languages

MDPI

Article

L1 Attrition *vis-à-vis* L2 Acquisition: Lexicon, Syntax–Pragmatics Interface, and Prosody in L1-English L2-Italian Late Bilinguals

Mattia Zingaretti [1,2,*], Vasiliki Chondrogianni [1], D. Robert Ladd [1] and Antonella Sorace [1]

[1] School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh EH8 9AD, UK; v.chondrogianni@ed.ac.uk (V.C.); d.r.ladd@ed.ac.uk (D.R.L.); a.sorace@ed.ac.uk (A.S.)
[2] School of Education, Language and Psychology, York St John University, York YO31 7EX, UK
* Correspondence: m.zingaretti@yorksj.ac.uk

Abstract

Late bilingual speakers immersed in a second language (L2) environment often experience the non-pathological attrition of their first language (L1), exhibiting selective and reversible changes in L1 processing and production. While attrition research has largely focused on long-term residents in anglophone countries, examining changes primarily within a single L1 domain, the present study employs a novel experimental design to investigate L1 attrition, alongside L2 acquisition, across three domains (i.e., the lexicon, syntax–pragmatics interface, and prosody) in two groups of L1-English L2-Italian late bilinguals: long-term residents in Italy vs. university students in the UK. A total of 112 participants completed online tasks assessing lexical retrieval, anaphora resolution, and sentence stress patterns in both languages. First, both bilingual groups showed comparable levels of semantic interference in lexical retrieval. Second, at the syntax–pragmatics interface, only residents in Italy showed signs of L1 attrition in real-time processing of anaphora, while resolution preferences were similar between groups; in the L2, both bilingual groups demonstrated target-like preferences, despite some slowdown in processing. Third, while both groups showed some evidence of target-like L2 prosody, with residents in Italy matching L1-Italian sentence stress patterns closely, prosodic attrition was only reported for residents in Italy in exploratory analyses. Overall, this study supports the notion of L1 attrition as a natural consequence of bilingualism—one that is domain- and experience-dependent, unfolds along a continuum, and involves a complex (and possibly inverse) relationship between L1 and L2 performance that warrants further investigation.

Keywords: first language attrition; second language acquisition; late bilingualism; classroom-based language learning; lexicon; syntax–pragmatics interface; prosody

Academic Editors: Lewis Baker, Laura Domínguez and Glyn Hicks

Received: 13 October 2024
Revised: 12 July 2025
Accepted: 29 August 2025
Published: 4 September 2025

Citation: Zingaretti, M., Chondrogianni, V., Ladd, D. R., & Sorace, A. (2025). L1 Attrition *vis-à-vis* L2 Acquisition: Lexicon, Syntax–Pragmatics Interface, and Prosody in L1-English L2-Italian Late Bilinguals. *Languages*, *10*(9), 224. https://doi.org/10.3390/languages10090224

1. Introduction

1.1. The Emergence of Attrition in Late Bilingualism

The study of first language (L1) attrition in late-sequential bilingualism has gained significant momentum over the past four decades (Schmid, 2016). Instances of L1 attrition have been documented in various domains (for a comprehensive overview, see Schmid & Köpke, 2019), including the lexicon (e.g., Linck et al., 2009; Schmid & Jarvis, 2014), syntax and its interfaces (e.g., Tsimpli et al., 2004; Chamorro et al., 2016a), phonology and phonetics (e.g., Mennen, 2004; van Maastricht et al., 2016; de Leeuw et al., 2018), and even in the perception of emotions and emotionally-charged language, such as hate speech (e.g., Kim & Starks, 2008; Zingaretti et al., 2024). It has been proposed that attrition can be

understood as existing on a continuum, ranging from transient, online difficulties to more permanent alterations in underlying linguistic representations (Schmid & Köpke, 2017a, 2017b, 2019). This study specifically investigates attrition in the form of selective changes in real-time comprehension and production (reflecting processing difficulties rather than the loss of linguistic representations), which we interpret as arising in our data as a natural consequence of bilingualism, rather than being due to cross-linguistic influence (CLI) from the L2. Importantly, the understanding of attrition effects on a continuum implies that attrition effects can manifest at any stage of bilingual development, influenced by multiple factors, including proficiency, length of residence, age of L2 onset, and the context of L2 acquisition—e.g., through immersion or classroom-based instruction (Schmid & Köpke, 2017a, 2017b; Kupisch et al., 2017; de Leeuw, 2017). Therefore, under this view, "every bilingual is an attriter" (Schmid & Köpke, 2017a, 2017b). While this claim remains controversial, space precludes a full discussion here (but see responses to Schmid & Köpke, 2017a). Crucially, however, the implications for this study are that different types of bilingual speakers, including both classroom-based and immersed bilinguals, can exhibit signs of attrition, albeit to different degrees, and that susceptibility to attrition may vary across linguistic domains. To explore this, the present research expands the typical scope of L1 attrition research in late bilingualism by addressing three under-researched areas: (1) investigating L1 attrition, alongside L2 acquisition, across multiple linguistic domains; (2) comparing long-term residents in an L2-speaking country to classroom-based L2 learners in their L1 country; and (3) examining attrition with English as the L1.

1.2. Investigating L1 Attrition Alongside L2 Acquisition Across Multiple Domains

Much of the existing research has examined the attrition of linguistic domains in isolation (cf. list of studies in 1.1., among others), possibly due to the lack of an "…agreed-upon measuring stick" to compare different linguistic systems (Schmid & Köpke, 2009, p. 212). Yet, investigating L1 attrition in multiple domains within the same study allows us to assess whether different areas of language may be affected within the same individuals, and, if so, to what degree. While doing so, in this study, we also aim to examine the relationship between L2 acquisition and potential L1 attrition of the same linguistic properties by targeting three different domains: the lexicon, the syntax–pragmatics interface, and prosody. The following sections review the relevant literature for each linguistic domain in turn.

1.2.1. The Lexicon

The lexicon is often considered the first and most "vulnerable" or "sensitive" domain to undergo attrition within a linguistic system (Schmid & Köpke, 2009, p. 211; cf. Weinreich, 1953; Andersen, 1982; Weltens & Grendel, 1993; Köpke, 2002). There are at least two reasons that make this assumption convincing: (a) the lexicon contains a considerably larger number of items compared to, for instance, phonetic or morphological systems, and (b) the network of lexical items is less interconnected than others (e.g., the phonological inventory), thus being better able to tolerate change without yielding overall restructuring (Schmid & Köpke, 2009). For bilingual speakers who have fully acquired their L1 before learning an L2, lexical attrition can manifest itself as follows: disfluency, by means of increased pauses, hesitations, repetitions, and self-corrections during speech production (de Leeuw, 2007; Schmid & Jarvis, 2014); naming accuracy, in the form of difficulties in retrieving precise words, leading to lexical substitutions, circumlocutions, or semantic errors (Seliger & Vago, 1991; Jarvis, 2019); and lexical simplification, through reduced lexical diversity and sophistication, with a reliance on high-frequency words (Schmid & Jarvis, 2014).

