants). However, the inclusion of the former in the definition is still debated (cf. §1.2. for a discussion).

From the origin of the term, it is quite clear that HLs are especially relevant in Canada. In fact, the country's multicultural policies (Noels and Clément, 1998), and especially provincial policies, have led over the years to a rich linguistic and cultural mosaic, with Statistics Canada (2017) reporting that, as of 2016, over 7 million Canadian citizens speak an immigrant language at home. Canadians seem to value the maintenance and transmission of non-official mother tongues, given for example the many HL classes attended by children (Harrison, 2000), or the positive attitude they display toward HLs and bilingualism (Pérez-Leroux, Cuza and Thomas, 2011)[1]. The most "neutral" definition of a Heritage Language speaker is that of a person, part of a minority language community, who is exposed to both the minority variety and the majority variety of the community in which they live (Montrul, 2010). This is none other than an adaptation of the best-known definition of HLs, as reported by Polinsky and Kagan (2007): the one from Valdés (2000). Valdés focuses on the United States situation, identifying HL speakers as individuals subjected, during their development, to a language other than English, and who show some degree of bilingualism with both of those languages. Therefore, English represents the majority variety in Valdés' definition, but this can (and has been) easily modified to account for any other majority variety in an analogue situation.

From this introductory definition it is already clear that HL speakers are considered, most of the time and by many scholars, bilingual individuals. Explicitly including bilingualism in the definition made it possible to consider as part of the HL studies agenda also research conducted in the 1980s in Europe: the fact that the term "Heritage Language" was not used in the Old

[1] Cf. also §3.1., for a discussion on the Italian Heritage community of Toronto (ON), Canada.

Continent up until the late 90s does not mean Europeans did not experience or study such situations, as Kupisch (2013) underlines.

Another definition which seems inclusive of many aspects without implying judgments is provided by Rothman (2009): a HL is a minority language spoken at home or readily available to children. It is acquired by them naturalistically, but it is also potentially affected by the fact of being non-dominant: i.e., by reduced input, differences in formal education and literacy, and the influence of the majority language. In this definition, Rothman avoids expressing conclusions about HL grammars or formal aspects of the language.

The literature developed around HLs in the last twenty years has expanded the description by Montrul (2010) provided above, stressing different facets of the topic. This has led to diverse points of view on what is a HL, with the attention shifting among various aspects. Nagy (2015), for example, asserts that the study of HL has been conducted within two approaches, with little communication between them. She calls the first one the Experimental Approach. Representatives are Silvina Montrul or Maria Polinsky. Its focus is on the psycholinguistics and acquisition of HLs. Nagy, then, calls the second approach Comparative Variationist, since it is based on the homonymous sociolinguistic method. Representatives are Carmen Silva-Corvalàn or Naomi Nagy herself.

Definitions of HLs following the first approach usually include judgements about the speakers' performance, and therefore in this framework there has been a lot of research around their proficiency levels and the way their attainment diverges from other groups of speakers. Definitions of HL following the second approach, instead, tend to exclude any judgment about an end state. These definitions are not about how the language is spoken, but rather take into account the linguistic autobiography of speakers.

Researchers coming from the first approach usually move within a formal linguistic framework and/or an individual-cognitive framework (cf. §1.3.2.), while those coming from the second approach usually move within a sociolinguistic and ethnolinguistic constructivist-variationist framework. Therefore, each group has its corresponding views on what is a language, how to rate proficiency, or how acquisition works.

An important premise is to be made. Analyzing a linguistic group often involves comparing and contrasting it to another group or to a set of "expected realizations", i.e., a standard. In the case of HLs, who is the group we make comparisons with? This has been addressed as the *baseline problem*.

Polinsky (2018) draws conclusions on this matter, establishing that the main baseline for HLs is the language of the first-generation immigrants, the ones who provided the input for the

following generation – the Heritage speakers. As we can see, this position implies a narrower definition of HL speaker which includes only immigrant languages, excluding indigenous languages (cf. §1.2.). Moreover, it should be clear that the language of first-generation immigrants is not necessarily a standard variety. Using this kind of baseline seems particularly fitting for studies regarding the acquisition of HLs. However, the baseline group can change if the research has a different aim. For example, other viable comparison groups might be exchange students who are native speakers of the HL studied (Aalberse, Backus, Muysken, 2019), Homeland speakers, or other bilingual groups.

With "Homeland" we mean the territory originally occupied, and in this case still occupied, by an ethnic group. In the case of immigrant languages, it is the area that the immigrants came from. A comparison with the Homeland may allow for the matching of the regional and social features of the speakers. This choice of baseline fits well studies about patterns of variation and change, as the ones from the HLVC research group of Toronto (Nagy, 2011). Using a Homeland sample, nevertheless, should be done carefully: the researcher has to remember that in the Homeland too varieties of the same languages are organized in a complex diasystem. There is no homogenous "Homeland language". Therefore, they should be aware of what variety of the Homeland linguistic repertoire they are choosing, and of what its position on the axes of variation (diatopic, diastratic, diaphasic) is. Relying on the standard Homeland variety by default must be avoided.

As far as bilingual baseline groups are concerned, the idea is to compare bilingual HL speakers to bilinguals whose majority language is the HL. This choice seems suitable for studies focused on HL speakers' proficiency levels. In fact, monolinguals and bilinguals are not qualitatively the same (Grosjean, 2010; cf. §1.3.). When they are contrasted, a comparison between different sets of knowledge and processing capabilities is made. The results obtained may therefore be due to this basic difference between monolinguals and bilinguals, and not to a difference in the actual proficiency levels of the two groups (Aalberse, Backus, Muysken, 2019). Using a bilingual baseline eliminates the risk of mistaking the former for the latter, since the two groups share similar interferences and processing. This way, what could be measured instead is how proficiency varies due to prestige factors, society attitudes, and amount and type of linguistic input and output.

1.2. Ethno-sociolinguistic dynamics in Heritage Languages

By accepting the broadest definition by Montrul above, one agrees to consider not only immigrant languages as HLs, but also indigenous languages. This is the position of some scholars: for example, Fishman (2001, 2006) puts immigrant, minority, and indigenous languages in the same group, i.e., HLs, because they all compete with a dominant language and have speakers that connect with the language via ancestry. Aalberse, Backus and Muysken (2019) agree with him, adding that these three types of languages share similar language contact effects too. However, they report that the Canadian government defined HLs as mother tongues that are neither official languages, nor indigenous languages in a country; Nagy (2011) also identifies HLs as the languages spoken by recent immigrants and corresponding to their heritage. Probably the most important reason to exclude indigenous languages from the definition of HL comes from the position of Canadian First Nations communities: Cummins (2005) reports that they perceive their language not as a HL, and prefer to refer to it as indigenous or aboriginal language.

As a matter of fact, most of the cases analyzed in literature tend to exclude indigenous varieties. Benmamoun, Montrul, and Polinsky (2013), for instance, affirm that HLs are typically the ones spoken by second-generation immigrants. Within this meaning of HL speakers as children of immigrant parents, Kupisch (2013) highlights another unclear aspect, asking whether HL speakers must have both parents sharing the same heritage to be classified as such, or not.

Polinsky (2018) decided to join the two opinions and generalize by stating that a HL is the home/minority language of a bilingual who is dominant in the majority language of the society they live in[2]. Majority languages are spoken by dominant groups, therefore have an official status. They are normally standardized and provided with a written form, and represent prestige varieties. Minority languages, in contrast, are spoken by ethnic minorities, and in most cases have no official status. They may or may not have undergone standardization and the development of a written and formal register, but are typically not taught in local schools, and neither are they vehicles of the dominant culture and media, resulting in a restriction of the contexts in which they are used. Therefore, they represent lower prestige varieties.

It is important to highlight that what is a minority language in one country can be the majority language in another country: an example is Italian, which is a minority language in Canada for

[2] Note that this definition implies a sort of "deficit" – either grammatical or functional – in the HL. Cf. §1.3.2. for discussion about proficiency levels.

immigrants and their families, but the majority language in Italy. Despite having a written tradition, formal registers and a standard form, the lack of official status in Canada makes Italian an example of minority language in that context. Majority and minority languages can often be found in what has been called, following Ferguson (1959), situations of *diglossia*: in a simplified fashion, the presence of several languages, or varieties of the same language, with a clear social and functional specialization. While bilingualism just refers to the coexistence of two languages in the individual or the community *per se*, diglossia refers to their functional and hierarchical differentiation in the repertoire (Berruto, 1995).

Since minority languages are usually spoken by ethnic minority communities, a brief discussion about the connection between language, ethnicity and identity is needed.

First and foremost, the concept of ethnic group has to be defined. Hutchinson and Smith (1996) identify an ethnic group on six main features, with various degrees of incorporation among members and groups. An ethnic group usually exhibits:

1. A common proper name;
2. A myth of common ancestry or fictive kinship;
3. Shared historical memories;
4. One or more elements of a common culture, usually including religion, customs, or language;
5. A link with a homeland, whether or not the ethnic group still occupies the territory;
6. A sense of solidarity on the part of at least some of the group (Hutchinson and Smith 1996, 6-7).

Ethnicity, the core value of ethnic groups, has a fuzzy meaning. It can indicate the "essence" of the ethnic group, the sense of belonging to the ethnic group all members share, or what makes one ethnic group different from others (Tonkin et al., 1989).

Ethnicity has become a key aspect of the identity of groups and individuals. As for gender, sexual identity, or social class, for example, it fulfills the purpose of affirming one's self, but also of building and negotiating status, power, and prestige in societies (O'Reilly, 2001). Language, as we see in point 4, can be a key part of ethnicity, and thus of the process of constructing and expressing identity. This is why using the HL associated with one's ethnic group may represent a way to communicate the sense of belonging to the community, or a tie with ancestry. Sometimes, who you identify with or want to be identified with determines some of your linguistic choices: something that second, and especially third wave variationist sociolinguistics have taught us (Bayley, 2013; cf. §3.1.). Indeed, social constructivists have

presented the idea of identity as "as interactionally and discursively negotiated and accomplished and defined it through the concept of positioning[3]" (Abdi, 2001), with people having possibly multiple identities whose aspects are taken up in and through different discourses.

Examples of the bond between language ideology, choice, and ethnic identity may be seen in Wodak et al. (1999) or in Kroskrity (1992). The former analyzes the construction of national identity and the maintenance of the status quo through discourse in the Austrian context. The latter investigates the role of language ideologies in the formation of multiethnic cultural identities among the Arizona Tewa, a group who escaped the colonial oppression from Spain by leaving its homeland in New Mexico 300 years ago. Their ethnic identity and ancestral language were maintained thanks to language ideology and use.

Aalberse, Backus and Muysken (2019, 8) affirm that the typical HL speaker has a personal and emotional bond with the HL. In this sense, HL speakers are thought to form not only an ethnic, but also a language community, which shares linguistic norms (a vision promoted for example by Nagy, 2015). Nonetheless, this is not the case for *every* HL or HL speaker. Some of these varieties undergo a huge restriction of the contexts of use, being spoken only at home. This is one of the reasons why Montrul (2013b) is against the idea of considering HLs as different regional varieties, as we will further see in §1.4.3.

Given the potential relevance of ethnic factors, several research lines have investigated them. Examples are comparative variationist studies including among tested factors Ethnic Orientation (see the HLVC project; Nagy, Chociej and Hoffman, 2014; §3.3.), or studies based on ethnographic observations (for example, the Turchetta and Vedovelli 2018 project). Let us describe some of these analyses.

Studies from the HLVC project employ an Ethnic Orientation Questionnaire (EOQ). The idea is that speakers' engagement with their ancestral ethnic group (Noel, 2014) can be assessed and measured via the answers to a questionnaire. Behaviors, experiences, or judgements related to ethnicity may be predictors of linguistic variation (Nagy, Chociej, Hoffmann, 2014), so the answers are operationalized in order to test whether significant correlations with other variables show up. The EOQ they employ is adapted from Keefe and Padilla (1987), and consists of 37 questions grouped in sub-sets based on topic[4]. Featured subjects are language use, social

[3] "An event of identification in which a recognizable category of identity gets explicitly or implicitly applied to an individual" (Wortham, 2004, 166).

[4] They also tried the "Reference Group method", i.e., a division in subgroups based not on the topics of the questions, but on the people the questions refer to, in order to determine the role played by different kinds of social

networks[5], community of practices, attitudes toward cultural heritage, and discrimination (Nagy, Chociej, Hoffmann, 2014). Speakers' responses are scored on a three-point scale where 0 indicates orientation toward Canadian culture and 2 toward the Heritage culture; therefore, speakers "fall in one of three positions along a continuum from heritage-culture oriented to mainstream-culture oriented" (Nagy, Chociej, Hoffmann, 2014, 21). One crucial suggestion the authors offered is to adapt the EO criteria to the generation of speakers analyzed, since generational differences on how EO relates with linguistic variation were found for two linguistic variables – VOT (Nagy and Kochetov, 2013) and pro-drop (Nagy et al., 2011).

A key finding of many HLVC studies is that the ethnolinguistic vitality of the community and the EO of individuals do not show up as significant predictors of HL variation. For example, Nagy (2017a) investigates pro-drop in Heritage Italian and Heritage Faetar, an endangered language, and concludes that patterns of linguistic variation in both languages do not correlate with EO; in fact, the data shows no contact effects with English, but instead patterns of internal change. Similar results appear for other languages and variables (classifiers or case-marking; Nagy, 2017b), where no correlation between individuals' rate of use of more English-like forms and their EO appeared. The issue of relating EO scores to linguistic variation was slightly more difficult in the case of VOT. Nagy and Kochetov (2013) compared VOT in three HLs – Russian, Ukrainian, and Italian. At first glance, Russian and Ukrainian seemed to show a correlation between EO and English-like forms, i.e., the less a speaker reported using their HL, the more their VOTs resembled English VOT. However, this correlation was interpreted differently based on the high correlation of EO scores and generations, with Generation 1 speakers having higher EO scores than Generation 2 speakers, and Generation 2 speakers having higher EO scores than Generation 3 speakers. In fact, the authors state that "where there is socially-demarcated variation (by generation) in a language's VOT scores, we find that these values correlate to EOQ, and where there is no such social-marking, we find no correlation to EOQ" (Nagy and Kochetov, 2013, 33-34). Given that: a) Generation and EOQ scores are related, and b) Russian and Ukrainian's behaviors are different than Italian's, Nagy (2015) suggested that the increased VOT (more English-like) from one generation to the next in later generations of Russian and Ukrainian speakers could be interpreted as a language-internal change, rather than a contact-induced change.

networks. However, this grouping method revealed less consistent effects than the topic method (Nagy, Chociej, Hoffmann, 2014).

[5] cf. §1.4.1.; §3.2.

Overall, these results may be in some way surprising. First, “the expected relationships between patterns of linguistic variation and social factors relating to attitude and prestige are not upheld” (Nagy, 2018, 441); second, findings do not support the hypothesis of community-dependent changes based on contact with the Homeland and/or integration with the city – in this case, Toronto (Nagy, 2018). However, we could also look at these results under a positive light. Against much of the literature on HLs, we find support neither for the idea that minority languages tend towards simplification when in contact to majority languages, nor for transfer effects appearing in HLs (Nagy, 2017b). In fact, contact-induced changes would be supported by different patterns of use and/or attitude correlated to different rate of linguistic patterns (Nagy, 2018), but this does not seem the case.

An interesting study on ethnicity in Heritage Italian speakers is Di Salvo (2017), a research that is part of the larger ethnographic investigation of Turchetta and Vedovelli (2018). The study investigates two groups of first-generation Italian immigrants in Toronto – 20 speakers in total. The first group came to Canada between the end of World War II and the mid-1960s, and the second group during the 2000s. The author tests whether migration wave can be considered a sociolinguistic variable by comparing code-switching dynamics (from Italian to Canadian English) between the two groups, using speech samples from qualitative interviews in Italian.

Di Salvo analyzes the different underlying meanings, intentions, and messages conveyed by code-switching, and concludes that the two groups have different "cultural, social, and symbolic capitals" (Di Salvo, 2017, 91). Indeed, the first group is made of speakers who emigrated without English competence and in a period where Italians were still perceived as war enemies, resulting in disadvantaged/subordinate/marginalized positions. These speakers use English to convey their social progress and integration, and they interrupted the transmission of Italian to their descendants to encourage them to learn English instead, the symbol of opportunities. The second group is made of speakers who emigrated with qualifications already, a higher education, and English competence, and at a time when Italians weren't "enemies" anymore, but instead resources thanks to the "Made in Italy" brand success of the last years. These speakers integrated more easily, and do not need to prove their English competence, because it comes naturally with the position they occupy in the social fabric. Their use of English code-switching can convey the idea of English as "harsh" or "assertive" (Di Salvo, 2017, 85-86). Since Italian is an asset for these speakers, they have transmitted the language to their descendants, or intend to do so.

The conclusion is that the two groups show different connections to their ethnicity and different ways to live it and use it to build their identity. Nevertheless, the author stresses that the true sociolinguistic variable is not the migration wave itself,

> but rather the different symbolic and cultural capital of the migrants along with the changed conditions (political, cultural, linguistic, and economic) of the country of immigration. [...] The two groups are different not only in the number of years spent in the host country but also in a wider range of factors whose sociolinguistic relevance has been demonstrated (Di Salvo, 2017, 92).

Dressler (2010), instead, focuses on self-identification as a HL learner among six students of German at the post-secondary level in Western Canada. The definition used of HL learner is that of a student who has "a parent or grandparent who speaks/spoke German or ha[s] spent a significant part of their childhood in a country where German is spoken" (Dressler, 2010, 1). In the research some interesting aspects related to ethnicity emerge. For example, one student whose parents experienced German loss due to societal pressure describes herself as German Canadian by appealing to some cultural artifacts (Bartlett, 2007), such as "German work ethic", and so does another student who refers to German food and traditions. They both self-identify as HL learners. On the other hand, a student born in Romania and who spent 2 years in Germany before moving to Canada at the age of 10 refuses the definition. She does so not only because she does not think of herself as German, but also because she thinks that Germans would consider her as a foreigner/outsider – and therefore, due to an assumed exterior perception, rather than internal feelings. The author points out that the judgements the student refers to could be the product of a "a language ideology in Germany that promotes assimilation" (Dressler, 2010, 10), affecting immigrants and minority communities.

Guardado (2002) uses semi-structured interviews with parents to investigate Spanish maintenance and loss among children of Hispanic families in Canada. The author selected two families with at least one child over the age of six, fluent in English but with a deficiency in Spanish and/or reluctant to speak it, and two families where the child is fluent in both languages. It emerges from the study that the children who maintained Spanish had a strong Spanish identity and developed an emotional attachment to the Hispanic culture and family roots through their parents.

Guardado (2014) employs data from a 1.5-year ethnographic study, examining the discourses produced by 15 Hispanic families living in Metro Vancouver. Compared to other Critical Discourse Analyses, the author does not have the intention of analyzing negative-

oriented discourses, but rather of highlighting "the ways in which linguistic minorities discursively construct the continuation of their languages in the new generations" (Guardado, 2014, 3). The analysis reveals the recurrence of a variety of discourses regarding HL development, including the ones that the author defines Cohesiveness, Identity, Affect and Opposition. Cohesiveness is a discourse on the idea of HL development as a crucial way of promoting a sense of unity and understanding, both within in the family and in the ethnic community. Identity is a discourse about HLs as building blocks during the construction of one's identity (since they are linked to ethnic roots), and as tools to develop a healthy and proud sense of self. Affect is a discourse implying the HL as an emotional attachment or means to express emotion. Opposition is a discourse regarding the idea of HLs as political acts of resistance against assimilating forces. Furthermore, the author states that, out of 24 publications on HL development, 21 of them featured some version of the discourse Cohesiveness.

Abdi (2011) uses data from an ethnographic study on identity and positioning among 21 students in Spanish classes of a Canadian high school, including some Heritage Spanish students. The author stresses out how language ideology affected the perception of one of the students' ethnicity and identity. Indeed, this student was reluctant to speak Spanish in the classroom despite having Hispanic roots, living in a bilingual household, experiencing Hispanic culture, and being excellent in her literacy abilities in Spanish. As a consequence, the other students and the teacher did not recognize her as Hispanic and/or acknowledge her cultural heritage.

However, the ethnic and social aspects for some scholars are not enough to accurately describe HLs. If we considered just these parameters, as Kupisch (2013) points out, HLs would be no different than minority languages themselves, and their speakers undistinguished from other bilinguals. Aspects related to proficiency (§1.3.2.) and modalities of acquisition (§1.4.) have thus been taken into account by researchers interested in connoting HL speakers differently than minority language speakers – and therefore, who propose a narrower definition of HL speakers.

1.3. Bilingualism and proficiency in Heritage Languages

1.3.1. HL speakers as bilinguals

HL speakers have also been referred to as early bilinguals (acquisition of the two languages before puberty, Montrul 2013b), either simultaneous (acquisition of both languages starting at

birth) or sequential (acquisition of L2 after 3-4 years[6]). In the case of simultaneous bilinguals, Meisel (2011) explains that they have two L1s. In the case of sequential bilinguals, it is crucial for the identification of HL speakers that the HL has been the first in order of acquisition (Benmamoun, Montrul, and Polinsky, 2013). Note that the inclusion of bilingualism as a criterion of classification of HL speakers is not an agreed upon aspect. As we have seen in §1.1., the Canadian government, for instance, does not refer to the coexistence of more than one variety within a HL speaker. In many research or projects, monolingual individuals have been considered HL speakers based only on the fact that they have a mother tongue different than the official languages of the country they live in. An example is the HLVC project (Nagy, 2009; 2011; cf. §1.2., §3.3.), which studies variation and change within the HL communities of the Great Toronto Area based on a multi-lingual corpus of natural speech. Some of the first-generation immigrants recorded do not have a L2 (and, as we will see later, consequently have not undergone dominance shift).

In general, the theme of bilingualism brings about the need to better define terminology (some of which we have already used) and to clarify a few of its aspects affecting HL speakers' communities, if the goals is to provide a definition of HL involving bilingualism too.

Following Benmamoun, Polinsky and Montrul (2013), when talking about bilingualism we can distinguish:

a) First (L1) and second (L2) languages, based on the order of acquisition.
b) Primary and secondary languages, based on the rates of usage. The language which an individual learns first is not necessarily the one they use the most. This dichotomy implies unbalanced situations, where languages have different domains and contexts of use based on individual choices. However, these could be influenced in turn by the choices from the community, and by broader socio-political settings. Therefore, this dichotomy is potentially connected with the following one.
c) Majority and minority languages, based on social and political factors as described in §1.1. and §1.2.

Using these concepts, now it should be easier to give a more specific description of the HL speaker under the "bilingual" approach:

[6] However, there is no agreement about this age cutoff. Classifications of bilinguals depending on the age of acquisition of the L2 will be described in §1.4.2.

> A heritage speaker is an early bilingual who grew up hearing and speaking the heritage language (L1) and the majority language (L2) either simultaneously or sequentially in early childhood (that is, roughly up to age 5…), but whose L2 became their primary language at some point during childhood (typically after the onset of schooling). (Benmamoun, Polinsky and Montrul, 2013, 6-7)

The change in terms of dominance described above might affect language development, causing a decrease in confidence in the HL and thus an even greater decrease in usage. This can, in turn, lead to slower speech rates, or affect grammatical structures (Aalberse, Backus and Muysken, 2018).

However, this quote has two ambiguous elements. On the one hand, as Kupisch (2013) points out, if simultaneously bilinguals are considered, the definition should mention two L1s, and not an L1 and L2. If we consider simultaneous bilingualism and use the proper formulation of two L1s, the shift in terms of usage has to be described not in terms of order of acquisition, but of the majority/minority dichotomy: the L1 which becomes the primary language is also the dominant variety. On the other hand, even if bilingualism was to be included as a diagnostic for HL speakers, this definition assumes that dominance shift is crucial to gain the HL speaker status, and this is something some researchers do not agree upon (Nagy, 2015; Kupisch and Rothman, 2018).

1.3.2. Assessing HL speakers' proficiency

Seeing HL speakers as bilinguals has moved the debate towards the description of their proficiency level, its measurement and the factors determining it. As a preliminary note, Polinksy (2018) states that HL speakers usually display more mastery of the majority language rather than the HL; nevertheless, HL speakers are distinctively characterized by the fact that each of them shows different rates of overall proficiency, and that inside the HL system their proficiencies in different activities may vary (for example, having great aural skills, while lacking in writing abilities) (Montrul, 2013a). The fact that *every* HL speaker has a poorer proficiency in their HL (a perspective which Nagy (2005) refers to as the "deficiency perspective") should not be taken for granted, since there is so much variability.

Research on the levels of HL proficiency has usually employed concepts and methods used in L2 proficiency assessment –a symptom of the bias surrounding the perception of what HL are, i.e., non-native languages. Now, there are different approaches to what we measure when we measure L2 proficiency, more on the formal side or more on the social side. Each of these,

indeed, implies a different view of what we mean by "language". Let us briefly summarize them (Butler, 2013).

First, there are formal linguistic approaches to proficiency: here, evaluating L2 proficiency means measuring the Chomskyan (1965) competence of the speaker, the I-language (Chomsky, 1986), the knowledge of abstract grammatical rules and forms. The main method used to do so is to compare the grammar of L2 learners to the grammar of monolingual native speakers, using these as a baseline. This view has some criticisms: first, the so-called "monolingual bias", i.e., the idea that bilinguals are essentially two monolingual people in one and that they can be put in the same spectrum as monolinguals. Researchers like Grosjean (2010) have claimed that monolinguals and bilinguals are qualitatively different, and thus, that using monolinguals as a guideline to measure bilinguals' proficiency makes no sense. As mentioned earlier, this is the main reason why other bilingual groups may be used as a baseline to assess HL speakers' proficiency. Second, it has been largely acknowledged that language abilities are to a greater extent context dependent. This has led even formal linguists such as Liceras (2010) to reconsider the role played by E-language (Chomsky, 1986) – or performance – in the process of mastering an L2. These two points are summarized in Kupisch and Rothman's (2018) position regarding HLs: as Grosjean said, there is no sense in expecting HL speakers to sound like monolinguals, given the input they receive, their frequent lack of formal education in the HL and the pressure exerted by the majority language (Rothman, 2009). The basic social conditions that benefit monolingual natives simply are not there to make a comparison.

The other two main approaches to proficiency both imply a different view of linguistic competence. While the Chomskyan tradition only included grammatical knowledge in competence, and claimed that the contextual use of language was part of the performance, Hymes' (1972) idea of competence is that of a *communicative* competence, including grammatical and sociolinguistic knowledge. From these grounds, the individual-cognitive approach and the socio-contextual approach to proficiency developed.

The individual-cognitive approach stems from the idea of language acquisition as a process not innate and specialized, but rather data-driven and part of general cognitive mechanisms. The focus is not only on rules and structures, but also, for example, on vocabulary, comprehension, and metalinguistic abilities. However, L2 speakers' proficiency has been measured under this agenda in the same way as in formal linguistics: by comparing it to the "target linguistic performance" (Butler 2013, 119), i.e., that of monolingual native speakers. This brings the monolingual bias and the afore-mentioned critiques back up.

The second approach mentioned, the socio-contextual one, sees language as socially constructed, stressing its relational, communicative aspects, focusing on the social factors correlating to variation, and on language use. This is why in this approach measuring proficiency means measuring also sociolinguistic and pragmatic knowledge and their implementation. It stemmed in particular from Firth and Wagner (1997)'s article, and has made researchers reconsider the notions of native/nonnative and interlanguage. However, the main issues here are: how to account for all this variation in Second Language Acquisition? How to quantitatively test for these types of skills? How much of variation is systematic?

In the context of HL studies, note that it is implied, by using L2 proficiency assessment methods, that HL are L2s and/or behave as such, and by using the grammar of monolingual natives as a baseline for proficiency, that a native-like attainment in the HL has not been reached by HL speakers. This is a leitmotiv for the scholars following these frameworks, as we will see in §1.4. Kupisch and Rothman (2018) point out that this aspect is problematic. Here we see another proof of the fact that the concept of nativeness should be discussed, re-thought and agreed upon by the scientific community.

A final, fascinating, approach to proficiency that tries to reconcile the individual-cognitive one and the socio-contextual one is that of Larsen-Freeman (2007). The author uses ideas from ecology, chaos and complexity theory, exemplar theory (Goldinger, 1997) and usage-based positions (e.g., Bybee 2006). She describes language as a dynamic system, a "network of […] language-using patterns" (2007, 783-784), constantly reshaped depending on the environment in different spaces and times. Both social and cognitive factors influence the way Second Language Acquisition works. The result is that there is no endpoint: language acquisition never ends, and proficiency has no final goal to reach, like nativeness. As we will see in §1.4., some HLs researchers – e.g., Montrul - oppose this kind of view, rather supporting the concept of an acquisition process that begins and ends.

However, not all research has treated HLs as the same as L2s. Studies specifically focusing on HLs speakers' proficiency levels have been conducted and then reported, for instance, by Polinsky and Kagan (2007). It seems that scholars in this field all implicitly agree that the ability in mastering the HL can be modelled into a continuum. This is a point where the study of HLs meets creole studies, according to Polinsky and Kagan (2007). In creolistics, given the lexifier, there is a range of varieties, lying on a continuum, which are closer to the lexifier (acrolets) or farther from it (basilects) (Bickerton, 1975). Variation is analyzed by measuring the distance of a creole variety from the lexifier. Therefore, HLs variation could be analyzed by measuring the distance of the HL from the chosen baseline. The logical consequence is that the more one's

Heritage variety resembles the baseline, the more proficient the speaker is. Among the ways to measure the distance from the baseline, we find speech rate[7] (Polinsky and Kagan, 2007) and lexical knowledge[8], defined by Benmamoun, Polinsky and Montrul (2013) linguistic diagnostics for proficiency. However, there are also what they define biographical diagnostics, i.e., the ones that would be considered within the socio-contextual approaches to proficiency. Examples are the manner and length of exposure to the baseline (cf. §1.4.), or caretakers' attitudes towards the HL. In particular, for "manner" of exposure we mean the contexts and registers involving the use of the HL (connected with the political and social status of the language – cf. Montrul, 2010), and literacy in the HL.

Nevertheless, according to the more formal and cognitive views, adding bilingualism to the definition of a HL speaker is still not enough. Some scholars not only want to distinguish HL speakers from minority speakers (cf. §1.2.), but they also want to highlight differences with monolinguals and with complete and balanced bilinguals, further pursuing their idea of "non-native-like" HL speakers and identifying them with a subgroup of bilinguals. This is the point where Experimental approaches claim acquisition plays a role, and their definitions of HL speakers start to be narrower.

1.4. Acquisition in Heritage Languages

The discussion around acquisition stemmed from the consideration that HL speakers show "less mastery" of the HL, which thus becomes the weaker of the two languages available to the bilingual HL speaker (Benmamoun, Polinsky and Montrul, 2013). Taking this as valid, the motivation behind incomplete proficiency in the HL has been individuated by the formal and individual-cognitive linguists in acquisition dynamics. Thus, incomplete acquisition became, for this group of scholars, the criterion to distinguish HL speakers from monolingual speakers: Benmamoun, Montrul, and Polinsky (2013) use the under-development of the HL at age-appropriate levels as a diagnostic to identify them.

Of course, scholars from different approaches do not deny the role played by acquisition, especially given the minority-language status of HLs; but the difference from formalists lays in

[7] For a discussion about the cross-linguistic validity of this using speech rate as a metrics, see Roach (1998) and Pellegrino et al. (2011). Moreover, Polinsky (2012) reported the need to further examine the validity of this measure; see Nagy and Brook (2020) for the consequent empirical evaluation.

[8] As claimed by the individual-cognitive approach to proficiency. However, lexical knowledge has been demonstrated to be related with grammatical knowledge (Polinsky, 1997; 2000; 2006), providing a connection with the more formal approaches too.

refuting determinism and negative evaluations of the acquisition outcome. Generally, within a formalist framework, if you are a HL speaker, then you are not native-like in the way you acquired the HL, the way you have acquired it lacks something, and this incompleteness reflects in your grammar (interpreted as an implicit set of rules and structures). Within sociolinguistic frameworks, if you are a HL speaker, perhaps you have experienced a different setting for acquisition, but its outcome is not necessarily incomplete; it may or may not differ from the outcomes of a "baseline" acquisition process, but it is not defined as "lacking". Moreover, the focus of these types of research is on patterns or trends in speakers' linguistic behavior, rather than on their language acquisition (Nagy, 2015). Sometimes it cannot be determined whether a divergent behavior from a baseline is motivated by acquisition or not. Possible differences in the outcome of acquisition are interpreted not only (or mostly) as the result of cognitive, individual, linguistic mechanisms of re-elaboration or re-structuring of the input, but also as the result of social and/or ethnic factors, or of voluntary identity-building choices. In other words, while formalists have a more endogenous approach, sociolinguists tend towards a more exogenous approach (Willis, 2017)[9].

1.4.1. Sociolinguistic factors influencing HL acquisition

The factors potentially affecting HL acquisition, from a more sociolinguistic perspective, are summarized by Aalberse, Backus and Muysken (2018). They consist in the attitude of society towards bilingualism; the social network the language is acquired in; the family and educational situation the language is acquired in. Let us describe them.

First, societies can have different ways of seeing bilingualism. Lambert (1975) proposes a simple classification in additive versus subtractive bilingualism. While with the former society has an attitude that conceives the acquisition of another language as something enriching and worth doing, with the latter society disregards or devalues it, so that minorities feel pressured to put aside their minority language in favor of the majority language. This negative attitude may even bring about the complete replacement of the minority language with the majority one.

Second, HL acquisition is likely to be influenced by the social network the speaker is immersed in, using the Milroys'[10] meaning of social network. A social network is the sum of all the contacts a speaker has with other speakers. It is characterized primarily by denseness (the extent to which its members are in turn connected among them) and multiplexity (the

[9] Where "endogenous" here is used as language-internal, and not as syntax-internal.
[10] Milroy, L. 1980; Milroy, L. and Milroy, J. 1992.

overall quantity of relations inside the network), but Berruto (1995) adds the frequency and duration of the interactions between pairs of members, and the degree of centrality of the ego – the main speaker analyzed. The Milroys' works in Great Britain have demonstrated that social networks are great indicators of linguistic attitudes and behaviors, in particular in patterns of maintenance and/or innovation. Berruto (1995) asserts that social networks are possibly the only social factor explaining why people of the same age, gender, class, work status, and education level have different linguistic behaviors. Moreover, he also supports the idea that networks influence language acquisition, and even claims that the choice of the language to speak in multilingual settings depends on social networks too.

Third, elements like how much HL families visit their home country, the background of a HL speaker's caretakers, the presence or absence of siblings in the family, the possibility and will to educate HL children in their HL, have also been proposed as factors affecting acquisition. However, many of these are collinear factors, i.e., it is hard to tease them apart in a study to understand which one(s) really affect HL acquisition. Furthermore, they could influence the chronology of acquisition, determining if a HL speaker is a simultaneous or sequential bilingual (cf. §1.3.1.).

These three factors together may account for the "insufficient input" Polinsky (2013, 2) addresses as the cause behind restricted HL and therefore incomplete acquisition: they are all aspects related to the quality and quantity of linguistic input and output.

However, the typical factor affecting HL acquisition considered in the formal and individual-cognitive approach is the age of onset of bilingualism.

1.4.2. The age factor in HL acquisition

The general assumption when talking about language acquisition is "the earlier, the better" (Flege, 2007). The age at which a speaker started acquiring the language has been extensively used in literature as a criterion for the classification of bilinguals, and thus, of Heritage speakers according to some definitions.

First, we have a classification as early vs. late bilinguals. Early bilinguals learnt the L2 early in life (i.e., during childhood), late bilinguals late in life (i.e., after childhood). The age cutoff is frequently discussed; the age span set for "childhood" is not consistent. Heredia and Cieślicka (2014) report a series of periods proposed in literature; examples are the ranges 0-6, 0-13, or 0-16. As we have seen, Montrul (2013b) uses "puberty" as a more general threshold.

Then, another classification is as simultaneous vs. sequential bilinguals. Simultaneous bilinguals started learning the two languages at the same time within a strict time period, that Paradis (2001, 2010) sets to 4 years of age. Most commonly, however, they are described as individuals who acquired two languages simultaneously from birth, and that have two L1s. Sequential bilinguals are characterized by the fact that their L1 is established before the learning of L2 starts (Heredia and Cieślicka, 2014). The L2 acquisition process, nevertheless, has to start fairly early: for some after 3 or 4 years of age, or before school entry (Paradis, 2010), for others between 4 and 7 (Montrul and Potowski, 2007).

Meisel claimed that simultaneous bilinguals, child learners, and adult learners show qualitatively different trajectories of acquisition and attainments. Similarly, we have Schwartz's (2003) model of the domain-by-age, according to which each linguistic domain has different learning routes based on the age of onset.

The basic assumption which pushed researchers to take in such great consideration age, is the presence of maturational constraints in language acquisition, i.e., innate, biological mechanisms available at their full power just for a limited amount of time. The most popular hypotheses on the topic are the Critical Period or Critical Age Hypotheses (Birdsong, 1999; Lenneberg, 1967). We will discuss this position in §1.4.4. While it is fairly agreed upon that there are age effects in L2 and HL acquisition, there is no agreement on the existence of a critical period and its onset and offset, nor on whether age affects acquisition in a domain-specific way. Similarly, age-based constraints are not thought of in the same way by every researcher. While some authors conceive them as strong, inevitable, and biology-dependent limits, others conceive them as probabilistic limits to acquisition (Birdsong, 2005). Moreover, considering age as the most important variable when studying HL acquisition puts linguists in front of an important methodological challenge: the linear dependency among the age of onset, the length of exposure, and the chronological age of testing (Butler, 2013, 124).

As a matter of fact, research about the consequences of age of onset show mixed results. Kupisch (2018) analyzes several studies and concludes that there is little evidence supporting the age of onset hypothesis. Recent phonology works suggest that early age acquisition and simultaneity in bilingualism are "a necessary although not a sufficient requirement for nativelike or ultimate attainment" (Abrahamsson and Hyltenstam, 2009, 290). The most reasonable conclusion we can draw is that, if age played a role in acquisition, it would be a probabilistic one, and that age would be intertwined with collinear factors (Aalberse, Backus and Muysken, 2018).

1.4.3. HL acquisition trajectories: are they "incomplete"?

Let us consider a HL speaker who underwent dominance shift (exposed to two languages during the first years, then dominant in just one of them while still being in a partially bilingual environment), and let us keep in mind that some HL researchers consider HL speakers' acquisition incomplete. Put it this way, it is easy to understand why scholars debate whether it would be better to approximate HL speakers' learning trajectories to the ones of early L1 learners or adult L2 learners. Lleó (2017, 2) wonders "whether the acquisition of the HL will develop monolingual-like, as a second language, or as something else, and why". This issue is faced by Montrul (2010), who first describes the characteristics of L1 and L2 acquisition and then tries to place HL acquisition in the picture.

According to Montrul, the result is that HL acquisition has an intermediate status. It shares with L1 acquisition the early exposure, the richness of ecological input, and the early achievement of control over some language features (phonology and certain structures). It shares with L2 acquisition (according to formal approaches) developmental and transfer errors than persist even after correction and are typically fossilized, a variably proficient output, the crucial importance of motivation and affective factors, and, at least as Montrul states, the fact of being "*typically incomplete*" (Montrul 2010, 12).

However, the discussion on HL speakers' incomplete acquisition is very complex, and the idea has been (and is) more and more questioned. Not everyone agrees that HL speakers should be identified by incomplete acquisition, because not everyone sees their grammars as incomplete. In fact, we can claim that HL systems are consistent and organized, despite some of them being different from the baseline systems. Using the term "incomplete acquisition" risks sending a different message, implying that HL grammars (and therefore, HL speakers) have something missing, and that this potentially undermines the functioning of their language. Perhaps, one should not label as "incomplete" a variety that could simply be different from what we already know and/or is already established.

This is defined by Montrul (2013b) as a sociolinguistic view: scholars like Carreira and Potowski (2011), indeed, do not see a type of deficiency in the possible differences from the baseline, and even claim that HL speakers, in the end, simply speak a different regional variety. This is also coherent with Larsen-Freeman's (2007, cf. §1.3.2.) idea that language learning is a never-ending process. Kupisch (2013) also seems against including an end state such as "incompletely acquired" in the definition of HL speakers. Later, in Kupisch and Rothman (2018), she argues that "different" and "incomplete" should not be treated as synonyms; what

is mistaken for "incompleteness" in adults may be attributable to a difference in the quality of the input and to inaccessibility to formal education and/or literacy. Finally, Nagy (2011) also rejects the idea of treating HL speakers as individuals who feature a kind of linguistic deficit.

Nonetheless, Montrul (2013b) does not agree with the sociolinguistic position. She defends the concept of incomplete acquisition, asserting that acquisition must still be seen as a process having a beginning and an ending; the consequence is that, according to her, one can still talk about incomplete development of a language in relation to the endpoint. In particular, her objection towards defining HLs regional varieties consists of two main arguments. On the one hand, she mentions the presence of clear proficiency effects: "[T]hat is, heritage speakers with the lowest levels of proficiency in the language are the ones who have reduced vocabularies, basic word order, and make morphosyntactic errors with case, gender agreement and other morphology" (Montrul, 2013b, 178). On the other hand, varieties are used by groups of people to communicate, and their shape is modelled by this use, eventually converging on some linguistic patterns. Most HL speakers, according to Montrul, do not behave this way: as we also illustrated in §1.2., she argues that most of them use the HL only at home, and that HL children speakers use the majority language with peers and even siblings.

Since the idea and terminology of "incomplete acquisition" is discussed, authors proposed alternative terms. Kupisch and Rothman (2018) suggest "differential acquisition"; Polinsky (2018) "divergent attainment[11]". By using a different terminology, these researchers recognize that what has always been considered imperfect proficiency and/or attainment could simply be a deviation from the baseline, a diverse mechanism of construction of a grammar which allows for differences across individuals.

1.4.4. Bilingual models of acquisition: where HL grammars stand

HL grammars offer an occasion to test the different hypotheses about linguistic - and in particular, bilingual - acquisition in a unique sociolinguistic setting. The possible dominance-shift, i.e., the dynamic of switching from the L1 as primary and the L2 as secondary language to the L1 as secondary and the L2 as primary language raises basic linguistic questions: not only how bilingual acquisition works, but also how it is influenced by age, brain plasticity, previous knowledge, and external factors, such as context, amounts of input, and ethnic orientation.

[11] We will see how divergent attainment works in §1.6.1.

In this section, a brief description of the main bilingual models of acquisition will be provided to understand what the possible lines of research on HLs are.

Montrul (2010) summarizes the three leading families of theories on bilingual acquisition. They approximately correspond to the approaches to proficiency, and therefore language, illustrated in §1.3.2. Moreover, the first two might roughly correspond to the Experimental Approach mentioned by Nagy (2015), while the third to the Comparative Variationist approach. Let us report them:

1. Formal linguistic perspective. This view derives from the Chomskyan tradition and takes as valid the notion of Universal Grammar (UG) (Chomsky, 1986). Generative linguists hypothesize that acquisition is innate and governed by a specific learning device. One important claim is that the access to this device and to UG is subjected to a time limit. In other words, these linguists embrace the ideas of CPH or CAH, Critical Period Hypothesis or Critical Age Hypothesis (Birdsong, 1999; Lenneberg, 1967; cf. 1.3.). To simplify, the concept is that there is a window of time during which, if exposed to adequate and rich stimuli, a child naturally acquires native-like language without effort; if exposure begins after this timespan, instead, the acquisition will be incomplete. The main biological explanation usually presented is a loss of neuroplasticity; the generative linguistic explanation is the loss of access to the innate principles building the Language Acquisition Device (Chomsky, 1969).

   There are critiques to the hypothesis of a critical period for acquisition, such as Flege's (2007) or Hakuta's (2001). For example, this putative critical period still lacks an agreed upon and defined onset and offset. Supporters of the hypothesis should also provide evidence of qualitative differences in learning inside vs. outside the critical period, with a visible discontinuity in ability in correspondence of the putative critical period (Flege, 1995). Finally, they should provide evidence of resistance to environmental variation when the speaker is within the critical period.
2. Cognitive and neurolinguistic perspective. This view, differently from the previous one, does not see acquisition as an innate mechanism, but rather a part of other general cognitive processes. As a consequence, it also refuses the concept of UG. It distinguishes implicit and explicit learning and knowledge (Paradis, 2004; 2009): implicit learning is carried out unconsciously, and implicit knowledge can be accessed and executed quickly and automatically. Explicit learning is voluntary and reached with effort; since it is conscious, explicit knowledge can be verbalized. Paradis' (2004)

position regarding second language learning is that, while children usually use implicit knowledge when learning their native language, L2 learners resort to explicit knowledge, developing an implicit knowledge of the L2 only slowly and over time. Therefore, this approach still echoes the formal generative one, in terms of timing. Some scholars, indeed, support the idea that an adult L2 learner has to use a different cognitive system because the original, implicit, cognitive mechanisms for language gets lost during childhood. In other words, as for Bley-Vroman UG is accessible only during a limited time span, after which the speaker has to use general strategies, for DeKeyser (2000, 2003) implicit learning and knowledge are available only during childhood, and after that the speaker has to use explicit learning.

Flege (2007) is critical of this cognitive position, as he is of the "critical period", suggesting care when talking about any sort of maturational constraint, and providing counterexamples, such as close similarities between speakers who started learning the L2 after childhood and native speakers. This contrasts with the idea that early exposure is a sufficient condition for the access to UG, or to implicit learning.

3. Emergentist perspective. This view sees language as an emergent structure that gradually develops from how general cognitive abilities and the environment interact with linguistic and communicative input. Acquisition is carried out by extracting statistical regularities and patterns and by generalizing from the input. One key question is what the precise appearance of this structure is: for some scholars it is a network of constructions (Goldberg, 1999), for others memorized processing routines (O'Grady, 2005), yet for others local chunks of memorized associations (Ellis, 2009). The other fundamental issue on this topic is: if the language learning mechanism is so general and shared among individuals of all ages, then why do early L2 learner and post-puberty L2 learner outcomes differ? A hypothesis is to correlate these divergences to different types of inputs and to the interaction with the L1 system, like Flege (2007) argues; another one supposes differences between children and adults not in the mechanism itself, but in the parsing of the input and the units of the input the mechanism applies on. Aslin (2017), for example, reports four constraints on statistical learning that we could imagine being relevant in an emergentist model of L2 acquisition too: attention, perceptual biases, prosody, and familiarity of stimulus.

In conclusion, variable input, and possible interaction between the HL and the majority language system seem to be well-established elements of HLs. Moreover, the fact that some

HL speakers switch from the L1 being the primary language to the L1 being the secondary language, also showing the effects of attrition (cf. §1.6.1.) later in life, means that we need an acquisition theory that is dynamic and accounts for the re-shaping of already established knowledge. This pushes us to consider Flege's critiques as reasonable. The elaboration of a comprehensive and global model of HL acquisition, however, is outside the scope of this research. Therefore, we will narrow it down and try to discuss a viable theory of sound acquisition only, since this work focuses on sound production. In the following section, given our agreement with his critiques, we will consider Flege's model of bilingual sound acquisition in the context of HLs.

1.5. The sound system of Heritage Languages

This section will be dedicated to the ways HL speakers sound, perceive sound, and organize and represent sound knowledge. From now on, when we consider both the phonetic and phonological systems of a HL grammar, we will us the expression "sound system". We will use the specific term when treating just one of the two. §1.5.1. will illustrate Flege's model of bilingual sound acquisition to see if it is suitable for HL sound systems. §1.5.2 will explore the theme of native-like accent: what having a native-like sound system means, what the favoring causes for native-like accent are, how much HL speakers are perceived as sounding native-like.

1.5.1. How the system forms: Flege's SML model for HLs

Flege focuses on speech acquisition, so on the way sounds are perceived, organized, and produced by sequential bilinguals. His model, initially proposed in 1995, is the Speech Learning Model (SLM). The main postulate is that the mechanisms behind language learning are not subjected to age-limits: they are able to work throughout the life span and for L2s too. The researcher, as mentioned before, is critical of positions centered around the idea of age and maturation, because he believes that age effects *correlate* with age, but are not *caused* by age. Actually, age effects could be caused by factors like the Age of Arrival of the L2 learner in a foreign land, or the amount and quality of L2 input.