Understanding how bilingual lexical access is shaped by L2 proficiency and L2 immersion is crucial for explaining which factors may contribute to L1 lexical attrition. A case in point is the study by Linck et al. (2009), where L1-English-speaking learners of L2-Spanish in immersion settings were compared with classroom-based learners of Spanish. Their findings revealed that, after only three months, immersed learners produced fewer words in their L1 (English) during a verbal fluency task, indicating reduced L1 accessibility, and showed diminished sensitivity to L1 lexical interference in a translation-recognition task, suggesting stronger L2 lexical–conceptual links and increased inhibition of the L1. These results meet the predictions of two complementary frameworks: the Revised Hierarchical Model (RHM) and the Activation Threshold Hypothesis (ATH).

On the one hand, the RHM (Kroll & Stewart, 1994) posits two separate lexicons for the L1 and L2, connected at the conceptual level. In the early stages of L2 acquisition, learners rely heavily on their L1 to access the meanings of L2 words, with strong links from L2 lexical items to L1 equivalents. As L2 proficiency increases, direct connections between L2 words and concepts strengthen, reducing dependence on the L1. This developmental trajectory predicts that less proficient bilinguals will be more sensitive to form-based (lexical) interference, whereas more proficient bilinguals will exhibit greater sensitivity to meaning-based (semantic) interference, as shown in Linck et al. (2009), as well as earlier supporting studies (e.g., Talamas et al., 1999).

On the other hand, the ATH (Paradis, 1993, 2004, 2007) argues that mental representations have activation thresholds that are lowered through frequent and recent use and raised through disuse and inhibition. In bilinguals, this means that L1 items that are used less frequently, especially when there is a more active and competing L2 counterpart, become increasingly difficult to access and are more likely to undergo attrition. This is particularly relevant in L2 immersion contexts, as shown in Linck et al. (2009), where reduced L1 use and increased L2 activation enhance inhibitory control mechanisms, raising the activation threshold for L1 words and contributing to slower lexical retrieval in the L1. Further support comes from Baus et al. (2013), who found that after four months of L2 immersion, bilinguals exhibited slower L1 lexical access, especially for low-frequency words, likely due to increased inhibition of the L1.

In short, the evidence reviewed here highlights that bilingual lexical access is affected by L2 proficiency levels and the context in which bilinguals operate. Specifically, links between L2 words and concepts are thought to get stronger with increasing levels of L2 proficiency, and the L2 immersion context is found to facilitate L2 acquisition by enhancing L1 inhibition. Ultimately, this results in enhanced lexical access in the L2 and, vice versa, reduced lexical access in the L1. However, the extent to which lexical attrition can co-occur alongside attrition in other linguistic domains within the same individuals remains unexplored—a gap that the present study aims to address.

1.2.2. The Syntax–Pragmatics Interface

One of the main frameworks in the study of L1 attrition of syntactic phenomena is the Interface Hypothesis (hereafter "IH"), put forth and refined over the years by Sorace and colleagues (cf. Chamorro & Sorace, 2019, for an overview). Building on Chomsky's Minimalist Program (cf. Chomsky, 2000), initial work (e.g., Tsimpli et al., 2004) only assumed a distinction between "narrow syntax" and "interface properties", where the latter are grammatical structures that carry semantic or pragmatic information, determined by "interpretable" features. Later work (e.g., Sorace, 2005; Tsimpli & Sorace, 2006), on the other hand, refines this view by distinguishing not just between narrow syntax and interface properties, but between two types of interfaces: "internal" interfaces, i.e., those between linguistic subcomponents (e.g., syntax and semantics), and "external" interfaces,

i.e., those between a linguistic subcomponent (e.g., syntax) and other cognitive systems (e.g., discourse). It is argued that external interface structures in particular lead to increased optionality in production and interpretation.

Two accounts have been proposed to explain the selective vulnerability of (external) interface structures under the IH: the "representational" account and the "processing" account (cf. Sorace, 2011). The representational account, supported by studies such as Tsimpli et al. (2004), proposes that attrition involves changes or restructuring in bilinguals' underlying grammatical knowledge, particularly affecting external interface structures due to interpretable features becoming optionally unspecified. For instance, Tsimpli et al. (2004) examined attrition of null and overt subject pronouns in Greek and Italian speakers living in an English-speaking environment. They found evidence that speakers tended to overuse overt pronouns and misinterpret preverbal subjects, which, according to the authors, is consistent with a representational restructuring of interpretable features like [topic-shift] and [focus], rather than merely processing difficulties. The processing account, on the other hand, proposes that it is bilinguals' ability to process and integrate contextual information, rather than their underlying grammatical representations, that is affected in L1 attrition. The latter account has been corroborated by studies in which L1-Spanish speakers living in the UK were found to be affected by attrition exclusively at the external syntax–pragmatics interface in overt subject pronoun interpretation (Chamorro et al., 2016a)—and not at the syntax–semantics interface in direct object marking (Chamorro et al., 2016b)—crucially only in online (eye-tracking) measures, but not offline (in a naturalness judgment task). Chamorro et al. (2016a) interpret their results as evidence that attrition only affects the ability to access knowledge, rather than knowledge representation itself. Under the processing account, this is due to strained processing abilities when computing more demanding properties at external interfaces; with regards to anaphora resolution, this results in the selection of the most easily accessible antecedent (i.e., the subject), and, consequently, in the attrition of overt subjects for speakers of L1-Greek and Italian (as in Tsimpli et al., 2004).

Beyond the IH, structures at the syntax–pragmatics interface have been argued to be vulnerable to CLI. Specifically, according to Hulk and Müller (Hulk & Müller, 2000; Müller & Hulk, 2001), CLI occurs when there is partial structural overlap between languages; that is, when language A allows two possible syntactic interpretations for a structure, whereas language B only allows one. Such conditions may give rise to CLI, with language B reinforcing one interpretation in language A. For example, non-null subject languages, like English, exclusively allow overt pronouns in finite clauses, while null-subject languages, like Italian, regulate pronoun distribution according to discourse context (e.g., overt pronouns signal topic-shift, while null pronouns maintain topic continuity). Thus, CLI is expected from the language with a single option (English) to the language with multiple options (Italian), potentially leading bilinguals to overextend overt pronouns (the option shared between the two languages) in Italian to contexts where null pronouns would be favoured, but not vice versa. Conversely, CLI is not predicted in situations of complete overlap or no overlap between the two languages.

However, based on evidence of overt pronoun overextension observed in speakers of two non-null subject languages, in both bilingual adults (e.g., Bini, 1993; Margaza & Bel, 2006; Lozano, 2006; Guido Mendes & Iribarren, 2007; De Prada Pérez, 2009) and children (e.g., Sorace et al., 2009; Michnowicz, 2015; and, more recently, Rodríguez-Ordóñez & Sainzmaza-Lecanda, 2018; Bonfieni et al., 2019), CLI alone cannot fully account for attrition phenomena in these cases. As already suggested by Sorace (2011, 2016), bilinguals' processing limitations at external interfaces are more likely responsible, given the increased cognitive load associated with managing two active languages simultaneously, which impacts

real-time integration of syntactic and pragmatic information. Nevertheless, there is counterevidence challenging the IH, with studies demonstrating vulnerability to attrition not only at external interfaces but also at internal interfaces, such as syntax–morphology in children (e.g., Montrul & Yoon, 2019; Van Dijk et al., 2022) and syntax–semantics in adults (e.g., Cuza, 2010). Furthermore, attrition of narrow-syntax properties has also been documented (e.g., Iverson, 2012; Domínguez & Hicks, 2016; Gürel, 2004; and, most recently, Baker, 2024), which runs contrary to the predictions made by the IH. Sorace (2020, p. 3) acknowledges these complexities, pointing towards the existence of "a continuum of conditions on syntactic realisation, ranging from more 'internal' to more 'external'"; yet, this reformulation complicates the creation of clear empirical predictions (Hicks & Domínguez, 2020).