> Early but not late bilinguals tend to inhabit an L2-rich environment. Early bilinguals are usually enrolled in schools in which the L2 is used as the sole language of instruction soon after their arrival. [...] Late bilinguals, on the other hand, tend to receive far less education in L2-speaking schools [...] and they tend to maintain the L1 as their dominant language. This means that AOA

> [Age of Arrival] tends to be positively correlated with amount of L1 use and negatively correlated with amount of L2 use. A number of studies have shown that a frequent continued use of the L1 is associated with relatively poor performance in the L2. [...] This demonstrates that language use exerts an important influence on L2 speech learning. (Flege, 2007, 13)

The SLM model stems from the assumption that the phonetic system continues to adapt during the course of life: it originally organizes itself based on L1 sounds, but when it encounters L2 sounds, it changes its shape by adding new categories or modifying the existing ones. The phonetic system presents phonetic categories, which are representations inside the long-term memory; they correspond to the language-specific aspects of speech sounds. Therefore, one strong postulate of the model is that for bilinguals there is a unique phonological space in which phonetic categories arrange themselves, but, when possible, the maximum contrast between L1 and L2 categories is strived for. This is one of the ways in which Flege diverges from the formal tradition: there are not L1 and L2 sound systems, but there is a unique sound system in which L1 and L2 categories eventually develop. The fact that the capacity of creating new phonetic categories stays intact does not mean that it is used for each new sound encountered: the interaction with the previous knowledge – and the already established categories – prevents this (Flege, 2007).

L1 phonetic categories, if one continues to use the L1, keep developing. While they do so, they become stronger representations capable of perceptually assimilating L2 sounds (although Flege himself underlines the need to test this hypothesis). An L2 sound judged by a learner as an instance of one of the sounds from an L1 category is "equated" with the L1 sound. It is this mechanism of equivalence that likely stops the speaker from creating new categories, and it is called category assimilation. However, judgments about L2 sounds may change during the course of life, and a difference between two sounds, originally equated, may eventually be perceived. Modification is in this case expected: the representation adapts by becoming a merged category of features coming from the input in both the L1 and the L2. "Depending on the nature of the input that has been received over the course of a bilingual's lifetime, the merged category may resemble more closely the long-term representation of either L1 monolinguals or L2 monolinguals" (Flege, 2007, 6).

On the contrary, an L2 sound judged by a learner as distant from the closest L1 sound is likely to be responsible for the creation of a new category. This mechanism is called category dissimilation. However, the process of creating new categories can lead to an “overcrowding” of the phonetic space as compared to monolinguals. Therefore, these categories tend to disperse to maintain phonetic contrast. This means that, in order to distance themselves, L1 and L2

categories may occupy different positions in the phonetic space compared to the ones occupied in a monolingual speaker. The result is that both the L1 and the L2 categories will be different from the ones possessed by monolinguals. Nevertheless, Flege questions whether the creation of a new L2 category must always be a process of category dissimilation: if an L2 sound is at a really great distance from the closest L1 category, is there the need to postulate the dispersion of the categories due to contrast-searching? Can an L2 category be formed without influencing the L1 closest category?

It must also be highlighted that both these mechanisms – category assimilation and dissimilation - are probabilistic.

This model is fascinating under many aspects. It considers the possibility of mutual influence and transfer during acquisition. It is usage-based. It gives credit to the speaker's perception of speech. It accounts for the importance of input and context. Nonetheless, we will see that, according to Polinsky (2018), HL grammars also feature universal principles of language design (cf. §1.6.1.). These do not depend on previous linguistic knowledge, but are general trajectories probably caused by motor or cognitive constraints – such as the simplification of articulatory gestures that results in sound reduction patterns. How would Flege's model account for these? Would they be considered as external to the acquisition process, as a subsequent phase of variation and change through use? Or could they simply coexist with the mechanisms described by Flege, as other forces playing a role in the way languages are acquired?

Furthermore, evidence about the "strength" of phonetic categories is needed. Flege hypothesizes that L1 categories, the first that have been established in a sequential setting, become stronger over time due to continuous use, and that as a consequence they become the attractor for L2 sounds. How would this mechanism work for the HL speakers, sequential bilinguals, who underwent dominance shift? Their HL categories are the first to be established, of course, but does their "strengthening" process interrupt when the HL becomes the secondary language? Does the attractor become the L2, established later but now primary language? If so, how does said "switch" work?

Finally, would there be a way for the model to fit a situation of simultaneous bilingual acquisition? If so, which of the two languages would become the assimilating force?

One position that could potentially be merged with Flege's or enrich it is the "good enough" hypothesis on phonological knowledge by Polinsky (2018). This hypothesis tries to scrutinize how phonological dissimilation works for HLs by appealing to functional reasons. The key idea is that dissimilation operates only when it is useful for the speaker to do so: "Dissimilation […] leads heritage speakers to emphasize the contrasts that are apparent […] and to ignore contrasts

that are weaker, do not play a distinctive role, or allow the two languages to converge without significant information loss" (Polinsky, 2018, 115).

1.5.2. "Native-like" accent: production and perception in Heritage Language speakers

Judgments about how HL speakers sound go in opposite directions: on the one hand, they are often described as sounding "native" – and this, once again, pushes us to ask ourselves: "What does *native* mean?". On the other hand, they are perceived by Homeland speakers as special, in some ways, and pinpointed as not monolingual natives (Polinsky and Kagan, 2007; Polinsky, 2018). Moreover, they sometimes have an accent – while L1 speakers are not expected to have a special one (Lleó, 2017).

This brings us to some questions. If "despite th[e] impressive [phonetic] advantage, heritage speakers do not pattern like native speakers; rather, they consistently get recognized as a separate group" (Polinsky, 2018, 116), maybe:

a) We should clarify assumptions like: "Heritage speakers are typically described as having good phonology" (Montrul, 2010, 6). What do we mean by "good phonology"?
b) What are the factors influencing native-like pronunciation and intonation?
c) Are there some finer-grained phonetic details in the sound system of HL speakers that are absent in the sound system of Homeland speakers? Or are the phonetic variants the same, but constrained/distributed differently between the two varieties? Why do Homeland speakers seem to notice a difference?

Polinsky (2018) states that HL speakers show a smaller range of variation in the production of sounds compared to the variation in morphology or syntax. Her hypothesis is that early language acquisition may positively affect phonetics more than other domains, even when the exposure has been short. In general, HL speakers, as simultaneous or early bilinguals, have advantages over L2 learners in phonology (cf. Montrul, 2012); in the cases of short exposure, HL speakers might have an advantage over L2 learners not in tasks of sound production, but rather of sound-contrast perception. Indeed, Polinsky continues, HL speakers are superior to monolingual natives in tasks of recognition of contrasts even in the dominant language: another sign of the "bilingual advantage".

However, HL speakers' proficiencies in phonetics and phonology can also vary depending on what parts of the sound system we are analyzing. They seem good and quite homogeneous in the production and perception of segmental phonology (Polinsky, 2018, 123; 153-158), while they show much more variability, patterns of change or even loss in the production and perception of stress and intonation (Polinsky, 2018, 147-153; 158-162).

It is generally assumed that, in order to achieve a native accent, the earlier the better, and that a foreign accent can hardly be overcome after puberty (Lenneberg, 1967). Studies have demonstrated that age-effects are clearly present: the main conclusion we can draw is that early bilinguals should have an accent advantage over L2 learners. As we saw for models of general language acquisition, the nature of these age-effects is still questioned. Some explain them as resulting from a loss of neural plasticity, others with input-related issues (having less input or more foreign-accented input). For example, Long (1990) even proposes that phonological acquisition is more constrained than syntax, and that it declines as early as age six. Nevertheless, there is evidence that speakers with an Age of Onset later than six achieve native-like pronunciation and intonation; at the same time, there is also evidence that speakers with an Age of Onset before six have a foreign accent (Kupisch et al., 2015). This goes against the hypothesis of a rigid critical period model of phonological acquisition.

Therefore, it seems that early exposure is not a guarantee of monolingual, native-like accent during adulthood. Even though HL speakers, as we have said, have advantages over L2 speakers in phonology, they do not sound completely native-like, based on the monolingual speaker's perception of native accent.

Moreover, the sound system acquired early in life is not a monolithic, unmodifiable entity. It can change throughout the lifespan (Kupisch et al., 2015). Consequently, one could also question if a native accent can be (partially) lost, after full acquisition, given changed input conditions: this is the case of individuals immigrating to foreign countries, or of HL speakers who gave up their minority language spoken at home as soon as they started school in the majority language. The study from Oh et al. (2003) shows that such a drastic change in the input of L1, combined with a huge increase in the input and use of L2 after age 6, affects the L1 accent. De Leeuw, Schmid and Mennen (2010) provide evidence that adult immigrants, living abroad for an average of 37 years, compared to Homeland monolinguals, have significantly more foreign accent. Nonetheless, judgments about their nativeness were unclear and heterogenous, showing that this L1 sound attrition is not an inevitable process. Therefore, something more than simply age or length of residence in the L2 country must play a role.

A question pertinent to HLs and that Lleó (2018) also implies is whether the sound system of simultaneous balanced bilinguals is different for simultaneous bilinguals who speak one majority language and one minority language. A study that dives into this topic is Kupisch et al. (2015). This research on foreign accent focuses on simultaneous bilinguals, sociolinguistic factors, and majority/minority languages dynamics. It also tests the role of Age of Onset. The researchers compare adult simultaneous bilinguals who acquired a language as the majority one, adult bilinguals who acquired the same language as the minority one, speakers who learnt that language as an L2 during adulthood, and adult monolingual speakers of that language. The accents of these groups are rated by native monolinguals as native or foreign.

Among the tested predictors we find sociolinguistic and affective factors, such as:

- Language preference;
- Frequency of contact and use during childhood (based on schooling and language practices at home);
- Frequency of use during data collection (based on social relations, work, place of residence);
- Type of schooling in the language studied;
- Age at the time of data collection;
- Length of residence in a country where the language studied is an official one (Kupisch et al., 2015, 131-132).

The results show that the simultaneous bilinguals' accents were evaluated as foreign less often when speaking the majority language compared to the minority language. The raters judged the situation with the majority language with the same, high, degree of confidence as with monolingual speakers.

For simultaneous bilinguals speaking the minority language, however, evaluations were unclear and displayed a higher degree of uncertainty by the raters. While "L1 monolingual speakers are predominantly classified as native, L2ers [L2 learners] predominantly as foreign, and 2L1ers [simultaneous bilinguals] in their majority language as native, […] 2L1ers tend to sound foreign in the minority language (55%), but sometimes pass as natives (11%) or receive varying judgments (34%)" (Kupisch et al., 2015, 139).

Among tested predictors, the variables which correlate most strongly with native accent are language preference, language use at the time of testing, and length of residence in the country where the analyzed language was spoken before age 19. The fact that the last factor showed up as significant is really relevant, because it means that sound systems not only need early

acquisition, but also a period of stabilization "with input from several different native speakers early in life" (Kupisch et al., 2015, 143). This is precisely the situation majority languages benefit from, as opposed to minority languages, whose contexts of use are usually restricted. Moreover, results showed that living in the analyzed language's country is more beneficial for a native-like accent if it happened during childhood, rather than adulthood, confirming the idea of age playing a role. Nonetheless, "current use and exposure can positively affect the accent even during adulthood" (Kupisch et. al, 2015, 145).

Another interesting study on this topic, contrasting monolinguals and two kinds of HL speakers, is Godson (2004), examining vowel production in Western Armenian. The study is based on the following hypothesis: HL speakers sound "off" to a Homeland ear because their sound systems might be different not only from the monolinguals' corresponding ones (Polinsky, 2018), coherent with Flege's SLM (1995, cf. §1.5.1.), but also from more balanced bilinguals' ones (Polinsky and Kagan, 2007).

Godson compares a group of Western Armenian and English bilinguals who switched to English as the dominant language before age 8 (so, the more "balanced" ones), a group of bilinguals who learnt English only during adulthood (the less balanced), and a group of Western Armenian monolinguals. She shows that:

a) The pronunciation of /a/, /ɛ/ and /i/ was closer to the English sounds for the first two groups, confronted with the third one. However, this was truer for the first group than for the second one, which had realizations between those of the first and the third group.
b) The first two groups pronounced /o/ and /u/ similarly, but in a different way than English and – except for /u/ - than Western-Armenian monolinguals.

The author concludes that the impact of a dominant language in sound systems is not a simple and straightforward process; it is a force that interacts with "universal tendencies, normal diachronic change, and sociolinguistic pressures" (Godson, 2004, 45). Furthermore, providing proof of the influence of English for both the first and second group, even to different degrees, she demonstrates that a dominant language has an effect on a minority one over a lifetime. Finally, she highlights a partially divergent behavior for the group of less balanced HL speakers.

On the same page, in some ways, is Polinsky (2018). She also considers, in HLs sound systems, both the importance of interference with the dominant language and universal tendencies. In fact, she reports a series of studies from which she concludes that HL sound systems are likely to undergo significant leveling and questions the causes. On the one hand, we can address phenomena of interaction with the dominant language sound system; on the

other hand, these changes are sometimes systematic, and not ascribable to transfer. Thus, Polinsky suggests as a possible cause ease of articulation. In a few words, our bodies try to conserve energy in the articulatory organs. They do so by preferring articulatory gestures that require less effort, and therefore, by preferring the sounds associated with these gestures. It is a mechanism involved in many patterns of sound reduction all over languages, and also in hypoarticulated speech (Lindblom, 1990). However, at the current time there is no systematic study analyzing the role played by ease of articulation in HLs (Polinsky, 2018).

1.6. The grammar of Heritage Languages

Now that we have a picture of the social and formal aspects regarding HLs, we can try to assess, when HLs do not maintain the same grammar as the baseline, how and why their grammars differ, and therefore, how much they tend to move towards shift (cf. §1.7.).

We could summarize the possible explanations about differences with the baseline into two families. These follow Nagy's (2015) subdivision between Experimental approaches to HLs and Comparative Variationist approaches to HLs. The first one sees in the differences between a HL grammar and the baseline grammar a type of deficit within the HL, and therefore argues that those differences reflect incomplete acquisition, transfer from the dominant language, or attrition. The second one, instead, argues that these possible differences are not necessarily a manifestation of a "lacking" system, but rather they could be interpretated also as evidence of the natural variation and change of languages. Moreover, they are not mandatory for HLs, since HL grammars can be comprehensive systems not structurally different from their Homeland counterparts.

We will first discuss why some HL grammars seem to diverge from the baseline according to Experimental approaches in section 1.6.1. In section 1.6.2., we will address the Comparative Variationist view on HL grammars and their patterns of maintenance, variation, and change. Finally, we will dedicate section 1.6.3. to the question of why the results of these two research lines attribute different characteristics to HL grammars.

1.6.1. The Experimental approach on divergent HL grammars

Let us begin with the first family of explanations – those lying under the Experimental approach.

Scontras, Fuchs, and Polinsky (2015) and Polinsky (2018) list as possible causes of difference between HL grammars and baseline grammars[12] three main factors: divergent attainment, also mentioned in §1.4.3.; transfer from the dominant grammar, and attrition. These act jointly, but differently in each individual considered. They will be each discussed in turn.

1. Divergent attainment is the systematic divergence or innovation from the baseline that can be found in HL grammars. It has been addressed before in literature as the result of "incomplete acquisition". Polinsky (2018, 28) claims that divergent attainment leads to the formation of a coherent grammar, whereas transfer and attrition "may be less systematic". It is ascribed by the author not only to reduced input, but also to incipient changes in the baseline and universal principles of language design. For incipient changes in the baseline, Polinsky means that attrition, variation and change in progress can be already present in the group of first-generation immigrants. This would consequently provide the child with a heterogenous variety, and trigger divergent attainment. With universal principles of language design, Polinsky points out how patterns of structure design found in HLs resemble the ones observed in creolistic studies, which have been interpreted by certain models as expressions of underlying and innate principles.
2. Transfer is a phenomenon which happens in situations of language-contact, such as L2 acquisition, creole genesis, or community multilingualism. However, we can see forms of transfer also in situations of a different kind of contact, as for the one between regional varieties of a language. Transfer consists in a particular interplay, where features or structures of one language are transferred to another. In the case of HLs, it is usually thought that the dominant language gives the linguistic traits, while the HL acts as a recipient – more or less like the lexifier for the creation of the creole. In fact, the term "transfer" is mostly used to speak about "unbalanced" situations, where the process is for the most part one-directional: from the most prestigious variety to the least prestigious one[13]. However, the fact that HLs are minority languages, and therefore less prestigious varieties, does not make it acceptable to reject any hypothesis of transfer from the HL grammar to the dominant one: as Polinsky (2018, 138) reminds us, "the interaction between any two languages is bidirectional", and Grosjean (1982) also

[12] With the authors using first generation immigrants as baseline groups.
[13] Mutual transfer, instead, is called "convergence" (McMahon, 1994), and requires the varieties involved in the process to be perceived as equal on a social and political scale. It interests communities of stable and widespread bilingualism. Convergence takes place, for example, in the creation of linguistic areas such as the Balkans.

suggests that the interferences between two languages are bi-directional (even with the dominant variety influencing the minority one to a greater extent). An example is provided by Hoffman and Walker (2010), who show evidence of substrate transfer in the Toronto English spoken by Italian and Chinese first-generation immigrants.

Furthermore, transfer from the dominant language is not necessarily the easiest or best explanation for differences in the HL grammar compared to the baseline. From the sociolinguistic perspective, transfer can be complex, since structural borrowing is, according to Thomason and Kaufman (1988), dependent on sociolinguistic parameters, such as the attitudes of the receiving language speakers towards the giving language.

3. Turning to attrition, let us give a definition. Given, as the benchmark, a native monolingual speaker of same age and in the same language development phase, the analyzed HL speaker shows attrition when their linguistic abilities or linguistic judgements are temporarily or permanently lost, as compared to the benchmark (Seliger, 1996). Attrition also implies that these linguistic skills, now lost, were at some point completely mastered. Attrition can cover an entire lifespan, calling into question the naive assumption that linguistic structures are stable once adulthood is reached.

   L1 attrition specifically refers to bilingualism situations. The logic behind L1 attrition is that if bilingual children show a particular structural property of the HL, and HL adult speakers do not, then that structure has been lost during the years. The extent of attrition seems to be inversely related to the age of onset of bilingualism – or the Age of Arrival in the foreign land, if one prefers Flege's (2007) point of view (cf. §1.5.1.): if a speaker leaves their Homeland country well after having fully mastered L1, they should undergo less and slower attrition (Polinsky, 2018). Should such a speaker have children, based on the timings of their possible attrition and the onset of the children's acquisition combined, they could provide their kids with an already "post-attrition" baseline, influencing the development of the kids' HL grammar (for example, as in the case of incipient changes in the input). This is one of the reasons why HL first-generation speakers' samples should be examined separately from Homeland speakers' samples.

As an alternative to the three causes proposed by Scontras, Fuchs, and Polinsky (2015) and Polinsky (2018) just described, Lleó (2017) imports from studies on simultaneous bilinguals a list of factors provoking divergence (originally from monolinguals, in the case of HLs from the

first-generation immigrants' baseline)[14]. These factors are delay, acceleration, and transfer (an aspect shared with Polinsky, 2018). According to the initial formulation, delay is the slowdown of the acquisition, and acceleration is the premature appearance of a property in the grammar. The notion of transfer has already been discussed. Lleó proposes to consider delay and acceleration as processes involving not the overall acquisition of the grammar, but rather the acquisition of a certain category or phenomenon. She also concludes that, while these two can be eventually overcome and do not have long-lasting consequences, transfer is far more incisive and resistant in the grammar, and therefore plausibly permanent.

Overall, it seems that, for Experimental approaches, the interpretation of the differences between HL grammars and baseline grammars is that of a simplification of the HL grammar. As a matter of fact, among many linguists a common leitmotiv is that HLs have eroded, are *necessarily* subject to interference from the dominant language, and are in conclusion simplified systems (Nagy, 2017a). This view is exemplified, for instance, in Fernández-Ordóñez (2012, 73-74), which also explicitly mentions the "loss of previously existing distinctions" in HLs. Any variation in HLs is interpreted under this idea (Montrul, 2008). Experimental results (as reviewed in Benmamoun, Montrul, and Polinsky, 2013) support this view of HLs as the defective counterpart of their Homeland variety.

The second family of explanations, however, starts from completely different assumptions and draws completely different conclusions.

1.6.2. The Comparative Variationist approach on divergent (?) HL grammars

This second research line usually employs comparative variationist methods to investigate patterns of maintenance, variation, and change in HLs – hence the name given by Nagy (2015). Starting from the idea of variation as a natural and constitutive element of linguistic systems (Weinreich, Labov, and Herzog, 1958), variation and change are actually included in language structure, and are considered to be dependent on speakers' choices, constrained by both social and linguistic factors. Therefore, these choices reflect the underlying grammatical system of the language (Bailey, 2013). Attrition, transfer, or "incomplete acquisition" can, of course, be sources of variation and change, but this must not be always the case, because "HLs are not simply impoverished versions of their respective full homeland languages" (Łyskawa and Nagy, 2019, 2). Furthermore, considering these phenomena as the only and mandatory

[14] These factors were originally proposed by Paradis and Genesee (1996) for syntax/morphology, but Lleó adopts and adapts them also for phonology.

explanations possible means claiming that every HL experience a certain amount of shifting, without actually taking in consideration measures of Ethnolinguistic Vitality or other ethnic and/or social factors.

As a consequence, the three phenomena mentioned above should not be automatically invoked as soon as variation in a HL is encountered; under the socio-variationist approach, hypotheses of transfer, attrition, and incomplete acquisition ought to be tested and supported by data and comparisons among: a) different social groups, because each of these phenomena implies some social aspect, like majority/minority dynamics, age, immigration wave, community, rates of language use and so on; b) different linguistic contexts, in order to investigate all the possible linguistic factors influencing variation.

Within the Comparative Variationist research on HLs, tested predictors usually include both social and linguistic independent variables. These serve – among others - one main purpose: understanding not only the absolute frequency or rate of occurrence of the dependent variable's variants, but also the grammar underlying them, i.e., the ranking of constraints regulating said frequencies/rates. This way, a finer-grained analysis of patterns of maintenance, variation, and change is possible. This approach has been crucial to highlight the systematicity of HL grammars (Łyskawa and Nagy, 2019).

Interestingly, many studies following this second approach to variation find few cases of systematic differences between HL grammars and Homeland grammars, hence excluding the effects of incomplete acquisition and/or attrition, and no or little correlation between language use/ethnic/social factors and linguistic behavior, excluding contact-induced effects. Examples may be found in Flores-Ferrán (2004); Nagy et al. (2011); Rumpf and DiVenanzio (2012); Nagy and Kochetov (2013); Flores, Rinke, and Azevedo (2017); Łyskawa and Nagy, (2019). Such studies, by examining larger portions of the population compared to Experimental studies, seem to overcome idiosyncrasies and to be able to distinguish between variation and actual change, showing us that HLs do not necessarily *have* to be different from a baseline.

1.6.3. Different results for different approaches

In her review of the results on VOT and null subjects in some HLs of Toronto, Nagy (2015) discusses the difference between the outcomes produced by Comparative Variationist studies and Experimental studies, asking why the former have shown successful patterns of transmission of the grammar, while the latter signs of incomplete acquisition or attrition – at

least, for pro-drop in HLs. The reasons she proposes, however, may be seen as not case-dependent ones, but rather quite general.

Some of the possibilities can be excluded precisely thanks to the multiple comparisons among different groups that constitute the core of the methodology of comparative variationist paradigm. First, the author rejects the hypothesis of being "too late" in Toronto to have caught a change, i.e., of having data on already incomplete/attrited first generations of immigrants, because their performances do not differ from those of the Homeland samples.

Second, the author rejects the hypothesis of being "too early" in Toronto to have caught a change – possibly, because the context of this city is supportive enough to have prevented the change from occurring so far, compared to other contexts. She is able to do so via cross-linguistic comparison of community-level factors, i.e., by contrasting the different demographic profiles of the HL investigated. Indeed, they differ for the length of time they have been established in Toronto, the ratio of mother tongue speakers to ethnic community members, and for the size of the community, but none of these factors relates to patterns of retention or change[15].

Third, the author rejects the hypothesis of having the dominant language – i.e., Toronto English – being influenced and changed by HLs, and therefore not able to "exert" its contact effects, thanks to a comparison with the data coming from the Hoffman and Walker's (2010) Contact in the City corpus, a corpus of the English spoken in Toronto by individuals of different ethnic origins.

The last two hypotheses about why the results from the two approaches diverge, instead, are not to be excluded. The second to last hypothesis regards the types of speakers participating in the studies. Indeed, Experimental studies usually employ one of the following types of speakers or a combination of them:

a) Children and/or adolescent in school;
b) Speakers from HL classes;
c) Speakers who fall under the narrower definition of HL speaker, i.e., who are defined with reference to their proficiency levels and are already considered non-native like.

On the other hand, the Comparative Variationist studies that Nagy refers to, i.e., those part of the HLVC project (Nagy 2011), employ the following types of speakers:

[15] With the partial exception of Russian and Ukrainian having smaller communities and behaving similarly in the way their VOT is evolving, a change that however seems not to be influenced by English.

a) Speakers ranging in age from 16 to 85 (mean: 48);
b) Speakers who have not been recruited through schools or institutions;
c) Speakers who fall under the broader definition of HL speaker, i.e., who are defined only with reference to their linguistic autobiography.

Finally, the last hypothesis is that the different methods used produce different results. Experimental tasks are usually set in formal/new/unsettling contexts, e.g., laboratories; they "target exactly the contexts where differential performance is anticipated" (Nagy, 2015, 324); their data is the result of somewhat artificial activities and controlled elicitations. Instead, Comparative Variationist approaches are usually based on naturalistic conversational recordings, collected in settings that are known to the speaker(s) or at least feel comfortable, and with the aim of extracting the most spontaneous production possible. Only by conducting multi-methods research on the same sample of speakers we will be able to empirically test Nagy's (2015) hypothesis; hence, we hope for a fruitful dialogue and collaboration between scholars from the two approaches in the future.

1.7. Maintenance and shift in Heritage Languages

The peculiar ethnic, social, and acquisition dynamics characterizing HLs might lead to several phenomena that imply different sorts of change. These changes can be analyzed under what Aalberse, Muysken and Backus (2019) refer to as the "scenario approach", typical of contact linguistics. This approach stems from the idea that specific socially determined language contact settings have specific linguistic outcomes (Aalberse, Muysken and Backus, 2019, 31). Examples of well-known scenarios are maintenance and shift (Thomason and Kaufman, 1988), codeswitching, attrition (cf. §1.6.1.), and levelling. They are mainly asymmetrical, i.e., the languages involved are not equal in regard to their social power. Considering these general settings, the more individual settings such as family habits or speakers' attitudes, and the typological features of the languages in contact, allows us to explain many differences between HL speakers, and the ways HLs may change. The starting point of a contact-induced change view, however, is that linguistic factors are secondary compared to the sociolinguistic history of the speakers (Thomason and Kaufman, 1988, 35).

Let us describe specifically two relevant scenarios: maintenance and shift.

If a community keeps using the HL as a vehicle for communication, we speak of language maintenance. Instead, shift consists in a consistent reduction, or abandon, of the use of the L1

by speakers (Schmid, 2011); if the community gives up the HL specifically for the majority language, we speak of language shift towards the majority language. Maintenance and shift can be modelled onto a continuum, since speakers can gradually decrease the domains of use of their HL, and possibly going through other scenarios (like attrition) as the process goes. Cases in which the HL is used only in informal in-group communication, and also in combination with the majority language, are still considered situations of maintenance. Indeed, maintenance often resolves into a strict specialization of functions between the languages involved, i.e., diglossia (cf. §1.2.).

As hinted above, social factors affect choices of maintenance and shift; for example, labor migrants shift to a majority language more quickly. Other factors are the size of the community, the length of the contact, how much the HL community is replenished with new immigrants, the support coming from institutions, and some factors we have already described earlier, such as family language use, networks, ethnic affiliation, attitudes, or communities of practice. However, we can add to our list also interlocutor effects, as Aalbert, Muysken and Backus (2019, 52) suggest. The three authors even report that this is consistently found out to be most important of the factors. In fact, if an act is frequent enough, it becomes routinized, i.e., its qualities become fixed. Therefore, if a conversation with a person is frequent, the language of the interaction becomes fixed. For example, families accustomed to speaking their HL at home, will probably continue to do so out of habit, unless some other, stronger factor counteracts this habit. One pattern of HL choice is that children use the majority language among them and the HL with parents, adults, and in cases of advanced shift, only grandparents.

When language choices are more oriented towards shift - and this is often thought to be the case for many minority languages (cf. §1.6.) - the language runs the risk of extinction. Advanced stages of shifting mean that a language is near death: this is a case where the language is referred to as "moribund", "obsolescent", or "relic" (Aalberse, Muysken and Backus, 2019, 46). While some scholars claim that it is the speakers' prerogative to have "the right to maintain and the right to shift" (Clyne, 2003, 68), one could argue whether under a situation of strong social pressure shift can really be considered as a choice, while still technically being so. This is why most sociolinguists perceive shift negatively, and try to propose strategies to influence language choices and avoid it.

Shift usually takes the form of the purposeful lack of transmission of the HL to the children. This is a frequent caretakers' strategy to ensure their kids more opportunities in societies that take in great consideration mastering an accent-less and proficient majority language, and that limit the possibilities of socio-economic advancement for those who do not meet these

standards. The product of a shift situation is a variety of the HL which shows what have been evaluated as evident signs of wear, such as severely limited vocabulary, very simplified grammars, and loss of registers (Aalberse, Muysken, Backus, 2019, 61). Nevertheless, only a careful investigation can distinguish between patterns of simplification from a baseline and pattern of variation and change.

The picture, however, is complicated by the fact that not only the factors influencing language choice interact, but they can also have opposite effects in different communities, facilitating maintenance in one and shift in another. A clear decision tree of what linguistic choices will be made can be constructed only in the most stable and strict diglossia situations, which are very rare (Aalberse, Muysken and Backus, 2019, 52). An attempt to combine the factors and to precisely quantify the degree of maintenance is provided by the Ethnolinguistic Vitality Theory (Giles and Johnson, 1987), which gives languages a Vitality Index score, and adds to the picture also the factor of perceived vitality: if the ethnic groups is perceived as highly vital, the salience of its identity increases, and speakers are more likely to accentuate group speech markers.

1.8. Are Heritage Language speakers native speakers?

What does "native speaker" mean?

According to Love and Ansaldo (2010), the term was first used in linguistics by Bloomfield (1935 [1933], 43) with the following meaning: a native speaker is an individual born and brought up from birth to speak a given language, and in a family where caretakers and/or other adults share the same experience with the same language. The language spoken by this native speaker is the native language, often synonymous with "mother tongue".

However, the concept of "nativeness" has not been agreed upon. Several definitions succeeded, up to the point that now we still know neither what "native speaker" means, nor who a native speaker is or what they know (Myhill, 2003). Despite this lack of clarity, the notion has been largely used by researchers who implied that everybody had a sense of what a native speaker is (Escudero and Smith, 2001), and who avoided its problematicity (Myhill, 2003). It has even entered our daily life (Doerr, 2009).

Some scholars essentially pursued a notion of nativeness consistent with Bloomfield's proposal. An example is Rothman and Treffers-Daller (2014, 95), who affirm that: "A native language is one that is acquired from naturalistic exposure, in early childhood and in an authentic social context/speech community". Other scholars, instead, have a different view on

who becomes a native speaker. For them, a native speaker is someone who achieves native attainment: "Native speaker competence is typically the result of normal first language acquisition in an environment where the native language is dominant in various contexts, and learners have extensive and continuous exposure to it and opportunities to use it" (Benmamoun, Montrul and Polinsky, 2013, 129).

Perhaps the second position is the one which most influenced the general perception of nativeness, since a native speaker is often conceptualized not only as an individual who started listening to and speaking their native language from birth, but also who possesses "a complete and possibly innate competence in the language" (Pennycook, 1994, 175). The notions of nativeness, dominance and proficiency, therefore, started to blur.

However, during the last decades a strong debate about nativeness spread in linguistics: the scientific community began to acknowledge how strongly connected this concept is with politics, power, and sociology. Myhill (2003), for example, claims that there is nothing empirical about being a native speaker: the concept is a social construct, and this forces us to analyze why we have constructed it this way, how we have applied it, and what the consequences are.

Pennycook (1994) is very eloquent in explaining how the general concept of "native speaker" is linked to three ideologies. The first is the direct correspondence between being a citizen of a nation-state and being a native speaker of the nation-state's national language. The second is the idea of languages as homogeneous and fixed entities spoken by homogenous communities – and it is precisely thanks to this ideology that a distinction between native and non-native is possible. The third is that the native speaker is by default the speaker with the highest competence in all the domains of their native language. Claiming that non-native speakers are less competent, or deficient in relation to native speakers, makes the definition of nativeness inherently political, according to Liu (1999).

However, we can easily demonstrate that these three ideologies at the foundation of the common idea of "native speaker" rely on wrong assumptions. Let us, for example, think of Italy. The nation-state's national language is standard Italian. People born and raised in Italy by Italian caretakers would therefore be prototypical native speakers of Italian. However, lots of individuals born in the previous century acquired as a first language an Italo-romance dialect; some of them never fully learnt Italian, others learnt it as late L2 learners, others as early bilinguals. Moreover, Italy's linguistic repertoire is complex and multifaceted, it is a multidimensional continuum of continua, originated by the crossing of all the axes of variation (Grassi et al., 1997). It is not homogeneous, nor is its speech community, since it involves

people from all regions and with different linguistic backgrounds. Each speaker may have different competences in different Italian varieties or registers. Would we say that someone dominant and highly proficient in their Regional Italian, and/or in *italiano popolare* (Berruto, 2006 [1987]), but not able to master the highest registers, to write or to speak in formal contexts, is not a native speaker of Italian? According to the third ideology implied in the common notion of native speaker, we should; but in reality, we would not.

This example shows us how much has often been taken for granted on this topic, and how many of our assumptions are biased. Love and Ansaldo (2010) push us to think of our ideas of "native speaker" and "normal language transmission" as products of a monolingual acquisition under normative pressure: in other words, idealizations. This idealized native speaker has become an authority in formal linguistics, the benchmark for the assessment of proficiency levels, the source of intuitions, judgments, and information, and all because of their naturalistic exposure and acquisition (not learning) of the language, something that generated implicit knowledge.

This is precisely why scholars started proposing alternatives for judging the proficiency levels of L2 learners. For instance, Cook (1999) suggests considering multicompetence, while House (2003) suggests evaluating English L2 proficiency in terms of *lingua franca*, and not of degree of proximity to the native speaker.

Multicompetence is defined as "the knowledge of more than one language in the same mind or the same community" (Cook, 2012, 3768); it is said to also concern the relationships between these languages in a mind or community, and to affect all cognitive processes, not language alone (Cook, 2016). It is a term intended to be, for multilinguals, as neutral and free from comparative evaluations with monolinguals as the term "competence" is for monolinguals (Cook, 1999). There is no point in judging L2 learners based on monolinguals' performances, because they are involved in different language and thought processes. As examples, Cook (1999, 192-193) reports for multicompetent speakers:

- On the one hand, the constant availability of the L1 and its communication strategies when using the L2; the possibility to codeswitch/codemix; the faster and more accurate performances in language-switching tasks compared to monolinguals;
- On the other hand, less efficient long-term memory of information gathered in lectures in the L2; values slightly below the L1 level of working memory span in the L2, at all stages of L2 performance; performances slightly below the ones from L1 monolinguals in tasks of object naming and following instructions to mark letters in words, and all

because, according to the author, multicompetent speakers have available more than one response to the same stimulus, slowing down reaction times.

The author concludes that “people who speak differently from some arbitrary group are not speaking better or worse, just differently”, and that “L2 users should be treated as people in their own right, not as deficient native speakers” (Cook, 1999, 194-195).

House's (2003) concept of English as a *Lingua Franca* (ELF) stems from the consideration that the number of non-native speakers of English is four times the number of native speakers (Graddol, 1997). It must be stressed that House specifically refers to situations where English is used to communicate between people who both have a different native language different than English, and therefore, that her ideas and conclusions should not be extended to other types of interactions, such as immigrant speakers using English with locals once they are in an Anglophone land.

ELF is not purpose-specific, a pidgin, or an interlanguage, but a versatile tool, a "language for communication", rather than a "language for identification" (House, 2003, 559). Its description should not be based on the concept of speech community (whatever model is considered), because speech communities are defined on the grounds of shared, stable, and homogeneous social and/or behavioral properties, while ELF communication interests individuals moving in and out of various contexts, each with a different form of participation. Therefore, House argues that its description should be based on the concept of community of practice. ELF interactions are seen as joint enterprises aimed to negotiate contents and/or linguistic forms (echoing Clark's model of language as joint project), and they do not have the purpose of building or showing identities. Consequently, since ELF speakers are not monolingual speakers, their English proficiency level must not be assessed via comparison with monolinguals, but with an "expert in ELF use, a stable multilingual speaker under comparable socio-cultural and historical conditions of language use, and with comparable goals for interaction" (House, 2003, 573).

It is easy to imagine situations where multicompetence or ELF would be useful tools to assess HL speakers’ proficiency in a much fairer and balanced way than simply comparing them to monolinguals: HL speakers potentially master more than one language (and/or culture), and interactions among members of different Heritage groups are likely to happen using English as a *Lingua Franca*[16].

[16] Of course, not every time a HL speaker uses English we are facing a language choice that is not for identification, but just for communication. As we have seen in Di Salvo (2017) (cf. §1.2.), using English may be a very precise

Alternatives for how to define "native speakers" also came up. Rampton's (1990) proposal is to give up the overly general term "native speaker" in favor of the specific notions of language expertise, language affiliation, and language inheritance. This way, the implications behind "native speakers" would be disambiguated, divided, and tackled by separate concepts. Language expertise refers to linguistic proficiency. Language affiliation refers to the attachment for or identification with a language, regardless of the membership to the group usually associated with it. Language inheritance refers to the birth of an individual into a family or community with a particular linguistic tradition, regardless of the individual's affiliation to that language or their self-claimed expertise in that language.

Escudero and Smith (2001) express the need for scientific rigor: their idea, differently from other scholars mentioned, is not to completely change the notion, but to impose some order. They suggest reformulating the concept using prototype theory. They also pose another interesting question about nativeness, i.e., whether native speakers can *become* non-native. The two authors list some limit cases of native speakers (i.e., people suffering from language impairment, pre-literate children, and non-literate people without schooling), and conclude that, in order to account for a multiplicity of situations while still recognizing that some features can be "more native" than others, a prototypical approach could be suitable[17].

It is from the idealized notion of native speaker that many opinions on HL speakers stem. If one considers the native speaker in terms of acquisition, attainment, proficiency, and education, then it is easy to understand why many HL speakers would not fit the definition. An example is Benmamoun, Polinsky and Montrul (2013), researchers who, as mentioned before, follow the "native attainment" criterion. Indeed, they describe prototypical native speakers as people who fully acquired their native language system; then, they describe L2 speakers as people who "typically exhibit persistent signs of non-target like acquisition in phonetics, phonology, inflectional morphology, semantics, syntax, and discourse/pragmatics" (Benmamoun, Polinsky and Montrul, 2013, 130). After establishing that native speakers and L2 learners are just the two extremes of a continuum of language attainment, they claim that HL speakers are

choice of the HL speaker to convey information about who they are, they have become, or they want to be perceived as.

[17] Nevertheless, the authors explicitly state that: "The prototype profile chosen by linguists and SLA researchers should be expressed purely in intralinguistic terms. This third suggestion implies that we should disregard any extralinguistic deviations that may play a role in a layperson's perceptions of nativeness mainly because of their subjective and misleading nature. […] [T]he prototype profile most appropriate for the linguistic investigator should only contain linguistic features and thus exclude as irrelevant features such as self-perception, perception by others, nationality, ethnicity or initial acquisitional environment" (Escudero and Smith, 2001, 283-285).

positioned in between these extremes. They are something "in the middle", neither native speakers, nor L2 speakers. They share features with both of them.

On a totally different page are Rothman and Treffers-Daller (2014). As said earlier, these authors instead follow the original Bloomfield criterion. They recall that HL speakers acquired the HL with ecological exposure, in a naturalistic and authentic setting, and during early childhood. Under the Bloomfieldian definition, they are undeniably native speakers of their HL. The authors also try to investigate why other researchers are resistant to define them as "true natives", and they conclude that the fallacy is in the confusion between the concepts of nativeness, dominance, and proficiency.

The problem lies in the fact that, by equating nativeness and perceived dominance, one falls into a circular argument, where if one is a native speaker, then they are dominant in their native language, but if one is dominant in a language, then they are perceived as native speakers of that language.

Since monolinguals are dominant in just one language, they have been considered by everyone native, without any doubt. Bilinguals have set new challenges, since their dominance is not a default product. Nonetheless, the most balanced and simultaneous ones have succeeded in gaining the status of native speakers. HL speakers, instead, "exist on a continuum of relative dominance in the heritage language" (Rothman and Treffers-Daller, 2014, 95), that prevents some scholars from conferring them the label. The authors' position is that the ends – the apparent competence outcomes – cannot justify the means. HL speakers are native speakers[18], regardless of any attrition they could have undergone. Nativeness should be separated from any sort of evaluation. In the terms proposed by Rampton (1990), HL speakers would show different rates of language expertise, but they would be identified by their language inheritance and higher rates of language affiliation.

Let us turn back to our initial proposal: that of re-thinking the concept of a native speaker.

First and foremost, we cannot deny that this notion has been politicized - as Liu (1999) suggested. It has affected the daily lives, opportunities, and jobs of people all over the world. Proofs may be found, for example, in the field of TESOL, Teaching English as a Second Language. I will report an extract from Canagarajah (1999) that provides a demonstration of this:

[18] Of their HL and of the dominant language, "if the majority language was acquired naturalistically in early childhood before or at the age of 4–6 years when schooling in the majority language is introduced" (Rothman and Treffers-Daller, 2014, 96).

> Surfing absentmindedly through the electronic forum for second language teachers (TESL-L) one day, I was struck by a desperate e-mail from a Korean graduate student. She said, "I am finishing my MA in TESL at the end of this semester at xxx University in Boston, and I hope to return to my country. But I cannot hope to find a teaching position back home. They don't hire non-native speakers. What are the prospects for finding jobs here in the United States? Can someone give me some clues about job openings here?". I knew the sad reality of the job market in the United States and considered it kindness not to reply. I could only imagine her consternation when even in the West, advertisement after advertisement confronts her with the fact that only those who are "native English speakers" or those with "native English competence" can apply for the available positions. Fresh from graduate school, certified with a Masters or a doctorate in applied linguistics, and groomed for a career in language teaching by a reputed university, the non-native ESL teacher often discovers a gloomy professional future. (Canagarajah, 1999, 77)

If politics, exclusion, and gatekeeping are not reasons enough to discuss the native speaker, as linguists let us recall that there are also theoretical and methodological reasons.

I personally agree with Rothman and Treffers-Daller (2014) when they say that there has been a conflation between “nativeness” and “dominance”. Psycholinguists, formalists, or sociolinguists are all interested in different aspects of language, have different views on what language is, and have different goals when they do research. The natural consequence is that they have all used different “native speakers” in their studies, and the concept became confused and fuzzy.

In other words, our needs, as linguists, are various and dependent on one’s framework and purpose of the research. Thus, a proposal like Rampton’s would make our job easier, allowing us to precisely pinpoint what kind of speaker we need for our goals, avoiding ambiguity.

If one is not willing to give up the term “native speaker”, maybe returning to a stricter definition of nativeness that excludes end-states of any sort is the best solution to be as inclusive as possible, and to speak of “multinativeness”, or individuals having more than one native language: HL speakers, for example.

The conclusion this chapter hopes to have offered is that HL speakers are a valuable resource to reach the goal of re-thinking nativeness. They allow us to explore the world of multilingualism by questioning assumptions, models and methods that have been used for decades, and to keep discussing and searching for answers.

2. Rhotics

Rhotics, or r-sounds, represent a particular class of sounds. Their first peculiarity is how the class has originated: while usually phonetic classes are built based upon articulatory or auditory properties, rhotics have been grouped based on the tendency diffused in languages of the world to write them with the characters derived from the Greek letter *rho*, such as "r" (Ladefoged and Maddieson, 1996; Wiese, 2011). Indeed, rhotics are not correlated with a single manner or place of articulation. There are many variants of r-sounds, and sometimes they coexist within the same language variety (Van Hout and Van de Velde, 2001). We could say that rhotics are characterized by articulatory complexity and flexibility; such features are considered by some scholars the cause of their tendency towards variation and change (van Hout and Van de Velde, 2001; Scobbie, 2006). As a matter of fact, not only are r-sounds involved in many phonological diachronic processes, but they are also often socio-indexed, showing particular social and/or geographical distributions. An example is the labiodental approximant [ʋ] rhotic in England (Docherty and Foulkes, 2001).

Rhotics in phonetics have sometimes been identified with the sounds denoted in the IPA system by r-like symbols (Wiese, 2011). These sounds are all meant to be voiced, and they include:

- (By manner) trills, taps and flaps, fricatives, approximants, and lateral flaps;
- (By place) alveolars, retroflexes, and uvulars (Wiese, 2011).

However, sounds missing in this classification have been defined rhotics nonetheless, such as some voiceless sounds, or post-alveolar, dental, and bilabial sounds (even though some of them, e.g., bilabials, are quite rare, and not everyone agrees on defining them as rhotics) (Ladefoged and Maddieson, 1996; Wiese, 2001; 2011). Consequently, Scobbie (2006) proposes to use the label "rhotic" for any oral lingual sonorant consonant that is not palatal, lateral, or labial. This analysis, nevertheless, would not be exhaustive either, as rhotic/rhotacized vowels do in fact exist. As we can clearly see, the articulatory variability within the class prevents us from giving a uniform definition of rhotics from this standpoint, and from excluding other potential sounds from the class based only on their articulatory properties (Wiese, 2011). The lack of articulatory homogeneity of rhotics is the reason why there is uncertainty regarding what the precise members of the category are.

However, an acoustic feature has been often indicated as the defining property of rhotics: the lowering of the F3. This is described as the acoustic correlate of tongue retraction and apicality, which are featured by most r-sounds (Celata, 2010). Furthermore, these two properties are also shared by retroflex sounds (Romito and Scuticchio, 2009), which are indeed characterized by a low F3 too, and we will see that rhotics and retroflexes are related in more than one way. Unfortunately, though, the lowered third formant cannot be considered the unifying property of the whole class. For instance, the Hausa and Arrernte retroflex approximants, the Czech fricative rhotic, and Swedish uvular r-sounds feature high third formants (Ladefoged and Maddieson, 1996).

Rhotics' hyper-variation presents us with a great number of challenges (Scobbie, 2006). For phonology, there is the issue of understanding the most-fitting abstract feature for the category, or the phonological behaviors that allowed certain sounds to group together, and of clarifying why the phonetic counterpart of the category has "orderly heterogeneity"[19]. For cognitive theories of sound representation, there is the issue of accounting for the formation of a potential category from such diverse inputs, or for the production of different exemplars based on a prototype. For sociolinguistics, there is the issue of describing and explaining the flexible social meanings that r-sounds tend to take (Spreafico and Vietti, 2013).

Not all these questions about rhotics have been answered. For example, there are still different proposals regarding the nature of the class, or the interaction between the phonetic and phonological interfaces, and there are many varieties whose patterns of variation and change of r-sounds still have to be studied.

We will first describe the sounds that have been classified as types of rhotics, and illustrate a phenomenon strictly connected to them – retroflexion (section 2.1.). Successively, we will explore the shared phonological behavior of rhotics in section 2.2. Section 2.3. will discuss rhotics in the relevant varieties, e.g., Canadian English and Italian. To conclude the chapter, we will debate what allows us to group rhotics into a class in section 2.4.