Overall, despite mixed evidence regarding whether interfaces alone are selectively vulnerable to attrition, the existing research reviewed here highlights structures at the syntax–pragmatics interface as particularly—though not exclusively—susceptible. To better disentangle the effects of CLI from processing demands, it is therefore important to test situations in which CLI would *not* be predicted. Thus, the present study aims to address this gap by examining an external interface structure (i.e., anaphora resolution) where CLI from the L2 (Italian, a null-subject language) to the L1 (English, a non-null subject language) would not be expected (following Hulk & Müller, 2000; Müller & Hulk, 2001) (see Section 1.5.2).

1.2.3. Prosody

An intriguing yet underexplored area is that of prosodic attrition, specifically with regard to intonation. While some studies have examined L1 attrition in terms of pronunciation, often measured through accent ratings (e.g., Mennen, 2004), relatively few have investigated prosodic features considered to involve grammatical knowledge rather than phonetic accuracy. In particular, languages are known to differ in the way they use prominence—also known as sentence stress—to signal differences between "new" and "given" words or phrases within an utterance (Bolinger, 1972; Ladd, 1980; Swerts et al., 2002). In English (and the Germanic languages generally), it is common to "deaccent" repeated words (e.g., Nooteboom & Terken, 1982; Horne, 1990; Hirschberg, 1993). Compare the following sentences:

(1) a. Alex never received a letter.
b. The candidates were notified by letter, but Alex never received a letter.

In (1a), the main sentence stress would normally be placed on *letter*, but in the second clause of (1b), we would expect it to occur on *received*, avoiding a sequence of two clauses both ending in stressed *letter*. This deaccenting is rarely found in most comparable contexts in the Romance languages, including Italian and Romanian (e.g., Ladd, 2008), Catalan and Spanish (e.g., Cruttenden, 1993), and French (e.g., Gussenhoven, 2004), and such "information structure" effects are often conveyed by changes of word order (such as right-dislocation, clefting, fronting, etc.). There is clear experimental evidence for this difference between Germanic and Romance languages (e.g., Swerts et al., 2002; Swerts, 2007), which has been shown to be subject to bidirectional prosodic transfer in L2 speakers (e.g., van Maastricht et al., 2016).

Specifically, van Maastricht et al. (2016) conducted a large-scale study involving 124 participants to examine bidirectional intonational transfer in a functional context. The study included L1-Dutch learners of L2-Spanish, L1-Spanish learners of L2-Dutch, and L1-Dutch and L1-Spanish control groups. Acoustic analyses of pitch accents revealed that both learner groups exhibited prosodic transfer from their L1 to their L2 while also acquiring L2 intonation patterns, with higher proficiency correlating with more "native-like" L2 intonation (cf. Swerts & Zerbian, 2010; Reichle & Birdsong, 2014). At the same time, both L2-Dutch and L2-Spanish learners showed evidence of transferring pitch accent

features from their L2 back to their L1, which was a surprising finding given that these were classroom-based learners who had only started learning the L2 at university, had at most six years of L2 experience, and used the L2 only sporadically.

The study by van Maastricht et al. (2016) provides evidence of bidirectional prosodic transfer and suggests that L1 prosodic attrition can occur even among classroom-based learners with moderate proficiency and limited L2 experience. Nevertheless, the extent to which prosodic attrition can manifest alongside attrition in other linguistic domains within the same speakers remains to be seen—an issue the current study seeks to investigate.

1.3. Comparing Long-Term Residents to Classroom-Based Learners

L1 attrition research has primarily focused on long-term immersed bilinguals who have lived in an L2-dominant environment for extended periods (e.g., Gürel, 2004; Tsimpli et al., 2004; Kaltsa et al., 2015; Chamorro et al., 2016a, 2016b; Gargiulo, 2020). However, less attention has been given to instructed bilinguals, who acquire and use the L2 extensively in formal educational settings while remaining in an L1-dominant environment. Despite maintaining frequent L1 use, these bilinguals still experience language co-activation and cross-linguistic influence (Schmid & Köpke, 2017a), raising the question of whether immersion and L1 disuse are necessary conditions for attrition effects to emerge, or whether linguistic competition alone can drive changes in the L1 (Yılmaz & Schmid, 2018).

Some key differences in the input that L2 learners typically receive in immersive versus classroom-based contexts also need to be acknowledged. First, immersed bilinguals typically receive a greater quantity of L2 input and, in turn, reduced L1 input compared to instructed bilinguals. Second, the quality of the input that immersed bilinguals receive tends to be richer and more naturalistic, encompassing varied registers and interactional contexts, in contrast to the often simplified or formal input encountered in L2 classrooms. Third, immersed bilinguals are more likely to be L2-dominant, while classroom-based learners are generally L1-dominant. Finally, immersed bilinguals are more likely to attain higher L2 proficiency than their instructed counterparts. All of these factors have been argued to increase the likelihood of L1 attrition (cf. Schmid & Cherciov, 2019, for an overview).

Although some studies have now documented L2 influence on the L1 even in instructed contexts (e.g., Kecskes, 1998; Kecskes & Papp, 2000, 2003; Cenoz, 2003; Cook, 2003; Długosz, 2021; Requena & Berry, 2021), relatively few studies have directly compared bilinguals in different input contexts. To our knowledge, the first to make this direct comparison were Linck et al. (2009), who examined lexical retrieval in immersed vs. instructed L1-English L2-Spanish bilinguals, reporting reduced L1 lexical access in immersion but no attrition-like effects in the instructed group (see Section 1.2.1). More recently, Martín-Villena (2023) examined advanced L1-Spanish L2-English bilinguals receiving frequent L2 exposure in a formal instructed setting in Spain. The study compared their production, interpretation, and processing of third-person singular subject referring expressions with that of immersed bilinguals in the UK and Spanish functional monolinguals. The results revealed that instructed bilinguals exhibited increased use of overt pronouns in topic continuity contexts, aligning with patterns previously reported in L1 attrition research (see Lozano, 2009; Martín-Villena & Lozano, 2020). Furthermore, the study found that both instructed and immersed bilinguals differed from functional monolinguals in their interpretation and processing of overt pronouns, suggesting that L1 attrition, at least at the syntax–pragmatics interface, can occur even in bilinguals who remain in an L1-dominant environment.

Given the limited research directly comparing bilinguals in different input contexts, further investigation is needed to determine whether attrition-like changes in instructed bilinguals parallel those of immersed bilinguals, and whether said changes may manifest

across linguistic domains. By directly comparing immersed and instructed bilingual groups in terms of lexicon, syntax–pragmatics interface, and prosody, this study offers a novel design that enables a more nuanced understanding of the role of input in L1 attrition across domains.

1.4. Examining Attrition in L1-English

Despite extensive research on L1 attrition, relatively few studies have examined attrition in L1-English speakers. One possible reason for this gap is the global status of English as a prestigious *lingua franca* (cf. Crystal, 2012; Jenkins et al., 2018). L1-English speakers, even when residing in foreign-language environments, often continue to use their L1 regularly for communication, work, and social interaction. Additionally, widespread English-language media and the prevalence of English-speaking expatriate communities may provide continued L1 exposure, potentially mitigating L1 attrition effects (cf. Zingaretti & Spelorzi, 2021). However, despite the assumption that maintained L1 exposure may result in less attrition, no clear or comprehensive conclusions can yet be drawn regarding the precise relationship between frequency of L1 use and attrition due to a complex interplay of external and individual factors, as well as methodological inconsistencies and limitations (Schmid, 2019).

Nonetheless, some research has shown that L1-English speakers are not immune to attrition, at least with regard to some linguistic domains. For instance, Porte (1999) investigated English language teachers living in Spain and found that prolonged exposure to "non-native" English (e.g., through teaching) led to changes in their L1, including lexical borrowing, grammatical restructuring, and increased tolerance of L2-influenced errors. Similarly, Dostert (2004) documented cases of English attrition among long-term expatriates in Germany, with participants reporting lexical retrieval difficulties and fluency issues when using their L1 in monolingual contexts. These studies suggest that even speakers of a dominant global language can experience attrition-like effects, particularly in contexts of reduced L1 input and increased L2 exposure. Attrition effects in L1-English have also been identified at the phonetic level. Sučková (2020), for instance, examined VOT in Anglophone expatriates residing in Czechia and found that their English phonetic patterns had shifted under the influence of Czech. Specifically, expatriates produced shorter VOTs for English voiceless plosives and exhibited increased prevoicing for voiced plosives, indicating L2-induced phonetic changes. These findings align with broader evidence of phonetic drift in L1 speakers immersed in an L2-dominant environment.

However, Gürel (2007) provides an important counterpoint, showing that syntactic attrition in English may not occur as readily as lexical or phonetic changes. Comparing L1-English speakers in Turkey to L1-Turkish speakers in North America, the study found no conclusive evidence of syntactic restructuring in L1-English despite extensive L2-Turkish exposure. The findings suggest that frequency of L1 use plays a crucial role: as predicted by the ATH (Paradis, 2007; cf. Section 1.2.1), attrition is unlikely unless L1 contact is significantly reduced.

Given these findings, further research is needed to explore how different linguistic domains of L1-English are affected under varying levels of L2 exposure and L1 use. This study contributes to the discussion by examining English L1 attrition across different linguistic domains in bilingual speakers in immersed vs. instructed settings.

1.5. The Current Study

As outlined above, despite attrition being widely documented in separate language domains, it is unclear whether multiple domains may be affected within the same individuals, and, if so, to what degree. It is also unclear whether changes in the L1 are related

to L2 acquisition within the same domain. Moreover, much of the existing research has focused on long-term residents immersed in an English L2 environment. As such, whether L2 speakers in other settings, such as classroom-based learners with English as their L1, are also susceptible to attrition remains underexplored. Thus, the present research focuses on examining attrition across the lexicon, syntax–pragmatics interface, and prosody in the L1 (English) and L2 (Italian) of two groups of late bilingual speakers: long-term residents in Italy vs. university students in the UK. By analysing these two groups, we aim to investigate not only how different L2-learning settings impact the L1 across different domains but also what happens to the L2 in this process.

1.5.1. Research Questions

In this work, we aim to answer the following research questions (RQs):

1. Is evidence of L1 (i.e., English) attrition found in different groups of late L2 (i.e., Italian) speakers (i.e., long-term residents in Italy vs. university students in the UK)? If so, does attrition affect the two groups to the same extent?
2. Does L1 attrition affect separate language domains (i.e., the lexicon, syntax–pragmatics interface, and prosody) within the same individuals? If so, does attrition affect these domains to different degrees[1]?

1.5.2. Predictions and Hypotheses

Given that the present study focuses on online instances of L1 attrition, specifically in late bilingual adults (whose L1 was fully acquired before the onset of bilingualism), we put forward the following predictions[2] and associated hypotheses:

- *Prediction 1*: Assuming Schmid and Köpke's (2017a, 2017b) contested claim that "every bilingual is an attriter" (as discussed in Section 1.1), we expect evidence of L1-English attrition for both groups, albeit to different degrees.
    - *Hypotheses*: Following Schmid and Köpke's (2017a, 2017b) continuum-based model of attrition, we hypothesise that long-term residents in Italy may show more extensive L1 attrition than university students learning Italian in the UK. This prediction is based on differences in the quantity and quality of the input the two groups receive, as well as likely differences in dominance and proficiency (see participant characteristics in Section 2.1), which may favour attrition in the immersed context over the instructed context, as discussed in Section 1.3. Specifically, long-term residents in Italy may show not only slower processing but also lower accuracy and greater divergence in preferences from L1 functional monolinguals[3] (i.e., controls) than university students in the UK in L1 comprehension/production. Attrition effects may, in turn, be more limited for university students in the UK—e.g., showing only slower processing, but not necessarily lower accuracy or highly divergent preferences from L1 controls in comprehension and production tasks. However, the overall degree of L1-English attrition found may be lower than that reported in previous research on L2-English speakers due to the prestige and lingua franca status of English nowadays. This potentially allows residents in Italy to access (and thus maintain) their L1 more easily, as well as making it less stimulating for university students in the UK to practice and better their L2, when compared to bilinguals of other L1s (e.g., L1-Italian, Spanish, French, etc.) in anglophone countries or university students learning L2-English in their respective L1 countries.
- *Prediction 2*: We expect that some[4] domains will be more susceptible to L1 attrition than others, with attrition anticipated to be more prominent in the lexicon compared

to the syntax–pragmatics interface or prosody. However, differences are also expected within these domains, depending on the group of speakers.

- *Hypotheses*: First, given that lexical access has consistently been shown to be one of the most vulnerable domains in L1 attrition (see Section 1.2.1), lexical attrition may be present in both bilingual groups; however, given that L2 immersion contexts have been found to increase L1 inhibition (Linck et al., 2009) as well as differences in input, lexical attrition may be more pronounced among long-term residents in Italy. Second, based on the predictions of Sorace and Filiaci's (2006) IH, attrition at the syntax–pragmatics interface (i.e., in anaphora resolution) is generally expected. However, based on Hulk and Müller's (Hulk & Müller, 2000; Müller & Hulk, 2001) assumptions of the directionality of CLI, attrition from Italian to English is not predicted, since English permits only overt pronouns, whereas Italian allows both overt and null forms conditioned by discourse (see Section 1.2.2). Thus, while attrition due to CLI is not predicted, attrition driven by processing demands may still remain possible, especially in the case of long-term residents due to the likely increase in L1 inhibition and differences in input as a result of L2 immersion. Third, given that residents in an L2-speaking country are exposed daily to L1-like prosody in a range of communicative settings, the quality and quantity of L2 input may gradually influence their L1 prosodic patterns. On the other hand, classroom-based students in their L1-speaking country typically receive more limited and formal input, and prosody is rarely a focus of explicit instruction, making prosodic attrition less likely in this group.

2. Materials and Methods

2.1. Participants

Participants were recruited on social media platforms (e.g., Facebook, Twitter/X, Instagram) and by word of mouth (at the University of Edinburgh, in other UK universities, and elsewhere). A total of 112 participants were recruited, divided into the following:

1. University students in their final year of Italian studies at different UK universities, born in the UK, and who grew up with English as their only L1—currently living in the UK (n = 27).
2. Long-term residents in Italy, born in the UK, and who grew up with English as their only L1—currently living in Italy after having emigrated there (and started learning Italian) at sixteen years of age or older (n = 27).
3. L1-English controls, born and living in the UK and fluent in no other language aside from English—with knowledge of an L2 at school level permitted, as long as it was not deemed fluent (n = 31).
4. L1-Italian controls, born and living in Italy and fluent in no other language aside from Italian—with knowledge of an L2 at school level permitted, as long as it was not deemed fluent (n = 27).