During the course of the chapter, we will use as a covert symbol "/r/" to refer to the r-phoneme of a given language, without specification of place or manner of articulation (e.g., "Italian /r/" will indicate the rhotic phoneme in Italian, without specifically referring to a trill, tap, approximant, or fricative realization). Specific realizations of the phoneme will be indicated either by their correspondent phonetic transcription (e.g., [ɾ] for the alveolar tap) or with a description (e.g., "the Spanish alveolar trill").

[19] "For phonological theories favoring underspecification and featural analysis, the problem is… to explain why the phonetic realization of this abstract /r/ is both systematic and arbitrary" (Scobbie, 2006, 338).

2.1. Types of rhotics

A good starting point to describe rhotics typologically is the study by Maddieson (1984), which summarizes their distribution in a sample of 316 languages. Let us start with the amount of rhotics in phonological systems of the world.

The author reports that 57.7% of the languages analyzed has one rhotic phoneme. While 23.3% has zero rhotic phonemes, we also have to consider that these languages are still reported to show rhotics as allophones of other phonemes. 16.1% of languages has two rhotic phonemes, and the percentage of languages with three rhotic phonemes is incredibly small - 2.5%. When two rhotic phonemes coexist within a single language, they are usually contrasted by manner, rather than place of articulation (Ladefoged and Maddieson, 1996). For example, in Spanish, intervocalic trills and taps contrast. In Hausa, a Chadic language, an alveolar trill contrasts with a post-alveolar flap or approximant; they both appear in a single or geminate form[20]. In the rare cases where three rhotic phonemes coexist, they are usually trills, taps or flaps, and approximants, while in Edoid languages the contrast happens in terms of voicing. In conclusion, the striking majority of languages has exactly one r-phoneme, and it seems that almost every language with three r-phonemes is located in Australia, such as Warlpiri.

Now, let us turn to the most frequent types of rhotics typologically. The most common one is the alveolar trill [r], which takes up the 47.5% of the total. Wiese (2011) compares this statistic with the one from UCLA Phonological Segment Inventory Database (sample of 451 languages), and shows the continuity between the two research, with the UCLA Database reporting a 40% of dental and/or alveolar voiced trills among rhotic phonemes of the world and the 20% of voiced alveolar flaps.

A summary of facts about the distribution of rhotics by Maddieson (1984, 82), based on his database, and reported by Wiese (2011), is presented in Table 1[21]:

[20] While it is straightforward to imagine how the geminate trill is realized - with a greater number of periods (cf. §2.1.1.), the geminate flap is a bit less simple. Ladefoged and Maddieson (1996, 237-238) describe it as "a relatively long retroflex approximant, marked by a low third formant, terminating with a brief more constricted phase".

[21] As for point c. of Table 1, “the overall number of r-sounds is smaller here, because the sources do not always report the manner of articulation (Maddieson 1984: 78). More generally, underspecified information on r-sounds may have led to a predominance of prototypical r-sounds in the database, i.e., of alveolar trills” (Wiese, 2011, 5).

a.	An r-sound is likely to be voiced.	308/316	97.5%
b.	An r-sound is likely to be dental or alveolar.	273/316	86.4%
c.	An r-sound is likely to be interrupted[22].	244/282	86.5%
d.	A retroflex r-sound is likely to be a continuant.	20/38	52.6%
e.	An approximant r-sound is likely to be retroflex.	15/28	53.6%
f.	A fricative r-sound is likely to be retroflex.	5/10	50.0%

Table 1: Quantitative generalizations for rhotics.

In the succeeding sections, we will briefly describe the main manners of articulation of rhotics: trills, taps and flaps, and fricatives and approximants; then, their connection with and realization as retroflexes. References to tongue-parts will follow Recasens' (1990) distinction into five regions: apex or tip, blade or lamina[23], pre-dorsum, dorsum, and root or post-dorsum.

2.1.1. Trills

Both Maddieson (1984) and Ladefoged and Maddieson (1996, 217) affirm that the most common form of rhotic across languages of the world is the trill. This typological criterion is what they use to define trills as the prototypical members of the rhotics class.

Trills are generated by the vibration of the active articulator against a passive articulator due to the pressure exerted by the air stream. An example of trill is the one produced by placing the tip of the tongue close to the alveolar ridge leaving an aperture. When the current of air passes through this aperture, aerodynamic conditions produce a pattern of closing and opening of the canal, with the tongue repeatedly beating against the alveolar ridge. The result, in this case, is the alveolar trill [r], found for example in Spanish or Standard Italian. Usually, apical trills are voiced sounds: while the phases of closure do not feature vocal cord vibration, the phases between two closures do so. It is important to underline that *all* trills are produced not by muscular contraction, but by the aerodynamic forces only: the active articulator is set in motion

[22] "Interrupted" is defined by Maddieson as a sound involving "an interruption of the flow of the air through the oral cavity" (Maddieson, 1984, 79). It is a slightly different concept than "non-continuant", since this feature instead refers to the involvement of a "full closure in the oral portion of the vocal tract" (Hayes, 2009, 78). Taps, flaps, and trills, which involve a very brief closure, are not unanimously defined as either continuant or non-continuant sounds by phonologists. This explains why Maddieson opted for a less problematic description and used the term "interrupted" when referring to trills, taps, and flaps. He uses the feature "continuant", instead, for approximant and fricative rhotics.

[23] According to Ladefoged and Maddieson (1996), the tip is the part that lays vertical when at rest, while the blade starts a few millimeters after the tip, with an extension of 1 to 1.5 cm behind the tip.

by the air flow, just like the vocal folds during the production of vowels (Ladefoged and Johnson, 2014). Therefore, trills represent a passive articulation.

An interesting aspect of trills is that they exert strong coarticulatory pressure on adjacent segments, while showing more resistance to coarticulation compared to taps and fricative rhotics (Recasens and Pallarès, 1999; Solé, 1999). This would be due to a more constrained tongue in trills than in taps, according to Recasens (1991).

In order to produce a trill, the air stream and the size of the aperture between the articulators must fall within certain limits, and the muscles must have a certain level of stiffness, or else the trill will fail, i.e., the vibration will not occur. In particular, apical trills require a conflict between the tongue tip, that has to be relaxed to vibrate, and the tongue sides and pre-dorsum, whose muscles have to be tense. This articulatory complexity may result in the loss of tip vibration, and in the alternation with fricatives, approximants, and taps (Celata, 2014). This seems to be valid for every type of trill: "Even in languages where a possible realization is a trill, not all speakers use a trill, and the speakers that do, have tap and approximant allophones as well as the trill" (Lindau, 1985, 161). Indeed, Barry (1997, 41) claims that trills, taps, and approximants are "articulatory siblings in different stages of development". Trills can also show subtler variation, such as their devoicing in prepausal position (Scobbie, 2006).

The precision required to produce trills makes them sounds that entail effort, i.e., more "difficult"; therefore, it is puzzling why apical ones are so statistically predominant that they are the prototype of rhotics (Wiese, 2011).

Each cycle of closure and aperture of a trill is called a period; single trills usually consist of two to five periods, with the first closure often slightly longer in duration than the others (Ladefoged and Maddieson, 1996). However, even a single closure can generate a trill, if the vibration is still originated by the air stream. On spectrograms, the phases of closure are light areas, since there is no/weak formant energy. The phases of aperture are described by Ladefoged and Maddieson (1996) as vowel-like in their acoustic structure, so on spectrograms they are dark areas, where the energy is concentrated in certain formant regions. Both the phases have a similar duration – 25 ms each.

2.1.2. Taps and flaps

Taps and flaps are rhotic consonants which feature just one closure, usually produced with the tip of the tongue against the upper articulators (i.e., they are generally apical). They can be distinguished from trills and stops.

While the vibration in trills is a consequence of aerodynamic forces, in taps and flaps the articulators are placed near together by a muscular contraction (active articulation) (Ladefoged and Maddieson, 1996). According to Recasens (1991), instead, taps and apicoalveolar trills show different degrees of constraint on lingual activity, and he claims that these two sounds are executed via different gestures. Taps do not require high articulatory control for the tongue body positioning, while trills do so and are generally more articulatory constrained. Therefore, taps would be more prone to coarticulatory effects compared to trills.

As for stops, the difference is that in taps and flaps there is not such a great buildup of air pressure behind the articulators.

Many linguists do not distinguish taps from flaps; Ladefoged, who originally proposed the distinction, continues to use it also in Ladefoged and Johnson (2014). Here, he describes flaps as movements from back to front, and taps as movements up and down. In flaps, at the beginning of the movement the tip of the tongue is already curled back – i.e., retroflex (§2.1.4.), and it goes back to its resting position behind the lower front teeth passing through the dental, alveolar, or post-alveolar region. In taps, at the beginning of the movement the tip of the tongue is not curled, therefore it simply moves up towards the dental or alveolar region and then down, returning to its resting position. As a consequence, it is logical to expect the realization of a tap or flap to be often dependent on the phonetic context in which the rhotic is to be found. The introduction of the tap/flap distinction could be partially traced back to the attempt of accounting for inter and intra-language variation in the realizations of this gesture.

Taps are present in Spanish as dental sounds, for instance, while flaps in American English are allophones of post-stress pre-vocalic alveolar stops (Ladefoged and Maddieson, 1996).

Given their similarity with trills, their acoustic structure does not differ from that of trills, but of course has just one phase of closure. Quilis (1981) reported for the Spanish dental tap a mean duration of the phase of closure of 20 ms.

2.1.3. Fricatives and approximants

Inside the family of rhotics we can also find sounds produced with no direct contact between articulators, just an approximation. The result is either a fricative or an approximant rhotic.

Fricative rhotics occur in the KiVunjo dialect of KiChaka as alveolar sounds, or in French, as uvular sounds. Scobbie (2006, 339) claims that "frication is a more typical characteristic of non-trilled uvular and velar /r/". Shosted (2008) suggests that uvular fricatives can easily develop from uvular trills: according to the author, since the articulation of a uvular trill is more

posterior than an apical articulation, the vocal fold vibration tends to be discontinuous, and the natural result is a devoicing and frication of the trill.

Approximant rhotics are, according to Scobbie (2006), extremely prone to variation. They occur in Southern British English, as alveolar sounds, or in German, as uvular sounds, for instance. A particular kind of approximant rhotic is the one found in American English, the so called "bunched r" (cf. §2.3.1.). Its peculiarity is the absence of raising of the coronal area. Its main acoustic feature is the lowering of the third formant. As for trills, approximant rhotics are generally voiced, but voiceless ones are featured in Scottish Gaelic or Konda (Dravidian family), for instance (Ladefoged and Maddieson, 1996).

Approximants rhotics are linked to vowels not only because of their similar acoustic and perceptual properties. Scobbie (2006) maintains that retroflex approximants preceded by high or front vowels show long transitions that can eventually lead to the replacement of the retroflex with a centering diphthong, establishing a correspondence between /r/ and schwa. On the other hand, non-high vowels followed by rhotic approximants may coalescence in a rhotacized vowel.

2.1.4. Retroflex rhotics

As stated at the beginning of this chapter, retroflexion is an especially relevant feature when discussing types of rhotics. Retroflexion and rhotics are related not only because a large amount of rhotics is indeed retroflex (see Table 1 in §2.1.), but also because the retroflexion of a segment due to a rhotic context is a very common process cross-linguistically, and it is the process through which new retroflex phonemes often enter a language. For instance, retroflexion of coronal stops followed by a rhotic is a feature of Calabrese dialect and, possibly, Calabrese Reginal Italian (cf. §2.3.2.2.), varieties that we will investigate in this thesis.

The traditional description of retroflexion is that of a type of articulation where the tongue tip bends backwards and up to some extent (Trask, 1996, 308). While usually the label "retroflex" is used to refer to a place of articulation between the postalveolar and palatal regions, it seems more appropriate to consider it a complex gesture or articulatory shape (Ohala, 1983). Just like rhotics, retroflex consonants also show a high level of articulatory variability (Celata, 2010). However, and differently from rhotics, there is much more consensus regarding the unifying acoustic feature of retroflexion. Its acoustic correlate is the lowering of the third formant, which has been considered the result of posteriority, and possibly velarization and lip

rounding (Celata, 2010). It is also a shared acoustic feature with North American English approximant rhotic [ɹ].

Hamann (2003) provides us with an extensive work on retroflex consonants. She proposes to give an articulatory definition of retroflexion based on four properties: apicality, posteriority, sublingual cavity, and retraction[24]. These are to be intended as the defining features of the prototypical retroflex, occurring in different degrees in different retroflex segments. According to the author, a segment with fewer than three retroflex properties, does not belong to the category of retroflexes. Moreover, the four properties defining retroflexion are interrelated, with each of them implying another one (or more than one).

As stated at the beginning of this section, many phonological processes involve retroflexes and rhotics together. An example is the frequent retroflexion of anterior coronals following an /r/ (/rt/ → [ʈ], Bhat, 1973), found in North-Germanic languages, in Australian languages, and Indo-Aryan languages, both in synchrony and diachrony. However, three clarifications have to be made:

- Retroflexion in a rhotic context can affect non-coronal segments too, as in Tibetan languages (Hamann, 2003);
- The influence of the rhotic is not one-directional: the rhotic can either precede or follow the target segment (Bhat, 1973);
- The relevant rhotic does not have to be retroflex to produce such an effect (Romito and Scuticchio, 2008).

Let us analyze this phenomenon more in detail.

The retroflexion of coronal stops followed or preceded by rhotics has been interpreted differently by scholars. Hamann (2003) gives two plausible explanations to the process, an articulatory one and a perceptive one. The former would be summarized in the following steps:

(a) rt > ɽt > ɽʈ > ʈ or tr > tɽ > ʈɽ > ʈ[25]

In stage 1 we can see the retroflexion of the rhotic - a frequent type of variation, since many languages allow a retroflex variant of their rhotic(s), that could have the purpose of enhancing

[24] The fourth property, retraction, is the most discussed. Retraction is defined by Bhat (1974) as the backing of the tongue body, generally towards the velum or the pharynx. Bhat's opinion is that retroflection and retraction must remain distinct: retraction does not occur only with retroflexion, and some retroflexes are not retracted (Lardil retroflex consonants and Badaga retroflex vowels). Nevertheless, Hamann claims that yes, retraction is not an exclusive property of retroflexes, but retroflexes in particular *have* to show retraction, since, in order to lift and vibrate the apex, the tongue dorsum also *has* to retract and raise against the pharynx (Celata, 2010).

[25] Where t represents every coronal stop and ʈ its retroflex counterpart.

the perceptual cue for the rhotic by further lowering the F3. In stage 2 there would be the assimilation of the retroflex trait to the adjacent segment due to coarticulation (especially regressive coarticulation, which is typically exerted by retroflex consonants - Celata, 2004). In stage 3 (discretionary) there would the dropping of the rhotic portion.

The latter explanation would be summarized in the following step:

(b) rt > ʈ or tr > ʈ

Such a hypothesis stems from the consideration that retroflexes and coronal rhotics share a low third formant. The low F3 of the rhotic can cause the listener to acoustically re-analyze the sequence, to allocate the low F3 feature on the adjacent segment, and to interpret it as a single retroflex sound (cf. misparsing, Ohala, 1995). The direct passage illustrated by (b), compared to (a), is preferred by Hamann (2003), because it would fit synchronic processes of sound change without intermediate steps. However, the second case (tr > ʈ) would be rarer. Indeed, Celata (2004; 2010) explains that in vowel transitions, the formant cues of retroflexion are more prominent in VC transitions than in CV ones; the same logic would apply to consonant clusters, with the consequence of retroflexion of a consonant preceded by /r/ being more frequent than the opposite. The same author critiques Hamann's hypothesis precisely for this: even though the two clusters, -rt- and -tr-, are similar, they are not completely the same, so there is no reason to postulate the same exact process of retroflexion for the two of them. As a matter of fact, they could be due to different causes: while a perceptive explanation seems more adequate for -rt- clusters, an articulatory explanation could fit -tr- clusters better. Moreover, for the -tr- cases there would be also other reconstructions other than Hamann's one that do not imply a step of assimilation, but rather of affrication, especially looking at Italo-Romance dialects' outcomes. Indeed, even though Bhat (1973) and Hamann (2003) do not mention it, retroflexion of coronal stops followed by a rhotic is a feature of many Italo-Romance dialects and Regional Italians, including the ones from the Calabrese area. Therefore, we will extensively talk about the characteristics and outcomes of this process in §2.3.2.2., during our description of Calabrese Italian rhotics.

2.2. Phonological behavior and lenition of rhotics

It is interesting that rhotics share phonological properties within languages. We will illustrate the main ones as follows:

- Rhotics tend to occupy certain places in syllable structure: they are often the second or third member of an onset cluster, or the first member of a coda cluster. Consequently, they seem to be attracted by the syllable nucleus, if not to occupy it. As a matter of fact, descriptions of the syllable include as one of its properties the fact that its sonority increases from the onset to the nucleus, and decreases from the nucleus to the coda (Nespor and Bafile, 2008). Since the striking majority of rhotics has a high level of intrinsic sonority[26], it seems that there is indeed a connection among the positions they tend to assume, their levels of sonority, and the structure of the syllable (cf. §2.4.).
- In many languages with more than one phonemic rhotic – i.e., Australian ones – these are prohibited from occurring word initially (Wurm, 1972).
- Rhotics are associated with other rhotics in allophonic and diachronic alternations, or in free variation. This is also attested by Walsh Dickey's (1997) survey in 53 languages of a sample of 70 languages from different families.

Rhotics have also often been grouped with laterals into the larger class of liquids. They even have a member in common, the lateral flap, which is perceived both as /r/ and /l/ (Ladefoged and Maddieson, 1996). The reasons underlying the creation of the category of liquids lie in:

1) Phonotactic similarities. Indeed, the distribution in the syllable we have described for rhotics is shared within sonorant consonants; however, since nasals feature the expulsion of the air through the nasal cavities, they can be kept apart, and the remaining sonorant consonants can be grouped into the liquid category (Ballard and Starks, 2004);
2) Patterns of alternation between rhotics and laterals, either synchronically, or diachronically (with rhotics replacing laterals overtime or vice-versa). Such a pattern is found in the Neapolitan Italo-Romance dialect (Ledgeway, 2009), where word-internal laterals have been replaced with rhotics (e.g., *salza* > *sarza*, "sauce"; *pullmanne* > *purmanne*, "bus", Ledgeway, 2009, 106);
3) The involvement in the same phonological processes, such as metathesis, vocalization, or dissimilation (Walsh Dickey, 1997) (especially in coda environments), across languages and both in synchrony and diachrony;

[26] According to Maddieson (1984) (cf. Table 1, §2.1.), 97.5% of rhotics in 316 languages is voiced. "This observation (as well as others put forward by Maddieson) looks very much like a statement on the segmental, and not prosodic, content of rhotics. However, it is obvious that there is a relationship between the feature of voice and the sonority rankings of sound classes, such that voiced sounds are generally inside voiceless sounds within a syllable. It is less clear whether the implication needs to be seen as an implication from voicing to sonority or vice versa. That is, rhotics might be overwhelmingly voiced *because* they are high on the sonority hierarchy, and not vice versa" (Wiese, 2001, 351).

4) Late and/or "incomplete" acquisition by both L1 and L2 learners (Proctor, 2011). We will see, for example, the large amount of "defective" /r/ variants in Italian (cf. §2.3.2.1.).

Since we mentioned rhotics' synchronic and diachronic alternations, it is worth describing how r-sounds are invested by one of the most common alternation patterns of languages: lenition.

Lenition has been defined as "synchronic alternations, as well as diachronic sound changes, whereby a sound becomes "weaker", or where a "weaker" sound bears an allophonic relation to a "stronger" sound. […] The core idea, as applied to consonants, is some reduction in constriction degree or duration" (Kirchner, 2013, 3). An alternative definition is "the failure to reach a phonetically specified target: articulatory undershoot or under-achievement. […] It is only if we view the process diachronically that we can see what looks like lenition[27]" (Bauer, 2008, 611; 615). This definition allows us to include in lenition also phenomena like [r]- or [l]-vocalization or vowel reduction, and to understand why passages from less complex to more complex articulations, such as the change from stops to fricatives[28], have been classified as lenitions. Moreover, it highlights the importance of context when trying to understand lenition: a sound change, like the voicing of a voiceless stop, may be lenition in an intervocalic context, but fortition in word-final position (Rennicke, 2015). Finally, it implies a universal mechanism causing lenition, i.e., gestural reduction through automation (Bybee, 2003) or articulatory effort minimization (Kirchner, 2013)[29] and, especially in coda contexts compared to onsets, asynchrony in articulatory gestures (Recasens and Farnetani, 1994; Lawson et al., 2018).

Therefore, several phenomena have been classified under the label of lenition, including spirantization (Honeybone, 2008): the reduction from a consonant such as a stop, an affricate, or a rhotic, to a fricative or approximant continuant.

There are some contexts which seem to encourage lenition, and others that apparently block it. The former include syllable codas, word-final positions, and intervocalic positions. The latter include word-initial positions, and onsets of stressed syllables (Kirchner, 2013). However, the position where the phenomenon occurs does not contribute to definitions of lenition: even if

[27] As it has been traditionally defined.
[28] With "complex" here referring to fricatives requiring greater muscular control for accurate articulation than stops do (Ladefoged and Maddieson, 1996).
[29] "The actual effort involved in a given utterance may vary with speech rate, loudness, the size of the speaker's jaw, the amount of air in the speaker's lungs, the presence of chewing gum in the speaker's mouth, etc. Certain aspects of speech production are undoubtedly sensitive to such idiosyncratic, token specific conditions; but other processes, including many documented lenition processes, are stable across tokens" (Kirchner, 2013, 29).

there are contexts that are more prone to host the phenomenon due to articulatory physiology, lenition might not appear in those contexts, and occur in "unexpected" ones (Bauer, 2008). As we described in §2.1., trills in particular, but also taps, are particularly complex segments, from an articulatory standpoint. Hence, many authors describe some patterns of rhotic variation as patterns of lenition, since it is plausible that complex segments undergo target undershoot and weaken. Under this account, variants of trills and taps could emerge as a user-friendly strategy of minimization of the higher articulatory difficulty (Jaworski, 2010). As a consequence, some scholars consider fricatives and/or approximants as lenited rhotic variants (Barry, 1997; Jaworski, 2010; Jaworski and Gillian, 2011; Gillian and Jaworski, 2014; Kirchner, 2013; Sebregts, 2014; Rennicke, 2015).

For instance, Gillian and Jaworski (2014) analyzed rhotic production in both Polish and Warlpiri speakers. If rhotic variation was due to reduction of the articulatory effort, i.e., it was acknowledgeable as lenition, then it would be caused by a universal mechanism and thus it should appear with similar features in languages as different as Polish and Warlpiri. As a matter of fact, Gillian and Jaworski found similar cross-linguistic patterns, and concluded that both the languages produce what can be defined as lenited rhotic variants – in particular, approximants rather than trills.

Thus, tapping, fricativization and approximantization of a trill are all regarded as possible lenition strategies (Jaworski, 2010). However, some of the aforementioned researchers also suggest that these realizations are linked as in a scale, and that they represent different degrees of rhotic reduction (Barry, 1997; Jaworski, 2010; Rennicke, 2015). The scale would go from the least sonorous to the most sonorous segment:

trill > tap > fricative > approximant

Empirical diachronic or apparent-time (cf. §3.1.) studies of rhotic variation and change could further support this type of analysis.

2.3. Distribution of rhotics across Canada and Italy

2.3.1. Rhotics in Canadian English

The classification of Canadian English (CE) is based on a series of isoglosses tied to variation in vowel systems (Trudgill and Hannah, 2008; Labov et al., 2006; Boberg, 2008). Therefore, we can assume that rhotics are not affected by such taxonomies; a simplification can be made by considering their behavior homogeneous in Canada.

Except for some outstanding distinguishing vowel phenomena, such as the Canadian Shift, the *Atlas of North American English (ANAE)* by Labov et al. (2006) describes CE as "essentially a North-American variety very similar to that spoken in the Midland and Western regions of the United States" (Labov et al., 2006, 216). Moreover, Wells (1982) explicitly states that the Canadian consonant system agrees with that of General American. We will consequently analyze rhotics in CE as the ones from the rhotic dialects of North American English (NAE), always keeping in mind that the huge number of immigrants, ethnic minorities, and HL speakers in our area of interest is an important source of variation (Dollinger and Clarke, 2012).

As stated by ANAE and Wells (1982), most of NAE is a rhotic variety of English. Indeed, English is divided into rhotic and non-rhotic varieties, based on whether historic postvocalic /r/s, in prepausal and preconsonantal position, are pronounced (Ladefoged and Maddieson, 1996). Wells (1982) specifies that there are additional environments where historical /r/ may be lost through a process of R Dissimilation. These are the cases of orthographic /r/s in unstressed non-final syllables adjacent to an /r/ belonging to another syllable, such as in the words "su**r**prise" or "gove**r**nor".

Consonantal /r/s in NAE are associated with many different articulatory realizations, varying intra and inter speaker. Based on Delattre and Freeman (1968) and Wells (1982), we can affirm that all of them are approximant and voiced realizations. In particular, two of the articulatory settings have been largely discussed by phoneticians: the "bunched" /r/ configuration and the "retroflex" /r/ configuration[30]. Studies have built on these two large categories, now reducing the number of settings by grouping similar ones together, now exploring narrower distinctions among shapes (examples are Espy Wilson et al., 2000; Mielke et al., 2010). Parts of the critiques to the bunched vs. retroflex dichotomy are due to the ambiguity of some shapes, which do not fit in either of the labels (e.g., "blade up"). The conclusion drawn from the debate is that /r/-related tongue movements represent an articulatory continuum (Westbury et al., 1998) of approximant realizations, whose endpoints are bunched and retroflex /r/ (Ong and Stone, 1998; Guenther et al., 1999; Espy Wilson et al., 2000; Zhang et al., 2003). As for the distribution of these two diffused articulatory settings, based on previous research[31], it seems that we find

[30] The simplest definitions of the prototypical retroflex and bunched /r/s are, for the former, a sound produced while lifting and curling the tip of the tongue; for the latter, a sound produced while keeping the tip down and lifting the dorsum. More in detail, NAE /r/s are realized using three constrictions: at the lips, at the mid-palatal region, and at the pharynx (Delattre and Freeman, 1968). The constriction at the lips is realized via various degrees of rounding; the one at the mid-palatal region also shows great variability, being realized with the apex, the blade, or the dorsum of the tongue; the one at the pharynx is reported to be, according to Delattre and Freeman, the main difference between American /r/s and British /r/s, where it should be absent.

[31] Delattre and Freeman, 1968; Ong and Stone, 1998; Mielke et al., 2010; Tiede et al., 2011.

retroflex and bunched /r/s when it is more "natural" for them to be[32], but such a straightforward relationship between /r/ allophonic patterns and phonetic naturalness has also been critically reviewed by Mielke et al. (2010; 2016)[33].

An issue that stands out immediately is how speakers are able to trace back all the different articulatory configurations to the phoneme /r/[34]. The main reason is that, curiously, articulatory variability corresponds to acoustic homogeneity. In fact, the most salient and consistent feature of AE /r/ is its low F3, almost correspondent to the value of F2[35], and with a value between the 60% and the 80% of the average vowel F3 (Zhou et al., 2008). Hagiwara (1995) affirms that this lowering consists in a drop of F3 below 2000 Hz for men, and a proportional drop for women, whose formants are generally higher. The low F3 is likely to be obtained thanks to the three simultaneous constrictions that characterize NAE consonantal /r/[36] (Guenther et al., 1999; Westbury et al., 1998)[37]. Another proof of the consistency of the F3 behavior is provided by Boyce et al. (1997). According to this research, the shape and duration of the F3 trajectory is not influenced by the preceding segments, despite their place of articulation. Therefore, the authors suggested that the many articulatory gestures used to realize /r/ in NAE all have the same purpose: achieving this consistent acoustic pattern. The choice among the possible articulatory settings for /r/ may be dependent on the requirements of the context and on individual anatomy. Guenther et al. (1999) agree. These positions imply the idea that listeners build abstract sound categories based on acoustic signals, more than gestural cues, and then, as speakers, they plan their gestures targeting this acoustic representation, adapting their

[32] I.e., speakers tend to use bunched /r/s in an already bunched context, and retroflex /r/s in contexts lacking competing tongue shapes.

[33] Mielke et al. (2010) highlight that naturalness is highly idiosyncratic and therefore not simple, generating complex patterns of conditioning environments. What is natural for one speaker could be unnatural for another (Mielke et al., 2016). Yes, they report a context-dependent distribution of retroflex and bunched /r/s, but they also claim that, if we think of the environments as a scale of compatibility with retroflexion, each speaker who uses both bunching and retroflection draws a boundary between bunched and retroflex differently, based on their specific motivations.

[34] Vice versa, Guenther et al. (1999) focus on the issue of speakers being able to *produce* such perceptually acceptable /r/s while varying so much their vocal tract shapes. They speculate that these articulatory settings are systematically related, involving the so-called trading relations (Guenther et al., 1999, 2855). The authors propose three ways of changing the vocal-tract shape in order to reach acoustic stability. Should one of them be made difficult by phonetic context, speakers would be expected to exploit the other two. Their study furnishes evidence supporting this idea, showing tradeoffs between the length of the front cavity and the length and/or area of the palatal constriction.

[35] Boyce et al., 1997; Westbury et al., 1998; Guenther et al., 1999; Espy Wilson et al., 2000; Twist et al., 2007; Mielke et al., 2010, Lawson et al., 2013.

[36] Delattre and Freeman, 1968; Boyce et al., 1997; Westbury et al., 1998; Espy Wilson et al., 2000.

[37] The acoustic difference between realizations - if existent - could lie in the highest formants; Zhou et al. (2008) observed that retroflex shapes have more frequency spacing between F4 and F5 than bunched shapes. However, the authors themselves (2008, 4467) highlighted that "the process of speech perception appears to largely depend on the pattern of the first three formants", suggesting that these differences could not be relevant for the human ear.

movements, and using strategies fitted for their individual anatomy and the phonetic and phonological context[38].

Crucially, the allophonic variation of /r/ in NAE not only has no acoustic correlate, but is also not perceptible by speakers. Twist et al. (2007) (and later, Mielke et al. 2016) affirm that it is precisely the perceptibility of an allophonic pattern that allows it to be conventionalized in a speech community, and no previous research indicated that this is valid for /r/ variability. As a consequence, the authors tested the perception of /r/ allophonic patterns and reported that speakers are only weakly aware of them. They also explain that this is the reason why speaker feel free to use different articulations: because not one of them comes with social awareness. Variation, since it is undetected by the speakers, is not socially imbued in this case, and may be much more heterogeneous and arbitrary. This would be why no dialectal patterns have been observed. In fact, a question that could come to mind is whether the different tongue shapes have a regional distribution across North America (i.e., they follow a dialectal pattern), but Boyce et al. (2015) showed that speakers from the Cincinnati region, OH, use a large range of tongue shapes. Their study confirms similar results from previous research. While usually idiosyncratic forms emerge due to *both* speakers' own biological features (e.g., anatomical and physiological ones) and personal reaction to social pressure (Celata, 2014), NAE /r/ seems to follow a different route.

A final remark that could be made regards the source of all information about NAE /r/: laboratory speech. For obvious reasons, in-depth analysis of acoustic or articulatory feature are based on controlled and somehow artificial speech, usually monological. It would be interesting to see how rhotics would behave in the context of informal and spontaneous dialogues, mainly for two reasons. First, such a context would probably generate a situation of hypoarticulation (Lindblom, 1990), where the effects of coarticulation and patterns of reduction are much more pronounced. Second, dialogue may trigger automatic or semi-automatic mechanisms of simplification of production and comprehension due to the interactive cooperation and/or negotiation between interlocutors (Pickering and Garrod, 2004).

A possible solution to the problem is to introduce in the phonetic laboratory the notion of social network (Milroy, 1980), as Celata (2014) proposed and operationalized. The idea would be to record in the laboratory the speech production of a small number of speakers during

[38] This position resembles Ohala's (1992) many-to-one mappings: a vocal tract configuration produces a specific acoustic profile, while an acoustic profile can be the result of more configurations. Therefore, listener/speakers could use different articulations for a phonetic object than those used by the speaker who they learnt the object from. Acoustic signals do not dictate the interpretation of segments, but are used to categorize them (Chabot, 2019).

dialogical interactions with people occupying different positions in their network, in order to catch not only more concrete interactions, but also different patterns of speech negotiation, depending on the relationship with the interlocutor.

Other than the consonantal /r/s, NAE also features r-colored vowels and the rhotic allophone mentioned in §2.1.2, i.e., the allophone of post-stress pre-vocalic alveolar stops (alveolar flap or tap [ɾ], as in "city" or "ladder"). However, since our focus is on consonantal /r/, we will not illustrate these further.

In summary, /r/s in CE are approximants articulated in various ways all sharing an additional constriction compared to British English. Acoustically, these realizations are characterized by the acoustic feature of a low F3 (as r-colored vowels - Lindau, 1978; Ladefoged and Johnson, 2014). Their distribution may follow patterns of idiosyncratic phonetic naturalness. None of them seem to be socially indexed. We will see how this potentially contrasts with Italian and Calabrese Regional Italian rhotics (cf. §2.3.2.1; §2.3.2.2.).

2.3.2. Rhotics in Italy

Within the multidimensional continuum of continua (Grassi et al., 1997) that describes the Italian linguistic repertoire, let us briefly distinguish Standard Italian (SI) from Regional Italian (RI). As is well known, it was the Italo-Romance dialect (e.g., derived from Vulgar Latin, therefore primary dialects according to Coseriu, 1980) Florentine that later became modern SI. RIs originated from the prolonged contact between the other Italo-Romance dialects and SI. Therefore, RIs are varieties of Italian on the diatopic axis (Crocco, 2017, 91) which may also vary along the other axes. Their identifying features mainly come from local dialects (Grassi et al., 1997). Nowadays, RIs are the current spoken Italian (Poggi Salani, 2010; Crocco, 2017); SI, instead, is a linguistic abstraction with no native speakers (Cerruti, Crocco, and Marzo, 2017, 7).

SI is the language codified by traditional accounts of Italian grammar, and whose morphologic, syntactic, and lexical traits are prescribed by textbooks and dictionaries. Those features come from written and formal Italian, and they are supposed to be shared by all Italian speakers despite their regional origin[39]. SI phonetic and phonological features are not taught in

[39] SI does not acknowledge a series of linguistic phenomena that originated in spoken informal language, and that overtime have become acceptable, and well-established within the whole community even in formal and educated speech and writing (Cerruti, Crocco, and Marco, 2017, 8). As a matter of fact, due to their prescriptivist nature, SI descriptions are slow and reluctant to change, with the result that not only do they miss some of the so-called "neo-standard" features (Berruto, 2012 [1987]; e.g., the use of the indicative mood at the expense of the subjunctive),

schools, but they have been codified nonetheless. Several pronunciation models followed during the years, starting with a norm based on a "cultivated Florentine pronunciation purified from local features" (Crocco, 2017, 92), continuing with a norm based on Roman pronunciation during the Fascist era, up to Canepari's (1999) model known as *modern neutral* Italian pronunciation. However, all these standard models are only followed by voice professional, while the rest of population uses pronunciations with different degrees of regionality and idiosyncrasy. We can think of the current standard pronunciation model as a norm that "represents the traditional reference and is recommended by several fundamental works on the pronunciation of Italian" (Crocco, 2017, 92). Hence, SI phonetics and phonology can be used as a starting point and/or a baseline for descriptions based on empirical observation of the speech community. These usually build on the SI system by adding more features - some of which are found nation-wide, while others coming from RIs. Examples are the approximant variants of /r/, that are featured by speakers in all parts of Italy according to empirical research – cf. §2.3.2.1., but are not acknowledged by SI, and the retroflex variants of /r/, that are featured just by some extreme-Southern speakers and that are not acknowledged by SI.

Section 3.2.1. will describe the main realizations of r-sounds in the whole peninsula. Then section 3.2.2. will focus on Calabrese RI and dialect, since our work is based on the analysis of speakers of Calabrese origin.

2.3.2.1. Rhotics in Italian

Let us begin with the descriptions of Italian rhotics based on traditional accounts – i.e., with the rhotics codified by SI. We will use the r-sounds from the standard pronunciation model to identify a baseline free from region-specific traits, and we will build on this baseline by considering empirical studies of rhotics as they are used in spontaneous speech by Italian speakers from different regions.

Italian contrasts singleton vs. geminate consonants; this applies to /r/ too, with minimal pairs such as *caro,* "dear" vs. *carro,* "cart". According to Ladefoged and Maddieson (1996), intervocalic /r/s are realized as apical trills featuring one or two contacts when singleton, and from five to seven contacts when geminate. Other authors, instead, describe the one or two

but they also include some anachronistic features that have fallen into disuse (e.g., the tri-partite system of demonstrative adjectives *questo, codesto, quello* ("this", "that" - far from the speaker, but close to the hearer, "that" - far from both participants), or mid-high and mid-low vowels being contrastive, with minimal pairs such as [ˈbɔtte] vs. [ˈbotte], "hit" vs. "barrel", and [ˈpɛska] vs. [ˈpeska], "peach" vs. "fishing").

contacts rhotic as a reduced trill, more specifically, as a tap (Bertinetto and Loporcaro, 2005; Canepari, 1999). Romano (2013) specifies that the timings of the phases of closure and opening of these two sounds can be the ones reported by Ladefoged and Maddieson (1996, 218) (cf. §2.1.1., §2.1.2.): 25 ms each. However, there is no consent regarding the distribution of trills and taps in SI. Table 2 summarizes different Italian authors' proposals, which are in some cases conflicting:

Author(s)	**Taps**	**Trills**
Canepari (1999)	In unstressed syllables; [ˈVrːɾV, VɾˈrV, (ˌ)VɾɾV, Vɾ(ˌ)ɾV]	In stressed syllables; [ˈVrːɾV, VɾˈrV]
Bertinetto (2010)	Intervocalic singletons	After a pause; in /rC/ clusters; geminates
Romano (2013)[40]	Intervocalic unstressed singletons; in the explosive phase of /rr/[41]	In heterosyllabic clusters /rC/ and stressed syllables (2 or 3-strike trill); geminates (5 strike trill)
Celata, Vietti, and Spreafico (2019)	Intervocalic singletons; word and utterance initially; in heterosyllabic clusters /rC/	In all other contexts; in heterosyllabic clusters /rC/[42]

Table 2: Distribution of taps and trills in Standard Italian according to different Italian scholars.

The main constant we can notice is the singleton tap vs. geminate trill contrast, but no empirical evidence is available to confirm this SI description. Studies analyzing intervocalic singletons and geminates do not include rhotics (Celata, Vietti and Spreafico, 2019)[43]. Moreover, for the tautosyllabic clusters /Cr/ or /CCr/, Farnetani and Kori (1986) claim that trills strongly reduced their duration compared to intervocalic contexts, and are usually realized as taps.

[40] This is the only study described by the author as "instrumentally tested".
[41] Romano considered single-strike rhotics as taps. He also highlighted that "they may have a closing phase longer than 50 ms, which is slightly (and suspiciously) higher than the one usually measured for taps in other languages" (Romano, 2013, 214).
[42] Heterosyllabic clusters can either have a trill, or its reduced version, e.g., a tap.
[43] The only exceptions could be Celata, Vietti and Spreafico (2019), who investigate this subject in two speakers of a mid-coastal Tuscan RI, and Celata and Costamagna (2012, 128), who only report duration data from an Umbrian speaker, stating that geminate rhotics are more than three times longer than singletons. These two studies use a limited sample and focused only on the Central region of Italy.

Trills and taps aside, a wide range of realizations characterizes actual, spontaneous, Italian /r/s. These are the realizations that SI models exclude from the phonetic description of Italian. As we claimed in §2.3.2., many of them can be specific features of RIs, but some are so diffused across regions and styles that they cannot be simply traced back to diatopic variation. Indeed, there is the large diffusion of "defective" variants - commonly known as "r moscia" (limp or lifeless r-sounds) (Romano, 2013). Under the lifeless r-sound label have been grouped very different articulations; unfortunately, this category has not been instrumentally tested and described in-depth yet. Nor is there agreement on its sociolinguistic status. While foreign scholars like Ladefoged and Maddieson (1996, 226) describe one type of lifeless r - the uvular trill - as a property of a prestige dialect, and Chambers and Trudgill (1998, 191) as a property of educated speech, in Italy, currently lifeless r's are considered "snobbish" or symbols of affectation, rather than education, by some, and pathological by others (Romano, 2013). This ambiguity is also depending on what sounds each individual considers a lifeless r, since the label includes such different gestures.

According to Canepari (1999), some lifeless r's are so diffused in a specific area that they should be considered typical of some regions, such as apical uvularized trills in different regions of North Italy, or retroflex fricatives for word-initial long trills in Sicily and southern Calabria, and in the same regions, post-alveolar stops or affricates for the -tr- and -dr- clusters. Romano (2013), however, is narrower in his definition on what sounds belong to the category. He states that, while the most common idea of lifeless r may be a uvular r, the most diffused variants in reality are labiodental approximants, often velarized or uvularized. Furthermore, according to the author, speakers "affected" by lifeless r "tend to occasionally allow the back articulation to prevail or to realise simple wavings between vowels, sometimes even yielding to no gesture traces at all" (Romano, 2013, 218). As a consequence, speakers may also show a variety of back r-sounds in Italy, which he provides a list of[44].

As the Italian actually used by speakers from different parts of the peninsula usually displays geographical variation, if one wanted to list the most frequent, spontaneous rhotics in Italy, analyses of corpora that include speakers from all regions would be needed. Unfortunately, at least to my knowledge, such a comprehensive study has not been conducted yet, with the exception of Vietti, Spreafico and Romano (2010). These three authors and Celata, indeed, have reported on different occasions[45] a lack of scientific literature about the actual status of Italian r-sounds, their regional realizations, and the distinction between singletons and geminates,

[44] For the full list of Italian back r-sounds, see Romano (2013, 220).
[45] See Vietti, Spreafico and Romano (2010), Romano (2013), and Celata, Vietti and Spreafico (2019).

further causing disagreement on the sociolinguistic status of some variants and the lack of a complete reference framework[46].

The aforementioned study from Vietti, Spreafico and Romano (2010) investigates the production of the /r/s in the word *frigorifero*, "fridge", from 8 speakers from 14 different cities of Italy. The authors found a 38% of apicoalveolar taps with a particular energetic profile that makes them be perceived as more intense compared to the ones in Iberian varieties (which are described as "normal" taps). They are still classified as taps based on the number of constrictions, i.e., just one. 6% of occurrences are instead "normal" taps.

Other variants found are:

- Uvularized, velarized, and pharyngealized approximants (5%) or obstruents (5%);
- (Lateralized) flaps in Venice only (North-East) (6%);
- Approximants (25%);
- Deletion/Omission (3.3%);
- Forms of vowel rhotacization (1.7%).

From this literature review, we are tempted to suggest that the most diffused and "non-defective" variants of /r/ in Italy are apical taps and approximants, with apical trills for heterosyllabic /rC/ clusters and geminate /r/. "Defective" variants are possibly socially indexed, but no agreement has been reached on their precise sociolinguistic status.

2.3.2.2. Rhotics in Calabrese Regional Italian and dialect

There are two striking aspects of the rhotics of Southern Calabrese RI and dialect. The first one is the geminate pronunciation of word initial /r/, as in *la ruota*, [la rˈɾwɔːta], "the wheel", especially in the province of Reggio Calabria (De Blasi, 2014). The second is that rhotics are often involved in processes of retroflexion. The retroflexion of the clusters -str-, -ntr-, -t(t)r-, -dr- (and also -ll-, which may have a rhotacized retroflex output) is a feature of some Calabrese dialects, as well as of other Italo-Romance dialects (from Apulia, Abruzzo, Campania, Sicily, Sardinia, Corsica and even Northern Tuscan) (Romito and Scuticchio, 2008; Celata, 2010).

[46] There is some more research about rhotics in specific RI or Italo-Romance dialects, such as Celata (2014), Celata, Meluzzi and Ricci (2016) and Meluzzi and Celata (2020) on /r/ in Sicilian RI and dialect; Spreafico and Vietti (2016) on /r/ in Bozen; Celata, Vietti and Spreafico (2018) on /r/ (and /rr/) in mid-coastal Tuscan RI; Spreafico et al. (2015) on /r/ in Western Tuscan RI. Since some Italo-Romance dialects and RIs (especially Southern ones) feature retroflexion of clusters involving rhotics, there are also studies specific to the matter, such as Sorianello and Mancuso (1998), Celata (2004), Celata (2010), and Romito and Scuticchio (2008).

Moreover, Celata (2004) states that geminate rhotics preceded by a vowel or a pause, including word initial ones, are also affected by retroflexion[47].

Nevertheless, apart from a handful of studies about the -ll- cluster in Calabrese dialects, Calabrese rhotics have not received much attention. Therefore, in order to describe the possible characteristics of rhotics and retroflexes in this area, we have to resort to sources investigating the varieties closest to Calabrese RI and dialect. From a dialectologist perspective, "RIs are divided into northern, central, southern, and Sardinian varieties (Pellegrini, 1977). The southern area is further grouped into high-southern and low-southern, with Calabria overlapping both. Our speakers come from parts of Calabria in the low-southern region" (Baird, Cristiano and Nagy, 2021, 2). Therefore, studies analyzing Sicilian varieties would fit our purposes, since they belong to the same group as Calabrese ones (i.e., the low-southern area). In fact, they also show the two phenomena we listed at the beginning of this section (Celata, 2010).

Obviously, collecting information about RI features based on the corresponding local dialect may introduce error, since we cannot know *a priori* which elements would be shared by the two varieties. However, "considering the corresponding dialect to glean information about the RI is justified, since, according to Telmon (1993, 96), the dialect is the clearest means to understand and interpret the RI" (Baird, Cristiano, and Nagy, 2021). Moreover, the general lack of studies on the relevant topic forces us to search for rhotics' analyses within the dialectological literature (where we may find patterns potentially transmittable to RI) when RI literature is not available.

Let us focus on the second salient phenomenon we listed, the retroflexion of consonant clusters involving an alveolar or dental (alveodental from now on) stop followed by a trill, such as [ˈʈˢɛːni] for the Italian [ˈtrɛːni], "trains" (Celata, 2010, 7).

Interestingly, as we can see from the example, in the Italo-Romance dialects that feature the phenomenon the retroflex realization actually corresponds to a monophonemic voiceless postalveolar affricate [ʈ(ː)ˢ] (Sorianello and Mancuso, 1998), and not to a biphonemic cluster made of a retroflex stop and retroflex rhotic – [ʈ(ː)ɽ]. This is an outcome that neither Bhat (1973) nor Hamann (2003) described, as we can see from the reconstruction hypothesis by Hamann illustrated in §2.1.4. (Celata, 2004, who also supposes the presence of such an output in other languages of the world).

[47] The process of retroflexion of these sounds would lead to a similar outcome postulated for the other clusters mentioned: a voiced retroflex fricative, considered postalveolar, [ʐ̟] (Ruffino, 1991; Sorianello and Mancuso, 1998).

Sorianello and Mancuso (1998, 147) explained this outcome by adding an ulterior step to the reconstruction proposed by Hamann:

(c) tr > tɽ > ʈɽ > ʈˢ

Step 4, involving the passage from the voiced retroflex rhotic [ɽ] to the voiceless retroflex fricative [ʂ], would be due to the frequent fricativization of retroflex rhotics.

The starting point of Hamann and Sorianello and Mancuso's reconstructions is the retroflexion of the post-consonantal rhotic. However, based on their explanations only, it is not clear why this process is not also triggered in all the other contexts, or why it does not involve non-coronal consonants, such as [p] or [k]. Celata (2004) tried to provide a more extensive explanation for the process, while also proposing another reconstruction. Let us illustrate her reasoning.

First of all, we have seen that, on the one hand, apical trills are particularly resistant to coarticulation, while exerting a great coarticulatory pressure on adjacent segments. On the other hand, they tend to be articulatorily reduced due to their many physiological constraints (cf. §2.1.1.). These reduction processes, that may have various outcomes (cf. §2.1.1.), are context sensitive: they are more diffused in syllable-final position (Ohala and Kawasaki, 1984; Recasens, 2004), as Farnetani and Kori (1986) confirmed for the Italian clusters /Cr/ and /CCr/, where not only the rhotic was shorter, but was also realized not as a trill, but as a tap.