Background information was collected through the Language Experience and Proficiency Questionnaire (Marian et al., 2007) administered via Qualtrics (Provo, UT). The rest of the experiment was carried out on Testable (Rezlescu et al., 2020), where bilingual participants were also asked to complete Amenta et al.'s (2021) LexITA vocabulary test to assess their L2 proficiency levels, as well as a reading-span (Waters & Caplan, 1996) task (in their respective L1s) to measure working memory capacity. Table 1 shows the main characteristics of the four participant groups. As can be seen in Table 1, residents in Italy scored markedly higher than university students in the UK on the LexITA vocabulary test (M = 69.9 vs. M = 32.2), with this difference found to be statistically significant, according

to Student's independent samples *t*-test, $t(52) = -6.68$, $p < .001$. As group and proficiency cannot be modelled together in the analyses (see Section 3), this between-group difference in L2 proficiency is highlighted here, given its relevance for interpreting the findings (as discussed in Section 4). It is also worth noting that the two groups differed significantly in age, with residents being substantially older than students ($M = 45.4$ vs. $M = 21.9$), as confirmed by Welch's *t*-test (accounting for unequal variances), $t(26.26) = -8.46$, $p < .001$. In contrast, no significant difference was observed in working memory scores measured by the reading-span task ($M = 49.7$ vs. $M = 46.4$), as indicated by Student's *t*-test, $t(52) = -1.35$, $p = .184$. Both age and working memory were included as covariates in the analyses.

Table 1. Characteristics of the participant groups.

Characteristic	University Students	Long-Term Residents	L1-English Controls	L1-Italian Controls
Number of Participants	27	27	31	27
Born in	UK	UK	UK	Italy
Living in	UK	Italy	UK	Italy
L1	English	English	English	Italian
L2	Italian	Italian	N/A	N/A
Gender [a]	♀ = 24 ♂ = 3 ⚲ = 0	♀ = 21 ♂ = 6 ⚲ = 0	♀ = 17 ♂ = 13 ⚲ = 1	♀ = 23 ♂ = 4 ⚲ = 0
Age (Years)	*M* = 21.9 *SD* = 1	*M* = 45.4 *SD* = 14.4	*M* = 27 *SD* = 6.2	*M* = 28.2 *SD* = 7.9
Age of L2 Acquisition (Years)	*M* = 17.6 *SD* = 1.3	*M* = 24 *SD* = 7.7	N/A	N/A
Length of Residence in L2 Country (Years) [b]	*M* = 0.1 *SD* = 0.3	*M* = 20.4 *SD* = 14.1	N/A	N/A
L2 Proficiency Score [c]	*M* = 32.2 *SD* = 24.1	*M* = 69.9 *SD* = 16.6	N/A	N/A
Working Memory Score (Number Recalled) [d]	*M* = 46.4 *SD* = 9.8	*M* = 49.7 *SD* = 8.3	*M* = 45.7 *SD* = 10.7	*M* = 47.7 *SD* = 5.5

[a] Gender is reported as follows: ♀(female), ♂(male), and ⚲(other/non-binary). [b] Two university students both reported spending one year in Italy prior to testing; hence, the mean (*M*) and standard deviation (*SD*) have been reported for students as well as long-term residents' length of residence. [c] L2 proficiency scored as *N* words correctly identified as being Italian words –2**N* of non-words wrongly identified as Italian words (out of 90 words). [d] Working memory scored as the number of correctly recalled to-be-remembered words (out of 60).

2.2. Materials

2.2.1. Lexicon: Translation-Recognition Task

To investigate the activation of L1 lexical information and access to L2 meaning, Sunderman and Kroll's (2006) translation-recognition task (TRT) was adapted from Spanish to Italian for the purposes of this study. In brief, participants saw an L2 (Italian) word followed by an L1 (English) word and had to judge whether the second word was the correct translation of the first by pressing a button (cf. Section 2.3 for full procedural details)[5]. The task included 120 trials for each participant, with 60 baseline trials (i.e., *yes* trials) and 60 target trials (i.e., *no* trials). Half of the *no* trials included related (+) distractors of three kinds (see Table 2 for an example):

1. Lexical neighbour (LN) distractors, where the second word presented (the L1 word; e.g., "amber") is similar to the first word (the L2 word; e.g., *albero*/"tree") in form.

2. Translation neighbour (TN) distractors, where the second word presented (e.g., "spree") is similar in form to the correct L1 translation (e.g., "tree" for *albero*).
3. Semantic (S) distractors, where the second word presented (e.g., "leaves") is similar in meaning to the correct L1 translation (e.g., "tree" for *albero*).

Table 2. Illustration of the distractors for *albero*/"tree" used in the translation recognition task.

Italian	English	Related (+)			Unrelated (−)		
		LN+	**TN+**	**S+**	**LN−**	**TN−**	**S−**
albero	tree	amber	spree	leaves	norms	hirer	agenda

Note. LN = lexical neighbour (distractor related orthographically to L2 word); TN = translation neighbour (distractor related orthographically to L1 translation); S = semantic (distractor related in meaning to L1 translation). Each participant saw only one of the six possible distractors for *albero*. Unlike the original task employed in Sunderman and Kroll (2006), no grammatical class condition was included. Since the TRT engages both L1 and L2, only bilinguals performed the task.

Following Sunderman's (2013) recommendations, the remaining *no* trials included unrelated (−) trials matched to the related (+) trials for length and frequency, with frequencies retrieved from the WaCky corpora (Baroni et al., 2009), using ukWaC for English and itWaC for Italian, respectively. The LN trials were controlled for onset code change and the Levenshtein distance from the L2 word, whereas the TN trials were controlled for onset code change and the Levenshtein distance from the L1 word (Levenshtein, 1966). *Yes* trial items were created separately and matched to *no* trials for frequency and length. Six lists were created to counterbalance the trials.

2.2.2. Syntax–Pragmatics Interface: Self-Paced Reading Task

To investigate the resolution of anaphora in both L1-English and L2-Italian, two self-paced reading tasks (SPRTs) were employed (one for each language). The experimental items used in the SPRTs were designed on the basis of those in previous studies on anaphora resolution (e.g., Bonfieni, 2018). The items in each SPRT included the following:

- Twelve items with (overt) pronoun ambiguity, possibly resolved either towards the NP1 or towards the NP2 (i.e., "ambiguous NP1/2").
- Twelve items with unambiguous (overt) pronoun resolution, half with forced NP1 resolutions (i.e., "unambiguous NP1") and half with forced NP2 resolutions (i.e., "unambiguous NP2").
- Forty-eight (English)/sixty (Italian) filler items[6].

Examples of ambiguous NP1/2 and unambiguous NP1 items are reported below for anaphora resolution (2–3) in both languages, where the English stimulus (b) is a translation of Italian (a).

- Examples of ambiguous NP1/2 stimuli for anaphora resolution:

(2) a. La nipote saluta la nonna sull'autobus. Lei è davvero ansiosa.
b. The granddaughter greets the grandmother on the bus. She is really anxious.