Now, the context we are considering for retroflexion has the following features: it is tautosyllabic; the rhotic occupies the second or third position of the syllable; it involves a stop; this stop has almost the same place of articulation of the rhotic. Therefore, the first two characteristics make this context particularly fitting for the reduction of the rhotic compared to other contexts.

Celata's hypothesis is that a gradual process of reduction caused a backing of the place of articulation of the trill, and a gradual change of its manner of articulation, from trill, to tap, to approximant, without necessarily phonologization of each step. While this segment lost its status of independent consonant, the first segment of the cluster kept its original intensity level, since it is a stop and occupies the onset of the syllable. A retroflex pronunciation may have appeared at any stage without affecting the phonological and/or perceptive features of the sequence (recall that retroflexes and apical rhotics share the low F3). Successively, the alveodental stop and apical approximant sequence undergoes a process of affrication, following a frequently attested phonetic pattern. This reconstruction is exemplified by the following sequence (Celata, 2004):

(d) [tr] > [tɾ] ~ [tɽ] > [tɹ] ~ [tɻ] > [tʂ]

Celata decided to not specify the precise moment where the place of articulation became post-alveolar, but this is not considered as a problem, since apicality and retraction are always present, at some rate, in the articulation of trills. Nevertheless, one could still ask why this process involved only the clusters with alveodental stops, and not the other ones. The answer, according to Celata, is in the process of blending, where two adjacent segments realized with close places of articulation merge into an intermediate realization, reducing the articulatory effort (Recasens, 1999).

This reconstruction differs from previous ones because it is not based on regressive assimilation, but on the affrication of the whole cluster; it focuses on the particular contextual factors that triggered reduction, and on the homorganicity of the cluster, which involves anterior coronals.

As we stated earlier, there is a lack of studies specific to Calabrese varieties. Therefore, for a quantitative analysis of the realizations of the relevant clusters we have to resort to Celata (2010), who investigated dialectal speech from five Sicilian speakers. Speech was collected from dialogues between speakers on topic suggested by the experimenter. /t(ː)r/ contexts were realized as [ʈ(ː)ʂ] in 69.2% of the cases, and as [t(ː)r] in the remaining 30.8%. Moreover, "according to the spectral data, [ʈ(ː)ʂ] exhibits essentially the same place of articulation as the other postalveolar consonant [t(ː)ʃ] […]. Among the four parameters proposed by Hamann (2003, 19) for characterizing the articulation of retroflex consonants […], apicality and posteriority appear to be relevant for Sicilian [ʈ(ː)ʂ]" (Celata, 2010, 23-24).

Excluding the discussion around retroflexion of consonantal clusters, there is no empirical evidence regarding the realization of intervocalic Calabrese /r/s – at least to my knowledge. The closest studies on the topic would be analyses of Sicilian varieties, such as Meluzzi and Celata (2020).

The two researchers investigated singleton and geminate intervocalic rhotics produced by a Sicilian target speaker and two interlocutors (Peer 1 and 2), and elicited from a sentence reading task and from map-task dialogues, both in Sicilian dialect and Sicilian RI. Let us illustrate the results for RI.

Concerning singletons, in the sentence reading task, the target speaker produced 91% apical taps, and 9% alveolar approximants and apical trills. He followed the same pattern in dialogues, as the two interlocutors did (even though Peer 2 showed 77% taps and a larger number of approximants).

As far as geminates are concerned, in the sentence reading task, the target speaker produced 90% trills and few fricatives[48]. He followed the same pattern in the dialogue with Peer 2, showing no taps. Peer 2 produced 53% trills, and 47% approximants. In the dialogue with Peer 1, instead, the target speaker dropped to 75% trills, and included in his production 12% taps. This difference in the production pattern of the target speakers has been interpreted by the authors as convergence to Peer 1, who produced 56% taps, 31% trills, and 13% approximants.

We stated in §2.3.1. that NAE realizations of /r/ do not seem to be socially indexed. In contrast, some realizations in Calabrese dialect and RI seem to be. There are several sources suggesting that the retroflex realization of clusters involving rhotics may be socially indexed in Sicilian dialects and RIs. No information regarding their sociolinguistic status in Calabrese varieties is provided. However, we could imagine that a similar situation may apply in Calabria. In Sicily, indeed, such retroflex pronunciations are extremely salient and recognized by speakers, and they represent "a marker of ethno-linguistic identity" (Celata, 2010, 8). Therefore, individuals - especially younger and more educated ones - may try to avoid it when speaking RI due to the *dilalia*[49] dynamics between dialects and SI, according to Celata (2010). Nonetheless, as we have seen, this study also shows that Sicilian speakers produce the majority of the relevant clusters as retroflex. Therefore, two explanations may be possible. On the one hand, it could be that the speech elicited was truly spontaneous and that speakers did not try to control their pronunciation. On the other hand, it might be possible that, despite trying, they are not able to avoid the phenomenon in most cases. This latter hypothesis would be influenced also by the peculiar dialect-Standard dynamics characterizing the South of Italy. As a matter of fact, this dialectic manifests in variable ways across the peninsula: Southern dialects would be much more vital than North-Western ones, for example, affecting everyday speech to a larger extent and possibly making it harder for speakers to avoid pervasive dialectal traits. Moreover, while dialects would be losing ground as exclusive L1s, they would be gaining it as second[50] L1s (with RI being the other L1) or L2s for the youngest Southern speakers – possibly also due to an agentive use of the dialect during interactions with peers (Meluzzi and Celata, 2020).

Dialect-standard dynamics, in these cases, are particularly ambiguous: the dialect is perceived both with a positive and negative connotation. It is part of the local identity and tradition, and therefore a possible emotional attachment to it may manifest, but it is also

[48] The authors collapsed under the label "fricative" voiced retroflex fricatives [ʐː] and two-constriction tap-fricative variants or spirantized rhotics [ɾ̝̊], since they are probably perceptually identical.
[49] Cf. Berruto (1995).
[50] With the term "second" used not to refer to a chronological order of acquisition, but to the second member of a couple of languages acquired and being hierarchically equal.

considered as "provincial" and socially subordinated, and therefore avoided. This would also be due to past institutional censorship of dialects that could still affect speakers' behaviors. The review of linguistic autobiographies of Calabrese teenagers by Nodari (2017) testifies this bivalent behavior. On the one hand, parents do not teach dialects and scold their children for using them; participants describe their dialect negatively, report that teachers punish those who use it at school, and feel shame in admitting knowing it. On the other hand, kids report using it among peers in informal contexts. This dichotomy would translate into what Auer (2005) describes as an increasing regionalization of local vernaculars, counterbalanced by a dialectization of the standard.

Especially in a constructivist view of language, speakers may try to build and/or declare their identity (including ethnicity) through speech and selecting variants, possibly variants that were originally a marker of geographical origin or social status. We could suppose that the quantity of vernacular traits featured in RI speech could also depend on how much "regionality" is valued across the community. While it seems that regionality is not ranked as a positive value in Calabria – but with a behavior that partially contrasts this view, we cannot assume that the same would apply to HL speakers of Calabrese origin living in Toronto, whose attachment to the Homeland and its ethnic and linguistic identity may vary.

To exemplify how the case of HL speakers may be challenging for sociolinguistics and dialect/Standard dynamics, let us briefly return to the subject of coronal cluster retroflexion.

Interestingly, NAE speakers might have the postalveolar affricates [ʧ] and [ʤ] in -tr- and -dr- clusters (Scobbie, 2006). There seems to be no evidence of social evaluation of the affricate variants. A comparison between the American and the Calabrese Italian realizations of -tr- could shed light on their articulatory and acoustic similarities. What seems clear, however, is that the two varieties at least share a similar behavior in similar contexts. For a speaker of both Calabrese Italian and NAE, this would mean a “simplification” of their sound system(s), since a certain region would be shared or would even overlap. In Flege’s SLM model (illustrated in §1.5.1.), a speaker of one of the two varieties would probably already have a space in their sound system dedicated to the phenomenon of -tr- and -dr- affrication. When first encountering a similar phenomenon in the L2 (or 2L1) variety, they could judge it as an instance of the L1 -tr-/-dr- affrication process, equating the phenomena of the two varieties, and letting them share the same region of the phonological space. Since speech acquisition is a never-ending process in Flege’s model, the perception of the two phenomena may change, and speakers could eventually start to perceive a difference between NAE -tr-/-dr- affrication and Calabrese Italian -tr-/-dr- affrication. In this case, their -tr-/-dr- affrication mental representation would change

to include features of the phenomena from both varieties. In order to verify such a hypothesis, a comparison between affricate -tr-/-dr- realizations of NAE, Calabrese Italian and bilingual speakers would be needed.

However, from a more sociolinguistic point of view, the representation and production of said affricated clusters could vary based on ethnic factors. We have argued that -tr-/-dr- affrication is likely to be negatively socio-indexed in Calabria. This would diminish the "strength" of affricated realizations within the sound system. The situation may be not the same for Calabrese Italian HL speakers.

On the one hand, HL speakers could have a different perception of locally marked traits than Homeland speakers, due to the different social contexts they experience. HL speakers with an emotional bond with their Heritage culture could feel the need to "perform" their ethnic identity through locally marked traits, while this need could be weaker among Homeland speakers, given the lack of contact with a distinct majority language. This could translate as higher rates of affrication among HL speakers with high Ethnic Orientation scores (cf. §1.2; §3.4.2.). If we think of it more in Flege's perspective, another possibility would be that speakers may try to differentiate -tr-/-dr- affrication between NAE and Calabrese Italian in order to keep distinct their different ethnic identities. In a way, they could want to say that Calabrese -tr-/-dr- affrication and NAE -tr-/-dr- affrication *are not the same thing*, because they belong to different cultures, ethnicities, identities. Therefore, Heritage speakers might exaggerate differences between the two realizations, preventing them from sharing the same sound system region[51].

On the other hand, HL speakers could be affected by an external source of influence not available to Homeland speakers: NAE. If -tr-/-dr- affrication was a pervasive phenomenon in NAE, speakers with a stronger orientation towards English culture and with higher rates of use of English (compared to Calabrese Italian), may produce more affricated -tr-/-dr- clusters than Homeland speakers, because their representation had been strengthened by contact with English.

These are just two examples of possible interaction between ethnic factors and speech production in a contact setting. Nevertheless, they demonstrate how the study of HL speakers provides both interesting challenges for sociolinguistics, and a different testing-field of hypotheses about the dialect/Standard interplay in the Italo-Romance context.

[51] For a discussion about real linguistic diversity vs. perceived, asserted, or built linguistic diversity, see Grandi (2020).

We have also conducted an exploratory analysis on -tr-/-dr- affrication among Homeland and Heritage Italian speakers, and Italo-Canadian speakers of English. Preliminary results will be illustrated in §3.6.4.

2.4. The nature of the class of rhotics

We opened this chapter by stating that the origin of the class of rhotics lies in non-phonological reasons, such as orthographic conventions, diachronic relationships, or IPA symbols. This has led us to have a conventional phonological category with ambiguous boundaries and a non-definite number of members, and left us with the problem of how to identify a criterion for membership. According to Chabot (2019), the reason for these issues is straightforward: since there is no correlated phonetical feature associated with rhotics, phonological theory has struggled and struggles to model them in representational, traditional, segmental, terms.

Indeed, we have described a wide range of articulatory behaviors for the sounds usually included in the class of rhotics, which therefore cannot be defined based on a unifying articulatory property shared among its members. Neither an acoustic property emerges as the clear unifying factor, as not all rhotics of the world share the proposed acoustic parameter, e.g., the low F3.

While some may doubt the existence of the phonological class itself, Chabot (2019), based on previous research, claims that there are good reasons to believe that rhotics do in fact represent a phonological real object. The question that remains is: how to describe it?

Lindau (1985) has suggested that the phonological class of rhotics has no physical correlate: the category emerges as a series of relations of Wittgensteinian family resemblance between sounds or subgroups of sounds. Such relations would also allow for points of connections between the category of rhotics and sounds belonging to other categories:

> Sounds with similar constriction locations are likely to have similar spectral properties whether or not they are 'rhotic'. Taps, flaps and trills all have similarities to stops because they all involve closure, and, indeed, often alternate with them. Fricative rhotics have obvious similarities to other fricatives. And so on. Although there are several well-defined subsets of sounds (trills, flaps, etc) that are included in the rhotic class, the overall unity of the group seems to rest mostly on the historical connections between these subgroups. (Ladefoged and Maddieson, 1996, 245)

Walsh Dickey (1997) also follows the family resemblance approach, arguing that this way, instead of having one phonetic property unifying rhotics, we would have a set of related

phonetic properties. Family resemblance would allow subclasses of rhotics – such as trills or taps – to be related to at least one other subclass in the network of shared properties (Wiese, 2001).

Nonetheless, the family resemblance approach has not been exempted from critiques: "For example, in the family resemblance model there is no direct connection between [ɾ] and [x] and yet speakers of Brazilian Portuguese must be able to relate both segments back to a single phonological object […]. Theoretically, the model is epistemologically problematic: the set of rhotics is assembled first, and then the phonetic resemblances are defined" (Chabot, 2019, 4); "it is hard to see how an unambiguous class of rhotics is established in this way" (Wiese, 2001, 340).

As a consequence, Chabot (2019) proposes to solve the problem differently, i.e., by applying Kaye's (2005) epistemological principle: "Phonological identity is only recoverable via observation of a segment's phonological behavior, both within the system and in phonological processing" (Chabot, 2019, 6). As a matter of fact, we have described in §2.2. a series of phonological patterns shared among rhotics. Chabot adds two other properties, stating that rhotics tend to be stable in phonological processes and diachronically, and then uses all these behaviors to define the phonological identity of the class itself. With "stability" he means that even if the phonetic realization may be subject to variation and change, phonotactical properties stay the same, with rhotics always functioning as sonorants – even when the realization is not sonorant. This means that variation and change would not alter the role rhotics occupy in the phonological system.

On the same page is Wiese (2001), who essentially agrees on the "stability" concept expressed by Chabot, but who goes one step forward, as he uses the distributional patterns of rhotics as the essence of the phonological class, but also tries to analyze what this means for traditional phonological theory. He proposes a prosodic and universal definition of rhotics, for which /r/ is represented by a point on the sonority scale that now we will provisionally define between laterals and glides. Indeed, he redefines the concept of sonority scale.

A traditional sonority scale is the one proposed by Clements (1990):

obstruent < nasal < liquid < glide < vowel

First of all, if we wanted to define the class of rhotics as the point they occupy *on the traditional scale*, then this point should never move. However, the manner - and therefore sonority in the standard sense - of r-sounds ranges from fricative to vocalic, exactly like the scale does, and this variation makes it impossible to place rhotics on the traditional scale. This

is why Wiese revises it, proposing instead that "the sonority hierarchy is nothing but an abstract ordering of points on a scale. The positions are defined not by their inherent segmental features (which seems impossible, at least in the case of /r/), but by nothing than their relative position in the scale" (Wiese, 2001, 356).

Under such a definition, rhotics would be represented as the position highlighted in bold in the following abstract sonority hierarchy (Wiese, 2001, 357):

P1 < P2 < P3 < **P4** < P5 < P6

This redefinition would also have the benefit of answering a frequent criticism to the sonority scale, i.e., its circularity. Indeed, on the one hand it seems that the positions occupied by segments within syllables derive from the sonority scale, but on the other hand, the sonority scale is constructed based on the observed sequential patterns of segments in syllables. "If we remove segmental features from the hierarchy, but leave the notion of relative ordering intact, large parts of that criticism do not apply anymore" (Wiese, 2001, 358).

Wiese demonstrates the validity of his idea using examples from French and German. Standard French /r/ is either a uvular approximant or a fricative. In French, we can have onsets made of some obstruent-sonorant clusters, but no obstruent-fricative clusters; however, we can have obstruent-fricative /r/ clusters. Therefore, French /r/, even if sometimes fricative, does not behave as a fricative in phonotactics, but as a sonorant, or rather, a sound in P4 position on the sonority scale proposed by Wiese.

There are two main problems related to Wiese's idea. The first would be finding a definition of rhotacized vowels, which would be rhotics and vowels at the same time. The author solves this issue by considering rhotacized vowels as retroflex vowels. The second would be how to treat rhotics in languages of the world that do not fit the common phonotactic pattern. Indeed, some initial r-sounds seems to behave differently; an example comes from Polish, where /r/ can appear word-initially and before obstruents, such as in [rt or [rv clusters ("[" denotes the left word boundary). Wiese solves this other issue by applying extrasyllabicity to these clusters, which would make them not analyzable using regular syllable-related principles. This solution, however, strongly relies on the idea of universal rules that, when revealed not to be universal, are claimed to have some *ad-hoc* exceptions in order to not be renounced. In the case of phonology, there seems to be a pattern of every exception to universal rules and processes proposed condensing at word-boundaries. Without denying that word-boundaries often show different behaviors than other prosodic positions, one might critique the tendency of universalist

theories to use word-boundaries almost as a scapegoat anytime a new universal proposal struggles with possible counterexamples.

In some models, however, the phonological issue of having variable phonetic forms, as for rhotics, would be simply avoided. Exemplar models applied to phonetics and phonology assume that our sound-knowledge descends from experience. Sound categories naturally emerge via clustering, starting from the storing of all the exemplars we are exposed to, together with their idiosyncratic details and semantic, contextual, social, and distributional information[52], and their consequent arrangement in a network organized on similarity relationships. Phonetic variability is not derived through rules or transformations, but it is directly stored. In a model such as Bybee's (2003), the storing strongly depends on frequency: every form is stored, but the more a form is used, the stronger it is memorized and the easier it is to access. The more frequent a form is, the more the motor activity necessary to pronounce it is repeated, the more automatized the process becomes, and the more frequent articulatory reduction occurs.

This way, we would not need a unifying phonetic feature allowing for the generation of the r-sound phonological class: it naturally creates based on relationships of similarity between some items. These similarities do not necessarily have to be acoustic or articulatory, but can just be distributional, for example. The result is a category whose limits are blurred. At the core, we would find the items that occur the most frequent, and that share the most properties - which makes them more similar and readily recognizable, i.e., "closer" in the network. At the periphery, there would be less frequent exemplars that are less similar to others, so whose features are not unambiguously rhotics, and that could also be confused with other sounds. The limits of the category would naturally blur into the limits of another category, with in-between items sharing some features (acoustic, and/or articulatory, and/or distributional, for instance). These items may be disambiguated via access to our experience, past knowledge, contextual information, social expectation, through comparison and matching of the ambiguous input to the most likely exemplar from the network (Vietti, 2012).

A strongly reduced /r/ in the Italian word *caro*, [ˈkaːɾo], “dear”, could be mistaken for a /v/, leading speakers to mistake the word *caro* for the word *cavo*, “empty”, in a sentence. However, speakers are able to disambiguate thanks to the idiosyncratic, semantic, and pragmatic information that was stored together with exemplars, and that allows us to know that, for instance:

[52] There is still no agreement about the precise amount and type of information that would be stored. For a discussion, see Vietti (2012).

a) the word *caro*, based on its meaning and contextual distribution, is more fitting in the relevant sentence than the word *cavo*;
b) the word *caro* is generally more frequent than *cavo*;
c) /r/ can occur in more or less reduced ways, some of which are perceptually close to /v/.

Obviously, an Exemplar approach to rhotics would still present some problems.

The first one is that such models would explain mechanisms of production and perception by speakers and listener treating them as language specific. Some scholars, instead, look for universally valid motivations allowing them to label rhotics as a natural class of sounds. In particular, an Exemplar model would fail to explain why the phonotactics of rhotics is so incredibly similar across languages, why some rhotics are typologically predominant, or why it is so frequent that rhotics, among many sounds, serve as social variables. It seems that Exemplar theories would explain the "how", but not the "why", and if one believes and searches for the "identity" or the "nature" of classes of sounds, they sure would not be the best models to resolve to[53].

As a final remark concerning rhotics' social meaning, as we hypothesized at the beginning of the chapter, maybe it is precisely the articulatory "elasticity" that rhotics tolerate without having speakers not recognize them or the words they appear in, compared to other sounds, the factor allowing for their frequent social indexing, but this explanation would simply move the question elsewhere. If this is true, then why do they vary so much without losing their "identity" as rhotics?

The most plausible yet temporary solution, in the end, seems a distributional-structuralist hypothesis, not so far from Chabot (2019) and Wiese (2001), but with two important caveats. First, the concept of structure is conceived neither as a pre-existing, superimposed, entity, nor as an explanation, but rather as an *explanandum*; second, distributional behavior is considered in a probabilistic, trending sense, without referring to universal impositions (and potentially, this type of solution could merge with the "how" described by Exemplar Theories). Therefore, perhaps rhotics are distinguished by the position they have a high probability to occur in, put in relation to the other positions and the sounds that tend to appear in them. In this sense, "rhotic" is a place: thus, its actual realization can vary, and in the case of inter-language variation, it does so because it relates to the specific phonological inventory and phonotactic patterns of the

[53] For detailed discussions about the limits or ambiguities of Exemplar theories, see Labov (2006) and Pierrehumbert (2006).

language considered. We have described this solution as temporary, however, for three main reasons:

1. It also fails to account for the typological predominance of certain types of rhotics;
2. It underlies the idea that certain sounds are not inherently rhotics, but they "become" ones only when they are inserted in a structure; the question then is, how can they be inserted in something that they also contribute to create, since we are refusing the notion of pre-existing structure?
3. It partially rests on a paradigm, structuralism, that has several criticalities. In his review of the problematic treatment of the notion of phoneme by European structuralism, Albano Leoni (2009) concludes that, for structuralism:

> Alla fine il fonema non è più né la rappresentazione psichica di un evento fisico, né una intenzione fonatoria, dunque individuale, ma è un non meglio definito elemento del quale si sa che svolge una funzione distintiva e che si colloca in una rete strutturata di relazioni con le altre unità, ma non si sa cosa sia[54]. (Albano Leoni, 2009, 100-101)

We agree with him: it seems that, by using a distributional-structuralist definition of rhotic, something "gets lost in translation".

[54] "In the end the phoneme is not either the psychic representation of a physical events, or a phonatory intention, thus individual, anymore, but it is an unspecified element of which we know that it performs a distinctive function, and that it positions in a structured network of relations with the other units, but we don't know what it is".

3. (r) in Heritage Italian

The present work explores patterns of variation and change in the pronunciation of (r) by Heritage Italian speakers living in Toronto. This research rests within the scope of the HLVC project, whose purposes, materials, and methods are replicated to ensure continuity and coherence across the project.

In a broader perspective, the first goal of this thesis project is to keep describing and analyzing HL speakers' communities, and the case of Italian in Toronto provides us with the chance of studying if (and how) variation and change differ when a majority, national, standardized language finds itself being a minority language in a tight contact situation with Canadian English. The second goal is to enrich the literature on rhotics in Italian: not only their production in spontaneous speech, but also their connection with social factors such as age, sex, ethnic orientation, and language use.

This chapter presents the case study. Section 3.1. will give an overview of the theoretical framework we employ, i.e., of the Comparative Variationist paradigm. Section 3.2. will describe the Italian community in Toronto, to give an idea of the social and political context our speakers are integrated in, and of the way they live their linguistic situation. In section 3.3. we will present a direct application of this approach, and the project this research falls within: the HLVC project and the HerLD corpus – which has been used for this thesis. Then, we will describe the case study we analyzed with this work: word-internal singleton (r). Section 3.4. will illustrate the research questions and hypotheses, and the methods used. Hence, it will report the procedure employed to select speakers and tokens, the variables considered and the way they have been coded, and the statistical methodology used for their analysis. Section 3.5. will present the results. Finally, section 3.6. will be dedicated to their discussion and cross-linguistic comparison with another Heritage language, with the final paragraph summarizing the conclusions we draw from the case study.

3.1. Theoretical framework

We describe in this section the theoretical paradigm our case study follows.

Languages are not unalterable entities: they are systems constantly changing and developing, and not only within the individual acquiring them, but especially in the community using and adapting them based on communicative needs (Magni, 2014).

As we know, languages have been studied from Saussure on using the diachronic or the synchronic approach. However, this dichotomy, a methodological simplification, has to be overcome especially when considering linguistic variation. Variation and change, indeed, affect both the diachronic and the synchronic dimension: they happen within time, space, and society, generating inter and intra linguistic variation (Magni, 2014).

Language variation occurs when a linguistic element, a variable, has more than one form, or variant (Gardiner and Nagy, 2017). Sometimes, variation stays stable over time, or forms keep alternating or competing[55]; other times, variation gradually becomes a change, i.e., one of the variants gradually replace the others, and the language goes from a status A to a status B in such a way that, at the end of the process, B represents the norm.

Until the mid-1960s, language change had been studied only in diachrony (e.g., with the historical-comparative method). Linguists fundamentally thought that one could not see the process of change, but just its outcomes. Therefore, any study had to formulate hypotheses in retrospect. Historical linguists, who were the most interested in the topic of change, usually examined written data from past moments of history, and so the modalities or causes of change remained unknown.

This paradigm changed with the publication of the paper by Weinreich, Labov, and Herzog (1958), *Empirical Foundations for a Theory of Language Change*, and of the following empirical works by Labov (1963, 1966) and (Bailey, G., 2003). The former established a new theoretical view on language change; the latter operationalized this theory into a new method of research, which was later addressed as variationist sociolinguistics. The roots of this change of perspective can be glimpsed in Philology, Dialectology, Sociology, and Anthropology (Tagliamonte, 2016).

Weinreich, Labov, and Herzog first considered the approaches by Paul and Saussure, which set the tone for the study of linguistics, and showed how, according to these two, variability and systematicity within languages excluded each other. In fact, after the creation of the saussurian dichotomy synchronic-diachronic, theories of language change had a different place than theories of language structure. Instead, the three authors introduced the concept of orderly heterogeneity: languages, both from a synchronic and diachronic view, possess a kind of

[55] See, for example, Aronoff (2016) for a review of competition in lexicon and morphology and Goldberg (2019) for competition among constructions.

variability that is not chaotic, and that can therefore be analyzed and "put in order". Variation and change are actually included within language structure: they are not random, but depend on speakers' choices constrained by both social and linguistic factors and that reflect the underlying grammatical system (Bailey, R. 2013). Language change is originated when a particular alternation in a subgroup of the speakers' community, i.e., variation, generalizes, becomes orderly differentiation, and spreads within the heterogeneous community as a whole, becoming the norm. This also means that not all variability turns into change, but that all change involves variability.

The three authors highlighted five problems that needed to be addressed within such a perspective:

1. Constraints: what are the limits to change (and variability)?
2. Transition: how are the intermediate states of changing made?
3. Embedding: what are the changes associated with an observed one, and how are they associated in a manner not ascribable to chance?
4. Evaluation: how does the community evaluate change, influencing its process?
5. Actuation, "perhaps the most basic: What factors can account for the actuation of changes? Why do changes in a structural feature take place in a particular language at a given time, but not in other languages with the same feature, or in the same language at other times?" (Weinreich, Labov, Herzog, 1958, 102).

Magni (2014) lists as the three main aspects any model of language change should account for:

1. Innovation: where and how do new variants emerge?
2. Propagation: how does this innovation succeed and spread?
3. Interpretation: how can we identify general principles or universal tendencies of language change?

Weinreich, Labov, and Herzog themselves claimed that the aspect of innovation – the actuation problem - still largely remained a mystery, and even today we hardly answer this question. According to Magni (2014), innovations are triggered by too many competing motivations - structural, functional, and cognitive.

As for the propagation problem, the sociolinguistic wave launched by Weinreich, Labov, and Herzog shed light on the social aspects which determined the success of the innovation, and its spread via social networks or language-contact situations; nonetheless, the propagation

of the innovation is also influenced by how well it functionally adapts to the pre-existing linguistic system (e.g., is there already a rule, category, pattern, or scheme in which the innovation would fit?) (Magni, 2014).

Weinreich, Labov, and Herzog's idea was later operationalized by Labov himself, and his work acted as a counterpoint to formalist generative methods. Labov's studies on Martha's Vineyard and New York City (1963, 1966) showed a way to quantify linguistic variation and to track down linguistic change during the process; in other words, he established a synchronic approach to language change (Bailey, G. 2003).

Change was observed when taking place thanks to Labov's innovative idea of the apparent-time construct. This hypothesis claims that if social and stylistic factors are constant, linguistic differences among generations of speakers would mirror real-life diachronic change. Labov then empirically examined how the apparent-time differences intersect with social and stylistic factors, and demonstrated that innovations enter the speech of a subgroup, and then gradually spread all over the community. So, while variation can mark class or sex differences, it can also mark linguistic change, and in this case, it correlates with age. The prototypical pattern of the change-in-progress, therefore, is the one whereby a variant has an increasing frequency down the age scale, from the oldest to the youngest generation (Chambers, 2003). The hypothesis of the apparent-time construct, however, relies on the assumption that idiolects stay stable through the course of adult life. Not all sociolinguistic situations conform to this conjecture. An example is age-grading, a situation where members of the speakers' community modify their speech at some point during their lives to conform with age-appropriate norms. It can potentially undo an incipient linguistic change (Chambers, 2003). Moreover, differences within age groups could also be due to linguistic features being restricted to certain groups: they would be not distinguished from change-in-progress variation (McMahon, 1994).

The study of variation and change, from these two cornerstones, has incorporated many other enriching ideas and constructs. Indeed, according to Eckert (2005), there have been three waves of sociolinguistics and therefore variationism.

Labov's pioneering studies and the following urban studies represent the first wave. It is characterized by the use of surveys and interviews across large populations and objectively predetermined socioeconomic categories. It also showed the correlations between the regional, informal, or ethnically distinctive variants and the basic socio-economic features of speakers (age, sex, class).

The second wave is characterized by a more ethnographic approach, for example relying on local social networks. Its goal is to use not pre-determined, objectively based, social categories,

but to extract them from internal indices, i.e., the perspective of the speakers analyzed. This wave is exemplified by Milroy's (1980) work in Belfast, or by Eckert's (2000) own work on adolescents in a school in Detroit, Michigan. For instance, Eckert used, in addition to "social classes", the two main social groups identified by the kids at school, i.e., jocks and burnouts.

The third wave is characterized by a community of practice (Lave and Wenger, 1991; Wenger, 2000) approach. This is defined by Eckert as follows:

> A community of practice is an aggregate of people who come together on a regular basis to engage in some enterprise [...]. In the course of their engagement, the community of practice develops ways of doing things - practices. And these practices involve the construction of a shared orientation to the world around them. [...] The individual [...] is tied into the social matrix through structured forms of engagement. (Eckert, 2005, 16)

This third wave is based on the notion of personal identity as something the individual gradually builds using linguistic variants and styles; then, these gradually become ways to express social meaning, i.e., the identity of groups, not individuals. This third wave is exemplified by Zhang's (2005) study of Chinese yuppies.

However, these three waves do not represent a chronological progression, but a question of emphasis on different aspects. Sociology, ethnography, cultural studies: these subjects and others help enrich our view of what sociolinguistics is and does, and how it could consider variation and change. The HLVC project, as we will describe in §3.3., makes use of notions and methods from all three waves, and aims to obtain insights into how HLs work by comparing the results from statistical models of language variables. This is why we can define it as an application of the comparative variationist framework (Poplack and Tagliamonte, 2001). In the words of Tagliamonte (2012, 163):

> First, the methodology of Variationist Sociolinguistics enables different influences on linguistic features to be disentangled through systemic examination of their behavior. Second, an approach that compares these findings across related data sets (dialects, social groups) permits inter- and intra-corpora similarities and differences to be interpreted.

3.2. The Italian community in Toronto

HLs in Canada, and especially in Ontario and Toronto, benefit from a particular ethnolinguistic vitality. First and foremost, they have a large presence in the country: the 2016 Census (Statistics Canada, 2017) indicates that 29% of the Ontario respondents and 47% of the

Toronto respondents report having a HL as a mother tongue. Moreover, we have discussed in §1.4.1 that society's attitude towards bilingualism influences HL acquisition and transmission: we mentioned Lambert's (1975) classification of additive vs. subtractive bilingualism, for example. In this sense, Canadians are more oriented towards the former, thus encouraging the use of non-official languages – as attested also by past and present official political actions. In Berry's (1998) framework too Canadians' orientation creates the integration or "mosaic" paradigm, where maintaining relationships with the larger society and simultaneously with one's identity is valued. These positive attitudes towards HLs expressed by both citizens and institutions can be seen as a response to the massive presence of immigrant citizens in Canada, which is the G8 country that has the highest percentage of them – and they are concentrated in Toronto (Di Salvo, 2017).

While the first groups of Italians that migrated to Canada date back to the end of the 19th century, a massive migration occurred at the end of World War II due to the stipulation of an agreement between the Italian and Canadian governments in 1951, with the aim of promoting Italian migration. These migrants came from a disadvantaged social and economic condition, often lacking competence in English and/or with a low level of literacy (Di Salvo, 2017); at their arrival, they encountered an educated society on the verge of economic growth (Turchetta, 2018). A second change in the migration dynamic occurred in 1967, when a new system was established based on the intention of favoring the migration of more qualified and professionally skilled individuals (Di Salvo, 2017).

Both migrants from the 20th century and their descendants, and new waves of highly educated migrants from the 21st century, are vehicles for the maintenance and transmission of Italian in Toronto. Other instruments of diffusion of the language are HL courses, but also:

- Radio and television programs;
- Locally produced Italian advertising;
- Editorial Canadian productions in Italian.

These means were especially relevant for the first generation of migrants of the post-war, and they also contributed to the diffusion of the national language among those migrants who were dialect (rather than Italian) speakers (Turchetta, 2018).

However, other means of diffusion of the language can be found. Cultural and social practices also contribute to the preservation of the language, as Turchetta (2018) highlights. As examples, she mentions religious collective representations and rituals; laic celebrations and events, often regionally marked or connected to sports; themed conventions and cultural events.

Shared social practices, indeed, are a way of expressing or preserving individual identities and possibly language choices.

From this brief summary emerges a great ethnolinguistic vitality for HLs in Toronto in general, but especially for Italian, one of the oldest migration languages of the city and the second one in terms of number of HL speakers in Toronto.

3.3. The HLVC project and HerLD corpus

An application of the comparative variationist framework (cf. §3.1.) to the study of HLs, and important research on HLs variation and change, is provided by the HLVC project, a project developed in the Greater Toronto Area (henceforth, Toronto), Canada.

Chapter 1 illustrated that we live in a multilingual world, where language choices contribute to constructing identities. It is starting precisely from these premises that the HLVC project was born (Nagy, 2009; Nagy, 2011; http://ngn.artsci.utoronto.ca/HLVC, accessed on November 12th, 2021). This project has built a multilingual corpus of natural speech in a range of HLs spoken in Toronto. The HLVC project wants to push variationist research beyond a monolingual approach, addressing multilingual speakers' full repertoires.

One of the main goals is to provide a unified and standardized methodology to analyze HL variation, so that inter-language comparisons and generalizations can be made. As a matter of fact, the pre-existing studies of HL variation in Canada used disparate methodologies, and Nagy (1996; 1997) already stressed the need to increase comparability across studies of different communities in order to increase our understanding of contact-induced change. Thus, the HLVC project meets the need to overcome the inconsistencies, and to contribute to theoretical developments. Its program proposes a variationist approach (cf. §3.1.), testing this paradigm with languages other than English, and able to also quantify degrees of similarity between grammars. This standardized framework aims to address the following research questions:

- Are cross-linguistic generalizations possible about the types of features, structures, rules, or constraints that are transferred earlier and more often? If so, what are they? Which ambient language(s) are they transferred from?
- How are social factors relevant? Which ones?
- Do the same (types of) speakers lead changes in both/all their languages? (Nagy, 2015, 312).

As Nagy (2015) explicitly affirms, variation within HLs is seen within the project as a natural feature of languages, and/or as a possible form of change, rather than as the outcome of failed acquisition or discarded as "errors" on an individual level.

The multilingual corpus for inter-generational, cross-linguistic, and diatopic (Heritage vs. Homeland varieties) comparison built within the project, called Heritage Language Documentation Corpus (HerLD), documents variation in ten HLs[56] using a corpus of spoken language, transcribed and time-aligned to audio recordings (.wav and .mp3), that is available online for research (Nagy, 2017). Each language is to be represented by a sample of 40 speakers. Each speaker has been chosen avoiding proficiency or fluency tests as inclusion/exclusion criteria: the aim was to "cast a wide net in order to [...] describe the range of performance of all types of speakers who meet the Canadian government's definition of a Heritage Speaker" (Nagy, 2015, 314). Speakers of each language are arranged into three generations of immigrants, and within each generation, the project has strived for a balanced sample in regard to age (ranging from 12 to 60+ years of age) and sex. Social class, a factor commonly included in variationist studies, has not been included as a social variable to investigate for the following reasons: on the one hand, class and status may change drastically on immigration; on the other hand, "it is not clear whether pre- or post-immigration status is more predictive of linguistic patterns, nor is it clear how to compare (pre-immigration) class across diverse countries of origin" (Nagy 2018, 433).

Speakers in the corpus are thus categorized as follows:

- Homeland: speaker has always lived in the corresponding Homeland, and has parents coming from the same area;
- Generation 1: speaker has lived at least their first 18 years in the Homeland and has lived in the Toronto for at least 20 years;
- Generation 2: parents qualify as Generation 1 (even if they might not be in the corpus) and speaker was either born in Toronto or arrived before age six;
- Generation 3: parents qualify as Generation 2 and speaker was born in Toronto.

The inclusion of Homeland speakers in the corpus is of crucial importance. As Thomason and Kaufman (1988, 111) highlight, in language contact research it is often impossible to trace the trajectory of an alleged contact-induced change due to the lack of pre-contact variety samples. Homeland speakers, even if today's world is globalized, experience less contact with

[56] Languages included are: Cantonese, Faetar, Hungarian, Italian, Korean, Polish, Portuguese, Russian, Tagalog, and Ukrainian.

English compared to people living in Toronto, and therefore their speech serves precisely the purpose described by Thomason and Kaufman. The HerLD corpus is therefore built so that it allows for different types of analyses: investigations of each language individually; comparisons between Homeland and corresponding Heritage varieties; comparisons across Generations of speakers of the same Heritage variety; cross-linguistic comparisons between different Heritage languages.

Speech data in the corpus is collected via guided conversations as relaxed as possible. These interviews are conducted according to standard sociolinguistic interview protocol (Labov 1984), in spaces selected by the participants - often at home – and guided by a fellow HL speaker. Each is about an hour long. They are recorded using a Zoom H4n digital recorder (44.1 kHz sampling rate) with an Audio Technica lavalier microphone.

There are three types of data in the corpus:

1. Sociolinguistic interviews with time-aligned orthographic transcriptions;
2. Elicitation tasks (First Words task, i.e., a picture-naming and story-telling task developed in Nagy's (1994) Faetar fieldwork, that takes about 10 minutes[57]);
3. Ethnically Oriented Questionnaires (EOQ). The EOQ, in particular, includes information about exposure to the HL and to English, language usage patterns, attitudes toward the HL and English and their related cultures, social networks, and community of practices (Nagy, Chociej, Hoffmann, 2014). The questionnaire, to which participants respond orally, produces a self-report on how speakers orient to their ethnicity and to their language(s), and its responses are then scored on a numeric scale. The EOQ was adapted from Keefe and Padilla (1987) by Hoffman and Walker (2010), and then used by the HLVC project, as mentioned in Chapter 1 (2). It consists of 37 questions grouped in sub-sets based on topic.

Transcription conventions are established for each language, and they are then conducted using the time-aligned transcription system ELAN (www.lat-mpi.eu/tools/elan; Wittenburg et al., 2006).

The data provided by the corpus is analyzed in three stages. Stage one consists of a monolingual analysis, i.e., the description of the variables in the language individually. Distributions of the variants are calculated and subjected to multivariate analysis, in order to see how they vary depending on generation, age, sex, EO, or linguistic factors. Patterns of

[57] This type of data provides more readily comparable data from each speaker (Nagy, 2018).

maintenance or divergence from the first-generation immigrants are assessed this way. Stage two consists of diatopic comparison - when possible - to the homeland varieties corresponding to the HL. Stage three consists of cross-linguistic and cross-community comparisons across HLs.

The HLVC project stands out for its approach to HLs and performance, for its awareness about topics such as nativeness and multilingualism, and for its goal of providing a reliable and replicable sociolinguistic methodology. The HerLD corpus is and will be an extremely useful resource for the documentation of these languages in their actual use, for linguistic analyses, but also for pedagogical purposes. An interesting feature of the project, indeed, is its goal of training students and speakers of the HLs to conduct linguistic research by involving them directly[58] (Nagy, 2017).

The HerLD Italian data that this work used targets speakers of a specific Italian area, given the huge range of diatopic variation of the peninsula, i.e., low-southern Calabria. The convergence of a group of Calabrese dialectal traits on a base still identifiable as Italian allows us to identify the variety of the recordings as Calabrese Regional Italian (RI), although these traits are present to different degrees across speakers.

De Blasi (2014, 110-111) reports as some of the phonetic features of Calabrese Regional Italian (RI: cf. §2.3.2.) (examples come from the HerLD corpus):

- Aspiration of -t-, especially following -r- or -n- (*certamente* [ˌʧɛɾtʰaˈmɛnthe], "surely");
- Retroflexion of the cluster -tr-;
- Gemination of intervocalic -b- and -g- (*cabina* [kabˈbiːna], "cabin", *cugino* [kudˈʤiːno], "cousin");
- Voicing of plosives after nasals (*tanto* [ˈtando], "a lot");
- Epenthesis in consonant clusters (*aritmetica* [ˌariteˈmɛːtika], "arithmetic");
- Stressed mid vowels pronounced as mid-low, [ɛ] [ɔ] (*certamente* [ˌʧɛɾtʰaˈmɛnthe], "surely").

These traits are featured in the interviews together with other phonetic, morphosyntactic and lexical traits of Calabrese dialects (such as *ci attualizzante* or "empathic ci" – *ci*, "there", before the verb *avere*, "to have", for emphasis; dialectal articles; dialectal words such as *ammucciare*, "to hide").

[58] Students who have participated as RAs are featured on the project's website: https://ngn.artsci.utoronto.ca/HLVC/3_2_active_ra.php, https://ngn.artsci.utoronto.ca/HLVC/3_3_former_ra.php (accessed on February 19th, 2020).

The following three extracts, from the Italian interviews of the HerLD corpus, illustrate phonetic details (examples come from Baird, Cristiano, Nagy, 2021, 2-3):

1. *Una volta una signora ha detto: "Voglio assaggiare il pranzo che avete fatto."*
[ˈuːna ˈvɔlt̪ʰə nə siɲˈɲoːɾa a dˈdʲettʰu ˈvɔʎʎ assadˈʤaːɾe i pˈpʰɾaɲ̪ts ke aˈveːt̪e ˈfatto]
"Once a lady told me: "I want to taste the meal you have prepared."' [I1F73A[59], 18:42]

2. *I genitori avevano un pezzo di terra, la campagna, ma io stavo al paese, andavo a insegnarmi l'arte di falegname, di muratore.*
[i ddʒenɪ̣ˈtoːɾə əˈvəːvno um ˈpɛtts di ˈtɛrrə la ɟ̑amˈpʰaɲɲa ma ˈiːo̤ ˈstaːvo al pʰăˈeːsə ənˈdaːva a iɲ̪tseɲˈɲarmɪ̣ l ˈartʰe di faleɲˈɲaːme di muɾaˈtɔːɾe̤]
"His parents had a piece of land, the farmland, but I stayed in town, I was learning the craft of woodworking, of bricklaying." [I1M62A, 21:28]

3. *Diciamo che mi sono… oltre allo studio fatto un pochino di casini là, no, mi sono occupato di vita d'associazione, ne avev- ho creato un'associazione studentesca.*
[diˈʤaːmo ɟ̑ə m̩ ˈsɔːnɔ ˈɔl͡t̪ʂ ˈallo ˈstuːdjo̤ o fˈfatt m poˈkʲiːnə dɪ̣ː dɪ̣ kaˈsiːnï lla nɔ mə ˈsɔːno okkʰˈpaːto dɪ̣ ˈvɪ̣ːta d asˌsoʃʃatˈʦjɔne n aˈveːo o kkrɛˈaːt ˈuːna assoʤəˈzjɔːnə studenˈtʰeska]
"Let's say I… In addition to studying, I've made a bit of a mess, there, you know, I've handled club life, I had- I've created a student club." [IXM35A, 20:44].

3.4. Hypotheses and methods

In our research, we followed the method described by Tagliamonte's (2012, 163) quote reported in §3.1. Phonetic realizations of word-internal singleton /r/ have been analyzed. /r/ in -tr- and -dr- onset clusters are not part of this analysis, since said clusters have been studied separately.

We refer to word-internal singleton /r/ as the variable investigated: (r). The symbol (r) will be used with two main purposes:

[59] Speakers are identified by a speakercode according to the following conventions: the first character of the speakercode identifies the HL ("I" stands for Italian); the second one the speaker's generation (X is for Homeland, then 1, 2 or 3); the third one the speaker's sex (M or F); the fourth and fifth ones the speaker's age; the final one (A, B, C, etc.) provides a unique identifier for otherwise identically labeled speakers.

1. To refer to the rhotic as a sociolinguistic variable, hence, when talking about its status and use in the relevant communities;
2. To refer simultaneously to the various rhotic realizations without necessarily implying an allophonic connection among them. In other words, to avoid assuming that, in Italian, /r/ is the underlying rhotic phoneme whose allophones are [r], [ɾ] and [ɹ]. In particular, we cannot be sure about the status of [ɹ], as in the Heritage community this variant might not be an allophone of the trill, but instead a phone borrowed from English.

The (r) variants we will consider, based on our review of the relevant types of rhotics in Calabrese RI and Canadian English (cf. §2.3.1.; §2.3.2.1.; §2.3.2.2.), are: the forms prescribed by Standard Italian descriptions, i.e., tap [ɾ] or trill [r], merged into a single level "taps&trills[60]"; the approximant [ɹ]; the fricative, for which we can use a symbol proposed by Celata (2014) and Celata, Meluzzi, and Ricci (2016) for the spirantized tap, [ɾ̝]. These symbols should be considered as references to the prototypical realization of the investigated sounds. Phones that use the same manner described by one of the symbols, but vary according to place, have been considered as belonging to the same prototype. For example, [r] indicates the alveolar trill; post-alveolar trills have not been distinguished by means of another symbol, and have been considered as instances of the same prototypical representation [r]. A difference in manner, e.g., an approximant rather than a trill, instead, means that we are looking at an instance of a different prototype, i.e., [ɹ].

3.4.1. Research questions and hypotheses

The present work investigates the following research questions:

- Since traditional descriptions of the Italian phonological system only mention taps and trills for rhotics, how often are rhotics *actually* realized in spontaneous speech? And how often do non-Standard variants such as approximants and fricatives appear?
- Are there any systematic patterns of variation in spontaneous speech concerning word-internal singleton (r)?
- What are the linguistic and social factors that constrain selection among the (r) variants? Are they different between Homeland and Heritage speakers? If not, are they ranked differently in terms of the strength of their effects?

[60] cf. §3.4.2. for the relevant discussion.

- Is it possible to identify the variable patterns as either resulting from a contact-induced change or a language-internal change?

The research questions will be analyzed in detail by testing the following hypotheses:

1) We argue that approximant and fricative variants are articulatory weaker variants of trills and taps that naturally emerge in contexts of relaxed speech, and therefore, we hypothesize that they will appear in both the Homeland and Heritage varieties.
2) It is hypothesized that approximant and fricative variants will be favored in coda position, a linguistic context often described to be particularly prone to have weaker/lenited variants (cf. Lawson et al., 2018; Kirchner, 2013; Rennicke, 2015).
3) We hypothesize that the proportion between taps&trills and the more reduced variants will be similar between the two communities. Heritage languages, since they are native languages, preserve all the intrinsic features of spontaneous speech (e.g., hypoarticulation), just like Homeland varieties do.
4) Since we are claiming that the presence of these variants in Heritage speakers is related to natural language-internal variation, just like in the Homeland, we hypothesize that no Generational effects will be found. This, and the lack of a community effect (i.e., no Homeland vs. Heritage effect), will be evidence that Heritage speakers are not transferring phonetic detail regarding (r) from the majority language (English), nor simplifying their grammar.
5) We hypothesize Heritage speakers not only to maintain similar percentages of variants as the Homeland, but also the same distributional pattern as Homeland speakers. The lack of an increase of the contexts favoring approximants in Heritage speakers will be evidence that not even the distributional patterns of Canadian English are being transferred (since Canadian English produces [ɹ] as a consonantal rhotic in every context).
6) While we expect the Heritage community as a whole to maintain the Homeland's grammar, we claim that individual-level differences can be explained by Ethnic orientation. We hypothesize that speakers (within a generation) with higher EO scores will produce more Italian-like variants – taps&trills – compared to those with lower EO scores.

3.4.2. Speaker and token selection and variables

A sample of 29 speakers (16 males and 13 females), ranging from 19 to 75 years of age (mean = 45.2; sd = 16.4), has been drawn from the Italian data of the HerLD corpus. Speakers analyzed come from the Homeland, Generation 1, and Generation 2 groups.

14 speakers have been selected from the Homeland group, 8 males and 6 females. The age range within this sample is 19-64, mean = 38.0, sd = 17.0.

15 speakers have been selected from the Heritage group: 7 for Generation 1 (4 males and 3 females), 8 for Generation 2 (4 males and 4 females). The age range for Generation 1 is 59-75, mean = 66.0, sd = 6.7. The age range for Generation 2 is 42-57, mean = 49.37, sd = 5.26.

The following table summarizes the distribution of speakers according to Generation, Sex, and age groups, in order to provide a clearer picture:

Generation	**Sex**	**18-30**	**30-58**	**59-75**
Homeland	Males	IXM19A, IXM19B, IXM28A	IXM35A, IXM47A, IXM52A	IXM61A, IXM64A
	Females	IXF22A, IXF22B, IXF24B, IXF24C	IXF35A, IXF51B	
Generation 1	Males			I1M61A, I1M61B, I1M62A, I1M75A
	Females			I1F59A, I1F71A, I1F73A
Generation 2	Males		I2M42A, I2M49A, I2M52A, I2M53A	
	Females		I2F44A, I2F45A, I2F53A, I2F57A	

Table 3: Distribution of speakers selected from the HerLD corpus.

Occurrences of (r) to analyze have been drawn from words in a limited list. This list contains the 20 most frequent words containing word-internal singleton /r/ in the Italian corpus. This choice has been made for two reasons: first, to make sure there were enough tokens of every word by every speaker selected; second, to mitigate possible frequency effects. The words from the list, in their citation-form, do not contain -tr- or -dr- onset clusters (which have been looked

at separately), but /r/ in other linguistic contexts: simple and complex onset and coda positions, stressed and unstressed syllables, and monosyllabic words. 40 to 60 tokens of these words were extracted from the sociolinguistic interviews of the HerLD corpus and coded for each speaker (mean of tokens per speaker = 57.68; sd = 5.18). The total number of tokens considered is 1555.

The following table presents the list of words and the number of tokens per word analyzed, in the Homeland and Heritage samples, and in total. Words are listed from top to bottom based on their frequency in the corpus. The target occurrence of /r/ in the word is bolded.

Word	**Tokens in Homeland**	**Tokens in Heritage**	**Total**	**% of sample**
*Pe**r** (for/to)*	56	54	110	7%
*E**r**a (he/she/it was)*	55	69	124	8%
*Pe**r**ché (why/because)*	66	47	113	7%
*Pe**r**ò (but)*	57	39	96	6%
*Semp**r**e (always)*	40	54	94	6%
*E**r**ano (they were)*	26	55	81	5%
*Allo**r**a (then)*	21	37	58	4%
*Lo**r**o (they/them/their)*	27	47	74	5%
*Fa**r**e (to do)*	39	56	95	6%
*Lavo**r**o (job)*	26	57	83	5%
*Anda**r**e (to go)*	18	37	55	4%
*Pu**r**e (also)*	18	36	54	3%
*P**r**ima (before)*	36	45	81	5%
*Genito**r**i (parents)*	30	42	72	5%
*E**r**o (I was)*	22	40	62	4%
*Rico**r**do (I remember)*	25	35	60	4%
*Fo**r**se (maybe)*	31	36	67	4%
*Pa**r**te (part)*	33	30	63	4%
*P**r**oprio (really/own)*	24	40	64	4%
*Anco**r**a (still)*	25	24	49	3%
Total	**675**	**880**	**1555**	

Table 4: Distribution of tokens per word.

Each token has been coded as the dependent variable and for a series of independent variables. Some factors were coded in ELAN, others in Excel during a following step, when the tokens were consolidated in a dataframe for analysis in Excel.

The dependent variable in our study is the realization of (r). For this, three levels have been considered (cf. §3.4.): Fricative, Approximant, or Tap&trill. This factor has been coded mostly auditorily[61], but visual inspection in Praat (Boersma and Weenink, 2021) helped with unclear cases. Let us explain the way these levels have been coded more in detail.

The variants tap and trill have been considered under a single label, Tap&trill, for several reasons. First, the high variability rate of spontaneous speech generally causes low rates of articulatory accuracy (e.g., weakening of consonants) (Voghera, 2001). Trills and taps are not consistently produced in the same way by the same speaker or across different speakers, and they are not as specified as we would expect from experimental speech, for example. Second, as we described in §2.1.1., and tied to the first reason, trills can also be produced with a single closure, making these realizations difficult to distinguish from taps simply by their waveform and/or spectrogram. We lack the necessary tools to determine whether the vibration of the rhotic was generated by the air stream or by muscular contraction, which would allow us to distinguish between the two realizations. Therefore, trying to classify them only by subjective impression would have introduced arbitrariness in the coding, given their perceptual similarity especially in fast speech.

To code Approximants and Fricatives, we applied some of the directions Celata, Meluzzi, and Ricci (2016) gave for their coding of Sicilian Italian rhotic realizations, given the similarities between low-southern Calabrese and Sicilian varieties (cf. §2.3.2.2.).

Approximant variants were classified as such first based on the non-perception of a phase of closure, even short, and then based on a formant structure distributed over the whole segment. Some of the Approximant realizations acoustically resembled the Canadian English [ɹ] more than others, with a sudden and drastic lowering of the third formant; some others were more vowel-like, and impressionistically close to schwa – in accordance with claims about approximant rhotics, cf. §2.1.3. However, many cases of alleged Approximants were not straightforward: the auditory perception of a brief closure contrasted with the sonagram, which did not show it. Celata, Meluzzi, and Ricci mention several instances in their corpus where it was difficult for them to distinguish between the two variants. These cases could have been classified by the authors as approximants, based on visual inspection of waveforms and

[61] Tokens have been carefully listened to using AudioTechnica ATH-M20X headphones. The author is responsible for any mistake or bias in the coding, since there has not been any checking by other coders.

sonograms, but they proved to be taps with a very short phase of closure after EPG analysis. Moreover, it must be recalled that the acoustic signal of the data coded for this research is not laboratory-like, hence, noise can compromise its cleanliness and hide some of the silences (closures) on the sonagram. For these two reasons, where a closure was perceived auditorily, the token was classified as a Tap, rather than as an Approximant.

Fricative variants were identified based on the auditory perception of a clear friction sound, visually corresponding to intense aperiodic noise. Frication, according to Celata, Meluzzi and Ricci, could be either a feature of the whole rhotic segment, giving rise to a prototypical fricative rhotic with no constriction at the beginning, or a simplification of the vocoid of the opening phase of a tap, which was devoiced up to the loss of its formant structure[62]. The latter case would therefore represent a rhotic composed of two separate acoustic and articulatory phases, as confirmed by EPG palatograms, and has been classified by Celata (2014) as a "spirantized tap" [ɾ̥̞]. In their sub-corpus of Sicilian Italian read speech, this realization was extremely frequent in coda position (39/43 occurrences), after a high vowel and before a voiceless obstruent. Since the prototypical fricative rhotic and the spirantized tap do not seem to be perceptually different according to the authors, and given the above-mentioned difficulty to detect phases of closure by visual inspection in some of our data, we have not distinguished between the two cases.

The following examples from our corpus show waveforms and sonagrams of tokens coded as Tap&trill, Approximant, or Fricatives:

[62] It should be noted that Loporcaro connected the creation of friction in the rhotic to the hypothesized devoicing the consonant undergoes after the apocope of the following vowel. According to him, final vowel deletion in the Altamura (Apulia) dialect causes the devoicing of the previous segment, which, in the case of /r/, translates as turbulence (Loporcaro, 1988).

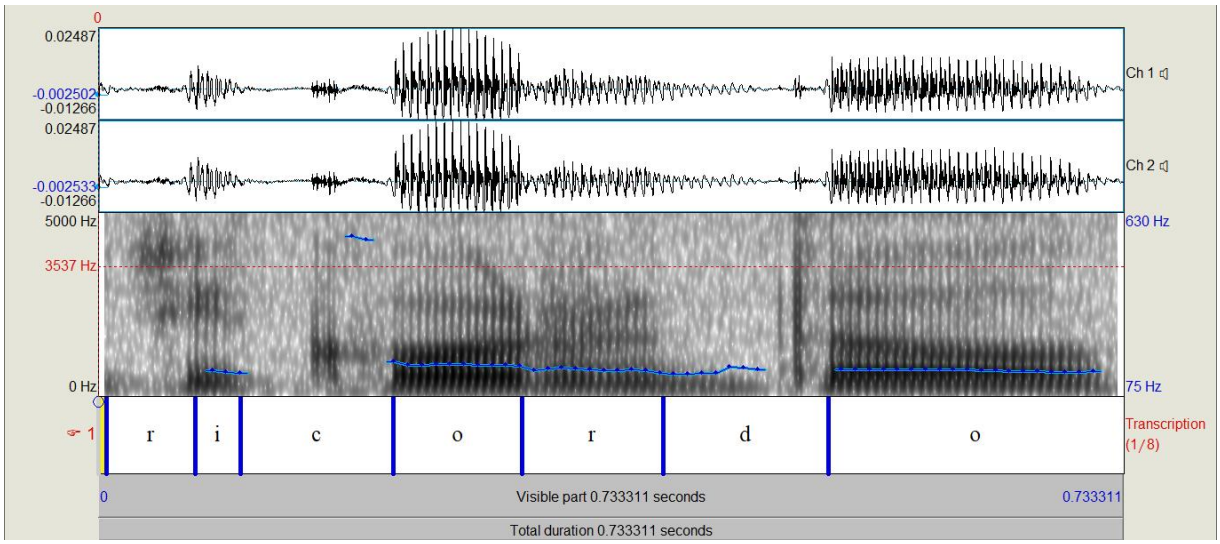

Fig 1: Visual example of a token of the word *ricordo*, "I remember", coded as Tap&trill, from speaker I2F45A.

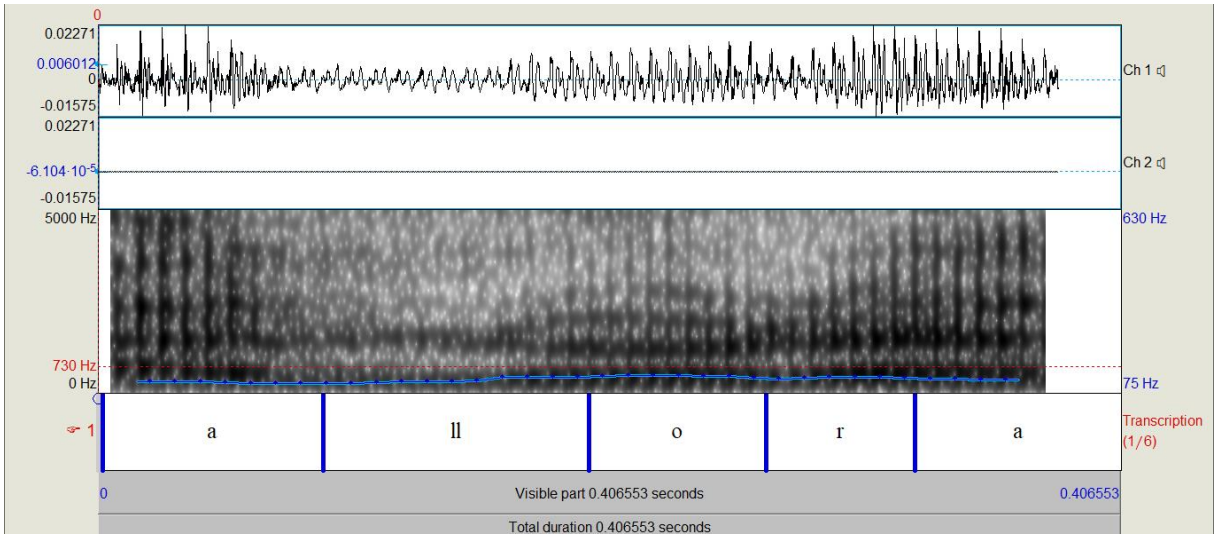

Fig. 2: Visual example of a token of the word *allora*, "then", coded as Approximant, from speaker I2M42A.

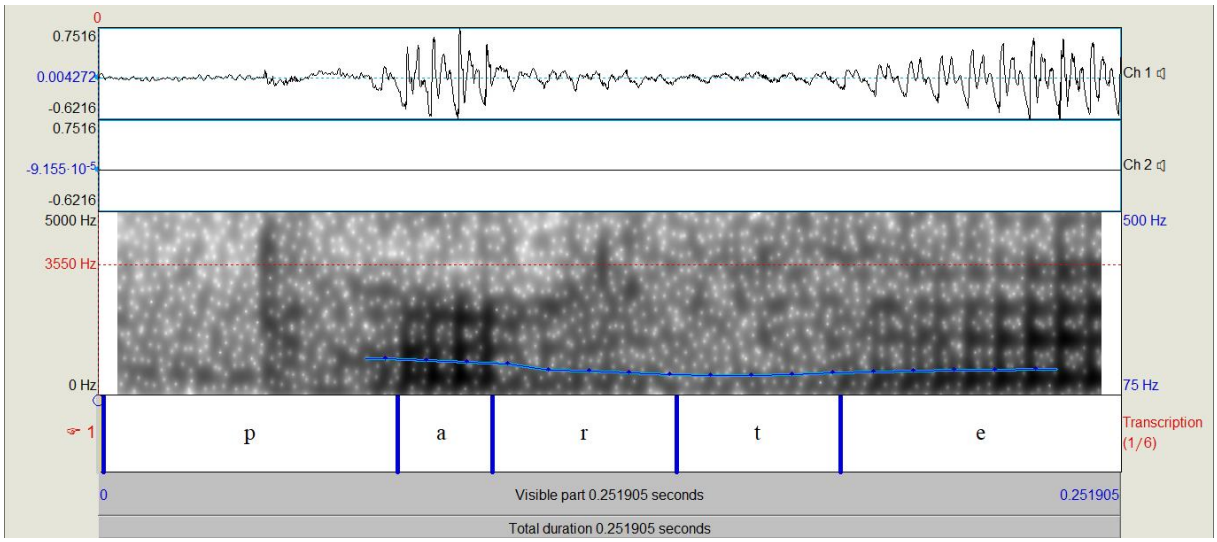

Fig. 3: Visual example of a token of the word *parte*, "part", coded as Fricative, from speaker IXM52A.

As eight of the tokens considered were too ambiguous to code for the dependent variable, they were excluded from the dataset. This brought the number of tokens analyzed from 1555 to 1547.

The independent variables considered can be divided into social factors and linguistic factors.

Social factors are Speaker, Sex, Age, Decade, and Generation. Speaker indicates the individual who produced the tokens, and is coded using the speakercode from the HerLD corpus; hence, it has 27 levels. Sex has two levels (Male and Female). The label "Sex" has been used since our coding reflects the sex the interviewer perceived the speakers to be, and not the gender speakers identify with – which in our corpus we have no access to (cf. §3.6.2.). Age is the only continuous factor, ranging from 19 to 75. Decade is the corresponding binned factor to Age. It has 6 levels: 1-2 stands for speakers in their teens and 20s, and so on in the same fashion for levels 3, 4, 5, 6, and 7. Generation (defined with respect to the speaker's family's immigration history) is a factor with three levels (Homeland, Generation 1, and Generation 2). This factor can be collapsed into Homeland vs. Heritage speakers by merging the levels Generation 1 and Generation 2, to compare the two communities. Some of these variables, such as Age and Decade, are clearly collinear, therefore extreme caution has been used during the subsequent modeling to avoid including collinear factors within a model. Lastly, Heritage speakers have been coded for another social variable: Ethnic Orientation (EO) scores.

As mentioned in Chapter 1 (cf. §1.2.), the HLVC project employs the speakers' answers to an Ethnic Orientation Questionnaire (EOQ) to measure their engagement with the ancestral

ethnic group. The questionnaire has different topic-based subsets, and is available at http://ngn.artsci.utoronto.ca/pdf/HLVC/short_questionnaire_English.pdf (accessed on January 06th, 2022). Each answer from a speaker is thought of as an expression of their position along a continuum from Canadian ethnicity to Italian ethnicity. This concept is operationalized by treating their answers as scores: an Italian-oriented answer is scored 2, a Canadian- or English-oriented answer is scored 0, and a mixed response is scored 1. EO scores can be implemented in the model in different ways: for instance, they can be treated as a single continuous factor, whose values are the average overall score for each speaker, or one can have multiple continuous EO factors whose values are the average score of a single subset for each speaker (cf. Nagy and Kochetov, 2013; Baird, Cristiano, and Nagy, 2021). Another possibility is to have multiple EO factors, where each corresponds to a subset, but to consider them binomial variables. Speakers whose average score for an EO subset is above the median are coded having "high" EO for that subset, while speakers whose average score for a EO subset is equal or below the median are coded having "low" EO for that subset (cf. Umbal and Nagy, 2021). This method would allow capturing non-linear correlations between EO scores and the frequency of a variant, but also combining factors involving EO in a Mixed Effects model.

In this work, we will use the latter method. Hence, we selected two subsets of the EOQ to test: Ethnic Identification (EO ID, questions A1-5), and Language Use (EO LU, questions B1-C4). These are the questions associated with each subset:

- Ethnic Identification:
  - A1. Do you think of yourself as Italian, Canadian or Italian-Canadian?
  - A2. Are most of your friends Italian?
  - A3. Are people in your neighbourhood Italian?
  - A4. Are the people you work with Italian?
  - A5. When you were growing up, were the kids in your school Italian? Were your friends? The kids in your neighbourhood?
- Language Use:

  B. Language:
  1. Do you speak Italian? How well? How often?
  2. If no: Can you understand Italian?
  3. Where did you learn Italian? At home? In school?
  4. Do you prefer to speak Italian or English?

5. Do you prefer to read and write in Italian or English? Do you read Italian magazines and newspapers? — Which ones?
6. Do you prefer to listen to the radio or watch TV in Italian or English?

C. Language choice:

1. What language does your family speak when you get together?
2. What language do you speak with your friends?
3. What language do you speak when you're talking about something personal? When you're angry?
4. Did/do you speak to your parents in Italian? Your grandparents?

Each subset corresponds to an independent factor. Each factor is coded with two levels, High or Low. Since EO scores were available for 14/15 Heritage speakers, models including EO scores as a factor have been tested separately. The following table summarizes the social factors and their levels:

Factor	**Levels**
Speaker	20 speakercodes (IXF35A…)
Sex	Male Female
Age	Continuous (19-75)
Decade	1&2 3 4 5 6 7
Generation	Homeland Generation 1 Generation 2
EO ID	High Low
EO LU	High Low

Table 5: Social factors analyzed and their levels.

Linguistic factors considered are Lexical word, Lexical class, Previous phone, Following phone, Position in the syllable, and Stress of the syllable. Lexical word simply indicates the lexical item the token of (r) comes from, and thus has 20 levels corresponding to the words from our list (cf. Table 2). Lexical class aims to provide a basic classification of the selected Lexical items into categories, employing as levels Noun, Verb, (personal) Pronoun/Adjective, and the macro-group Adverb/Conjunction/Preposition. While the labels used come from traditional descriptions of Italian grammar, the groupings were realized considering more recent critiques to traditional approaches (D'Errico et al., 2016; Colombo and Graffi, 2017).

Nouns and Verbs seem universal categories, and there are few doubts on their classification. Their members naturally cluster together and are quite distinct from other clusters, as the inductive distributional map proposed by D'Errico et al. (2016, 124) shows. Adverbs, Conjunctions and Prepositions are all non-variable lexical classes, but it has always been challenging to distinguish their members in Italian grammar. In our coding, we collapsed them together for the following reasons:

a) Colombo and Graffi (2017) suggest that several of their members are actually shared, and that these classes are related not only for their non-variability;
b) Traditionally defined Adverbs and Conjunctions cluster in the same area in the map by D'Errico et al. (2016);
c) The label "Preposition" in our corpus would only apply to tokens of the Lexical Item *per*, "for"/"to", a word also used to introduce infinitive subordinates (D'Errico et al., 2016, 129), just like some Conjunctions do (e.g., *Ho scritto un'e-mail per ricevere risposta*, "I've wrote an e-mail to receive an answer").

By exclusion, we considered the group (personal) Pronoun/Adjective as a separate level. Its only member is *loro* ("they"/"them"/"their"/"theirs"). Nevertheless, this lexical item is too different in terms of distribution, syntactic behavior, and semantics to be merged with another level – and Colombo and Graffi (2017) also support the independence of the class of personal pronouns in Italian.

Previous phone and Following phone indicate how the phones adjacent to the rhotic have been realized. Levels for Previous phone are Obstruent and Vowel (no token was preceded by a sonorant). Levels for Following phone are Obstruent, Sonorant, and Vowel. No tokens were preceded or followed by approximants[63]. The factor Position in the syllable is coded for the

[63] Initially, this factor was coded with an additional level: sentence final, for the cases when the rhotic was not followed by any segment. However, during the statistical modeling this level was excluded, and the tokens coded

following three levels: Onset, Coda (always simple), and Nucleus (for possible cases of syllabic [ɾ̩])[64]. The factor Stress of the syllable is coded for three levels: Unstressed, Stressed, and Monosyllabic word[65].

It must be highlighted that all the linguistic factors have been coded taking into account the phonetic realizations of the token and its context, not the syllable-structures and the phonemes of citation forms. Therefore, they were all coded the same way as the dependent variable: auditorily, with the aid of visual inspection in Praat. This choice has been made to avoid making assumptions about the mental representation of the relevant words by speakers.

As an example of what the difference between our coding and a citation form-based coding is, consider a token of (r) in the word *erano*, [ˈɛːɾa.no], "they were." Coding based on citation form is shown in the left side of Table 6 below. However, some of the tokens of the word *erano* were realized by deleting the vowel following (r) ([ˈɛɾ.no], leading to the coding shown on the right. Indeed, the deletion of the vowel following (r) generates a -rn- sequence, which is interpreted in Italian as heterosyllabic: hence the "coda" and "stressed" codings.

Factor	**Citation form-based coding /ˈe.ra.no/**	**Realization-based coding [ˈɛɾ.no]**
Type of previous phone	Vowel	Vowel
Type of following phone	Vowel	Sonorant
Position in the syllable	Onset	Coda
Stress of the syllable	Unstressed	Stressed

Table 6: Comparison between a coding based on citation-forms and a coding based on realizations using a token of the word *erano*, "they were", as an example.

In our dataset, the lexical items that have a linguistic coding that may differ from a citation form-based coding can be divided into two groups:

as sentence final were not considered, since their amount (n = 7) was not enough for a balanced statistical analysis, and neither it was linguistically appropriate to merge this level with another one.

[64] Originally, this factor had one more level, since it distinguished between simple and complex onsets (i.e., *andare*, [anˈdaːɾe], "to go", and *proprio*, [ˈprɔːpɾjo], "really"). However, since this distinction proved to be completely useless in the following statistical analyses, the two levels have been collapsed into just one onset level.

[65] Originally, this factor had one more level, since it distinguished between pre-stressed and post-stressed unstressed tokens (i.e., *perché*, [perˈke], "why"/"because", and *erano*, [ˈɛːra.no], "they were"). However, since this distinction proved to be completely useless in the following statistical analyses, the two levels have been collapsed into just one unstressed level.

1. The first group is made of the words *allora*, *andare*, *era*, *fare*, *genitori*, *lavoro* followed by consonant-initial words, and *erano*: words where /r/ in a citation-form would be in a sequence **(C)V**.rV.C(C)V[66], i.e., in the onset of a post-stress syllable (e.g., *allora mandava*, [alˈloːɾa manˈdaːva], "then he sent"). Some occurrences of these words were realized by deleting the vowel following (r): since in Italian /r/ is never the first member of a complex onset, our interpretation is that the deletion led the (r) to occupy the coda position of the stressed syllable (in a structure such as **(C)Vr**.C(C)V) (e.g., in *allora mandava*, "then he sent", after the apocope of /a/ in *allora*: al.lor.man.da.va, [alˈloɾ manˈdaːva]).
2. The second group is made of the words *perché*, *prima*, *forse*, *proprio*[67], and *sempre* (when followed by consonant-initial words, e.g., *sempre felice*, [ˈsɛmpɾe feˈliːʧe], "always happy", where /r/ in a citation-form would be either in a CrV.C(C)V or in a CVr.CV sequence. Some occurrences of these words were realized by deleting the vocalic segment adjacent to (r); our interpretation is that the deletion caused (r) to become the nucleus of the syllable (e.g., in a reduced *forse*: [ˈfr̩se], "maybe").

These decisions, i.e., interpreting the (r) tokens from group 1 as coda (r)s or those from group 2 as syllabic consonants are grounded in acoustic, typological, and lexical motivations:

1. The acoustic-based definition of syllable is that of a unit made of a group of phones surrounding an intensity peak – the nucleus. The most desirable unit is the one where the acoustic intensity (in more phonological terms, the sonority; Ladefoged, 1975) increases from the beginning of the unit to the peak, and decreases from the peak to the end of the unit. Syllables such as rCV, with C being an obstruent, are particularly less desirable acoustically, because they represent an intensity, or sonority, fall (Berent et al., 2007); hence, we interpret these structures as having /r/ as the coda of the previous syllable, and the CV as a sonority rise onset. As for syllabic /r/s, they have been judged as nuclei because they represent the sonority peak within the unit, i.e., the most intense segment in the proximity. It must be noted, however, that Italian phonology does not

[66] With C standing for consonant, r for the rhotic, V for vowel or approximant - and the bold for the stress, when it is relevant.
[67] To facilitate the explanation, we will assume that the glide in *proprio*, [ˈprɔ.pɾjo], "really", can be represented in the following sequences as a vowel.

include consonantal nuclei – but it does not consider that clusters as the ones we have found in spontaneous speech would arise either[68].

2. Typological surveys (Greenberg, 1978) suggest that cross-linguistically, sonority plateaus (e.g., rCV, with C being a sonorant) are less frequent than sonority rises, and that sonority falls are even less frequent than sonority plateaus. The low typological frequency of plateaus and falls, and the fact that sonority falls always imply the existence of sonority plateaus in a language, further convinced us to avoid interpreting the relevant structures as these types of sonority profiles: they would be rare profiles, and not previously described in the Italian literature.
3. There are no words in the Italian lexicon starting with rCV sequences. The absence of a given cluster in word-initial position is sometimes considered as evidence that the cluster is not a complex onset featured in the relevant language (Nespor and Bafile, 2008).
4. Lastly, it would have been inconsistent to code for Preceding and Following phone based on the actual productions, and for Position in the syllable and Stress of the syllable based on the phonological features of citation-forms.

To summarize, the following table lists the linguistic variables considered, with corresponding examples:

Linguistic factor	**Levels**	**Example**
Lexical word	cf. Table 4	
Lexical class	Verb Noun Pronoun/Adjective Adverb/Conjunction/Preposition	*ero*, "I was" *genitori*, "parents" *loro*, "they/them/their/theirs"; *forse*, "maybe"; *però*, "but"; *per*, "for"/"to"
Previous phone	Obstruent Vowel	[pɾ]*oprio*, "really" *all*[ɔɾ]*a*, "then"
Following phone	Obstruent	*fo*[ɾs]*e*, "maybe"

[68] However, consonant conglomerates including an /r/, and generated after the deletion of one or more vowels in Neapolitan dialect, have been proposed to be syllables having /r/ as the nucleus due to its higher intrinsic sonority by Albano Leoni (2015).

Linguistic factor	Levels	Example
	Sonorant	e[ɾn]o, “they were”[69]
	Vowel	*fa*[ɾe], “to do”
	Sentence final	semp[ɾ], “always”[70]
Position in the syllable	Onset	*ge.ni.to.**r**i*, “parents”; *p**r**i.ma*, “before”
	Coda	*pe**r**.ché*, “because”
	Nucleus	[ˈfɽ.se], “maybe”[71]
Stress of the syllable	Unstressed	*e**r**o* [ˈɛːɾo], “I was”
	Stressed	*ricordo* [ɾiˈkɔ**r**do], “I remember”
	Monosyllabic word	*per* [peɾ], “for”/”to”

Table 7: Linguistic factors analyzed and their levels.

After having coded every token and every variable, the Excel spreadsheet was consolidated as a dataframe and exported as a tab delimited text file, in order to be able to upload it in the program Rbrul (Johnson, 2009), both for descriptive statistical analysis and for Mixed Effects modeling. This file contains one token per row and one variable (dependent and independent) per column.

3.4.3. Statistical procedure

Data has been analyzed using Mixed Effects models (as suggested by the consolidated methodology of many LVC studies), a choice that has gained success among sociolinguists in the last decade.

From the '70s up to ca.2010, quantitative variationist research mostly relied on VARBRUL (Cedergren and Sankoff, 1974) and GoldVarb (Sankoff, Tagliamonte, and Smith, 2005) (Tagliamonte, 2006), computer programs originally born as a reflection of the variable rule concept introduced by Labov (1969). However, in recent years more and more authors -

[69] This example represents a reduced form of the word *erano*, where the deletion of /a/ caused (r) to be followed by the sonorant.
[70] This example represents a reduced form of the word *sempre* in sentence final position, were the deletion of final /e/ caused (r) to become the last segment before the end of the sentence.
[71] This example represents a reduced form of the word *forse*, where the deletion of /o/ caused (r) to become the syllabic nucleus.

especially from outside the field of linguistics (Tagliamonte, 2012) - highlighted the inadequacy of the procedure at the heart of these programs - multiple logistic regression with fixed effects only - when addressing language variation[72] (Díaz-Campos and Dickinson, 2019).

The main issue with multiple logistic regression models is their inability to manage grouping, nesting, and imbalance in the data, as Johnson (2014) explains. Data used in naturalistic sociolinguistic research, indeed, usually consists of multiple observations of a variable across both the same speaker and different speakers, and while a said word can be represented by just a few tokens, another word may be represented by a huge number of tokens from the same study. Therefore, both the speaker-level and the word-level are affected by a grouping structure. Moreover, since sociolinguists usually analyze differences between groups of speakers or words, the fact that group-level variables like sex have a nesting relationship with both the speakers and the words represents a problem that must be taken into consideration, because speakers and words may show idiosyncrasies: Guy (1980), for example, describes intra-speaker variation (individual variation in the rate of use of linguistic forms), and usage-based models use by-word variation as a foundational claim - a type of variation that could be affected also by external factors (Baayen and Milin, 2010).

These elements were not considered in the "traditional" quantitative variationist research, because VARBRUL and Goldvarb employ logistic regression analysis using fixed-effects[73] (predicting factors) only. Regression analysis, indeed, assumes (among other things) that the analyzed observations are all independent: we have seen that this is almost always not the case, with many observations coming from the same speaker. Therefore, the effects of the grouping of tokens (or variants) around a single speaker are completely lost, and speakers who are outliers are able to potentially falsify the outcomes, making the group-variable they come from significant, when it otherwise would not be (Brezina and Meyerhoff, 2014; Drager and Hay, 2012). As Baker (2010, 56) explains, "when grouping together a large number of speakers we can overlook differences within groups, which may have a skewing effect on our results". This happens precisely because the programs lack some kind of easy-to-use tool to control for potential outliers among speakers or words and reduce their group-level influence, since they only employ fixed-effects.

[72] This example represents a reduced form of the word *sempre* in sentence final position, were the deletion of final /e/ cau
sed (r) to become the last segment before the end of the sentence.
like sex or speech style, and that are presumed to be replicable in other studies.

This is the main reason why Johnson (2009) decided to introduce a new version of the variable rule program: Rbrul, available at http://www.danielezrajohnson.com, an open-source interface written in R (R Core Team, 2012). Unlike its predecessor GoldVarb, Rbrul uses mixed effects modelling, a particularly flexible tool for analyzing grouped data and repeated measures (Pinheiro and Bates, 2000) which makes use of both fixed and random effects. Random effects are factors sampled from a population (or at least a larger set), are usually not replicable in further studies, and can be shared by multiple fixed effects (Drager and Hay, 2012; Johnson, 2014). In a Mixed Effects model, each level of a random effect receives a coefficient that makes it unique, and that is able to weight each level differently. For example, a potential outlier would be assigned a lighter weight, reducing its capacity of influencing the overall significance of a group it belongs to. The model is structured so that it can account for inter-speaker and group-level significance simultaneously, but capturing group effects "only when they are strong enough to rise above inter-speaker variation" (Johnson, 2009, 365): a huge improvement from simple regression models, where it is against the assumption of independence of variables to include both the individual and a category label associated with them.

There are two types of random effects: random intercept and random slope. Speakers are usually treated as random intercepts. Random intercepts can be useful tools to interpret the individual behavior compared to the population: indeed, when a random effect is included in the model, the model also produces a model's intercept, which reflects its best guess at how the population would behave. By-speaker random intercepts, instead, represent the individuals' adjustment to the model's intercept, and they are calculated based on how much unexplained variation there is for each level of the random effect. Therefore, they show how and how much the speaker diverges from an "expected behaviour" (Drager and Hay, 2012).

Usually, in sociolinguistic research employing Mixed Effect models, speakers and lexical items are treated as random effects, but more generally, in cases of nesting, the nested factor must be a random effect, and the nesting factor must be a fixed effect (Johnson, 2014).

While Mixed Effects models may still be inaccurate when handling small samples or imbalanced data (Eager and Roy, 2017), ever since some authors suggested using them (Baayen, Davidson and Bates, 2008; Barr et al., 2013), they became more and more diffused (even in corpus linguistics: Gries, 2015, or in psycholinguistics to control Type I error rates, since "Rbrul's Type I error rate stays close to the theoretical value of 0.05 in many situations where GoldVarb's greatly exceeds it" - Johnson, 2009, 365), gradually replacing also ANOVA tests. While it is desirable to have many tokens in each level of the random variable to provide robust results (Austin and Leckie, 2018; Meyerhoff, 2015), Schweinberger (2022) reports that

several studies (Bell, Ferron, and Kromrey 2008; Clarke 2008; Clarke and Wheaton 2007; Maas and Hox 2005) show how small numbers of observations per random variable level do not negatively affect the robustness of the fixed-effects coefficients. The minimum number of observations per random effect variable level is therefore 1.

This research employed Mixed Effects models using Rbrul. Other than allowing for Mixed Effects modelling, Rbrul improves on GoldVarb by also handling continuous predictors and continuous dependent variables (Johnson, 2009); however, non-continuous dependent variable must be binary. As described in §3.4. and §3.4.2., our dependent variable – (r) – had three levels, that therefore had to be somehow merged in order to get a binomial variable. Since different mergings implied different linguistic assumptions, two analyses were conducted: in the first one, the levels Fricative and Approximant were collapsed, in order to study the distribution of Reduced variants of (r) compared to non-Reduced variants of (r) - Tap&trill. In the second one, instead, the levels Fricative and Tap&trill were collapsed, in order to study the distribution of more Italian-like variants of (r) compared to more English-like variants of (r).

On the one hand, considering Fricative and Approximant realizations as more reduced forms of Taps&trills comes from claims in literature (cf. §2.1.1.) suggesting that a "failed" trill could result in a fricative or an approximant, due to a missed complete constriction between the articulators. More support comes from diachrony, with some scholars suggesting for different languages historical relations between taps and approximants, and from studies on lenition, that suggest fricatives and approximants as lenited – thus, weaker – variants of trills and taps (cf. §2.2.).

On the other hand, considering Fricative and Tap&trill realizations as more Italian-like is due to the fact that they are both realizations attested in Italy, with Fricative rhotics being potentially relevant in the low-southern area (Celata, Meluzzi, and Ricci, 2016; cf. §2.3.2.1.; §2.3.2.2.), while the only (consonantal) rhotic realization attested in Canadian English is the Approximant, according to our previous literature review (§2.3.1.).

Step-up/step-down comparison of models determined the best-fitting model for a given set of factors included. Different best-fitting models for different sets of fixed factors included were compared, to assess which factors improved the model more. This procedure was followed to avoid including collinear factors together in a same model; for instance, the factors Age and Decade have always been tested separately. Hence, multicollinearity was prevented as much as possible, but it was always still controlled by checking the VIFs (the Variance Inflation Factor,

a measure of how much a predictor correlates with another one included in the model), and by excluding models featuring VIFs > 0[74] (Schweinberger, 2022).

Combined factors such as Sex.Generation, Following phone.Stress of syllable, or Position in the syllable.Stress of the syllable have also been tested with three main purposes:

1. To either identify constraint differences between groups, or to examine how patterns regarding linguistic features may be affected by other linguistic features;
2. To avoid incomplete information, i.e., a combination of variants (or levels) that does not occur in the data: in other words, if one wanted to include two separate fixed effects in the models, one should first check whether there are empty cells in a cross-tabulation between these two variables, and if so, use a combined factor instead of the two factors by themselves, to avoid unreliable models;
3. To check whether a fixed factor, not/hardly significant by itself, is instead part of a significant interaction with another fixed factor (Schweinberger, 2022).

Moreover, levels within a factor having really close factor weights and similar properties (e.g., "Sonorant" and "Obstruent") may have been merged (e.g., into "Consonant") to improve the model by simplifying the of levels.

The final best-fit model is the one showing the factors that account for the most variance in the dataset, while at the same time optimizing (reducing) the number of variables and levels in the set of independent variables. It is identified by comparing the AICC scores of the models and selecting the lowest one. Models were constructed both for the dataset as a whole, and for each generation, to assess the factors that constraint the selection among the (r) variants in each group. The significance threshold was set at $p < 0.05$ for every model, and, as Johnson (2014) and Schweinberger (2022) suggest, Speaker and Lexical item word were treated as Random effects in every model.

After having assessed which factors are significant using Mixed Effects models, a one-level model with only those factors was run. The numeric results in §3.5. all come from the one-level analyses.

Results will be presented as tables listing:

1. The significant factors of the best model, i.e., the predictors;
2. The logodds corresponding to each level of the factor. Logodds have the same purpose as factor weights (see point 5), but instead of employing a range from 0 to 1, they use a

[74] The boundary is set so low because there is no clear consensus about the maximum VIF value acceptable.

logarithmic scale. Since different science communities or software users may be more familiar with logodds than factor weights, we include them;

3. The number n of tokens for each level of the factor;
4. The percentage of tokens featuring the application value, compared to all tokens, for each level of the factor (e.g., % of Approximant&Fricative tokens compared to all tokens);
5. The factor weights corresponding to each level of the factor. Factor weights are a measure of probability of the application value occurring in the context indicated by the level. They range from 0 (disfavoring the application value) to 1 (favoring the application value) (Tagliamonte, 2006). They are listed from highest to lowest, so at the top of the hierarchy we will find the contexts most favoring the appearance of the relevant variant(s);
6. The range of the factor. The range is a measure of relative strength of the predictor, and it is obtained by subtracting the lowest factor weight from the highest factor weight of the factor. By comparing the ranges of the significant factors of a model, we can identify the strongest and weakest constraints (hence *relative* strength).

In order to evaluate our hypotheses, we will use the three lines of evidence formulated by Poplack and Tagliamonte, designed to show cross-generational similarities or differences (2001, 92):

1) Statistical significance: which factors are statistically significant, and which are not?
2) Relative strength: which factor is most significant (largest range) or least (smallest range)?
3) Constraint hierarchy: how are levels ranked within factors?

These comparisons are used to identify changes across grammars of different groups (Tagliamonte, 2006). Various scenarios (cf. §1.7.) can be identified by using this methodology, including grammar simplification or maintenance across groups.

3.5. Results

3.5.1. Distributional results

As mentioned in the previous section, the final number of tokens of (r) analyzed was 1547. Let us first present the distribution of the three (r) variants investigated across all speakers:

	Taps&trills	Approximants	Fricatives
Raw frequency	1106	360	81
% of tokens	72%	23%	5%

Table 8: Distribution of (r) variants in the data (n = 1547).

Immediately, Taps&trills stand out as the most frequent variants, while Fricatives are quite rare. The distribution of our dependent variable levels across Generations is shown in Table 9 and Figure 4.

(r) variants	**Generations**		
	Homeland (n = 668)	Generation 1 (n = 419)	Generation 2 (n = 460)
Taps&trills	463	313	330
Approximants	161	81	118
Fricatives	44	25	12

Table 9: Distribution of (r) variants across Generation of speakers.

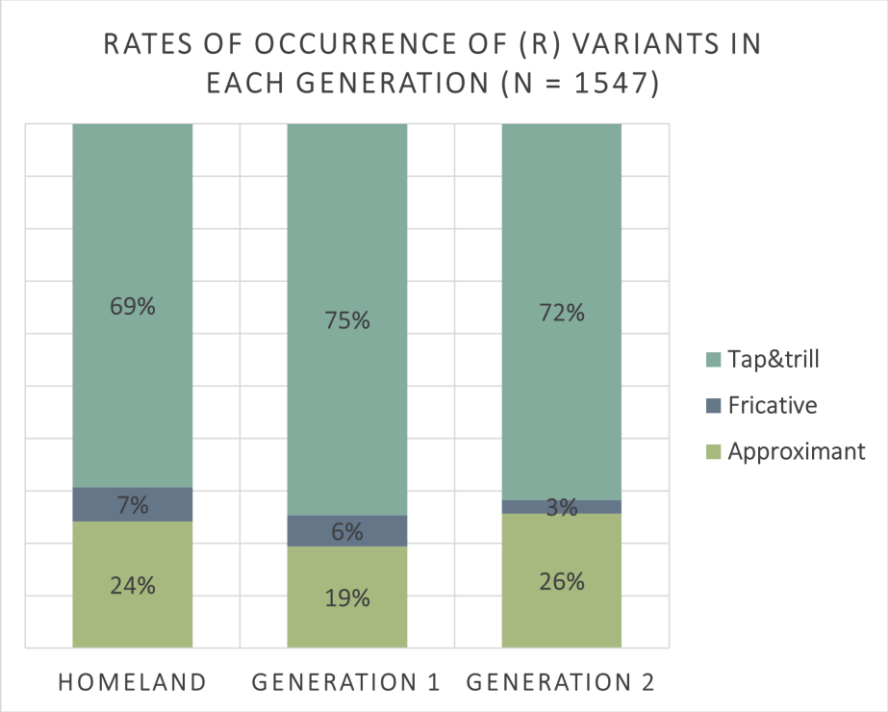

Fig. 4: Percentage of (r) variants across Generation of speakers (n = 1547) (the percentage refers to the column total).

Table 9 and Fig. 4 show that generations maintain a similar proportion between Taps&trills (more than two thirds of the variants produced) and the other variants. However, the proportion between Approximants and Fricatives is not similar as across groups: we can notice a reduction of the number of Fricative variants produced going from Homeland to Generation 2.

On the other hand, Fig. 5 demonstrates how the rates of occurrence of Approximants and Fricatives do not correlate strongly with Age: the data lacks any age-related linear increase or decrease of the rates of occurrence. The slope of the line for Tap&trill variants suggests an age-related change, with older speakers producing fewer Taps&trills than younger speakers. While this distributional result will not be reflected in a Mixed Effects model employing all speakers (cf. §3.5.3., Table 18), models analyzing each generation separately will, indeed, see this Age pattern confirmed (Tables 19, 20, 21).

Moreover, in Fig. 5 each symbol represents the raw frequency of use of one variant for one speaker. Hence, the graph also serves the purpose of showing how consistently the preference for Taps&trills (already suggested by Tables 7, 28, and Figure 4) is replicated at the individual level.

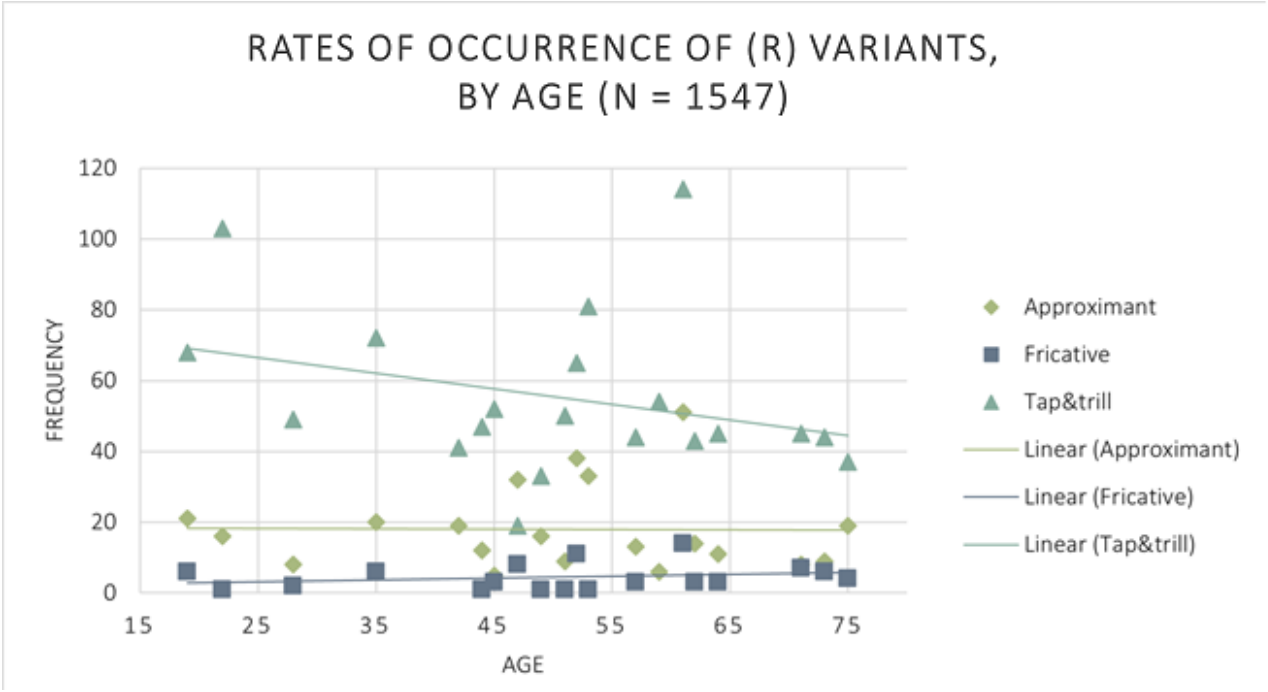

Fig. 5: Frequency of production of (r) variants based on speakers' Age.

Instead, a striking difference emerges for speakers' Sex across all Generations:

(r) variants	Sex	
	Males (n = 909)	*Females (n = 638)*
Taps&trills	581	525
Approximants	271	89
Fricatives	57	24

Table 10: Distribution of (r) variants in the whole dataset, based on speakers' Sex.