- Examples of unambiguous NP1 stimuli for anaphora resolution:

(3) a. La nipote saluta il nonno sull'autobus. Lei è davvero ansiosa.
b. The granddaughter greets the grandfather on the bus. She is really anxious.

As can be seen from the examples above for anaphora resolution, in target items (2a) and (2b), the pronoun *lei*/"she" is ambiguous, as it could refer either to the NP1 *la nipote*/"the granddaughter" or to the NP2 *la nonna*/"the grandmother". On the other hand, in unambiguous items (3a) and (3b) the pronoun *lei*/"she" should preferably refer to

the NP1 *la nipote*/"the granddaughter" because of the gender of the antecedent. Forced resolutions such as these were counterbalanced throughout the items, so that half would be disambiguated towards the NP1 (as in the examples 3a–b above) and the other half would be disambiguated towards the NP2. Therefore, the possibility that a pronoun may refer to a third, unmentioned individual was not taken into account in the present experiment. Moreover, because the pronoun is in the subject position of an independent sentence—thus, not in a structural relationship with potential antecedents (e.g., subject or object in a preceding clause)—its resolution is not related to syntax, as such, but rather discourse.

2.2.3. Prosody: Picture-Naming Task

To investigate sentence stress patterns in both L1-English and L2-Italian, two picture-naming tasks (PNTs) were employed (one for each language). The PNTs were designed on the basis of the tasks employed in previous research (i.e., Swerts et al., 2002; Swerts & Zerbian, 2010; van Maastricht et al., 2016), taking pictures from the Multilingual Picture (MultiPic) database by Duñabeitia et al. (2022). Specifically, participants were presented with pairs of images depicting four possible referents (i.e., dragon/flower/pumpkin/tiger—*drago/fiore/zucca/tigre*) in four possible colours (i.e., blue/green/red/yellow—*blu/verde/rosso/giallo*). Participants were asked to describe the two images in each picture, producing sentences like *(a) blue dragon (and a) yellow pumpkin*, where the description of the second image was the "target NP" that might exhibit deaccenting depending on the experimental condition. The different combinations of pairs of colours and images were presented in a fully randomised order in three conditions:

1. Contrastive/contrastive (CC), where both the adjective and the noun describing the second image (i.e., the target NP) were absent from the description of the preceding image; e.g., a blue dragon preceded by a red pumpkin (32 items in total).
2. Contrastive/given (CG), where the first content word of the description of the second image contrasted with the first content word of the preceding NP, but the second content word was the same; to give an English example, a red tiger preceded by a blue tiger (16 items in total).
3. Given/contrastive (GC), where the second content word of the target NP contrasted with the second content word of the preceding NP, but the first content word was the same; to give an English example, a yellow flower preceded by a yellow pumpkin (16 items in total).

Note that, because of the difference in the adjective–noun word order between English and Italian, items from condition 2 (CG) in English (e.g., "a blue tiger and a red tiger", shown in Figure 1) belong to condition 3 (GC) in Italian ("*una tigre blu e una tigre rossa*"), and vice versa. In both languages, that is, there are 16 items in each condition, but the items belong to different conditions depending on the language. In CG pairs (condition 2), it was expected that L1-English speakers would deaccent the repeated word and that L1-Italian speakers would not. In the other two kinds of pairs (CC and GC, conditions 1 and 3), more subtle phonetic effects on the first content word were assumed to be possible, but no deaccenting of the second content word was expected. Consequently, the reporting of the results was limited to noting the presence or absence of "adjustment" to the prominence of the second content word.

Figure 1. Example of a contrastive/given (CG) pair in English: a red tiger (target item) preceded by a blue tiger, eliciting the response "(a) blue tiger (and a) red tiger". This same item belongs to condition GC in Italian because of the difference in the adjective–noun word order between Italian and English.

2.3. Procedure

Due to COVID-19 restrictions, participant recruitment and data collection were carried out entirely online. Monolingual participants completed all tasks in one session, while bilingual participants completed the tasks first in one language, and after six days, in the other language. The order was counterbalanced across participants, except for the TRT, which was administered last during the L2 session to bilingual participants only due to the nature of the task, which engages both the L1 and the L2. For all tasks, participants were instructed to complete the experiment in a quiet room, turn off any devices, and carefully read on-screen instructions. Responses were collected via a keyboard or microphone as appropriate, and reaction times (RTs) were recorded for the TRT and the SPRT only.

The TRT consisted of reading words shown on the screen: an Italian word would appear in the middle of the screen for 400 milliseconds, followed by a short inter-stimuli interval of 100 milliseconds. A second word, in English, would then appear in the middle of the screen. Participants were asked to press 1 if the English word was the correct translation of the previous Italian word, or 0 if the English word was not the correct translation. After ten practice trials, the task began, with a break allowed halfway through as advised on their screen.

The SPRT consisted of reading sentences presented in chunks, such as (1–2) above, as they were presented on the screen using a stationary window paradigm (Just et al., 1982). Participants advanced through the sentence by pressing the space bar after reading each chunk. At the end of each sentence, participants were asked to indicate which NP the pronoun referred to by pressing either 1 or 0. For instance, after reading sentence (2a) above, participants would be asked *Chi è ansioso/a?* ("Who is anxious?") and they would then have to press either 1, corresponding to *la nipote*/"the granddaughter", or 0, corresponding to *la nonna*/"the grandmother". Eight practice trials were followed by the main task, with a break allowed halfway through.

The PNT[7] consisted of describing the colour and noun of the items appearing on the screen in pairs, by saying them out loud; e.g., "(a) blue tiger (and a) red tiger" (as for the example in Figure 1 above). Pairs of items were presented in turn, and participants were instructed to press next only after having described the item on the screen. The experiment began after nine practice trials, with a break allowed halfway through.

It is important to note that, originally, all PNT recordings were to be submitted to three local judges at the University of Edinburgh: an expert would judge all recordings, whilst the other two would judge recordings in one language each for validation. Specifically, the experts were to judge categorically whether relative prominence on the second noun phrase was produced as having been adjusted or not. Trials with uncertain judgments were to be flagged for reconciliation. Due to speaker variability (e.g., list intonation—that is, producing all items as a single prosodic sequence with rising pitch on each instead of treating each item separately; see, e.g., Morrill et al., 2014), the judges would not specify which acoustic cue was used to assess relative prominence (e.g., intensity, F0, etc.). Having two sets of judgments per language was expected to mitigate difficulties in objective evaluation. However, COVID-19 restrictions impacted the recruitment of judges. As a result, one local expert at the University of Edinburgh (co-author D.R.L.) judged all recordings, and a

second expert at Radboud University Nijmegen (Prof. Carlos Gussenhoven) assessed 50% of the English recordings for validation.

In the judgment of the data, some unforeseen difficulties involving the nature of the PNT arose. Specifically, over half of the recordings had degraded acoustic cues due to hesitations and increasing speech rates. In the Italian task, issues with grammatical gender were common, especially among the students. Additionally, participants tended to adopt faster speech rates or fall into list intonation as the task progressed. Rather than discarding substantial data, the judges made binary decisions on whether relative prominence had been adjusted, despite degraded cues. This way, only four recordings were excluded due to incompleteness or poor quality. In English, two separate analyses were conducted, one using the first expert's judgments alone and the other with reconciled judgments. Since the results were similar, offering reassurance as to the reliability of the judgments, Section 3 reports the findings on the first expert's judgments for both English and Italian.