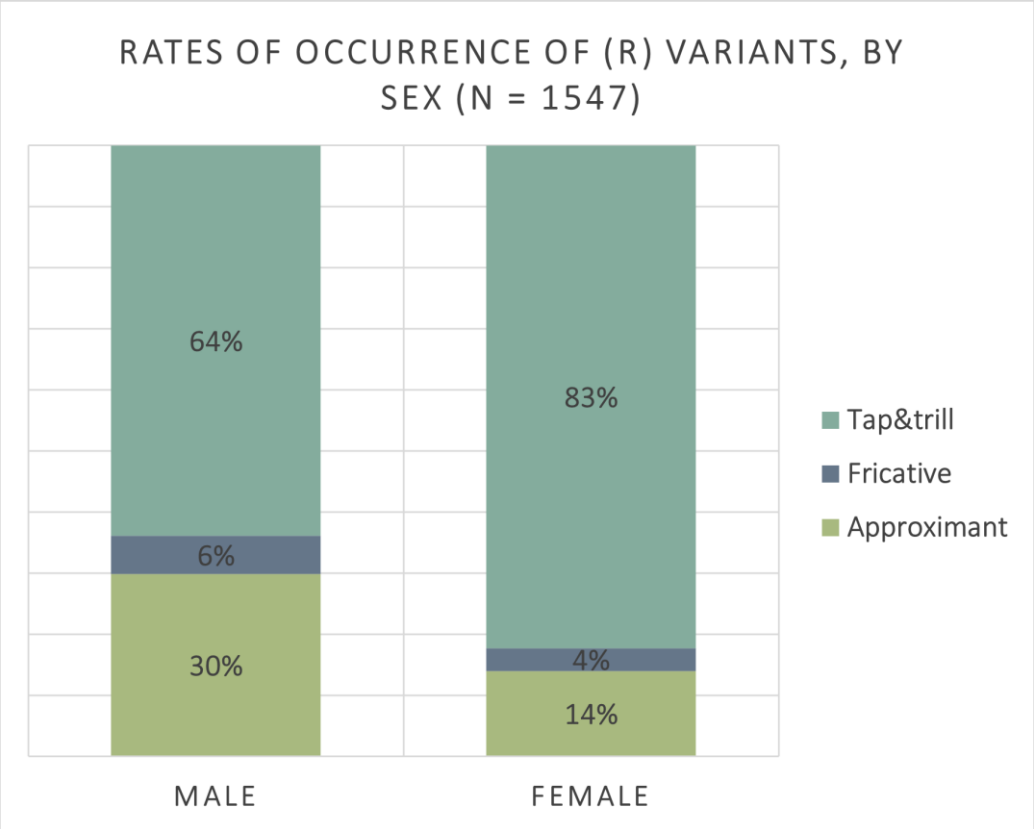

Fig. 6: Percentage of (r) variants based on speakers' Sex (n = 1547) (percentages refer to column totals).

Table 10 and Fig. 6 show that women across all groups produce far fewer Approximant and Fricative variants compared to men. We next check whether the behavior of a single Generation is skewing the overall distribution; perhaps it is either the males or the females from one Generation only who are responsible for this imbalance in the data.

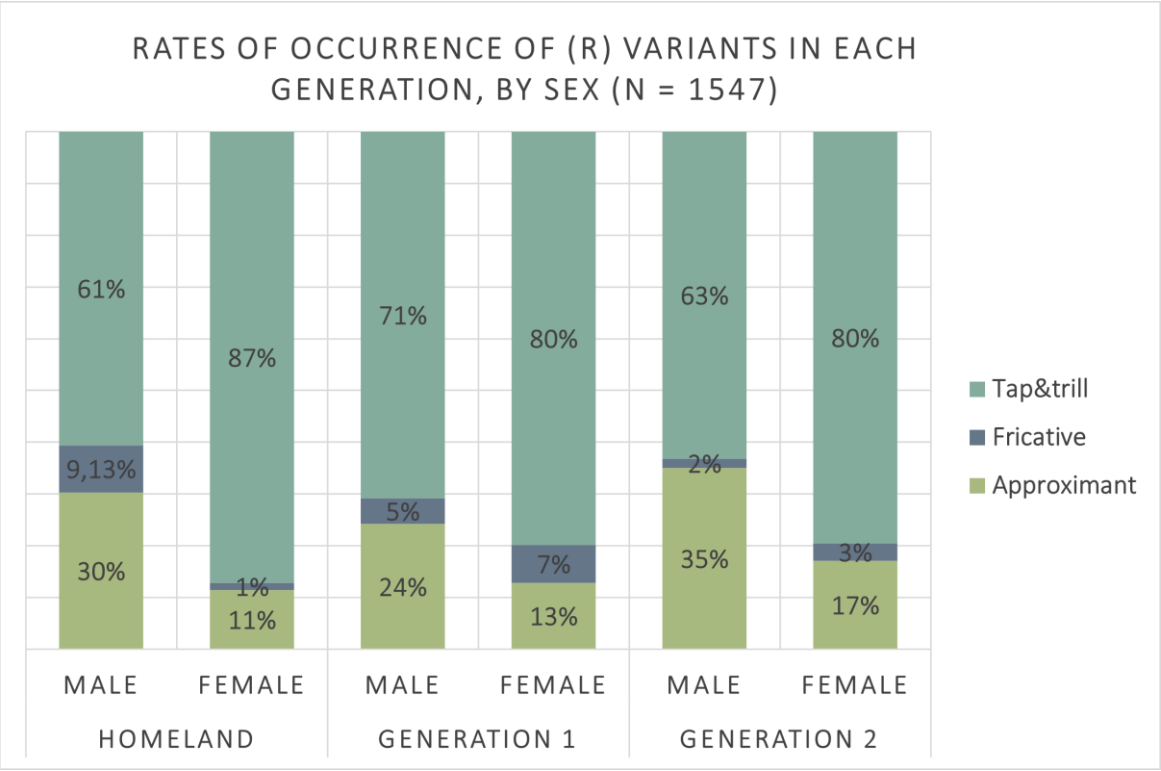

Fig. 7: Percentage of (r) variants per Generation and Sex (percentages refer to column totals).

Fig. 7 shows that a Sex-based difference appears consistently in every Generation group; however, it is slightly less pronounced in Generation 1.

Another interesting distribution is that of the different (r) realizations across the 14 Heritage speakers whose EO scores were available:

(r) variants	**Ethnic Identity**		**Language Use**	
	High (7 speakers)	*Low (7 speakers)*	*High (7 speakers)*	*Low (7 speakers)*
Taps&trills	322	288	322	288
Approximants	74	109	72	111
Fricatives	23	12	25	11

Table 11: Distribution of (r) variants based on Heritage speakers' EO scores (n = 829).

It must be specified that speakers who ranked a given score for EO ID are not the same who ranked the same score for EO LU (e.g., the 7 speakers who were High in ID are not the same who were High in LU). Hence, it is quite striking how the variants are distributed similarly

between the two different measures. In this exploration, the behavior of speakers with a certain level of Italian ethnic identity corresponds to the behavior of speakers with a certain level of Italian language use.

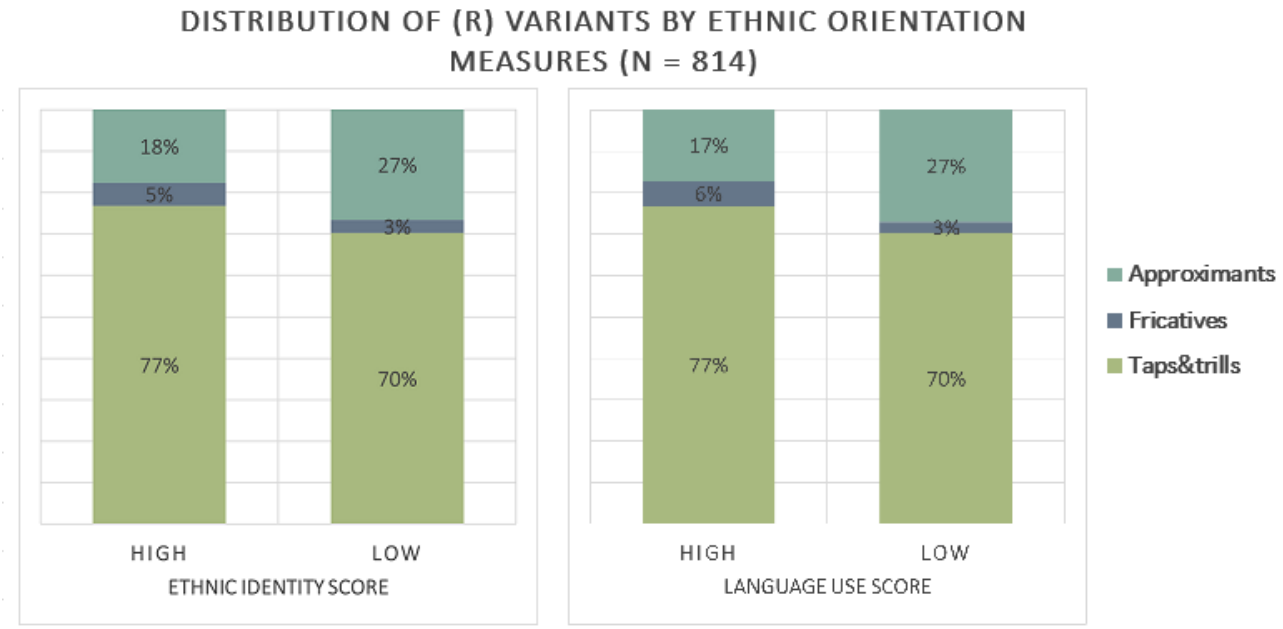

Fig. 8: Percentage of (r) variants based on EO scores (percentages refer to column totals).

Moreover, the EO measures seem to correlate with the percentage of Approximant variants produced, i.e., speakers oriented towards Italian ethnicity&language (High EO LU and High EO ID) produce fewer Approximants (English-like variants). This is visible in Fig. 8 above: by visual inspection, we could also say that, if these speakers use a weaker form of (r), this has more chances to be a Fricative than an Approximant compared to speakers with low EO scores. Therefore, we expect the EO scores to play a role in determining the output of the binary choice between English-like and Italian-like variant.

We will first present the best-fitting Mixed Effects models for the Reduced (Approximant&Fricative variants) vs. the non-Reduced (Tap&trill variants) comparison, and then for the English-like (Approximant variants) vs. Italian like (Fricative&Tap&trill variants) comparison. Finally, we will contrast models including EO scores to determine if Ethnic Orientation is a predictor of English-like (r) variants.

3.5.2. Reduced vs. non-Reduced: Mixed Effects models' results

Let us first contrast the outputs of the binary choice Reduced vs. non-Reduced.

While the total number of coded tokens was n = 1547, tokens originally coded with the level Sentence final for the factor Following phone had to be excluded, as mentioned in §3.4.2. This brought our total number of analyzed tokens to n = 1540.

Lexical word varies from 9-58% of Reduced forms; Speaker varies from 3-68% of Reduced forms.

Models fit via step-up/step-down comparison indicate that the strongest predictors of production of Reduced (r) variants in the whole dataset are Position in the syllable and Sex (table 9). Other factors tested in the model were not significant. In particular, for the factor Sex, Male speakers favor Reduced variants compared to Female speakers: we will see that this effect will stay the same across all models. As for the factor Position in the syllable, the level Coda favors Reduced variants, while Onset equally disfavors Reduced variants.

As the visual inspection in §3.5.1. suggested, neither Generation (even when merging Generation 1 and Generation 2) nor Decade were significant.

Predictor	**Logodds**	**n**	**% Reduced tokens**	**Factor weight**	**Range**
Position in syllable (p < 0.001)					*34*
Coda	0.692	431	41%	0.67	
Nucleus	0.006	23	26%	0.50	
Onset	-0.699	1086	24%	0.33	
Sex (p < 0.001)					*28*
Male	0.584	903	36%	0.64	
Female	-0.584	637	18%	0.36	

Table 12: Best-fitting Mixed Effects model with Reduced (Approximant&Fricative) as application value and non-Reduced (Tap&trill) as non-application value, for all speakers (n = 1540).

We next compare models for Homeland, Generation 1, and Generation 2 speakers, in order to assess differences across these groups, either in the significance of the conditioning factors or in their ranking. Let us remember that, since these models employ fewer tokens than those for all speakers, fewer factors may emerge as significant.

Please also note that, for the following sub-analyses, a few tokens (n= 10 in the Homeland; n = 9 in Generation 1; n = 4 in Generation 2) had to be excluded from the models: tokens coded as Nucleus for the factor Position in the syllable. In fact, these tokens constantly skewed the analyses, since there was too few of them in each subgroup to providing statistically reliable results. The following table illustrates the distribution among (r) variants of the Nucleus tokens in each subgroup:

	Nucleus tokens			
	Tap&trill	Approximant	Fricative	*Total*
Homeland	7	1	2	*10*
Generation 1	8	0	1	*9*
Generation 2	2	1	1	*4*
Tot.	*17*	*2*	*4*	

Table 13: Coding of the Nucleus (r)s across the different groups of speakers.

The strongest predictors for Reduced (r) variants among Homeland speakers are the combined factor Sex.Decade and Position in syllable (Table 14). The combined factor let us see that younger speaker of both sexes disfavor Reduced variants. The effects of Gender and Position in Syllable are as in the dataset as a whole.

Predictor	**Logodds**	**n**	**% Reduced tokens**	**Factor weight**	**Range**
Sex.Decade (p < 0.01)					*52*
Male.3, 4, 5, 6	1.324	292	48%	0.70	
Male.1, 2	0.101	147	22%	0.53	
Female.5	-0.446	59	17%	0.39	
Female.2, 3	-0.979	157	12%	0.27	
Position in syllable (p < 0.05)					*17*
Coda	0.34	213	33%	0.58	
Onset	-0.34	442	29%	0.42	

Table 14: Best-fitting Mixed Effects model with Reduced (Approximant&Fricative) as application value and non-Reduced (Tap&trill) as non-application value, for Homeland speakers (n = 655).

For Generation 1, the strong predictors are again the combined factor Sex.Decade and Position in the syllable (Table 15). The group follows previous trends, with Males favoring reduction more than Females, and older speakers favoring reduction. A model including Age as a continuous factor in place of Decade has been tested; however, it led to a mismatch between the step-up and the step-down procedures. As in the preceding models, Coda contexts favor Reduced variants.

Predictor	Logodds	n	% Reduced tokens	Factor weight	Range
Sex.Decade (p < 0.01)					*52*
Male.7	1.062	60	38%	0.74	
Male.6	0.101	175	26%	0.53	
Female.7	0.062	113	25%	0.52	
Female.5	-1.226	58	10%	0.23	
Position in syllable (p < 0.001)					*47*
Coda	1.026	116	51%	0.74	
Onset	-1.026	290	16%	0.26	

Table 15: Best-fitting Mixed Effects model with Reduced (Approximant&Fricative) as application value and non-Reduced (Tap&trill) as non-application value, for Generation 1 speakers (n = 406).

For Generation 2, the strongest predictors are the combined factors Stress of syllable.Position in syllable and Sex.Decade. The linguistic factor shows an inner ranking within Coda contexts: as a whole, they favor reduction compared to Onset contexts (Monosyllabic tokens are all tokens of the word *per*, i.e., tokens of Coda (r)), but some Coda contexts favor it more than others. It must be specified that all tokens of the level Unstressed.Coda come from instances of *perché*, "why"/"because". The effect of Sex.Decade is as in previous models.

Predictor	Logodds	n	% Reduced tokens	Factor weight	Range
Stress of syllable.Position in syllable (p < 0.001)					*43*
Stressed.Coda	1.075	53	59%	0.75	
Monosyllabic	0.185	29	38%	0.55	
Unstressed.Coda	-0.490	20	25%	0.38	

Predictor	Logodds	n	% Reduced tokens	Factor weight	Range
Onset	-0.771	354	23%	0.32	
Sex.Decade (p < 0.01)					*36*
Male.5	0.766	109	40%	0.68	
Male.4	0.290	108	33%	0.57	
Female.5	-0.327	120	23%	0.42	
Female.4	-0.729	119	18%	0.33	

Table 16: Best-fitting Mixed Effects model with Reduced (Approximant&Fricative) as application value and non-Reduced (Tap&trill) as non-application value, for Generation 2 speakers (n = 456).

Results from the Reduced vs. non-Reduced comparison are summarized as follows, with the predictors having the most consistent behavior listed first.

Factor	All speakers	Homeland	Generation 1	Generation 2
n	*1540*	*655*	*406*	*456*
Position in syllable	√	√	√	
Sex.Decade		√	√	√
Stress of syllable.Position in syllable				√
Sex	√			

Table 17: Significant predictors (indicated with "√") in Mixed Effects models comparing Reduced vs. non-Reduced (r) variants, for each speaker group.

Table 17 shows us the maintenance of both the linguistic factor across generations (since Position in syllable is present also in Generation 2, even if combined), and of the social factor Sex.Decade. The only exception to this is in the "All speakers" analysis, where Sex is significant by itself. Relative strength of the constraints, and constraint hierarchies (Poplack and Tagliamonte, 2001), will be discussed in §3.6.

3.5.3. English-like vs. Italian-like: Mixed Effects models' results

We will now present the results from the analyses that employ the second way of collapsing the levels of the dependent variable, i.e., that contrast the choice between English-like variants (Approximants) vs. Italian-like variants (Fricatives and Taps&trills). However, since the amount of Fricative realizations found in our data was quite low, compared to the other two realizations, results from these models might not be radically different from the ones in §3.5.2. We could have the same significant factors, but with different strengths or with different rankings. Even if the outcomes of the models in §3.5.2. and §3.5.3. might be similar, we still wanted to try both the collapsing strategies to avoid any bias and to make sure that particular constraints or effects are not lost due to the use of just one of the collapsing strategies.

For the following models, one speaker who categorically produced just Italian-like variants (IXF35A) had to be excluded from the dataset, bringing the analyzed number of tokens to n = 1501. This speaker patterned similarly to the other Female speakers of the dataset, especially Homeland Females, who also disfavor English-like variants, based on our visual inspection of Fig. 7, §3.5.1.

Lexical word varies from 8-46% of English-like forms; Speaker varies from 14-54% of English-like forms.

Models fit via step-up/step-down comparison indicate that the strongest predictors of production of English-like (r) variants in the whole dataset are the combined factor Following phone.Stress of the syllable and Sex (Table 18). Crucially for this binary interpretation of the dependent variable, again, Generation was not significant, even when merging Generation 1 and 2. Decade also was not significant. The lack of any "Heritage-related" effects and the sex related variation follow the trends from the §3.5.2. analyses.

Predictor	**Logodds**	**n**	**% English-like tokens**	**Factor weight**	**Range**
Following phone.Stress of syllable ($p < 0.001$)					*68*
Sonorant	1.714	38	68%	0.85	
Monosyllabic	0.238	107	31%	0.56	
Obstruent.Stressed	0.002	182	29%	0.50	
Vowel	-0.364	1064	23%	0.41	
Obstruent.Unstressed	-1.590	110	8%	0.17	

Predictor	Logodds	n	% English-like tokens	Factor weight	Range
Sex (p < 0.001)					*25*
Male	0.511	903	30%	0.63	
Female	-0.511	598	15%	0.38	

Table 18: Best-fitting Mixed Effects model with English-like (Approximant) as application value and Italian-like (Tap&trill&Fricative) as non-application value, for all speakers (n = 1501).

The hierarchy of levels within the linguistic factor shows us a nuanced situation. The analysis in §3.5.2. suggested that Coda contexts favor reduction. In our corpus, all Coda tokens are also followed by consonants, while tokens followed by consonants can either be Coda or (the small number of) Nucleus tokens. All Onset tokens are followed by vowels, but the opposite is once again not true: tokens from *per* may be followed by vowels, and those are either Coda or Nucleus tokens. However, in the ranking shown in Table 18 all tokens from *per* are grouped as Monosyllabic, no matter what the following phone is[75]. This means that the level Vowel coincides with Onset tokens only, and the other levels with Coda or Nucleus tokens.

Table 19 below clarifies this situation by showing the correspondences between the levels featured in the Following phone.Stress of the syllable factor, and the Position in syllable of (r).

Following phone.Stress of syllable	Coda		Onset		Nucleus		*Row total*
	n.	% English-like	n.	% English-like	n.	% English-like	
Sonorant	35	71%	-	-	3	33%	*38*
Monosyllabic	104	31%	-	-	3	0%	*107*
Obstruent.Stressed	179	0%	-	-	3	33%	*182*
Vowel	-	-	1064	23%	-	-	*1064*
Obstruent.Unstressed	96	9%	-	-	14	0%	*110*
Column total	*412*		*1066*		*23*		*1501*

Table 19: Crosstabulation of the tokens from the various levels of the Following phone.Stress of the syllable factor, and their Position in syllable, across all speakers (n = 1501).

Consequently, the ranking of levels in Table 18 is another way to express that:

[75] See footnote 22 for a discussion.

1. Most of the Coda (r)s favor the English-like variant, with some of them favoring it more than the others;
2. All Onset (r)s disfavor it;
3. Some Coda and Nucleus (r)s strongly disfavor the English-like variant, even more than Onset (r)s: the Obstruent.Unstressed ones (all from the word *perché),* i.e., the only pre-stress occurrences of (r) we have.
4. When the following phone is a Sonorant, the Stress of the syllable is not a relevant factor[76];
5. When the token is Monosyllabic, the Following phone is not a relevant factor[77].

We next compare models for Homeland, Generation 1, and Generation 2 speakers. As in the previous section, Nucleus tokens were excluded to provide more robust analyses.

Table 20 below indicates that the strongest predictors of English-like (r) variants for Homeland speakers are the combined factors Following phone.Stress of the syllable, Lexical Class, and Sex.Decade. Females as a whole disfavor English-like variants. The lack of a significant Age intersection with Females is due to the narrower age range within the Homeland Female group. As we can see from our speakers' table (Table 3, §3.4.2.), four of them are in their 20s, one in her 30s, and one in her 50s. On the other hand, Homeland Males also include speakers in their 40s and 60s. In fact, from Table 20 we can see that younger Males disfavor English-like variants compared to older Males. We also see the Lexical factor play a role for the first time, with the class Verb most favoring the approximant. For the phonological factor, we see once again how some Coda contexts (Coda (r)s followed by Sonorants or in a Monosyllabic word) favor English-like variants and other do not (Coda (r)s followed by Obstruents), but differently than Table 18, Onset (r)s – i.e., followed by Vowels – are higher in the ranking[78].

[76] Indeed, we collapsed the levels Sonorant.Stressed and Sonorant.Unstressed because they had almost identical Factor weights.
[77] This was confirmed in a separate Mixed Effects model involving just the Monosyllabic tokens – i.e., tokens of *per*. It did not find the Following phone to be a predictor. Of course, this could be due to the small number of *per* tokens, but visual analysis confirmed this result. As a further proof, when running the analysis corresponding to Table 18 with Monosyllabic.Sonorant, Monosyllabic.Obstruent, and Monosyllabic.Vowel as separate levels, they had really close Factor weights. These elements suggested that the feature governing the behavior of *per* tokens is not their Following phone, but their Stress – hence they were grouped based on the latter, and not the former (e.g., as in a merging like "Obstruent.Stressed+Monosyllabic").
[78] For the merging involving the Sonorant levels, and the Monosyllabic levels, see footnotes 21 and 22.

Predictor	Logodds	n	% English-like tokens	Factor weight	Range
Following phone.Stress of syllable (p < 0.001)					*82*
Sonorant	2.287	11	0.818	0.908	
Monosyllabic	0.254	53	0.283	0.563	
Vowel	0.006	421	0.283	0.502	
Obstruent.Stressed	-0.192	75	0.200	0.452	
Obstruent.Unstressed	-2356	56	0.036	0.087	
Lexical class (p < 0.05)					*34*
Verb	0.520	175	0.371	0.627	
Pronoun/Adjective	0.244	25	0.280	0.561	
Adverb/Conjunction/Preposition	0.163	332	0.235	0.541	
Noun	-0.926	84	0.119	0.284	
Sex.Decade (p < 0.05)					*33*
Male.3, 4, 5, 6	0.715	292	0.366	0.671	
Male.1, 2	-0.078	147	0.190	0.481	
Female	-0.637	177	0.141	0.346	

Table 20: Best-fitting Mixed Effects model with English-like (Approximant) as application value and Italian-like (Tap&trill&Fricative) as non-application value, for Homeland speakers (n = 616).

For Generation 1, the stronger predictors are the combined factor Sex.Decade and Position in syllable (Table 21). Males favor English-like variants more than Females, and within each Sex, older speakers favor English-like variants more than younger ones. The linguistic factor indicates that Coda contexts favor English-like variants more than Onset ones. The use of finer-grained factors that would allow us to identify a ranking within Coda contexts, as in Tables 20 and 18, produced worse-fitting models.

Predictor	Logodds	n	% English-like tokens	Factor weight	Range
Sex.Decade (p < 0.05)					*41*
Male.7	0.963	60	0.317	0.724	
Male.6	0.178	175	0.223	0.544	
Female.7	-0.360	113	0.150	0.411	
Female.5	-0.780	58	0.103	0.314	
Position in syllable (p < 0.01)					*34*
Coda	0.702	116	0.336	0.669	

Predictor	Logodds	n	% English-like tokens	Factor weight	Range
Onset	-0.702	290	0.145	0.331	

Table 21: Best-fitting Mixed Effects model with English-like (Approximant) as application value and Italian-like (Tap&trill&Fricative) as non-application value, for Generation 1 speakers (n = 406).

For Generation 2 (Table 22), the stronger predictors are the combined factors Following phone.Stress of syllable and Sex.Decade. While the social factor does not suggest anything different from the previous analyses, the linguistic factor gives us some new information. Compared to the Homeland's ranking (Table 20), Obstruent.Stressed tokens favor English-like variants much more, and Vowel tokens (Onset tokens) less. The highest and lowest ranked levels stay the same. Hence, this hierarchy seems a one step further towards the tendency to have Coda contexts favoring English-like variants, compared to Homeland speakers – something that would be cohesive with Generation 1's behavior (Table 21).

Predictor	Logodds	n	% English-like tokens	Factor weight	Range
Following phone.Stress of syllable (p < 0.001)					*73*
Sonorant	1.570	11	64%	0.83	
Obstruent.Stressed	0.756	42	45%	0.68	
Monosyllabic	0.444	29	38%	0.61	
Vowel	-0.590	354	22%	0.36	
Obstruent.Unstressed	-2.180	20	5%	0.10	
Sex.Decade (p < 0.01)					*38*
Male.5	0.744	109	38%	0.68	
Male.4	0.419	108	32%	0.60	
Female.5	-0.312	120	20%	0.42	
Female.4	-0.851	119	14%	0.30	

Table 22: Best-fitting Mixed Effects model with English-like (Approximant) as application value and Italian-like (Tap&trill&Fricative) as non-application value, for Generation 2 speakers (n = 456).

Results for the English-like vs. Italian-like comparison are summarized as follows, with the predictors having the most consistent behavior listed first.

Factor	All speakers	Homeland	Generation 1	Generation 2
n	*1501*	*616*	*406*	*456*
Following phone.Stress of syllable	√	√		√
Sex.Decade		√	√	√
Sex	√			
Lexical class		√		
Position in syllable			√	

Table 23: Significant predictors (indicated with "√") in Mixed Effects models comparing English-like vs. Italian-like (r) variants, for each speaker group.

Generally, the same factors play a role in all three groups of speakers. Table 23 shows us the maintenance of the social factor Sex.Decade across generations. The only exception to this is in the "All speakers" analysis, where Sex is relevant by itself. It also shows us the maintenance of the phonological factor across generations. While Generation 1 could seem an exception, we explained that Following phone.Stress of the syllable is a factor actually showing an inner hierarchy among tokens in Coda position; as a result, we can state that Generation 1 patterns as the other groups, but simply in a more general way. Lastly, this table highlights how the Lexical factor is significant only for Homeland speakers.

3.5.4. Mixed Effects models including EO scores

Lastly, we want to test whether Ethnic Orientation relates to the production of (r) variants among heritage speakers. Since EO scores are a measure of affinity to Italian or Canadian culture and Italian or English language, we are interested to check if they are involved in the binary choice between English-like vs. Italian-like variants (while keeping in mind that all variants appear in Homeland and Heritage varieties at similar rates). The following models are based on a sample of 14 (out of the 15 available) Heritage speakers, and also exclude Nucleus tokens.

The strongest predictors for the 14 Heritage speakers are the combined factor Following phone. Position in syllable, Sex, EO LU, and EO ID.

Predictor	Logodds	n	% English-like tokens	Factor weight	Range
Following phone.Stress of syllable(p < 0.001)					*57*
Sonorant	1.662	23	70%	0.84	
Monosyllabic	0.197	48	35%	0.55	
Obstruent.Stressed	0.092	102	35%	0.52	
Vowel	-0.958	603	18%	0.28	
Obstruent.Unstressed	-0.933	38	16%	0.27	
Sex (p < 0.001)					*22*
Male	0.443	404	29%	0.61	
Female	-0.443	410	16%	0.39	
EO.ID (p < 0.01)					*12*
Low	0.265	408	27%	0.57	
High	-0.265	406	18%	0.43	
EO.LU (p < 0.05)					*10*
Low	0.192	408	27%	0.55	
High	-0.192	406	18%	0.45	

Table 24: Best-fitting Mixed Effects model with English-like (Approximant) as application value and Italian-like (Tap&trill&Fricative) as non-application value, for 14 Heritage speakers (n = 814).

The linguistic factor is the same we saw for Generation 2 (and Homeland), with an almost identical ranking; the factor Sex follows the now usual trend of Male speakers favoring English-like variants. In addition, both our measures of Ethnic Orientation are significant, with speakers having lower levels of orientation towards Italian ethnicity&language favoring English-like variants. We note, however, that the strength of the effects of these EO scores (illustrated by the ranges) are considerably smaller than for the other factors discussed in this section.

3.6. Discussion

We begin by comparing the outputs of the two different analyses we conducted, in order to see which collapsing strategy provided the most cohesive results. Table 25 joins tables 17 and 23 to present the comparison. Results regarding the analyses of all speakers, and numbers of tokens, have been excluded for clarity purposes. Ranges indicate the relative strength of predictors in each group.

Predictors	Homeland	Generation 1	Generation 2
Reduced vs. non-Reduced analysis			
Sex.Decade	√	√	√
Ranges	52	52	36
Position in syllable	√	√	
Ranges	17	47	
Stress of syllable.Position in syllable			√
Range			43
Italian-like vs. English-like analysis			
Sex.Decade	√	√	√
Ranges	33	41	38
Following phone.Stress of syllable	√		√
Ranges	82		73
Position in syllable		√	
Range		34	
Lexical class	√		
Range	34		

Table 25: Comparison of the significant predictors (indicated with "√") and their ranges for each speaker group, according to two different collapsing strategies of the dependent variable.

Analyses employing one or the other binary choice are not so different from one another, in terms of significant constraints, as is illustrated in table 25. First of all, both analyses show the same social pattern across generations: the Sex.Decade effect is the same. Then, the first analysis shows the consistency of the phonological factor Position in syllable across groups; this is more or less mirrored by the second analysis. As we said at the end of §3.5.3., Following phone.Stress of the syllable is a factor showing an inner hierarchy among tokens in Coda position – and Position in syllable is, indeed, the significant predictor for Generation 1. Therefore, we can say that both analyses reveal a similar phonological pattern across generations, where the position of (r) in the syllable stays the most important condition affecting (r) realizations, but, for some groups, it interacts with other phonological aspects (such as stress) to produce more specific patterns/hierarchies. The only difference between the two analyses lies in the detection of the Lexical class predictor for Homeland speakers only in the second method.

As a reminder, the Reduced vs. non-Reduced analysis distinguishes Approximants&Fricatives vs. Taps&trills. The Italian-like vs. English-like analysis distinguishes Approximants vs. Fricatives&Taps&trills.

Since the occurrences of Approximants are a subset of the occurrences of Approximants&Fricatives, one would expect the behavior of the English-like variant to be subjected to the same general rules of the Reduced variants, but perhaps more specific. This would be the case in a language-internal process. On the other hand, in a process influenced by contact with English, Approximants would follow different rules than Fricatives, since Approximants, but not Fricatives, would be transferred from English. The possibility of having two different outcomes from the two collapsing strategies due to English influence, was exactly the reason why we explored both binary choices.

Given that table 25 reveals a good deal of similarity across the outcomes from the two methods of dividing the dependent variable, our results support the idea of Approximants&Fricatives being (r) variants primarily characterized by being Reduced, more than they support an English-like vs. Italian-like distinction. As a matter of fact, we said that the general phonological leitmotiv of both analyses is a position-dependent constraint; it simply becomes more "nuanced" in the English-like vs. Italian-like comparison. The second analysis does not show a completely different grammar regulating the realization of Approximants, compared to Approximants&Fricatives.

For this reason, we will explore in detail the Reduced vs. non-Reduced results, which are also more pertinent to our hypotheses. Results from the other comparison provide additional information about how specific processes unravel.

Let us first discuss the meaning of the phonological predictors (Position in syllable and Stress of syllable.Position in syllable). Then we will discuss the social predictor (Sex.Decade). Finally, we will evaluate our original hypotheses.

3.6.1. Influence of phonetic context on (r)

The significance of the phonological constraints affecting (r) indicates the importance of suprasegmental aspects of consonant reduction. It empirically reinforces the idea of a form of unity among the segments adjacent to an acoustic peak – i.e., of a syllable – and of articulatory consequences (reduction) of said unity.

Our hypotheses relied on the assumption that Approximant and Fricative (r)s, in spoken Italian, are lenited variants of Taps&trills, emerging due to the reduction of the articulatory effort that can characterize some spoken registers (cf. §2.2.).

We based this claim on several lines of evidence:

1) There is no description of such variants in traditional (or Standard) phonological accounts of Italian, meaning they have no counterparts in written Italian, and they are not taught as part of the repertoire to voice professionals, actors, phonologists, or in schools.
2) They appear, however, as (r) variants in studies analyzing spoken Italian from different parts of the nation, indicating that either speech rate or casualness may play a role.
3) They also appear in Italian as "defective" or "lifeless" (r) variants (Romano, 2013; cf. §2.3.2.1.), i.e., the articulatory variants used when speakers are not able to produce a trill, probably due to the effort and precision it requires.
4) Some authors suggest that approximant and fricative rhotics are articulatorily weaker variants of taps and trills[79]. In terms of target undershoot, both approximants and fricatives could be realized by missing the full closure(s) required to produce taps and trills, i.e., by reducing the magnitude of the gesture required (cf. §2.1.1.).
5) Phenomena of lenition are so frequent and pervasive in fast and/or casual speech (Warner, 2012) that there is no reason why rhotics should not be affected by them. As lenition includes spirantization, a passage from Taps&trills to Approximants or Fricatives would be an example of it.

Our results confirm our basic assumption. We linked the production of Approximant and Fricative (r) variants to universal mechanisms and tendencies, and not to language-dependent phonological representations (e.g., the realization of Approximants being a product of the contact with the rhotics of Canadian English), and in fact said variants appear in every group we analyzed, in similar proportions, and constrained by similar phonological factors. This similarity among groups lies in the fact that not only every generation has the same significant predictor, but also that its inner ranking does not change. As a matter of fact, all generations of speakers favor Reduced variants in Coda contexts, compared to Onset contexts. This is also true for Generation 2 speakers, where the first three levels of the ranking are all Coda contexts, and the last level is the Onsets. However, in this case Coda contexts differ in how much they

[79] Barry, 1997; Jaworski, 2010; Jaworski and Gillian, 2011; Gillian and Jaworski, 2014; Kirchner, 2013; Sebregts, 2014; Rennicke, 2015.

favor reduction based on a suprasegmental feature, with Stressed Codas being on top of the ranking, followed by monosyllables and Unstressed Codas. Please note that a monosyllable such as *per*, in Italian phonology, is considered to be unstressed; hence, contrary to popular belief, we find a prosodically prominent unit (Stressed Codas) to be more prone to weakening than its non-prominent counterpart. However, it must be highlighted that it is not the whole unit that we are claiming to be reduced, but just one of its portions: Coda (r), a segment following the most prominent part of the unit, i.e., the vowel nucleus. Hence, the behavior of (r) Nuclei may be of interest to this discussion.

If (r) Nuclei tend to be reduced, then the (r) pattern can be seen as stronger than general suprasegmental tendencies. If they do not, prosodic aspects overcome the pattern. Table 13, describing the (r) variant associated to (r) Nuclei in our dataset, shows how the majority (17/23) of the Nuclei tokens was coded as Taps&trills. Therefore, the tendency to not weaken the most salient part of the unit is respected: even if the nucleus is a consonant, it still stays the salient part of the syllable, and thus is not reduced.

In conclusion, it cannot be a coincidence that Approximant and Fricative variants are systematically favored by Coda contexts, the very contexts that have been so often connected to lenition.

It is also interesting that the relative strength of the phonological constraint is smaller in the Homeland sample, compared to that of Heritage speakers. This means that the gap between how much Codas favor Reduction, compared to Onsets, gets bigger among Heritage speakers: they intensify a pattern already present in the Homeland. This is a hint of grammar boosting: according to Flores and Rinke's (2020, 25) "heritage speakers may boost and further develop tendencies of language (internal) evolution inherent to variable phenomena".

The next question might be whether there is a rationale underlying the choice of the Reduced variant or not. For some varieties it has been assumed that target undershoot turns taps and trills into fricatives, and then fricatives into approximants (Rennicke, 2015). Our results, however, do not suggest such a three-stage process of reduction, but instead a (partial) differentiation of reduction outcomes based on phonetic context.

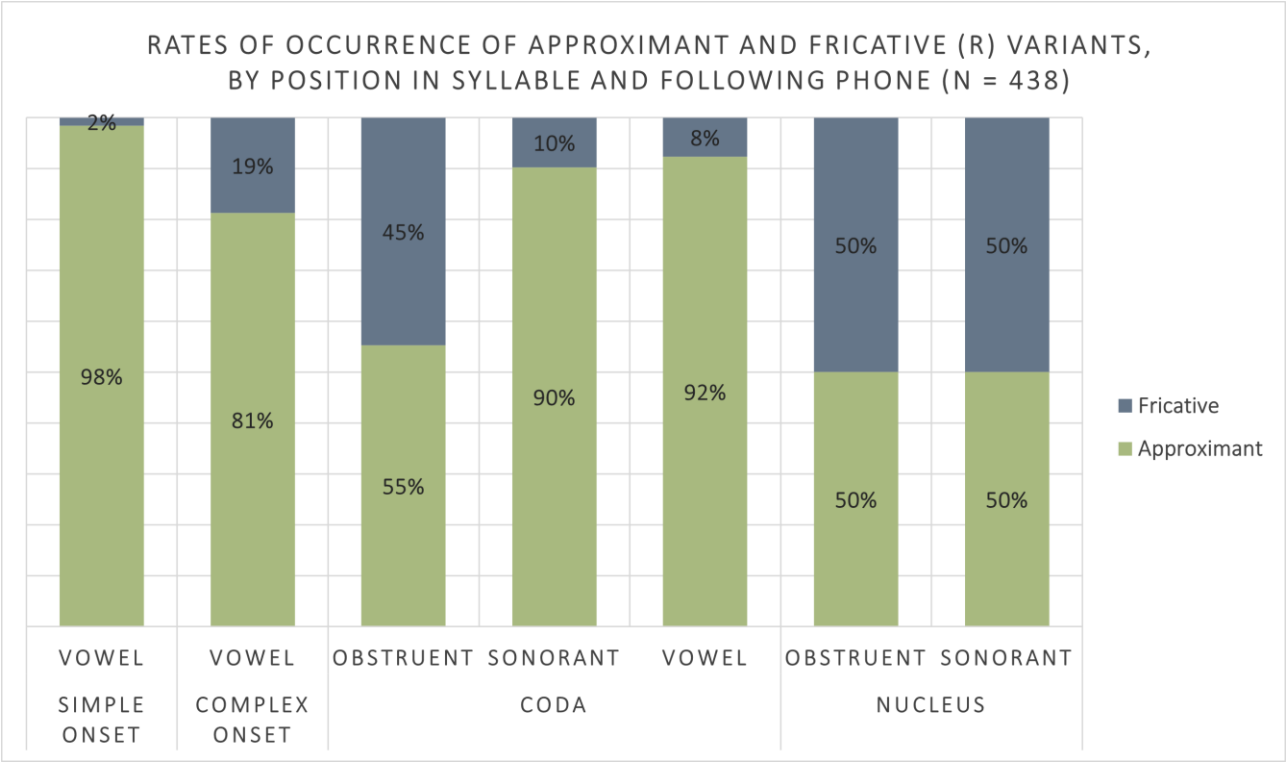

Fig. 9: Distribution of Reduced (r) variants across all speakers, based on the position they occupy in the syllable and the following phone (n = 438).

In Fig. 9, we distinguish Simple onsets and Complex onsets so that we implicitly have information also about the Previous phone. That is, simple onset tokens are always intervocalic, while complex onset tokens are always preceded by Obstruents; Coda tokens are, by definition, preceded by Vowels.

Based on the distribution shown in Fig. 9, and as we already know from Table 9 in §3.5.1., there is a preference to reduce Taps&trills to Approximants, compared to Fricatives. However, when Fricatives *do* emerge, it is in a specific context: when they are followed by Obstruents. In contrast, the context where they appear the least is the intervocalic one, i.e., Simple onsets. Hence, the trend for Approximants and Fricatives is to emerge where it is "easier" for them to appear based on their features. Fricatives are acoustically noise, so they are close to the prototype of consonant; Approximants, instead, share many acoustic and articulatory features with vowels, and hence are more like them. As a result, in vowel-like contexts we have more Approximants, in obstruent-like contexts the amount of Fricatives increases.

While it is preferable to have a sequence of distinct sounds to facilitate perception (cf. Obligatory Contour Principle; Leben, 1973; Goldsmith, 1976), the tendency to ease the job of the listener decreases in hypoarticulated speech, where the preference is to ease the job of the

speaker (Albano Leoni and Maturi, 2018) by reducing effort. In this case, shifting from an articulatory vowel-like setting to an obstruent-like setting (or vice versa) would require more effort than keeping the same setting.

The difference between Homeland and Generation 2 in the hierarchies of Following phone.Stress of syllable, in the English-like vs. Italian-like analysis (Approximant vs. Tap&trill&Fricative), also supports our argument of a context-dependent distribution of Reduced variants. Homeland speakers produce more Fricatives than Generation 2 speakers (44 vs. 12), as Table 9 (§3.5.1.) shows. Hence, in the Homeland case, the numbers are high enough to show us the difference between Approximant and Fricative-like contexts we described above. The two Coda contexts at the bottom of their ranking – Obstruent.Stressed and Obstruent.Unstressed – are there not because they are "less reduced", contradicting the hypothesis of a lenited Coda tendency, but because they are reduced more as *Fricatives*. On the contrary, Generation 2 speakers have: a) crucially, fewer chances to produce Fricatives, since they produce fewer Obstruents after (r); b) so few Fricatives that it is harder to detect the context-dependent effect visible in the Homeland. This is illustrated in Table 26 below.

	Homeland		**Generation 2**	
Stress of syllable.Following phone	*n. of Reduced tokens*	*% Fricatives*	*n. of Reduced tokens*	*% Fricatives*
Monosyllabic	**19**	**21%**	**11**	**0**
Unstressed	**112**	**16%**	**82**	**8%**
Obstruent	16	81%	5	80%
Sonorant	6	0	5	0
Vowel	90	6%	73	4%
Stressed	**72**	**28%**	**36**	**14%**
Obstruent	32	50%	24	21%
Sonorant	3	0	3	0
Vowel	37	11%	9	0
Total	*203*	*21%*	*130*	*9%*

Table 26: Rates of Fricative (r) variants produced by Homeland vs. Generation 2 speakers, based on syllable stress and following phonetic context (n = 333)

In contrast, Generation 1 speakers, who only show the Position in syllable effect, distribute Fricative variants in a more homogeneous way across contexts, compared to the other groups. This would be why the only constraint to emerge for them is the more general Coda constraint.

The idea of variable sounds emerging in contexts already having similar (and thus favoring) features is not new. It has basic coarticulatory explanations, and has been featured, for example, in the interpretation of the distribution of some segments in American English given by Mielke et al. (2010), as we discussed in §2.3.1. These authors concluded that speakers tend to use bunched /r/s in already bunched contexts, and retroflex /r/s in contexts lacking competing tongue shapes. This is also valid for vowels: retroflection tends to show with back vowels, and with /a/ more than /u/ (Ong and Stone, 1998; Tiede et al., 2011). It is implied that retroflex and bunched /r/s are found where it is more "natural" for them to be (with the concept of "naturalness" being further discussed in Mielke et al., 2016). The same concept could be applied here.

Nevertheless, we cannot discard the hypothesis of a three-stage process of reduction (Tap&trill > Fricative > Approximant) based on this information only. Said idea would be supported by younger speakers favoring Approximants rather than Fricatives more than older speakers do. This would be due to younger speakers being generally more advanced in linguistic changes, compared to older speakers. From our analyses, we only know that *older* speakers favor Approximants rather than Fricatives *and* Taps&trills, or Approximants and Fricatives rather than Taps&trills. We do not have access to a comparison between Approximants and Fricatives only, because the amount of Fricatives in our data is too little to allow for statistical analysis.

For this reason, we could interpret the choice between the two Reduced variants in our data in two ways. We know that one of the two outcomes seems to be losing ground in Heritage speakers: while Reduced variants as a whole are present with the same ratio as in the Homeland, Fricative outcomes decrease from Homeland to Generation 1, and from Generation 1 to Generation 2. Therefore:

1. Heritage speakers are boosting the language-internal process of lenition, but the way they are doing it points to an influence from English: they prefer more and more to reduce with Approximants (English-like variants) rather than Fricatives. If Reduced variants were selected *just* based on phonetic context, as we have argued earlier, this would mean that Approximants are being extended also to contexts more suited for Fricatives due to a contact-effect from English.

2. Alternatively, if Reduced variants were distributed along a scale of reduction (Tap&trill > Fricative > Approximant), Heritage speakers would favor Approximants more than Fricatives because they are in a more advanced stage of lenition, coherently with our claim of Heritage boosting.

From this first discussion of the results, we can summarize:

- There is a synchronic pattern of alternation between lenited variants, suggesting a language internal process, in both Heritage and Homeland speakers;
- This process is favored by a prosodically weaker position, i.e., the Coda position;
- This process is stronger in Heritage speakers, given the higher ranges of their "Coda constraint" compared to the Homeland, hence they are possibly boosting lenition;
- This process has two possible outcomes, either an Approximant or a Fricative rhotic, which tend to distribute according to context: Fricatives increase where the following phone is an Obstruent, meaning that, most of the time, we find Approximants where the following phone is a Vowel;
- One of the two outcomes, however, seems to be losing ground in Heritage speakers: while Reduced variants as a whole are present with the same ratio as in the Homeland, Fricative outcomes decrease from Homeland to Generation 1, and from Generation 1 to Generation 2. This could point at an English influence, or at further evidence supporting lenition boosting by Heritage speakers. Future comparison between just Approximant and Fricative realizations - either using the apparent-time construct (cf. §3.1.), or diachronic data – could shed light on this matter.

3.6.2. Influence of sex on (r)

One preliminary disclaimer is needed. The use of the label Sex in our variables, and the discussion we are going to present regarding the linguistic behavior of Male and Female speakers, is not here to deny a constructionist/performative account of gender (Butler, 1990; Omadjohwoefe, 2011), or to encourage an essentialist or two-culture approach to it (Meyerhoff, 2014; Meyerhoff and Ehrlich, 2019). Since this research belongs to a bigger project, it followed its general methodology, which consists in coding speakers based on the sex the interviewer perceived them to be. This is not to reject the speakers' gender identities, but only due to the large amount of data the project deals with, and its main focus - language variation and change across multiple groups and multiple times. Being a quantitative study, dealing with lots of

different cultures, and not focusing specifically on gender practices, the use of interviewer judgements was the easiest and most convenient solution.

We are conscious of the fact that the dichotomy "men and women" is simplistic and obsolete (Eckert and McConnell-Ginet 1992), that there are multiple ways to be a man or a woman (or male and female: see Monro, 2005), and that there are lots and lots of shadings in-between or outside the scope of the dichotomy itself (Monro, 2019). As a matter of fact, we were skeptical of the possibility that Sex was a significant factor in our research (as our hypotheses show). Yet, it is - even as a binary variable, not based on self-reports. Therefore, we will try to explain why it is so by relying on previous literature; this makes use of both the notions of men&women and males&females. We cannot claim that the speakers we coded as Males identify as Men, and those we coded as Females identify as Women, so we cannot be 100% sure that the literature describing gender behaviors fits our data, but in order to provide some kind of reason behind the Sex pattern we found, we will introduce this forced simplification.