2.4. Design Measures and Predicted Outcomes

Given the multiplicity of tasks, design measures and predicted outcomes are detailed below for each on the basis of the literature reviewed in Sections 1.2–1.4, and the predictions and hypotheses outlined in Section 1.5.2:

1. TRT: For accuracy (1 = accurate, 0 = inaccurate), it was expected that long-term residents in Italy (henceforth "residents") would be less likely to respond accurately in S+ trials compared to S- trials, as a result of the increased semantic interference reported in speakers with higher L2 proficiency. Conversely, university students in the UK (henceforth "students") were anticipated to show lower accuracy in TN+ trials compared to TN- trials, as a result of the heightened lexical interference reported in speakers with lower L2 proficiency. As for RTs, S+ trials were expected to result in longer RTs for residents, whereas students were predicted to have longer RTs for TN+ trials.
2. SPRT: For preference (1 = NP1, 0 = NP2), no significant differences were anticipated between the groups in L1-English (i.e., no L1 attrition predicted), due to the assumed directionality of CLI: from English (the language with one option) to Italian (the one with two options), but not vice versa. On the other hand, in L2-Italian, it was expected that ambiguous NP1/2 (target) trials would be more likely to be resolved as NP1 for both bilingual groups—especially students—when compared to L1-Italian controls, due to the influence of L1-English. As for RTs, no CLI-induced group differences in RTs were anticipated in L1-English; however, some general slowdown in English may still be possible, particularly among residents due to the likely increase in L1 inhibition and differences in input as a result of L2 immersion. In L2-Italian, both bilingual groups—especially students—were expected to take longer on ambiguous NP1/2 trials compared to L1-Italian controls.
3. PNT: For relative prominence (1 = adjusted for repetition, 0 = not adjusted), residents were expected to be less likely to adjust prominence in L1-English compared to L1-English controls and students in CG trials, due to potential L2-to-L1 intonational transfer. In L2-Italian, it was expected that students would be more likely than both residents and L1-Italian controls to adjust prominence in CG trials, due to possible L1-to-L2 transfer.

2.5. Statistical Procedures

All statistical analyses were conducted using jamovi (version 2.3), an open-source, R-based software (The jamovi project, 2022), with models implemented via the GAMLj module (Gallucci, 2019). The models were fitted separately for each language in the SPRT

and PNT[8]. For all tasks, mixed-effects models were employed to account for participant- and item-level variability. Following Barr et al. (2013), a maximal random effects structure was used, with random intercepts and slopes for both participants (subjectID) and items (itemID). Where model convergence could not be achieved, random correlations were removed, with the model complexity being reduced stepwise, excluding terms that accounted for the least variance first. Post hoc comparisons were Bonferroni-adjusted to control for multiple comparisons. Working memory score, as measured by the reading-span task, and age were included as covariates in the RT analyses (after being decorrelated using R; R Core Team, 2022) as both variables have been shown to influence processing speed in bilingual language comprehension (cf. Salthouse, 1996; Payne et al., 2014). We also attempted to include L2 proficiency, as measured by the LexITA test, in the models; however, the inclusion of this variable resulted in convergence issues and was therefore excluded from the final models. In both the TRT and SPRT, RTs were trimmed to exclude extreme values below 300 ms or above 3000 ms (cf. Ratcliff, 1993). In the SPRT, RTs were also residualised using word length (i.e., number of letters per segment) to account for individual variation between participants (cf. Marinis, 2010); moreover, pre-critical segments (regions 1–4) were analysed to ensure there were no group or condition differences prior to the critical region. No significant effects emerged in the pre-critical regions; therefore, only critical and post-critical regions (5–7) are reported in the main results[9]. Descriptive statistics and model specifications are provided in each results subsection, alongside model visualisations; full model outputs, together with data and materials, are available on the OSF project page.

3. Results

For the sake of clarity, we report the results for each task separately, and then discuss them altogether in relation to our RQs in Section 4.

3.1. TRT Results

An overall descriptive table containing the mean RT in milliseconds and percent accuracy for both bilingual groups for all trials is reported in Table 3. In line with Sunderman (2013), while the analysis for accuracy includes both correct and incorrect trials, RTs are only analysed for correct trials.

Table 3. Mean RT (ms) and percent accuracy in the translation-recognition task.

	University Students in UK		**Long-Term Residents in Italy**	
	RT	**Accuracy**	**RT**	**Accuracy**
Yes Trials	738	78.2%	775	94.1%
No Trials				
Lexical Neighbour				
Related	788	96.2%	859	96.6%
Unrelated	733	98.9%	777	99.3%
Translation Neighbour				
Related	777	93.3%	843	94.1%
Unrelated	726	97.0%	815	100%
Semantic Distractor				
Related	847	79.8%	1000	87.1%
Unrelated	691	98.5%	797	98.1%

Note. Lexical neighbour distractors are related orthographically to the L2 word; translation neighbour distractors are related orthographically to the L1 translation; semantic distractors are related in meaning to the L1 translation. Accuracy percentages include both correct and incorrect trials, while RT milliseconds are only computed for correct trials.

The logistic mixed model fitted to predict accuracy (1 = accurate, 0 = inaccurate) included group (students, residents), condition (LN+, LN−, S+, S−, TN+, TN−, *yes*[10]), and the group × condition interaction. The results showed a main effect of condition ($p < .001$), no main effect of group ($p = .52$) and a significant interaction between group and condition ($p < .001$). To explore this interaction, within-group comparisons revealed that responses in S+ trials were significantly less likely to be accurate than in S- trials for both students ($OR = .020$, $SE = .016$, $z = -4.37$, $p = .001$) and residents ($OR = .020$, $SE = .022$, $z = -3.58$, $p = .031$). Other within-group comparisons for TN+/− and LN+/− trials were not statistically significant ($ps > .05$). The interaction was driven by the fact that, between groups, only in *yes* trials were students' responses significantly less likely to be accurate than residents' responses ($OR = .110$, $SE = .023$, $z = -10.60$, $p < .001$). Predicted values are illustrated in Figure 2.

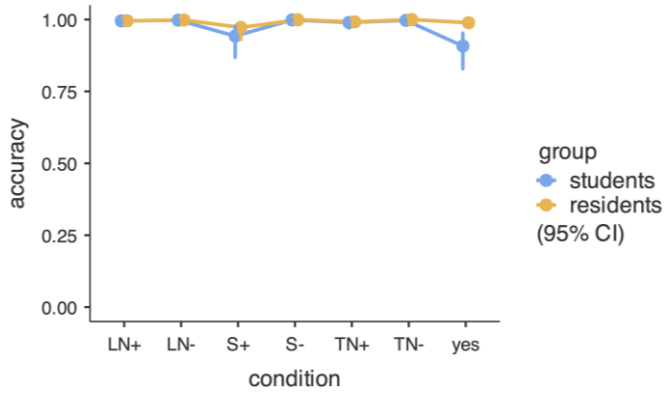

Figure 2. Predicted probability of accuracy (1 = accurate, 0 = inaccurate) in different conditions of the translation-recognition task for university students of Italian in the UK and long-term residents in Italy. LN = lexical neighbour distractor; TN = translation neighbour distractor; S = semantic distractor. +/− signs = related/unrelated trials. Error bars = 95% confidence interval (CI).