The study of language and gender can be said to begin with Lakoff's (1975) *Language and Woman's Place* (Eckert and McConnell-Ginet, 2013), launching decades of rich debate around the topic. From the many interesting notions and ideas developed in the field, we can summarize here two of them that are relevant to our purposes.

The first one is Labov's (1990) principles about gender and language variation. According to Labov, "For stable sociolinguistic variables, men use a higher frequency of nonstandard forms than women […]. In change from above, women favor the incoming prestige form more than men […]. In change from below, women are most often the innovators" (Labov, 1990, 210; 213; 215)[80]. Hence, he highlighted a "gender paradox", i.e., the fact that women behave both more conservatively and more innovatively than men at the same time (but in different parts of their language). However, Labov, seem to mix the concepts of prestige, standard, and innovation, in a way that suits more "classic" situations, but that needs further analysis, for example in contact situations[81]. Eckert (1989; 1998; 2000) answered Labov's paradox by demonstrating that different girls use different variants for different social purposes. In other words, it "is a paradox only if we assume it is the same women in each case, but this is not necessarily so" (Meyerhoff, 2014, 94).

[80] Labov defines change from below as "the basic form of linguistic change that operates within the system, below the level of social awareness. These include the systematic sound changes that make up the major mechanism of linguistic change" (Labov, 1990, 215).

[81] After all, it has already been highlighted how social factors may behave in different ways in minority languages, compared to dominant ones, and in language contact situations (Stanford and Preston, 2009).

The motivation for the first two principles proposed by Labov can be partially found, according to the author himself, in the concept of symbolic capital and market (Bourdieu and Boltanski, 1975), which leads us to our second point of discussion. According to Trudgill (1972), the respect of prestige norms by women could be considered as a need to mark their prestige in society; men, on the contrary, are freer to escape them because they have access to more ways to mark status than linguistic symbols. Gordon (1997) further specifies that women feel "pressured" to maintain standard forms to avoid negative social evaluation, since their social position usually makes them more vulnerable to criticism (Deuchar, 1988). As a matter of fact, social psychologists suggest that men do not pay the social cost of using non-standard forms (Meyerhoff, 2014).

Obviously, this concerns forms and features that are above the level of consciousness, since they are evaluated and stigmatized by speakers (Eckert, 1998). However, we would suggest that maybe the concept can be applied to more general speech modalities too, and that women's "meticulousness" has unconscious consequences. The "care" women put into avoiding negative judgement may be connected to them being more attentive to the way they talk, which perhaps results in an unconsciously more hyperarticulate speech style even in informal contexts where hypoarticulation is expected. The lower value of language as social prestige symbol, and the more "laid back" approach to non-standard features of men, could result in lower levels of "care" and attention, causing men to be freer to hypoarticulate in informal contexts. However, no empirical evidence that I know of exists supporting or denying this hypothesis.

In our study, we find a particular situation. On the one hand, we have lenited variants of (r), that we suggest are being caused by target undershoot, one of the main features of hypoarticulated styles. Lenition as a change would be a change from below (cf. fn. 5), with lenited variants being the innovation – and hence, according to Labov's third principle, they should be favored by women. It is a phenomenon naturally emerging in speech due to physiological and articulatory reasons, hence its presence, to different extents, in all speakers. Nevertheless, if we assume as true our hypothesis of a higher degree of "carefulness" affecting women's levels of hypoarticulation, we understand why, in this case, it is the Males that are employing the innovative variants more. This is especially true for Homeland speakers, whose productions have no reason to be tied in with ethnic motivations.

On the other hand, we are also studying a group of Italians abroad – the Heritage speakers, with different levels of attachment to Italian ethnicity and language. Such a group, especially being in contact with a majority language and non-Italian people, could have become aware of some features of its HL, up to the point that they became hallmarks, or symbols, of their

ethnicity and language. A change or a pattern of stable variation regarding these features, then, would be above the speakers' consciousness. This would mean, according to Labov's principles 1 and 2, that women should prefer more prestigious variants. However, there is no clear "prestige form" to be preserved by women, since prestige would depend on ethnic evaluation. Speakers with different ethnic orientations might consider as more prestigious either the forms coming from the majority or the minority language: it all depends on the specific contact dynamic. We cannot establish *a priori* which forms are more prestigious. We can only hypothesize it based on what we know of the ethnic community.

As Eckert (1989) suggested, social factors – including ethnic orientation - are entrenched in variation and should be considered in interaction. In our case, we argue that Heritage speakers are not conscious of a pattern of lenition, but could be aware that some variants are specifically Italian forms, compared to Canadian English. They would evaluate them as the more prestigious or the less prestigious forms based on their relationship with the Italian and Canadian ethnicities; consequently, they could enhance them or not based on the way they want to present their ethnic identity.

In other words, one account of our data is this: because lenition is below the speakers' consciousness, the lenited, innovative variants Approximants and Fricatives should be favored by Females, while in our data they are favored by Males both among Homeland and Heritage speakers. However, because Italian-like variants may be above the speakers' consciousness in the Heritage group, the (supposedly) prestigious variants Taps&trills(&Fricatives) should be favored by Females, and they are.

In our data, every social group features a Sex-related pattern. However, we also saw that the Sex-involving factor has a higher range in Homeland speakers (52) than in Generation 2 speakers (36), meaning that the rhotic-related difference between Sexes is bigger in the Homeland. It is possible that the difference in social pressure between men and women, leading to unconscious control of levels of hypoarticulation, is greater in Homeland speakers than in later generation of Heritage speakers. These two groups (Homeland and Generation 2), indeed, experience different contexts, job markets, join in different communities of practices, and are exposed to gender or sex related stereotypes to a different extent.

As a matter of fact, Italy, being a Mediterranean country, has more prominent stereotypical gender roles and ideologies than other Western regions (Tager and Good, 2005; Hunt et al., 2016; Pistella et al., 2017). "[T]here is a conservative tendency, placing more conformity pressure on men and women to develop and enact what they consider gender-appropriate roles" (Mosso et al., 2013, 14). The Global Gender Gap Report (World Economic Forum, 2006; 2015;

2021) gives to a vast selection of countries an index evaluating the country's performance in closing the gender gap on a 0-to-1 scale (with 0.00 = inequality, 1 = equality), and ranks the considered countries. The following table contrasts indices and ranks from Italy and Canada in three years:

Year of the Report	**Italy**	**Canada**
2006	Rank: 77 (out of 115 countries) Index: 0.646	Rank: 14 (out of 115 countries) Index: 0.716
2015	Rank: 41 (out of 145 countries) Index: 0.725	Rank: 30 (out of 145 countries) Index: 0.740
2021	Rank: 63 (out of 156 countries) Index: 0.721	Rank: 24 (out of 156 countries) Index: 0.772

Table 27: Comparison between the Gender Gap rankings and indexes of Italy and Canada, in three different years.

From these data, we can say that it is indeed possible that the social pressure, and the gender-based differences in need to perform carefully when using language, are greater in Italy than in Canada. This would explain why Sex-based differences are stronger in the Homeland sample than in the Generation 2 one.

We have now described what type of pattern we think we are looking at (lenition, below the level of consciousness, with the innovation being reduced variants), why we think it presents to such different extents in Males and Females (due to a possible more "careful" style employed by women in response to their use of language as symbolic capital), and why this distinction between Males and Females is stronger in the Homeland sample (sex and gender-related pressure is higher in the Homeland).

However, the role played by ethnicity among Heritage speakers is still not completely clear.

We would like to hypothesize that, for (some) Heritage speakers, a level of consciousness – even if low - around the Italian-like variants has developed or is starting to develop, and that therefore, we see patterns that can be accounted for by Labov's "above awareness" principles 1 and 2 *within* a natural "below awareness" pattern of lenition.

While Reduced variants are still featured among Heritage speakers, and the mechanism producing them stays below their awareness, the fact that EO factors were significant in an English-like vs. Italian-like comparison (cf. §3.5.4.) is empirical evidence that there is a connection between being more oriented towards Italian ethnicity and producing more Italian-

like variants. If this association were conscious, then the Sex-based difference in Heritage speakers would assume a new meaning. It would be due not only to the "carefulness" issue we exposed earlier, but also to women (with high EO scores) preserving the variants they feel are more prestigious[82] more than men with high EO scores do, as Labov prescribed.

Figure 10 below shows, from a distributional point of view, the consistency of the EO-related effect in both Sexes. Mixed Effects models using either the combined factor Sex.EO LU or Sex.EO ID maintain the same hierarchy shown in Figure 10 (Males.Low favor English-like variants the most, followed by Males.High, Females.Low, Females.High).

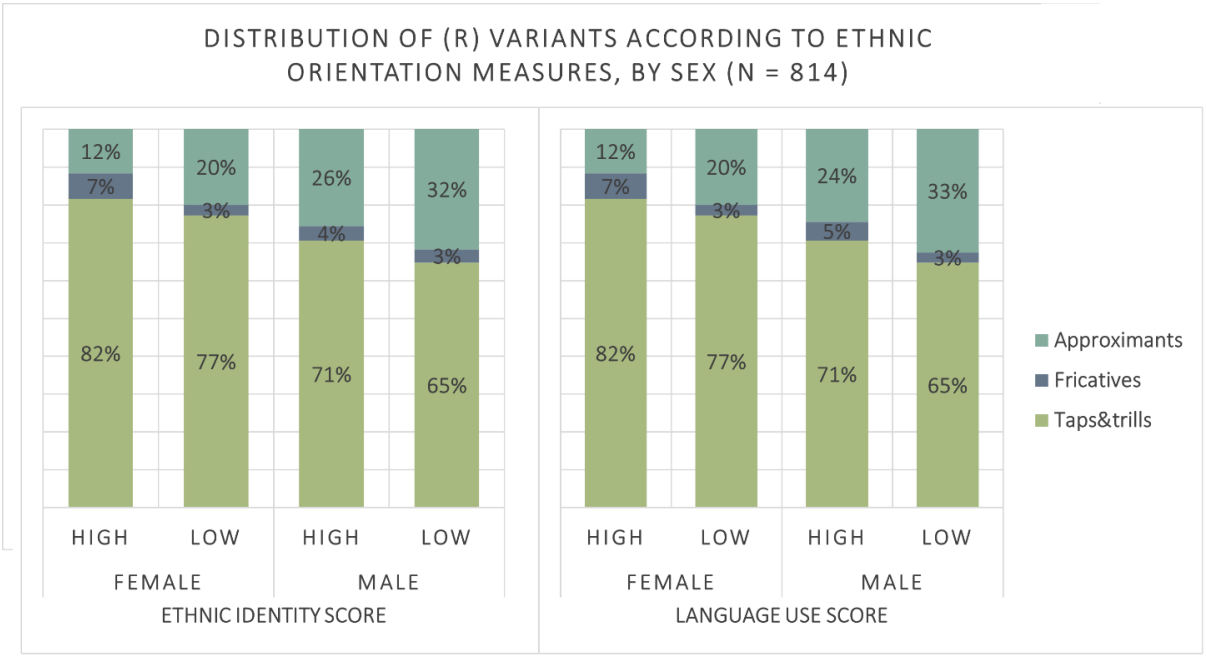

Fig. 10: Distribution of (r) variants across speakers of different EO scores, by sex (n = 814).

We are referring to speakers with higher EO scores in order to be more conservative in our hypothesis that women with high EO scores preserve the prestigious variants more than men with high EO scores do. While Italian-like variants might be perceived as more prestigious by every Heritage speaker, regardless of their EO, a safer assumption would be that it is more likely among those who share a strong ethnic bond with Italianness, and that have preserved it in a foreign land. Figure 10 supports our idea by presenting the impact of EO scores within Sexes.

[82] i.e., those which they have a stronger ethnic bond with.

The EO questions we used regarding Language Use also include questions about language *preference* (questions B4, B5, B6). These questions would be more suitable to investigate prestige-evaluation of linguistic variants. If we only consider these questions, and label speakers as either High or Low based on their scores, 12/14 speakers would classify the same way they did using the whole LU subset. The two "deviant" speakers, instead, would classify as High, when they were Low using the whole LU subset. Thus, we may safely interpret the EO LU score as representing a preference for Italian.

There are several ways to support the assumption that Italian culture and language are considered more prestigious than Canadian ones by Heritage Italians (with higher EO scores) living in Toronto. In the case of this community, we observed what a long history it has in Canada, how big and well-established it is, and how supported it is from social groups and institutions, without stigma (cf. §3.2.). There are many attractions related to Italian culture (just to cite two examples, food and fashion), and "Made in Italy" is now an international brand in a globalized job market (Di Salvo, 2017), where knowing Italian can be a highly valuable skill to obtain certain types of employment. Furthermore, anecdotally, many Italian communities abroad express pride in their ethnicity. Therefore, valuing Italian-like variants as more prestigious than English-like ones, because they allow speakers to present their ethnic identity immediately, does not seem implausible.

In conclusion, we suggest that awareness regarding the ethnic value of variants could be available to both Sexes, given that speakers with higher EO scores produce more Italian-like variants in both Sex groups, but that the higher rates of production of Italian-like variants in Females with high EO scores could be due not only to their more "careful" speech style, but also to them favoring more prestigious variants compared to Males. Future investigations distinguishing between the effects of care in speech style and of prestige evaluation could test this proposal.

3.6.3. Hypotheses evaluation

In order to evaluate our hypotheses, we will use the three lines of evidence formulated by Poplack and Tagliamonte reported in §3.4.3. and designed to show cross-group similarities or differences (2001, 92). We repeat them as follows:

1) Statistical significance: which factors are statistically significant, and which are not?

2) Relative strength: which factor is most significant (largest range) or least (smallest range)?
3) Constraint hierarchy: how are levels ranked within factors?
4) Does this order reflect the direction predicted by one or the other of the hypotheses?

For convenience, the hypotheses are repeated here, above the discussion of each.

- **Hypothesis 1**: We argue that approximant and fricative variants are articulatorily weaker variants of trills and taps that naturally emerge in contexts of relaxed speech, and therefore, we hypothesize that they will appear in both the Homeland and Heritage varieties.

  Hypothesis 2: It is hypothesized that approximant and fricative variants will be favored in coda position, a linguistic context often described to be particularly prone to have weaker/lenited variants (cf. Lawson et al., 2018; Kirchner, 2013; Rennicke, 2015).

  Not only the distribution of variants across speakers endorses the Hypothesis 1, but the phonological predictors found via statistical modeling suggest the validity of Hypothesis 2 too. The leitmotiv of the results is the fact that the Position of (r) in the syllable is a statistically significant predictor in each group. This factor always presents a hierarchy where Codas outrank Onsets. Hence, Codas systematically favor Reduced variants in each generation. For this reason, and as thoroughly discussed in §3.6.1., we can state that Approximant and Fricatives are lenited rhotic variants of Italian.

- **Hypothesis 3**: We hypothesize that the proportion between taps&trills and the more reduced variants will be similar between the two communities. Heritage languages, since they are native languages, preserve all the intrinsic features of spontaneous speech (e.g., hypoarticulation), just like Homeland varieties do.

  Hypothesis 4: Since we are claiming that the presence of these variants in Heritage speakers is related to natural language-internal variation, just like in the Homeland, we hypothesize that no Generational effects will be found. This, and the lack of a community effect (i.e., no Homeland vs. Heritage effect), will be evidence that Heritage speakers are not transferring phonetic detail regarding (r) from the majority language (English), nor simplifying their grammar.

  Hypothesis 5: We hypothesize Heritage speakers not only to maintain similar percentages of variants as the Homeland, but also the same distributional pattern as

Homeland speakers. The lack of an increase of the contexts favoring approximants in Heritage speakers will be evidence that not even the distributional patterns of Canadian English are being transferred (since Canadian English produces [ɹ] as a consonantal rhotic in every context).

This group of hypotheses is indeed true. Table 9 and Fig. 4 in §3.5.1. support the first claim from a distributional point of view. We have 70% Taps&trills and 30% Approximants and Fricatives in the Homeland sample and 73% Taps&trills and 27% Approximants and Fricatives in the Heritage sample. Statistical analysis shows that Generation is never a significant predictor, that the set of constraints affecting the production of Reduced variants stays stable across groups, and that the constraint hierarchies stay stable across groups. These results are particularly striking, because they can be interpreted as proof of the absence of contact-induced effects. Neither the prevalence of the approximant variant, nor its different distributional pattern in English, have influenced Heritage Italian rhotics in this regard.

This group of hypotheses suggested that the Homeland grammar would be maintained by Heritage speakers. In fact, we see no Homeland factors lost by Heritage speakers, and the "Coda effect" stays uniform across groups. However, if we compare the relative strength of the phonological factor, we can affirm that, more than maintenance, there seems to be a boosting effect. The range of the phonological factor is 17 for Homeland speakers, but 47 and 43 for Generation 1 and Generation 2 speakers: Heritage speakers are further developing a tendency of language internal evolution. In other words, Homeland speakers already reduced more in Codas than in Onsets, but Heritage speakers reduced Codas even more, and Onsets even less (while overall maintaining a steady rate of reduction).

- **Hypothesis 6**: While we expect the Heritage community as a whole to maintain the Homeland's grammar, we claim that individual-level differences can be explained by Ethnic orientation. We hypothesize that speakers (within a generation) with higher EO scores will produce more Italian-like variants – Taps&trills – compared to those with lower EO scores.

This also proved to be true, based on the Results from §3.5.4. The EO analysis reveals the significance of both the Ethnic Orientation measures we used – Language Use and Ethnic Identity, confirming the weight that ethnicity can have in the construction of an identity and in the consequent linguistic behaviors. It must be noted that this pattern is not about the *rates* of Italian-like variants vs. English-like variants in the Heritage community, since they stay similar to those from Homeland speakers (and we see no community effect in the model employing all speakers). This pattern is about a social constraint "splitting up" this stable ratio of variants between Low EO speakers and High EO speakers. Nevertheless, the EO factors both have quite low relative strengths, meaning that they don't impact (r) production as much as the phonological constraint. This means that most of the rhotic production depends on suprasegmental features, as expected in a lenition pattern; the weight of ethnicity, in comparison, is marginal. We could suggest that this is a consequence of ethnic and conscious evaluation of (r) variants being present for some speakers, at least, but only to a low degree, otherwise the EO factors would have had higher relative strengths.

Nevertheless, this case study also showed us unexpected results, such as the Sex-based difference. We provided a possible explanation based on differences in hypoarticulated styles between Males and Females, but future research might investigate whether there is a distinction between Males and Females specifically targeting rhotics, and the way they are perceived and evaluated by speakers based on Sex (or gender).

Another unexpected result, from the Italian-like vs. English-like comparison, is the significance of the Lexical Class predictor just in the Homeland group. This could be interpreted as grammar simplification among Heritage speakers, since they lost one constraint from the Homeland grammar; therefore, we will briefly discuss it.

Given all the other evidence against grammar simplification and for the boosting of a language-internal process among Heritage speakers, it is more coherent to interpret the loss of the lexical factor in the same light.

The hierarchy of levels of Lexical class lets us see how Verbs favor Approximant variants compared to other categories. Although we attempted to avoid frequency effects by selecting the most frequent lexical items in our corpus, it is possible that some traces can still emerge, and that this is one of them.

One possible explanation, indeed, lies in the fact that Verbs are the most frequent category in spoken Italian (Voghera, 2017). The *Vocabolario di base* (De Mauro 1991) is the set of words which comprise the base of the Italian lexicon. While the *Vocabolario* includes lots of nouns, many of them are used sporadically; however, it includes fewer verbs, but they are used more frequently. According to Voghera (2017, 145): "Questa tendenza è ancor più accentuata nel parlato a causa del fatto che generalmente quando parliamo tendiamo a preferire un nucleo relativamente ristretto di lemmi" ("This tendency is even more accentuated in speech due to the fact that, generally, when we speak we tend to prefer a relatively small set of entries").

If we hypothesize that the more frequent the word, the more it tends to undergo reduction (Bybee, 2003), it might be possible that the extreme frequency of Verbs in Italian skewed the data, making the Verb category more prone to having a Reduced (r) variant, i.e., Approximants. It is not completely clear which base unit should be considered to calculate a frequency effect (Lexical entry? Lexical form? Phoneme? Something else?), as it is not clear how items are stored in our memory (Labov, 2006); however, the lexically regularity (i.e., spreading across lexical classes) of sound changes (including lenition) has often been attributed to the frequency of some lexical categories:

> Even regular sound change goes through a period of variation when the change is more advanced in some contexts than in others. […] The hypothesis is that a word or phrase changes earlier if it occurs often in the favorable conditioning environment (Bybee, 2002; Raymond & Brown, 2012). As words
> of different morphological structure or lexical class may occur more or less often in the conditioning context, the sound change may occur earlier or later under what appear to be grammatical conditions. (Bybee, 2017, 274)

As we are claiming that Heritage speakers are boosting lenition, it is possible that they lost the Lexical class conditioning because they are in a more advanced stage of lenition than Homeland speakers. While in Heritage speakers the phenomenon has already spread to the whole lexicon, for Homeland speakers it is still lexically constrained.

As for the fact that Lexical class was significant only in the Italian-like vs. English-like comparison, and not in the Reduced vs. non-Reduced comparison, a possible answer could be as follows. The tokens favoring Approximants the most in the Lexical class factor (i.e., tokens from Verbs) feature intervocalic (r). Hence, if such a segment has to reduce, it is more plausible that it does so by becoming an Approximant due to bidirectional coarticulation, rather than a

Fricative[83]. Therefore, when we collapsed Fricatives with Taps&trills, and left the level Approximant by itself, the lexical factor was able to emerge, while it could not when Fricatives were merged with Approximants.

We are now able to answer our last research question, which asked whether it is possible to ascribe singleton (r) variation in Heritage Italian to contact-induced or language-internal dynamics. The response is: mainly the latter, but also possibly the former. While the production of non-Standard rhotic variants is the result of a lenition process, the ways this process unravels reveal:

a) A possible form of consciousness among heritage speakers regarding the more Italian-like status of some variants, that can only have developed thanks to the contact with English;
b) A possible preference for the lenited variant which is acoustically the same as the Canadian English rhotic.

In the next section, we discuss evidence from another HL that bolsters these interpretations.

3.6.3. Cross-linguistic comparison: Heritage Tagalog (r)

Umbal and Nagy's (2021) paper about (r) in Heritage Tagalog also falls within the scope of the HLVC project, employing the same methodology and corpus we used in the present research. Tagalog (r) is usually realized as either a tap or a trill, but an approximant variant is also reported in the Homeland, and it has been attributed to contact with English in the Philippines - especially among middle-class speakers. There is evidence that it has gained a positive social meaning. All variants are also used by Heritage Tagalog speakers. Hence, the purpose of the article is to tease apart contact-induced effects from language-internal ones in the production of taps&trills vs. approximant variants.

Umbal and Nagy report that Homeland speakers have a higher rate of use of approximants than Generation 1 speakers, but that Generation 2 speakers use approximants at a higher rate than Generation 1 speakers. Moreover, speakers who are more oriented to the use of English language (vs. Tagalog) also show increased use of approximants, while orientation towards Filipino ethnicity (vs. Canadian) does not play a role.

[83] There is evidence already that in Florentine Italian, voiceless stops reduce to approximants in the context V_{V, liquid, glide}, and to fricatives elsewhere (Kirchner, 2013).

While transfer from English does not seem the right explanation for the difference between Homeland and Generation 1, one could be tempted to suggest it for the difference between Generation 1 and 2. However, the authors highlight that the effect of the phonotactic constraints remain robust across groups and that Heritage speakers remain affected by them. This is evidence against transfer, because, if transfer was happening, then speakers should have also adopted the distributional constraints of English. Therefore, Umbal and Nagy suggest the possibility of a language-internal process of lenition from taps&trills to approximants, acting in Generation 2 speakers at higher rates - since Heritage speakers may boost language-internal developments. Further evidence of such a process is provided by the fact that the approximant variants tend to appear in medial coda and word final contexts, i.e., positions particularly suited for lenition. Finally, the authors do not find any sex-based difference, suggesting that either the process is still in its early phase, or that it is not related to prestige dynamics. However, since there is evidence of prestige evaluation of approximant variants among Tagalog speakers, the former hypothesis is the most plausible.

If we compare the Italian (r) pattern to the Tagalog one, we can see some similarities and differences.

First, in both languages the innovative variants are already present in the Homeland grammar. However, while they are consciously evaluated in Homeland Tagalog, they apparently stay below the level of awareness of speakers in Homeland Italian.

Second, the rates of the variants in Tagalog vary across groups; instead, in our study the distribution of variants is stable.

Third, ethnic identity and sex do not play a role in Tagalog (r), while they do in Italian (r). This suggests that, differently from Tagalog, there is evidence for the indexicality of (r) in Italian.

Fourth, both studies found a pattern of reduced variants in coda position and used this as evidence to support the idea of (r) lenition. Given the wide range of possible realizations of rhotics, language-internal variation inside the rhotic system is almost to be expected, and there is no need to attribute it to language contact.

Fifth, both studies found hints of boosting of language-internal developments by HL speakers, providing further support to Flores and Rinke's (2020, 25) claim that "heritage speakers may boost and further develop tendencies of language (internal) evolution inherent to variable phenomena".

Homeland Italian and Homeland Tagalog represent languages in different social contexts, especially in that contact with English, in Italy, has been almost null in the past and, even if it

has increased with globalization, it remains low in comparison to the situation of Homeland Tagalog. However, approximant rhotic variants, starting from the prototypical rhotic being a tap or a trill, emerged in both situations, supporting the idea that contact is not necessarily needed to cause variation inside the rhotic system. Furthermore, the fact that they can be attributed to lenition in both languages, thanks to their phonotactic pattern, reinforces the already rich literature describing lenition as a universal mechanism of languages, that has its roots in likewise universal articulatory tendencies to minimize effort.

The two studies provide further evidence of how HLs are dominated by the same mechanisms as any other non-Heritage language, and that they can be mastered by HL speakers reaching high proficiency. The maintenance of phonotactic constraints across groups, indeed, shows that HL speakers fully acquire (this aspect of) the HL and replicate its structural features. This is also the case with a preliminary analysis of -tr- and -dr- clusters affrication (cf. §3.6.4.), where HL speakers use a locally marked feature at rates similar to the Homeland's, without transferring elements from English (such as the higher frequency of affricated forms). Hence, HL speakers can be regarded as native speakers. Contact with a majority language and culture can play a role, but this role is not always – or not simply – that of providing transferred forms. In these two studies, indeed, it operates *within* the language-internal process. In the case of Tagalog, it might have been the trigger to boost the lenition process in Generation 2; in the case of Italian, it might have caused an ethnic evaluation of variants that intertwines with lenition, and it might orient speakers in favoring one lenited variant more than another.

3.6.4. An exploratory analysis: -tr- and -dr- clusters

Hypotheses regarding no transfer of phonetic detail from English have also been tested in an exploratory analysis of -tr- and -dr- clusters.

-tr- and -dr- clusters have been separated from the word-internal singleton /r/ analysis because they feature a particular behavior. As we illustrated in §2.3.2.2., their retroflection and possible affrication (e.g., [tr] > [t͡ʂ]) is a characteristic of Calabrese dialect and Calabrese Regional Italian. These clusters are also variably pronounced as affricates in Canadian English (cf. §2.3.2.2.). Therefore, the situation regarding these sounds is similar to that of word-internal singleton /r/. Word-internal singleton /r/ has an attested approximant variant in (Regional) Italian, while the approximant is the main realization for speakers of Canadian English. Thus, for Heritage speakers, English could be a source of influence, causing an increase in rate of use of the approximant realizations, which are marginal in the Homeland variety. In the same

fashion, -tr- and -dr- clusters have an attested affricated variant in (Regional) Italian, and the affricate realization is also common in English, which could be a source of influence as well.

Thus, an exploratory analysis on these clusters has been conducted. It investigated whether -tr- and -dr- were realized as affricated or non-affricated in Italian, both in Homeland and Heritage speakers, and in the English spoken by the Italo-Canadian community of Toronto.

Variety	**Speakers**[84]	
Homeland Italian	Males	IXM19A; IXM19B; IXM28A; IXM35A; IXM47A; IXM52A; IXM61A; IXM64A
	Females	IXF22A; IXF22B; IXF35A; IXF51B;
Heritage Italian	Males	I1M61A; I1M61B; I1M62A; I1M75A; I2M42A; I2M52A; I2M53A
	Females	I1F59A; I1F61A; I1F71A; I1F73A; I2F44A; I2F45A; I2F57A
Italo-Canadian English	Males	E3M19A; E3M21A; E2M51A; E1M63A
	Females	E3F21A; E3F21B; E3F21C; E2F45A; E2F58A

Table 27: Distribution of speakers used for the exploratory analysis of -tr- and -dr- clusters. Data come from the HerLD corpus and the Contact in the City corpus (Hoffman and Walker, 2010).

Preliminary results showed that:

1. -tr- and -dr- are pronounced as affricates 98% of the time (n = 200) in English, by members of the Italo-Canadian community: the source of influence could thus be defined as "strong".
2. Affricate realizations in the Homeland Italian data appear 14% of the time (n = 230).
3. Crucially, in the Heritage Italian data they appear 17% of the time (n = 416).

Homeland and Heritage Italian speakers maintained similar rates of production of affricate realizations. A preliminary Mixed Effects model did not find a statistical difference between Homeland and Heritage speakers. This means that Heritage speakers keep producing this locally marked feature at rates similar to those of Homeland speakers, and it suggests that they

[84] Homeland and Heritage Italian speakers have been referred to by using the speakercode convention of the HerLD corpus (cf. fn. 5). The Italo-Canadian English speakers codes follow the same convention. They are distinguished here by an initial "E" in place of "I"; their Generation has been deduced from their Ages of Arrival.

are not being influenced by English in that matter. Future research employing a larger number of tokens, a more nuanced speaker sample, and evaluating the impact of phonological variables, may confirm these impressions.

In conclusion, empirical results from the word-internal singleton /r/ analysis, and preliminary results from the exploratory -tr- and -dr- clusters analysis, demonstrate a cohesive behavior of Heritage speakers. Heritage speakers reproduce the Homeland grammar, without simplification or transfer from English.

3.6.5. Final remarks

Chapter 1 illustrated how often Heritage languages are assumed to tend towards a simplification of their grammar due to transfer from the majority language (Flege, 2015), attrition, or incomplete acquisition (Polinsky, 2018). In Chapter 2, we discussed patterns of articulatory reduction, attested in synchrony, that in diachrony eventually modify the phonetic inventory, e.g., lenition, and that are frequently found across languages (Magni 2014). Hence, Heritage Languages are expected to follow these trends: they could simplify their grammar for the reasons addressed by the Experimental approach (cf. §1.6.1.), they could maintain the same grammar found in a Homeland sample, boost it, or they could, rather than simplify their grammars, introduce new language-internal changes, absent in a Homeland sample. Comparative variationist works in the past have shown many cases of maintenance of the Homeland grammar (Nagy 2015).

By taking a comparative variationist approach, using a tested methodology (cf. §3.3.) and ecologically valid data, we found empirical evidence of Heritage speakers lacking the incomplete grammars they are so often described as having, regarding the production of rhotics in at least three different linguistic and social contexts.

Conclusion

This thesis explored the realization of rhotics in Heritage Calabrese Italian, spoken by the Heritage Italian community of Toronto (ON), Canada.

According to traditional phonological descriptions (Bertinetto and Loporcaro, 2005; Canepari, 1999), Italian rhotics can either be trills ([r]) or taps ([ɾ]). Empirical research, however, shows that they can also be realized as approximants ([ɹ]) in spontaneous speech (Vietti, Spreafico and Romano, 2010; Romano, 2013). Moreover, a fricative rhotic variant ([ɾ̝̊], (Celata, 2014; Celata, Meluzzi, and Ricci, 2016) appears in Low-Southern Regional Italians and Italo-romance dialects but has not received much attention. On the other hand, rhotic varieties of North American English, such as Canadian English, use an approximant variant ([ɹ]) as their consonantal rhotic (Westbury et al., 1998).

Therefore, this research investigated whether patterns of variation or change, regarding rhotics, could be identified in Homeland and Heritage Italian. Frequent assumptions about Heritage speakers suggest that the majority language they are in contact with is a source of transfer (Flege, 2015). Hence, in our case, it was possible that the frequency and distribution of [ɹ] in English was influencing Heritage Calabrese Italian, causing an increase of the Approximant variant, or a less-constrained distribution across contexts, compared to Homeland speakers.

Alternatively, the Approximant and Fricative variants diffused among Homeland speakers could be lenited variants of Taps&trills (Barry, 1997; Sebregts, 2014). That is, it was possible that Heritage speakers were not affected by contact-induced effects, but rather maintain or boost a language-internal phenomenon of lenition – a weakening process that can act either in synchrony or diachrony.

Our investigation sought empirical evidence supporting either possibility.

Using the HerLD corpus (Nagy, 2011), we coded and analyzed 1555 occurrences of word-internal, singleton /r/, elicited from 29 speakers coming from both a Homeland sample and a Heritage sample (two generations) of Calabrese Italian speakers. We first found the distribution of rhotic variants within each group, and then the grammar(s) underlying their realization, via Mixed Effects models.

Our hypotheses argued that Approximant and Fricative variants were articulatory weaker variants of Taps&trills, that naturally emerge in contexts of relaxed speech. Being lenited

variants, we expected them to be favored by coda position, a context prone to lenition. Since we linked the presence of Approximant and Fricative variants to natural language-internal variation, we also hypothesized that no difference between the Homeland and Heritage community, nor across generations, would be retrieved. This would be interpreted as the primary evidence of no transfer of phonetic detail regarding (r) from English. We further advocated for the absence of simplification of the Heritage grammar by hypothesizing that Heritage speakers did not extend the Approximant variant to new contexts, maintaining the same constraints as the Homeland speakers.

Results confirm that the three variants considered are all present both among Homeland and Heritage speakers, with similar frequencies, and with the same grammar (i.e., set of constraints). The position of (r) in the syllable is a statistically significant predictor of Approximant and Fricative variants in every group we considered, and it reveals that coda contexts systematically favor said variants in each generation. No generational or community differences were ever found. No Homeland constraints were lost by Heritage speakers. These results are proof of the absence of transfer from English, contact-induced effects, or grammar simplification among Heritage speakers. However, based on a comparison of the relative strengths of factors across generations, we claim that Heritage speakers are boosting the lenition process already going on in the Homeland: Heritage speakers are further developing a tendency of language internal evolution, as Flores and Rinke (2020) suggested.

Moreover, we anticipated ethnic factors to play a role in the linguistic behavior of Heritage speakers. Hence, we also tested the impact of two Ethnic Orientation (EO) measures on (r) – Language Use and Ethnic Identity. We hypothesized that speakers (within a generation) with higher EO scores produced more "Italian-like" variants (Taps&trills, Fricatives), compared to those with lower scores. This hypothesis also proved to be true, since EO factors are significant predictors of Italian-like variants. This confirms the weight that ethnicity can have in the construction of an identity and in the consequent linguistic behaviors.

Nevertheless, the investigation also provided unexpected results, such as a sex-related pattern, maintained across all generations, constraining (r). Males systematically favor Approximant and Fricative – thus, reduced – variants, compared to Females.

We hypothesized that the different uses of language as symbolic capital (Trudgill, 1972; Gordon, 1997) between men and women could translate as unconscious control of the levels of hypoarticulation. In our case, Males would employ a less "careful" hypoarticulated (Lindblom, 1990) style than Females, explaining why they use innovative variants (Approximants and

Fricatives) more – while, according to Labov's (1990) third principle, they should be favored by women in a change from below.

However, sex-based differences regarding (r) were found to be stronger among Homeland speakers than Heritage speakers. Since differences in sex and gender-related pressure and expectations are likely to affect linguistic use, we hypothesized that the Canadian social context, being more advanced in terms of reduction of the gender gap and gender-related stereotypes, compared to Italy (World Economic Forum, 2006; 2015; 2021; Pistella et al., 2017), would push Heritage speakers to shape their linguistic behavior accordingly.

We also entrenched the sex-related pattern with the ethnic one. While the mechanism producing reduced (r) variants stays below the speakers' awareness, the fact that EO was significant for Heritage speakers is empirical evidence of a connection between being more oriented towards Italian ethnicity and producing more Italian-like variants. If this association were conscious, then the sex-based difference in Heritage speakers would be due not only to the different "careful" levels, but also to women with high EO scores preserving the variants they feel are more prestigious more than men with high EO scores do, as in a change from above – Labov's principles 1 and 2. Future investigations aimed at distinguishing between the effects of care in speech style and of prestige evaluation could confirm or disconfirm this proposal.

Furthermore, we compared our results to those coming from an analogous investigation of rhotic production among Tagalog Heritage speakers (Nagy and Umbal, 2021). We saw that the two Heritage languages behave similarly, as they are boosting a lenition process of the Homeland employing the same phonotactic constraint (i.e., reduction of rhotics in coda position).

Similar evidence of grammar maintenance also appears in a preliminary analysis we conducted of rhotics in -tr- and -dr- onset clusters, where no influence from an English source, and thus from a majority language, seemed to appear among Heritage speakers.

In conclusion, this thesis served multiple purposes. Given that some information about rhotics is attested for the nearby Sicilian Regional variety, but not for the Calabrese one, we enriched the literature about rhotics in Italy, by analyzing their production in the Calabrese Regional variety. Then, we provided empirical evidence, based on spontaneous speech data and quantitative analysis, of a situation where Heritage speakers do not have "incomplete", "divergent", "simplified" grammars. Contact with a majority language does not necessarily cause attrition or transfer (Benmamoun, Montrul, and Polinsky, 2013; Polinsky, 2018). Heritage speakers are able to acquire phonetic details coming from their Heritage language, to

reproduce them during adulthood, and to pass them to the next generation of Heritage speakers. Furthermore, they are capable of acquiring not only linguistic constraints, but also social ones such as sex-based patterns of production, suggesting that social factors play, indeed, a role in the construction of grammars (Hymes, 1972; Firth and Wagner, 1997; Larsen-Freeman, 2007).

There is no empirical reason to deny Heritage speakers the title of "native speakers". We should disconnect the concept of nativeness from those of monolingualism, innate competence, dominance, and proficiency (Myhill, 2003; Love and Ansaldo, 2010; Rothman and Treffers-Daller, 2014).

Patterns of variation and change that can be found in Heritage languages should not be automatically interpreted as evidence of them not being native speakers, because variation and change are intrinsic parts of languages. Only by analyzing them as they occur in real interactions, we will be able to understand how speakers – *every* group of speakers - spontaneously and actively build and shape grammars.

Bibliography

Aalberse, S. Backus, A. Muysken, P. 2019. *Heritage Languages. A language contact approach*. Amsterdam/Philadelphia: John Benjamins.

Abrahamsson, N. Hyltenstam, K. 2009. Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59, 249-306.

Abdi, K. 2011. 'She really only speaks English': Positioning, language ideology, and heritage language learners. *The Canadian Modern Language Review*, 67 (2), 161-190.

Albano Leoni, F. 2009. *Dei suoni e dei sensi. Il volto fonico delle parole*. Bologna: il Mulino.

Albano Leoni, F. 2015. < Carmniell o'srngar >. Osservazioni sulla ortografia ingenua del napoletano e sulle sue possibili implicazioni fonetiche. In Iannàccaro, G. Guerini, F. Dal Negro, S. (Eds.), *Elaborazione ortografica delle varietà non standard. Esperienze spontanee in Italia e all'estero*, 51-78. Bergamo: Bergamo University Press/Edizioni Sestante.

Albano Leoni, F. Maturi, P. 2008. *Manuale di fonetica*. Rome: Carocci.

Aronoff, M. 2016. Competition and the lexicon. In Elia, A. Iacobino, C. Voghera, M. (Eds.), *Livelli di analisi e fenomeni di interfaccia. Atti del XLVII congresso internazionale della società di linguistica Italiana*. Rome: Bulzoni Editore.

Aslin, R. N. 2017. Statistical learning: a powerful mechanism that operates by mere exposure. *Wiley Interdisciplinary Reviews: Cognitive Science*, 8 (1-2), e1373.

Auer, P. 2005. Europe's sociolinguistic unity, or: A typology of European dialect/standard constellations. In Delbecque, N. Van der Auwera, J. Geeraerts, D. (Eds.), *Perspectives on Variation: Sociolinguistic, Historical, Comparative*, 7-42. Berlin: Mouton de Gruyter.

Austin, P. C. Leckie, G. 2018. The Effect of Number of Clusters and Cluster Size on Statistical Power and Type I Error Rates When Testing Random Effects Variance Components in Multilevel Linear and Logistic Regression Models. *Journal of Statistical Computation and Simulation*, 88 (16), 3151-3163.

Baayen, R. H. Davidson, D. J. Bates, D. M. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.

Baayen, R. H. Milin, P. 2010. Analyzing reaction times. *International Journal of Psychological Research*, 3 (2), 12-18.

Bailey, G. 2003. Real and Apparent Time. In Chambers, J. K. Trudgill, P. Schilling-Estes, N. (Eds.), *The Handbook of Language Variation and Change*. Oxford, UK: Blackwell.

Bailey, R. 2013. Variationist Sociolinguistics. In Bayley, R. Cameron, R. Lucas, C. (Eds.), *The Oxford Handbook of Sociolinguistics*, 1-26. New York (NY): Oxford University Press.

Baird, A. Cristiano, A. Nagy, N. 2021. Apocope in Heritage Italian. *Languages*, 6 (3), 120.

Ballard, E. Starks, D. 2004. Liquids: Laterals and Rhotics or Much More? *Proceedings of the 2004 Conference of the Australian Linguistic Society*, 1-9.

Baker, P. 2010. *Sociolinguistics and Corpus Linguistics*. Edinburgh, UK: Edinburgh University Press.

Barr, D. J. Levy, R. Scheepers, C. Tily, H. J. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 6 (3), 255-278.

Barry, W. J. 1997. Another R-tickle. *Journal of the International Phonetic Association*, 27, 35-45.

Bartlett, L. 2007. Bilingual literacies, social identification, and educational trajectories. *Linguistics and Education*, 18, 3-4, 215-231.

Bauer, L. 2008. Lenition revisited. *Journal of Linguistics*, 44, 605-624.

Bayley, R. 2013. Variationist Sociolinguistics. In Bayley, R. Cameron, R. Lucas, C. (Eds.), *The Oxford handbook of sociolinguistics*. Oxford, UK: Oxford University Press.

Bell, B. A. Ferron, J. M. Kromrey, J.D. 2008. Cluster Size in Multilevel Models: The Impact of Sparse Data Structures on Point and Interval Estimates in Two-Level Models. *JSM Proceedings: Section on Survey Research Methods*, 1122-1129. Alexandria (VA): American Statistical Association.

Benmamoun, E. Montrul, S. Polinsky, M. 2013. Heritage Languages and Their Speakers: Opportunities and Challenges for Linguistics. *Theoretical Linguistics*, 39 (3-4), 129–181.

Berent, I. Steriade, D. Lennertz, T. Vaknin, V. 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition*, 104 (3), 591-630.

Berruto, G. 2006 [1987]. *Sociolinguistica dell'italiano contemporaneo*. Rome: Carocci.

Berruto, G. 1995. *Fondamenti di sociolinguistica*. Rome: Laterza.

Bertinetto, P. M. Loporcaro, M. 2005. The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. *Journal of the International Phonetic Association*, 35 (2), 131-51.

Bhat, D.N.S. 1973. Retroflexion: An Areal Feature. *Working Papers on Language Universals*, 13, 27-67.

Bhat, D.N.S. 1974. Retroflexion and Retraction. *Journal of Phonetics*, 2, 233-237.

Bickerton, D. 1975. *Dynamics of a creole continuum*. Cambridge (MA): Cambridge University Press.

Birdsong, D. (Ed.) 1999. *Second language acquisition and the critical period hypothesis*. Mahwah (NJ): Lawrence Erlbaum Associates.

Birdsong, D. 2005. Nativelikeness and non-nativelikeness in L2A research. *International Review of Applied Linguistics*, 43, 319-28.

Bley-Vroman, R. 1989. The logical problem of second language learning. In Gass, S. Schachter, J. (Eds.), *Linguistic perspectives on second language acquisition*, 41-68. Cambridge, UK: Cambridge University Press.

Bley-Vroman, R. 2009. The evolving context of the fundamental difference hypothesis. *Studies in Second Language Acquisition*, 31, 175-198.

Bloomfield, L. 1935 [1933]. *Language*. London: Allen and Unwin.

Boberg, C. 2008. Regional Phonetic Differentiation in Standard Canadian English. *Journal of English Linguistics*, 36 (2), 129-154.

Boersma, P. Weenink, D. 2021. *Praat: doing phonetics by computer* [Computer program]. http://www.praat.org/

Bourdieu, P. Boltanski, L. 1975. Le fétichisme de la langue. *Actes de la recherche en sciences sociales*, 4, 2-32.

Boyce, S. Espy-Wilson, C. Y. 1997. Coarticulatory stability in American English /r/. *The Journal of the Acoustical Society of America*, 101 (6), 3741-3753.

Boyce, S. Tiede, M. Espy Wilson, C. Groves-Wright, K. 2015. Diversity of tongue shapes for the American English rhotic liquid. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*, n. 1041. Glasgow, UK: The University of Glasgow.

Brezina, V. Meyerhoff, M. 2014. Significant or random? A critical review of sociolinguistic generalisations based on large corpora. *International Journal of Corpus Linguistics*, 19 (1), 1-28.

Butler, Y. G. 2013. Bilingualism/Multilingualism and Second-Language Acquisition. In Bhatia, T. K. and Ritchie, W. C. (Eds.), *The Handbook of Bilingualism and Multilingualism*, Second Edition, 109-135. Oxford, UK: Wiley Blackwell.

Bybee, J. 2003. *Phonology and Language Use*. Cambridge, UK: Cambridge University Press.

Bybee, J. 2006. From usage to grammar: The mind's response to repetition. *Language*, 82, 711-733.

Bybee, J. 2017. Grammatical and lexical factors in sound change: A usage-based approach. *Language Variation and Change*, 29, 273-300.

Campbell, F. Gick, B. Wilson, I. Vatikiotis-Bateson, E. 2010. Spatial and temporal properties of gestures in North American English /r/. *Language and Speech*, 53, 49-69.

Canagarajah, A.S. 1999. Interrogating the "Native Speaker Fallacy": Non-Linguistic Roots, Non-Pedagogical Results. In George, B. (Ed.), *Non-native educators in English language teaching*. New York (NY): Routledge.

Canepari, L. 1999. *MaPI - Manuale di Pronuncia Italiana*. Bologna: Zanichelli.

Carreira, M. Potowski, K. 2011. Commentary: Pedagogical implications of experimental SNS research. *Heritage Language Journal*, 8 (1), 134-51.

Cedergren, H. Sankoff, D. 1974. Variable rules: Performance as a statistical reflection of competence. *Language*, 50 (2), 333-355.

Celata, C. 2004. Analisi del processo di retroflessione dei nessi con vibrante nei dialetti romanzi. *Quaderni del Laboratorio di Linguistica*, 5, 2004-2005.

Celata, C. 2010. Rhotic retroflexion in Romance. Acoustic data for an articulation-driven sound change. In Recasens, D. Sanchez Miret, F. Wireback, K. (Eds.), *Experimental phonetics and sound change*. München: Lincom Europa.

Celata, C. 2014. Per una sociofonetica sperimentale delle lingue romanze: la variabile (r) a Modica. *Atti del XIII Convegno Internazionale della SILFI (Società Internazionale di Linguistica e Filologia Italiana)*, Modica.

Celata, C. Meluzzi, C. Ricci, I. 2016. The sociophonetics of rhotic variation in Sicilian dialects and Sicilian Italian: corpus, methodology and first results. *Loquens*, 3 (1).

Celata, C. Vietti, A. Spreafico, L. 2019. An articulatory account of rhotic variation in Tuscan Italian. Synchronized UTI and EPG data. In Gibson, M. Gil, J. (Eds.), *Romance Phonetics and Phonology*, 91-117. Oxford, UK: Oxford University Press.

Cerruti, M. Crocco, C. Marzo, S. 2017. On the development of a new standard norm in Italian. In Cerruti, M. Crocco, C. Marzo, S. (Eds.), *Towards a New Standard: Theoretical and Empirical Studies on the Restandardization of Italian*, 3-28. Boston: De Gruyter Mouton.

Chambers, J. K. 2003. Patterns of Variation including Change. In Chambers, J. K. Trudgill, P. Schilling-Estes, N. (Eds.), *The Handbook of Language Variation and Change*. Oxford, UK: Blackwell.

Chambers, J. K. Trudgill, P. 1998. *Dialectology*. Cambridge: Cambridge University Press.

Chabot, A. 2019. What's wrong with being a rhotic? *Glossa: a journal of general linguistics*, 4 (1), 38, 1-24.

Chomsky, N. [1965] 1969. *Aspects of the Theory of Syntax*. Cambridge (MA): MIT Press.

Chomsky, N. 1986. *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.

Clark, H. H. 1973. The language-as-fixed-effect fallacy: a critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 12, 335-359.

Clarke, P. 2008. When Can Group Level Clustering Be Ignored? Multilevel Models Versus Single-Level Models with Sparse Data. *Journal of Epidemiology & Community Health*, 62 (8), 752-758.

Clarke, P. Wheaton, B. 2007. Addressing Data Sparseness in Contextual Population Research: Using Cluster Analysis to Create Synthetic Neighborhoods. *Sociological Methods & Research*, 35 (3), 311-351.

Clements, G. N. 1990. The role of the sonority cycle in core syllabification. In Kingston, J. Beckman, M. E. (Eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, 283-333. New York (NY): Cambridge University Press.

Coates, J. Cameron, D. (Eds.), 1989. *Women in Their Speech Communities*. London/New York (NY): Longman.

Colombo, A. Graffi, G. 2017. *Capire la grammatica. Il contributo della linguistica*. Rome: Carocci.

Cook, V. 1999. Going beyond the native speaker in language teaching. *TESOL Quarterly*, 33 (2), 185-209.

Cook, V. 2012. Multi-competence. In Chapelle, C. (Ed.), *The encyclopedia of applied linguistics*, 3768-3774. Chichester, UK: Wiley-Blackwell.

Cook, V. 2016. Where Is the Native Speaker Now? *TESOL Quarterly*, 50 (1), 186-189.

Coseriu, E. 1980. 'Historische Sprache' und 'Dialekt'. In Göschel, J. Ivić, P. Kehr, K. (Eds.), *Dialekt und Dialektologie: Ergebnisse des internationalen Symposions 'Zur Theorie des Dialekts'*, 106-122. Wiesbaden: Steiner.

Crocco, C. 2017. Everyone Has an Accent. Standard Italian and Regional Pronunciation. In Cerruti, M. Crocco, C. Marzo, S. (Eds.), *Towards a New Standard: Theoretical and Empirical Studies on the Restandardization of Italian*, 89-117. Boston: De Gruyter Mouton.

De Blasi, N. 2014. *Geografia e storia dell'italiano regionale*. Bologna: il Mulino.

DeKeyser, R. 2000. The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22, 499-534.

DeKeyser, R. 2003. Implicit and explicit learning. In Doughty, C. Long, M. (Eds.), *The handbook of second language acquisition*. 313-348. Malden (MA): Blackwell.

Delattre, P. Freeman, D.C. 1968. A dialect study of American r's by x-ray motion picture. *Linguistics*, 4A, 29-68.

De Leeuw, E. Schmid, M. Mennen, I. 2010. Perception of foreign accent in native speech. *Bilingualism: Language and Cognition*, 13 (1), 33-40.

De Mauro, T. 1991 [1980]. *Guida all'uso delle parole*. Rome: Editori Riuniti.

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

D'Errico, M. Grandi, N. Paternesi Meloni, S. Tamburini, F. 2016. Induzione di categorie grammaticali e lessicali. In Dedè, F. (Ed.), *Categorie grammaticali e classi di parole. Statuto e riflessi metalinguistici*, 115-137. Rome: Il Calamo.

Deuchar, M. 1988. A pragmatic account of women's use of standard speech. In Coates, J. Cameron, D. (Eds.), *Women in their speech communities*, 27-32. London: Longman.

Díaz-Campos, M. Dickinson, S. 2019. Using Statistics as a tool in the analysis of sociolinguistic variation: a comparison of current and traditional methods. In Rei-Doval, G. Tejedo-Herrero, F. (Eds.), *Lusophone, Galician, and Spanish Linguistics. Bridging Frames and Traditions*, 205-226. London: Routledge.

Di Salvo, M. 2017. Heritage language and identity in old and new Italian migrants in Toronto. In Di Salvo, M. Moreno, P. (Eds.), *Italian Communities Abroad: Multilingualism and Migration*, 74-95. Newcastle upon Tyne, UK: Cambridge Scholars.

Docherty, G. Foulkes, P. 2001. Variability in (r) production: Instrumental perspectives. In Van der Velde, H. Van Hout, R. (Eds.) *'r-atics. Sociolinguistic phonetic and phonological characteristics of /r/*, 27-43. Bruxelles: ILVP.

Doerr, N. M. 2009. Introduction. In Doerr, N. M. (Ed.), *The Native Speaker Concept: Ethnographic Investigations of Native Speaker Effects*, 1-10. Berlin/New York (NY): Mouton de Gruyter.

Dollinger, S. Clarke, S. 2012. On the autonomy and homogeneity of Canadian English. *World Englishes*, 31 (4), 449-466.

Drager, K. Hay, J. 2012. Exploiting random intercepts: Two case studies in sociophonetics. *Language Variation and Change*, 24 (1), 59-78.

Dressler, R. 2010. "There is no space for being German": Portraits of Willing and Reluctant Heritage Language Learners of German. *Heritage Language Journal*, 7 (2), 162-182.

Eager, C. D. Roy, J. 2017. *Mixed models are sometimes terrible*. Austin (TX): Linguistic Society of America.

Eckert, P. 1989. The Whole Woman. *Language Variation and Change*, 1, 245-267.

Eckert, P. 1998. Gender and sociolinguistic variation. *Language and gender: A reader*, 64-75.

Eckert, P. 2000. *Linguistic Variation as Social Practice*. Cambridge, UK: Cambridge University Press.

Eckert, P. McConnell-Ginet, S. 2013. *Language and gender*. Cambridge, UK: Cambridge University Press.

Eckert, P. 2005. *Variation, convention, and social meaning*. Paper presented at the Linguistic Society of America, Oakland (CA), January 5.

Ellis, N. 2009. Optimizing the input: Frequency and sampling in usage-based and form focused learning. In Doughty, C. Long, M. (Eds.), *The handbook of language teaching*, 139-159. Malden (MA): Blackwell.

Escudero, P. Smith, M. S. 2001. Reinventing the native speaker: or 'What you never wanted to know about the native speaker so never dared to ask.'. *EuroSLA Yearbook*, 1 (1), 275-286.

Espy Wilson, C. Boyce, S. 1994. Acoustic differences between 'bunched' and 'retroflex' variants of American English /r/. *The Journal of the Acoustical Society of America*, 95, 2823-2823.

Espy Wilson, C. Boyce, S. Jackson, M. Narayanan, S. Alwan, A. 2000. Acoustic modeling of American English /r/. *The Journal of the Acoustical Society of America*, 108, 343-356.

Fernández-Ordóñez, I. 2012. Dialect areas and linguistic change. In De Vogelaer, G. Seiler, G. (Eds.), *The dialect laboratory: Dialects as a testing ground for theories of language change*, 73-106. Amsterdam: John Benjamins.

Ferguson, C. 1959. Diglossia. *Word*, 16, 325-340.

Firth, A. Wagner, J. 1997. On discourse, communication, and (some) fundamental concepts in SLA research. *Modern Language Journal*, 81, 285-300.

Fishman, J. 2001. 300-plus years of heritage language education in the United States. In Kreeft Peyton, J. Ranard, D. McGinnis, S. (Eds.), *Heritage Languages in America. Preserving a national resource*, 81-97. Washington (DC): Center for Applied Linguistics & Delta Systems.

Fishman, J. 2006. Acquisition, maintenance and recovery of heritage languages. In Vald´es, G. Fishman, J. Ch´avez, R. P´erez, W. (Eds.), *Developing minority language resources: The case of Spanish in California*, 12-22. Clevedon, UK: Multilingual Matters.

Flege, J. E. 1995. Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, 92, 233-277.

Flege, J. E. 2007. Language contact in bilingualism: Phonetic system interactions. *Laboratory phonology*, 9, 353-381.

Flores, C. Rinke, E. 2020. The relevance of language-internal variation in predicting heritage language grammars. *Bilingualism: Language and Cognition*, 23, 25-26.

Flores, C. Rinke, E. Azevedo, C. 2017. Object realization across generations. A closer look on the spontaneous speech of Portuguese first and second generation migrants. In Domenico, E. (Ed.), *Complexity in acquisition*, 178-205. Cambridge, UK: Cambridge Scholars.

Flores-Ferrán, N. 2004. Spanish subject personal pronoun use in New York City Puerto Ricans: can we rest the case of English contact? *Language Variation and Change*, 16, 49-73.

Gardiner, S. Nagy, N. 2017. Stable Variation vs. Language Change and the Factors that Constrain Them. *University of Pennsylvania Working Papers in Linguistics*, 23 (2), 10.

Gathercole, V. C. M. Thomas, E. M. 2009. Bilingual first-language development: dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition*, 12, 213-37.

Giles, H. Johnson, P. 1987. Ethnolinguistic identity theory: A social psychological approach to language maintenance. *International Journal of the Sociology of Language*, 68, 69-100.

Gillian, E. Jaworski, S. 2014. A Discussion of Walpiri and Polish rhotics. *Unity in Diversity*, 2, 133-148.

Godson, L. 2004. Vowel production in the speech of Western Armenian heritage speakers. *Heritage Language Journal*, 2 (1), 45-70.

Goldberg, A. 1999. The emergence of the semantics of argument structure constructions. In MacWhinney, B. (Ed.), *The emergence of language*, 197-212. Mahwah (NJ): Erlbaum.

Goldberg, A. E. 2019. *Explain me this: Creativity, Competition, and the Partial Productivity of Constructions*. Princeton: Princeton University Press.

Goldinger, S. 1997. Words and voices: Perception and production in an episodic lexicon. In Johnson, K. Mullennix, J. (Eds.), *Talker variability in speech processing*, 33-66. San Diego: Academic Press.

Goldsmith, J. 1976. *Autosegmental Phonology*. PhD dissertation, MIT.

Gordon, E. 1997. Sex, Speech, and Stereotypes: Why Women Use Prestige Speech Forms More than Men. *Language in Society*, 26, 47-63.

Graddol, D. 1997. *The future of English?* London: British Council.

Grandi, N. 2020. La diversità inevitabile. La variazione linguistica tra tipologia e sociolinguistica. *Italiano LinguaDue*, 12 (1), 416-429.

Grassi, C. Sobrero, A. Telmon, T. 1997. *Fondamenti Di Dialettologia Italiana*. Bari: Laterza.

Greenberg, J. H. 1978. Some generalizations concerning initial and final consonant clusters. In Moravcsik, E. A. (Ed.), *Universals of human language*, vol. 2, 243-279. Stanford (CA): Stanford University Press.

Gries, S. 2015. The most under-used statistical method in corpus linguistics: multi-level (and mixed-effects) models. *Corpora*, 10 (1), 95-125.

Grosjean, F. 1982. *Life with Two Languages*. Cambridge (MA): Harvard University Press.

Grosjean, F. 2010. *Bilingual: Life and Reality*. Cambridge (MA): Harvard University Press.

Guardado, M. 2002. Loss and maintenance of first language skills: Case studies of Hispanic families in Vancouver. *Canadian Modern Language Review*, 58 (3), 341-363.

Guardado, M. 2014. The Discourses of Heritage Language Development: Engaging Ideologies in Canadian Hispanic Communities. *Heritage Language Journal*, 11 (1), 1-28.

Guenther, F. H. Espy Wilson, C. Boyce, S. Matthies, M. L. Zandipour, M. Perkell, J. S. 1999. Articulatory tradeoffs reduce acoustic variability during American English /r/ production. *The Journal of the Acoustical Society of America*, 105, 2854-2865.

Guy, G. R. 1980. Variation in the group and the individual: the case of final stop deletion. In Labov, W. (Ed.), *Locating language in time and space*, 1-36. New York (NY): Academic Press.

Hagiwara, R. 1995. Acoustic realizations of American /r/ as produced by women and men. *UCLA Working Papers in Phonetics*, 90, 1-187.

Hakuta, K. 2001. A critical period for second language acquisition? In Bailey, D. Bruer, J. Symons, F. Lichtman, J. (Eds.), *Critical Thinking about Critical Periods*, 193-205. Baltimore (MD): Paul H. Brookes.

Hamann, S. 2002. Retroflexion and Retraction revised. In Hall, T.A. Pompino-Marschall, B. Rochoń, M. (Eds.), *ZAS working papers*, 28, 13-25.

Hamann, S. 2003. *The phonetics and phonology of retroflexes*. PhD dissertation, University of Utrecht.

Harrison, B. 2000. Passing on the language: Heritage language diversity in Canada. *Canadian Social Trends*, 58, 14-19.

Hausmann, R. Tyson, L. D. Zahidi, S. (Eds.), 2006. *The Global Gender Gap Report 2006*. World Economic Forum, Geneva, Switzerland.

Hausmann, R. Tyson, L. D. Zahidi, S. (Eds.), 2015. *The Global Gender Gap Report 2015*. World Economic Forum, Geneva, Switzerland.

Hausmann, R. Tyson, L. D. Zahidi, S. (Eds.), 2021. *The Global Gender Gap Report 2021*. World Economic Forum, Geneva, Switzerland.

Hayes, B. 2009. *Introductory Phonology*. Malden (MA)/Oxford: Wiley-Blackwell.

Heredia, R. R. Cieślicka, A. B. 2014. Bilingual Memory Storage: Compound-Coordinate and Derivatives. In Heredia, R. R. Altarriba, J. (Eds.), *Foundations of bilingual memory*, 11-39. New York (NY): Springer.

Hoffman, M., Walker, J., 2010. Ethnolects and the city: ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change*, 22, 37-67.

Honeybone, P. 2008. Lenition, weakening and consonantal strength: tracing concepts through the history of phonology. In Carvalho, J. B. d. Scheer, T. Ségéral, P. (Eds.), *Lenition and Fortition*, 9-93. Berlin: Mouton de Gruyter.

House, J. 2003. English as a lingua franca: A threat to multilingualism? *Journal of Sociolinguistics*, 7 (4), 556-578.

Hunt, C. J. Fasoli, F. Carnaghi, A. Cadinu, M. 2016. Masculine self-presentation and distancing from femininity in gay men: An experimental examination of the role of masculinity threat. *Psychology of Men & Masculinity*, 17 (1), 108-112.

Hutchinson, J. Smith, A.D. 1996. *Ethnicity*. Oxford, UK: Oxford University Press.

Hymes, D. H. 1972. On communicative competence. In Pride, J. B. and Holmes, J. (Eds.), *Sociolinguistics: Selected Readings*, 269-293. Harmondsworth: Penguin.

Jaworski, S. 2010. Phonetic Realisations of the Polish Rhetoric Intervocalic Position: A Pilot Study. *Annales Neophilologiarum*, 4, 125-140.

Jaworski, S. Gillian, E. 2011. On the phonetic instability of the Polish rhotic /r/. *Poznań Studies in Contemporary Linguistics*, 47 (2), 380-398.

Johnson, D. E. 2009. Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and linguistics compass*, 3 (1), 359-383.

Johnson, D. E. 2014. *Progress in regression: Why natural language data calls for mixed-effects models*. Self-published manuscript.

Keefe, S. Padilla, A. 1987. *Chicano Ethnicity*. Albuquerque (NM): UNM Press.

Kirchner, R. M. 2013. *An effort-based approach to consonant lenition.* London/New York: Routledge.

Kupisch, T. 2013. A new term for a better distinction? A view from the higher end of the proficiency scale. *Theoretical Linguistics*, 39 (3-4), 203-214.

Kupisch, T. 2018. 2L1 simultaneous bilinguals as heritage speakers in morphosyntax and phonology. Introduction: Recent developments in early bilingualism. *Bilingualism*, 21 (4), 653-655.

Kupisch, T. Barton, D. Hailer, K. Kostogryz, E. Lein, T. Stangen, I. Van de Weijer, J. 2015. Foreign accent in adult simultaneous bilinguals. *Heritage Language Journal*, 1 (2), 123-150.

Kupisch, T. Rothman, J. 2018. Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, 22 (5), 564-582.

Labov W. 1963. The social motivation of a sound change. *Word*, 19, 273-309.

Labov W. 1966. *The social stratification of English in New York City*. Washington (DC): Center for Applied Linguistics.

Labov, W. 1969. Contraction, deletion and inherent variability of the English copula. *Language*, 45 (4), 715-762.

Labov, W. 1984. Field Methods of the Project on Linguistic Change and Variation. In Baugh, J. Scherzer, J. (Eds.), *Language in Use: Readings in Sociolinguistics*, 28-53. Englewood Cliffs: Prentice Hall.

Labov, W. 1990. The intersection of sex and social class in the course of linguistic change. *Language variation and change*, 2 (2), 205-254.

Labov, W. 2006. A sociolinguistic perspective on sociophonetic research. *Journal of phonetics*, 34 (4), 500-515.

Labov, W. Ash, S. Boberg, C. 2006. *The Atlas of North American English.* Berlin: Mouton de Gruyter.

Ladefoged, P. 1975. *A course in phonetics*. New York (NY): Harcourt Brace Jovanovich.

Ladefoged, P. Cochran, A. Disner, S. 1977. Laterals and trills. *Journal of the International Phonetic Association*, 7, 46-54.

Ladefoged, P. Johnson K. 2014. *A course in phonetics*. Cengage learning.

Ladefoged, P. Maddieson, I. 1996. *The sounds of the world's languages*. Oxford, UK: Blackwell.

Lakoff, R. T. 1975. *Language and Woman's Place*. New York (NY): Harper & Row.

Larsen-Freeman, D. 2007. Reflecting on the cognitive-social debate in second language acquisition. *The Modern Language Journal*, 91 (Focus Issue), 773-787.

Lave, J. Wenger, E. 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge, UK: Cambridge University Press.

Lawson, E. Scobbie, J. Stuart-Smith, J. 2013. Bunched /r/ promotes vowel merger to schwar: An ultrasound tongue imaging study of Scottish sociophonetic variation. *Journal of Phonetics*, 41, 198-210.

Lawson, E. Stuart-Smith, J. Scobbie, J. M. 2018. The role of gesture delay in coda/r/weakening: An articulatory, auditory and acoustic study. *The Journal of the Acoustical Society of America*, 143 (3), 1646-1657.

Leben, W. 1973. *Suprasegmental Phonology*. PhD Dissertation, MIT.

Ledgeway, A. 2009. *Grammatica diacronica del napoletano*. Tubinga: Max Niemeyer Verlag.

Lenneberg, E. 1967. *Biological Foundations of Language*. New York (NY): John Wiley & Sons, Inc.

Liceras, J. M. 2010. Second language acquisition and syntactic theory in the 21st century. *Annual Review of Applied Linguistics*, 30, 248-69.

Lindau, M. 1978. Vowel features. *Language*, 54 (3), 541-563.

Lindau, M. 1985. The story of r. In Fromkin, V. A. (Ed.), *Phonetic Linguistics,* 157-168. Orlando (FL): Academic Press.

Lindblom, B.E.F. 1990. Explaining phonetic variation: A sketch of the H&H theory. In Hardcastle, W.J. Marshal, A. (Eds.), *Speech production and speech modelling*, 403-439. Dordrecht: Kluwer Academic Publishers.

Liu, J. 1999. From their own perspectives: The impact of non-native ESL professionals on their students. In Braine, G. (Ed.), *Non-native educators in English language teaching*, 159-176. Mahwah (NJ): Lawrence Erlbaum.

Lleó, C. 2018. Aspects of the phonology of Spanish as a heritage language: From incomplete acquisition to transfer. *Bilingualism: Language and Cognition*, 21 (4), 732-747.

Long, M. H. 1990. Maturational constraints on language development. *Studies in Second Language Acquisition*, 12, 251-285.

Loporcaro, M. 1988. *Grammatica storica del dialetto di Altamura*. Pisa: Giardini.

Love, N. Ansaldo, U. 2010. The native speaker and the mother tongue. *Language Sciences*, 32 (6), 589-593.

Łyskawa, P. Nagy, N. 2019. Case marking variation in heritage Slavic languages in Toronto: Not so different. *Language Learning*, 70, 122-156.

Maas, C. JM. Hox, J. J. 2005. Sufficient Sample Sizes for Multilevel Modeling. *Methodology*, 1 (3), 86-92.

Maddieson, I. 1984. *Patterns of sounds*. Cambridge, UK: Cambridge University Press.

Magni, E. 2014. *Linguistica storica*. Bologna: Patròn editore.

McMahon, A. M. S. 1994. *Understanding language change*. Cambridge, UK: Cambridge University Press.

Meisel, J. 2011. *First and Second Language Acquisition*. Cambridge, UK: Cambridge University Press.

Meluzzi, C. Celata, C. 2020. Style and addressees in the selection of dialectal vs. standard phonetic forms: Sicilian rhotics. *Quaderns d'italià*, 25, 137-154.

Meyerhoff M. 2014. Variation and gender. Ehrlich, S. Meyerhoff, M. Holmes, J. (Eds.), *The Handbook of Language, Gender and Sexuality*, 87-102. Malden (MA): Wiley-Blackwell.

Meyerhoff, M. 2015. Turning variation on its head: Analysing subject prefixes in Nkep (Vanuatu) for language documentation. *Asia-Pacific Language Variation*, 1 (1), 78-108.

Meyerhoff, M. Ehrlich, S. 2019. Language, Gender, and Sexuality. *Annual Review of Linguistics*, 5, 455-475.

Mielke, J. Baker, A. Archangeli, D. 2010. Variability and homogeneity in American English /r/ allophony and /s/ retraction. *Laboratory Phonology*, 10, 699-730.

Mielke, J. Baker, A. Archangeli, D. 2016. Individual-level contact limits phonological complexity: Evidence from bunched and retroflex /ɹ/. *Language*, 92 (1), 101-140.

Milroy, L. 1980. *Language and social networks*. Oxford, UK: Blackwell.

Milroy, L. Milroy, J. 1992. Social Network and Social Class: Toward an Integrated Sociolinguistic Model. *Language in Society*, 21 (1), 1-26.

Monro, S. 2005. Beyond Male and Female: Poststructuralism and the Spectrum of Gender. *International Journal of Transgenderism*, 8 (1), 3-22.

Monro, S. 2019. Non-binary and genderqueer: An overview of the field. *International Journal of Transgenderism*, 20 (2-3), 126-131.

Montrul, S. 2008. *Incomplete acquisition in bilingualism. Re-examining the age factor*. Amsterdam: John Benjamins.

Montrul, S. 2010. Current issues in heritage language acquisition. *Annual Review of Applied Linguistics*, 30, 3-23.

Montrul, S. 2012. Is the heritage language like a second language? *EUROSLA Yearbook*, 12, 1-29.

Montrul, S. 2013a. How "native" are heritage speakers? *Heritage Language Journal*, 10 (2), 153-177.

Montrul, S. 2013b. Bilingualism and the Heritage Language Speaker. In Bhatia, T. K. and Ritchie, W. C. (Eds.), *The Handbook of Bilingualism and Multilingualism*, Second Edition. 168-189. Oxford, UK: Wiley Blackwell.

Montrul, S. Potowsky, K. 2007. Command of gender agreement in school-age Spanish-English bilingual children. *International Journal of Bilingualism*, 11, 301-328.

Mosso, C. Briante, G. Aiello, A. Russo. S. 2013. The Role of Legitimizing Ideologies as Predictors of Ambivalent Sexism in Young People: Evidence from Italy and the USA. *Social Justice Research*, 26 (1), 1-17.

Myhill, J. 2003. The native speaker, identity, and the authenticity hierarchy. *Language Sciences*, 25 (1), 77-97.

Nagy, N. 1994. Lexical change and language contact. *Penn Review of Linguistics*, 18, 117-132.

Nagy, N. 1996. *Language contact and language change in the Faetar speech community*. PhD Dissertation, University of Pennsylvania.

Nagy, N. 1997. Modeling contact-induced language change. *University of Pennsylvania Working Papers in Linguistics: A Selection of Papers from NWAVE 25*, 4 (1), 399-418.

Nagy, N. 2011. A multilingual corpus to explore geographic variation. *Rassegna Italiana di Linguistica Applicata*, 43 (1-2), 65-84.

Nagy, N. 2015. A sociolinguistic view of null subjects and VOT in Toronto heritage languages. *Lingua*, 164, B, 309-327.

Nagy, N. 2017a. Documenting variation in (endangered) heritage languages: how and why? In Hildebrandt, K. A. Carmen, J. Wilson, S. (Eds), *Documenting Variation in Endangered Languages. Language Documentation & Conservation Special Publication*, 13, 33-64. Honolulu: University of Hawai'i Press.

Nagy, N. 2017b. Cross-cultural approaches: Comparing heritage languages in Toronto. *Penn Working Papers in Linguistics*, 23 (2), 12.

Nagy, N. Aghdasi, N. Denis, D. Motut, A. 2011. Pro-drop in Heritage Languages: A cross-linguistic study of contact-induced change. *Penn Working Papers in Linguistics*, 17, 2.

Nagy, N. Brook, M. 2020. Constraints on speech rate: A heritage-language perspective. *International Journal of Bilingualism*, 19.

Nagy, N., Chociej, J. Hoffman, M. 2014. Analyzing Ethnic Orientation in the quantitative sociolinguistic paradigm. In Hall-Lew, L. Yaeger-Dror, M. (Eds.), *Special issue of Language and Communication: New perspectives on the concept of ethnolect*, 35, 9-26.

Nagy, N. Kochetov, A. 2013. Voice Onset Time across the generations: A cross-linguistic study of contact-induced change. In Siemund, P. Gogolin, I. Schulz, M. Davydova, J. (Eds.), *Multilingualism and Language Contact in Urban Areas: Acquisition - Development - Teaching – Communication*, 19-38. Amsterdam: John Benjamins.

Nagy, N. Umbal, P. 2021. Heritage Tagalog Phonology and a Variationist Framework of Language Contact. *Languages*, 6, 201.

Nespor, M. Bafile, L. 2008. *I suoni del linguaggio*. Bologna: Il Mulino.

Nodari, R. 2017. Autobiografie linguistiche di adolescenti calabresi. *Studi Italiani di Linguistica Teorica e Applicat*a, XLVI, 519-540.

Noels, K. Clément, R. 1998. Language in education: bridging educational policy and social psychological research. In Edwards, J. (Ed.), *Language in Canada*, 102-124. Cambridge, UK: Cambridge University Press.

Noels, K. 2014. Language variation and ethnic identity: A social psychological perspective. In Hall-Lew, L. Yaeger-Dror, M. (Eds.), *Special issue of Language and Communication: New perspectives on the concept of ethnolect*, 35, 88-96.

O'Grady, W. 2005. *Syntactic carpentry: An emergentist approach to syntax*. Mahwah (NJ): Erlbaum.

Oh, J. Jun, S. Knightly, L. Au, T. 2003. Holding on to childhood language memory. *Cognition*, 86, 53-64.

Ohala, J. J. 1983. *Aspects of Hindi phonology*. Delhi: Motilal Banarsidad.

Ohala, J. J. 1992. What's cognitive, what's not, in sound change. In Kellerman, G. Morrissey, M. D. (Eds.), *Diachrony within synchrony: language history and cognition*, 209-355. Frankfurt am Main: Peter Lang.

Ohala J.J. 1995. The perceptual basis of some sound pattern. In Connell, B. A. Arvaniti, A. (Eds.), *Papers in laboratory phonology, IV. Phonology and phonetic evidence*, 87-92. Cambridge, UK: Cambridge University Press.

Ohala J.J. Kawasaki, H. 1984. Prosodic phonology and phonetics. *Phonology Yearbook*, 1, 113-128.

Omadjohwoefe, O. S. 2011. Gender role differentiation and social mobility of women in Nigeria. *Journal of Social Sciences*, 27 (1), 67-74.

Ong, D. Stone, M. 1998. Three-dimensional vocal tract shapes in /r/ and /l/: A study of MRI, ultrasound, electropalatography, and acoustics. *Phonoscope*, 1, 1-13.

O' Reilly, C. C. 2001. Introduction: Minority Languages, Ethnicity, and the State in the European Union. In O' Reilly, C. C. (Ed.), *Language, Ethnicity, and the State, Volume One*, 1-19. Basingstoke/New York (NY): Palgrave.

Paradis, J. Genesee, F. 1996. Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition*, 18, 1-25.

Paradis, J. 2001. Do bilingual two-year-olds have separate phonological systems? *International Journal of Bilingualism*, 5, 19-38.

Paradis, M. 2004. *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins.

Paradis, M. 2009. *Declarative and procedural determinants of second languages*. Amsterdam: John Benjamins.

Paradis, J. 2010. The interface between bilingual development and specific language impairment. Keynote article for special issue with peer commentaries. *Applied Psycholinguistics*, 31, 3-28.

Pellegrino, F. Coupé, C. Marsico, E. 2011. A cross-language perspective on speech information rate. *Language*, 87 (3), 539-558.

Pennycook, A. 1994. *The cultural politics of English as an international language*. London: Longman.

Pérez-Leroux, A.T. Cuza, A. Thomas, D. 2011. From parental attitudes to input conditions. In Potowski, K. Rothman, J. (Eds.), *Bilingual youth: Spanish in English-speaking societies*, 149-176. Philadelphia: John Benjamins.

Pickering, M. J. Garrod, S. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Science*, 27 (2), 169-190.

Pierrehumbert, J. B. 2006. The next toolkit. *Journal of phonetics*, 4 (34), 516-530.

Pinheiro, J. C. Bates, D. M. 2000. *Mixed-effect models in S and S-PLUS*. New York (NY): Springer.

Pistella, J. Tanzilli, A. Ioverno, S. Lingiardi, V. Baiocco, R. 2017. Sexism and attitudes toward same-sex parenting in a sample of heterosexuals and sexual minorities: The mediation effect of sexual stigma. *Sexuality Research and Social Policy*, 15, 1-12.

Poggi Salani, T. 2010. *Italiano regionale*. Enciclopedia dell'italiano I: 726-729.

Polinsky, M. 1997. American Russian: language loss meets language acquisition. In Brown, W. Dornisch, E. Kondrashova, N. Zec, D. (Eds.), *Formal approaches to Slavic linguistics*, 370-407. Ann Arbor (MI): Michigan Slavic Publications.

Polinsky, M. 2000. The composite linguistic profile of speakers of Russian in the US. In Kagan, O. Rifkin, B. (Eds.), *The learning and teaching of Slavic languages and cultures*. 437-465. Bloomington (IN): Slavica.

Polinsky, M. 2006. Incomplete acquisition. *American Russian. Journal of Slavic Linguistics*, 14, 161-219.

Polinsky, M. 2012. What linguistics can learn from heritage languages: Research questions and research methods. *Conference on formal approaches to heritage languages*, University of Massachusetts.

Polinsky, M. 2013. Heritage languages. In Aronoff, M. (Ed.), *Oxford Bibliographies Online: Linguistics*. New York (NY): Oxford University Press.

Polinsky, M. 2018. *Heritage languages and their speakers* (Vol. 159). Cambridge, UK: Cambridge University Press.

Polinsky, M. Kagan, O. 2007. Heritage languages: In the 'wild' and in the classroom. *Language and linguistics compass*, 1 (5), 368-395.

Poplack, S. Meechan, M. 1998. Introduction: How languages fit together in codemixing. *International Journal of Bilingualism*, 2 (2), 127-138.

Poplack, S. Tagliamonte, S. A. 2001. *African American English in the Diaspora*. Malden/Oxford: Blackwell.

Proctor, M. 2011. Towards a gestural characterization of liquids: Evidence from Spanish and Russian. *Laboratory Phonology*, 2 (2), 451-485.

Quilis, A. 1981. *Fonetica Acustica de la Lengua Espanola*. Madrid: Biblioteca Romanica Hispanica.

Rampton, B. 1990. Displacing the "native speaker": Expertise, affiliation and inheritance. *ELT Journal*, 44, 97-101.

R Core Team. 2012. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Vienna, Austria. [http://www.R-project.org/, accessed on February 07th, 2022]

Recasens, D. 1990. The articulatory characteristics of palatal consonants. *Journal of Phonetics*, 18, 267-280.

Recasens, D. 1991. On the production characteristics of apicoalveolar taps and trills. *Journal of phonetics*, 19 (3-4), 267-280.

Recasens, D. 2004. The effect of syllable position on consonant reduction (evidence from Catalan consonant clusters). *Journal of Phonetics*, 32, 435-453.

Recasens, D. Farnetani, E. 1994. Spatiotemporal properties of different allophones of /l/: Phonological implications. *Phonologica, 1992: Proceedings of the 7th International Phonology Meeting*, 195-204.

Recasens, D. Pallarès, M. D. 1999. A study of /ɾ/ and /r/ in the light of the 'DAC' coarticulation model. *Journal of Phonetics*, 27, 143-170.

Rennicke, I. 2015. *Variation and change in the rhotics of Brazilian Portuguese*. PhD Dissertation, University of Helsinki.

Roach, P. 1998. Myth 18: Some languages are spoken more quickly than others. In Bauer, L. Trudgill, P. (Eds.), *Language myths*, 150-158. Harmondsworth: Penguin.

Romano, A. 2013. A preliminary contribution to the study of phonetic variation of /r/ in Italian and Italo-Romance. In Spreafico, L. Vietti, A. (Eds.), *Rhotics. New Data and Perspectives*, 209-226. Bozen: Bolzano University Press.

Romito, L. Scuticchio, P. A. 2008. La retroflessione in alcuni centri della provincia di Catanzaro: verso una mappa regionale. In Romito, L. Lio, R. Galatà, V. (Eds.), *Proceedings del IV Convegno AISV*, 412-422. Torriana: EDK Editore.

Rothman, J. 2009. Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism*, 13 (2), 155-163.

Rothman, J. Treffers-Daller, J. 2014. A prolegomenon to the construct of the native speaker: Heritage speaker bilinguals are natives too! *Applied linguistics*, 35 (1), 93-98.

Ruffino, G. 1991. *Dialetto e dialetti di Sicilia*. Palermo: CUSL.

Rumpf, A. DiVenanzio, L. 2012. Null and overt subjects in Italian and Spanish heritage speakers in Germany. *Heritage Languages: Language Contact-Change-Maintenance and Loss in the Wave of New Migration Landscapes Workshop*, Wuppertal.

Sankoff, D. Tagliamonte, S. Smith, E. 2005. *GoldVarb X: a variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto.

Schmid, M.S. 2011. *Language attrition*. Cambridge, UK: Cambridge University Press.

Schwartz, B. D. 2003. Child L2 acquisition: Paving the way. In Beachley, B. Brown, A. and Conlin, F. (Eds.), *Proceedings of the 27th Boston University Conference on Language Development*, 26-50. Somerville (MA): Cascadilla.

Schweinberger, M. 2022. *Fixed and Mixed-Effects Regression Models in R*. Brisbane: The University of Queensland. [https://slcladal.github.io/regression.html, accessed on February 07th, 2022]

Scobbie, J. 2006. (R) as a Variable. *Encyclopedia of language & linguistics, Second Edition*, 337-344.

Scontras, G. Fuchs, Z. Polinsky, M. 2015. Heritage language and linguistic theory. *Frontiers in psychology*, 6, 1545.

Sebregts, K. 2014. *The sociophonetics and phonology of Dutch r*. PhD Dissertation, Utrecht University.

Seliger, H. 1996. Primary language attrition in the context of bilingualism. In Ritchie, W. Bhatia, T. (Eds.), *Handbook of Second Language Acquisition*, 605-25. New York (NY): Academic Press.

Shosted, R. K. 2008. An Aerodynamic Explanation for the Uvularization of Trills? *Proceedings of ISSP*, 8, 421-424.

Solé, M. J. 1999. Production requirements of apical trills and assimilatory behavior. *Proceedings of the XIV International Conference of Phonetic Sciences*, 487-490.

Sorianello P. Mancuso, A. 1998. Le consonanti retroflesse nel cosentino: analisi preliminare. In Bertinetto, P. M. Cioni, L. (Eds.), *Unità fonetiche e fonologiche: produzione e percezione. Atti delle VIII Giornate di Studio del GFS, Pisa 18-19 dicembre 1997*, 142-154. Venezia: Università Ca' Foscari.

Spajic, S. Ladefoged, P. Bhaskararao, P. 1996. The trills of Toda. *Journal of the International Phonetic Association*, 26 (1), 1-21.

Spreafico, L. Celata, C. Vietti, A. Bertini, C. Ricci, I. 2015. An EPG+UTI study of Italian /r/. *ICPhS 2015: 18th International Congress of Phonetic Sciences* (Glasgow, August 10th 2015 – August 14th 2015).

Spreafico, L. Vietti, A. 2013. Introduction. In Spreafico, L. Vietti, A. (Eds.), *Rhotics: New data and perspectives*, 9-19. Bozen: Bolzano University Press.

Stanford, J. Preston, D. 2009. *Variation in indigenous minority languages*. Philadelphia: John Benjamins.

Statistics Canada. 2017. *2016 Census: Immigrant languages in Canada*. [https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E, accessed on February 07th, 2022].

Tager, D. Good, G. E. 2005. Italian and American masculinities: A comparison of masculine gender role norms. *Psychology of Men & Masculinity*, 6 (4), 264-274.

Tagliamonte, S. 2006. *Analysing sociolinguistic variation*. Cambridge, UK: Cambridge University Press.

Tagliamonte, S. A. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Malden/Oxford: Wiley-Blackwell.

Tagliamonte, S. A. 2016. *Making Waves: The Story of Variationist Sociolinguistics*. Malden/Oxford: Wiley-Blackwell.

Telmon, T. 1993. Varietà regionali. In Sobrero, A. (Ed.), *Introduzione all'italiano Contemporaneo. La Variazione e gli usi*, 93-149. Rome/Bari: Laterza.

Thomason, S. G. Kaufman, T. 1988. *Language Contact; Creolization and Genetic Linguistics*. Berkeley: University of California Press.

Tonkin, E. MacDonald, M. Chapman, M. (Eds), 1989. *History and Ethnicity*. London: Routledge.

Trask, R. L. 1996. *A Dictionary of Phonetics and Phonology*. London: Routledge.

Trudgill, P. 1972. Sex, Covert Prestige and Linguistic Change in the Urban British English of Norwich. *Language in Society*, 1, 179-195.

Trudgill, P. Hannah, J. 2008. *International English: A guide to the varieties of Standard English*. London/New York: Routledge.

Tucker, G.R. 1999. *A Global Perspective on Bilingualism and Bilingual Education*. Washington, (D.C.): Center for Applied Linguistics.

Turchetta, B. Modelli linguistici interpretativi della migrazione italiana. 2018. In Turchetta, B. Vedovelli, M. (Eds.), *Lo spazio linguistico italiano globale: il caso dell'Ontario*. Pisa: Pacini.

Turchetta, B. Vedovelli, M., 2018. *Lo spazio linguistico italiano globale: il caso dell'Ontario* [The Global Italian linguistic space: the case of Ontario]. Pisa: Pacini.

Twist, A. Baker, A. Mielke, J. Archangeli, D. 2007. Are 'covert' /ɹ/ allophones really indistinguishable? *Penn Working Papers in Linguistics*, 13, 2.

Valdés, G. 2000. Introduction. Spanish for Native Speakers, Volume I. *AATSP Professional Development Series Handbook for teachers K-16*. New York (NY): Harcourt College.

Van Hout, R. Van de Velde, H. 2001. Patterns of /r/ variation. In Van Hout, R. Van de Velde, H. (Eds.), *'r-atics. Sociolinguistic, phonetic and phonological characteristics of /r/*, 1-10. Bruxelles: ILVP.

Vietti, A. 2012. Alcune riflessioni sulla teoria degli esemplari e la variazione linguistica. *Rivista italiana di dialettologia*, 36.

Vietti, A. Spreafico, L. Romano, A. 2010. Tempi e modi di conservazione delle r italiane nei frigoriferi CLIPS. In Schmid, S. Michael, S. Dieter, S. (Eds.), *La dimensione temporale del parlato*, 113-128. Torriana: EDK.

Voghera, M. 2001. Teorie linguistiche e dati di parlato. *Atti del Congresso internazionale delle SLI "Teorie e dati linguistici"*, Naples, October 1999.

Voghera, M. 2017. *Dal parlato alla grammatica*. Rome: Carocci.

Walsh Dickey, L. 1997. *The phonology of liquids*. PhD Dissertation: University of Massachusetts.

Warner, N. 2012. Methods for studying spontaneous speech. In Cohn, A. Fougeron, C. Huffman, M. (Eds.), *The Oxford Handbook of Laboratory Phonology, Oxford Handbooks Series*, 621-633. Oxford, UK: Oxford University Press.

Weinreich, U. Labov, W. Herzog, M. 1968. Empirical foundations for a theory of language change. In Lehmann, W. P. Malkiel, Y. (Eds.), *Directions for historical linguistics*, 95-195. Austin (TX): University of Texas Press.

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

Wenger, E. 2000. *Communities of practice*. New York: Cambridge University Press.

Westbury, J. R. Hashi, M. Lindstrom, M. J. 1998. Differences among speakers in lingual articulation for American English /ɹ/. *Speech Communication*, 26, 203-226.

Wiese, R. 2001. The phonology of /r/. In Hall, T. A. (Ed.), *Distinctive Feature Theory*, 335-368. Berlin/New York: Mouton de Gruyter.

Wiese, R. 2011. The representation of rhotics. In Van Oostendorp, M. Ewen, C. Hume, E. Rice, K. (Eds.), *The Blackwell companion to phonology*, vol. 1, 711-729. Malden (MA): Wiley-Blackwell.

Willis, D. 2017. Endogenous and exogenous theories of language change. In Ledgeway, A. Roberts, I. (Eds.), *The Cambridge Handbook of Historical Syntax*, 491-514. Cambridge, UK: Cambridge University Press.

Wittenburg, P. Brugman, H. Russel, A. Klassmann, A. Sloetjes. H. 2006. ELAN: A Professional Framework for Multimodality Research. In Calzolari, N. Choukri, K. Gangemi, A. Maegaard, B. Mariani, J. Odijk, J. Tapias, D. (Eds.), *Proceedings of Fifth International Conference on Language Resources and Evaluation*, 1556-1559. Genoa: European Language Resources Association.

Wodak, R. de Cillia, R. Reisigl, M. Liebhart, K. 1999. *The discursive construction of national identity*. Translated by A. Hirsch and R. Mitten. Edinburgh, UK: Edinburgh University Press.

Wortham, S. 2004. From good student to outcast: The emergence of a classroom identity. *ETHOS*, 32, 164-187.

Wurm, S. A. 1972. *Languages of Australia and Tasmania*. The Hague: Mouton.

Zhang, Q. 2005. A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in society*, 32, 431-66.

Zhang, Z. Boyce, S. Espy Wilson, C. Tiede, M. 2003. Acoustic Strategies for Production of American English "retroflex" /r/. *Proceedings of the 15th International Congress of Phonetic Sciences, 1125-1128*. Barcelona, Spain: Universitat Autònoma.

Zhou, X. Espy-Wilson, C. Boyce, S. Tiede, M. Holland, C. Choe, A. 2008. A magnetic resonance imaging-based articulatory and acoustic study of ‘retroflex’ and ‘bunched’ American English /r/. *Journal of the Acoustical Society of America*, 123, 4466-4481.

Appendix

Appendix 1: Italian speakers' sample

The table provides, for each Italian speaker analyzed, the following information:

- Their unique speaker code (ID);
- Generation, based on immigration history (X = Homeland; 1 = 1st generation; 2 = 2nd generation);
- Sex;
- EO ID and EO LU scores (H = High; L = Low)
- The number of (r) tokens, and the number and percentages (of the total) of each (r) variant (Taps&trills, Approximants, Fricatives) – for the speakers involved in the word-internal singleton /r/ analysis;
- The number of -tr-/-dr- tokens, and the number and percentage of affrication of said tokens – for the speakers involved in the -tr-/-dr- analysis.

ID	**Gen.**	**Sex**	**EO ID**	**EO LU**	**(r) tok.**	**Taps&Tr.**		**Approx.**		**Fric.**		**-tr-/-dr- tok.**	**Affrication**	
					n	***n***	***%***	***n***	***%***	***n***	***%***	***n***	***n***	***%***
IXF22A	X	F	/	/	60	51	85%	9	15%	0	0%	31	0	0
IXF22B	X	F	/	/	60	52	87%	7	12%	1	2%	7	0	0
IXF35A	X	F	/	/	39	38	97%	0	0%	1	3%	10	0	0
IXF51B	X	F	/	/	60	50	83%	9	15%	1	2%	17	6	35%
IXM19A	X	M	/	/	50	34	68%	13	26%	3	6%	20	2	10%
IXM19B	X	M	/	/	45	34	76%	8	18%	3	7%	8	0	0
IXM28A	X	M	/	/	59	49	83%	8	14%	2	3%	15	0	0
IXM35A	X	M	/	/	59	34	58%	20	34%	5	8%	40	14	35%
IXM47A	X	M	/	/	59	19	32%	32	54%	8	14%	14	0	0
IXM52A	X	M	/	/	59	33	56%	18	31%	8	14%	18	1	6%
IXM61A	X	M	/	/	59	24	41%	26	44%	9	15%	15	3	20%
IXM64A	X	M	/	/	59	45	76%	11	19%	3	5%	35	14	40%
I1F59A	1	F	H	H	60	54	90%	6	10%	0	0%	26	0	0
I1F61A	1	F	/	/	/	/	/	/	/	/	/	50	18	36%
I1F71A	1	F	H	H	60	45	75%	8	13%	7	12%	27	11	41%
I1F73A	1	F	H	H	59	44	75%	9	15%	6	10%	27	2	7%
I1M61A	1	M	H	H	60	43	72%	13	22%	4	7%	11	2	18%

ID	Gen.	Sex	EO ID	EO LU	(r) tok.	Taps&Tr.		Approx.		Fric.		-tr-/-dr- tok.	Affrication	
					n	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*n*	*%*
I1M61B	1	M	L	H	60	47	78%	12	20%	1	2%	25	0	0
I1M62A	1	M	H	L	60	43	72%	14	23%	3	5%	63	7	11%
I1M75A	1	M	L	H	60	37	62%	19	32%	4	7%	23	13	57%
I2F44A	2	F	L	L	60	47	78%	12	20%	1	2%	14	0	0
I2F45A	2	F	H	H	60	52	87%	5	8%	3	5%	19	1	5%
I2F53A	2	F	L	L	60	48	80%	11	18%	1	2%	40	1	3%
I2F57A	2	F	L	L	60	44	73%	13	22%	3	5%	25	0	0
I2M42A	2	M	H	L	60	41	68%	19	32%	0	0%	35	0	0
I2M49A	2	M	/	/	50	33	66%	16	32%	1	2%	/	/	/
I2M52A	2	M	L	L	55	32	58%	20	36%	3	5%	15	0	0
I2M53A	2	M	L	L	55	33	60%	22	40%	0	0%	16	3	19%

Appendix 2: English speakers' sample

The table provides, for each Italo-Canadian speaker analyzed, the following information:

- Their unique speaker code, Generation, and Sex, as in Appendix 1;
- Their EO score (calculated on the basis of the whole EOQ; H = High, L = Low);
- The number of -tr-/-dr- tokens, and the number and percentage of affrication of said tokens.

ID	Gen.	Sex	EO score	-tr-/-dr- tokens	Affrication	
				n	*n*	*%*
E1M63A	1	M	H	12	11	92%
E2F45A	2	F	L	31	29	94%
E2F58A	2	F	H	9	9	100%
E2M51A	2	M	H	35	35	100%
E3F21A	3	F	H	43	42	98%
E3F21B	3	F	/	22	22	100%
E3F21C	3	F	L	27	27	100%
E3M19A	3	M	H	26	25	96%
E3M21A	3	M	L	28	28	100%