The linear mixed model fitted to predict RT included group (students, residents), condition (LN+, LN−, S+, S−, TN+, TN−, *yes*), the group × condition interaction, age, and reading-span score. The results showed a main effect of condition ($p < .001$), no main effect of group ($p = .87$), and a significant interaction between group and condition ($p < .001$). To explore this interaction, within-group comparisons revealed that RTs were significantly longer in S+ trials than in S- trials for both students and residents: for students, the estimated difference was 175.22 ms ($SE = 34.1$, $t(1630.3) = 5.13$, $p < .001$); for residents, the estimated difference was 222.29 ms ($SE = 33.5$, $t(1524.4) = 6.64$, $p < .001$). Other within-group comparisons for TN+/− and LN+/− trials were not statistically significant ($ps > .05$). Between groups, no differences were found within conditions (all $ps > .05$). Additionally, neither age nor reading-span score significantly predicted RT (both $ps > .05$). Mean RTs by group and condition are illustrated in Figure 3.

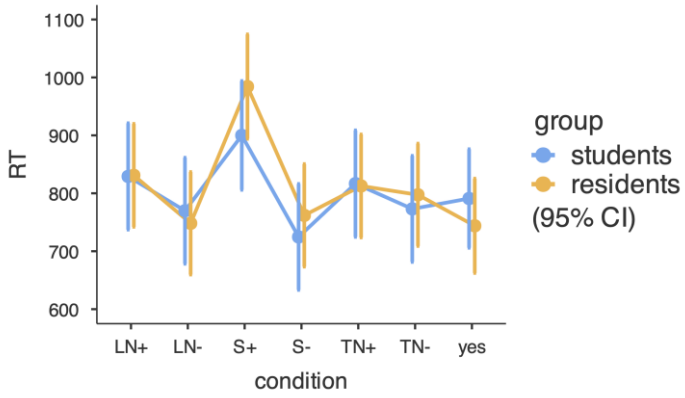

Figure 3. RT in different conditions of the translation-recognition task for university students of Italian in the UK and long-term residents in Italy. LN = lexical neighbour; TN = translation neighbour; S = semantic neighbour. +/− signs = related/unrelated trials. Error bars = 95% confidence interval (CI).

3.2. SPRT Results

To begin, mean percent preferences of all groups in the resolution of English and Italian anaphora are reported in Table 4. The table also includes the results of one-tailed one-sample Wilcoxon signed-rank tests, carried out to check whether groups' preference means were significantly different from chance level (50%) in the expected direction for each condition.

Table 4. Mean percent preferences in English and Italian anaphora resolution.

	University Students in UK			**Long-Term Residents in Italy**			**L1 Controls**		
	NP1	**NP2**	***p***	**NP1**	**NP2**	***p***	**NP1**	**NP2**	***p***
English									
Unambiguous NP1	97.5%	2.5%	***	99%	1%	***	97%	3%	***
Unambiguous NP2	2.5%	97.5%	***	1%	99%	***	3%	97%	***
Ambiguous NP1/2	55.5%	44.5%	*	53%	47%		57%	43%	**
Italian									
Unambiguous NP1	94.5%	5.5%	***	99%	1%	***	99%	1%	***
Unambiguous NP2	7%	93%	***	3%	97%	***	1%	99%	***
Ambiguous NP1/2	43%	57%		42%	58%		41%	59%	***

Note. "L1 controls" refer to English monolinguals in the English self-paced reading task and Italian monolinguals in the Italian task, respectively. Asterisks indicate a significant difference in the expected direction from chance level (50%) on one-tailed one-sample Wilcoxon signed-rank tests of means: * = $p < .05$, ** = $p < .01$, and *** = $p < .001$.

In Table 4 it is interesting to note that whilst clear NP1 preferences show in the resolution of ambiguous NP1/2 trials for L1-English controls (with a mean 57% preference for NP1 resolutions in said trials), the significance of the same NP1 preference is lower in the case of students (namely, 55.5%). In the same context, residents perform roughly at chance level (53%). By contrast, in Italian resolutions, the three groups seem to perform rather similarly, reporting comparable rates of NP2 preference in ambiguous NP1/2 trials (i.e., 59% for L1-Italian controls, 58% for residents, and 57% for students).

The logistic mixed model fitted to predict preference in English included group (students, L1-English controls, residents), condition (unambiguous NP1, unambiguous NP2, ambiguous NP1/2), and the group × condition interaction. The results showed a main effect of condition ($p < .001$), no main effect of group ($p = .84$), and no significant interaction between group and condition ($p = .46$). To unpack the main effect of condition: across groups, pronouns in ambiguous NP1/2 trials were significantly less likely to be resolved as referring to NP1 than in unambiguous NP1 trials ($OR = .015$, $SE = .010$, $z = -6.40$,

p = < .001), and significantly more likely to be resolved as NP1 than in unambiguous NP2 trials (OR = 141.13, SE = 99.81, z = 7.00, p = < .001). Predicted values are illustrated in Figure 4.

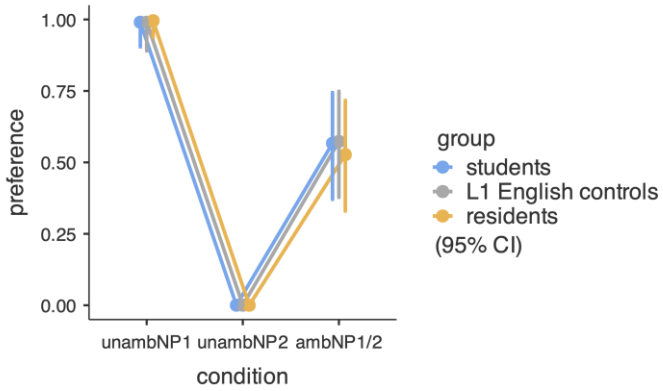

Figure 4. Predicted probability of preference (1 = NP1, 0 = NP2) in different conditions of the English self-paced reading task for university students of Italian in the UK, long-term residents in Italy, and L1 (English) controls. UnambNP1 = unambiguous NP1 trials; UnambNP2 = unambiguous NP2 trials; AmbNP1/2 = ambiguous NP1/2 trials. Error bars = 95% confidence interval (CI).

The logistic mixed model to predict preference in Italian included group (students, L1-Italian controls, residents), condition (unambiguous NP1, unambiguous NP2, ambiguous NP1/2), and the group × condition interaction. The results showed a main effect of condition (p < .001), no main effect of group (p = .66), and a significant interaction between group and condition (p = .002). Within-group comparisons revealed that, for all three groups, pronouns in ambiguous NP1/2 trials were significantly more likely to be resolved as referring to NP1 than in unambiguous NP2 trials: students (OR = 12.99, SE = 7.70, z = 4.33, p < .001), L1-Italian controls (OR = 191.63, SE = 215.17, z = 4.68, p < .001), and residents (OR = 37.75, SE = 26.81, z = 5.11, p < .001). Pronouns in ambiguous NP1/2 trials were also significantly less likely to be resolved as NP1 than in unambiguous NP1 trials: students (OR = .025, SE = .015, z = −5.92, p < .001), L1-Italian controls (OR = .002, SE = .002, z = −5.51, p < .001), and residents (OR = .002, SE = .003, z = −5.33, p < .001). The interaction can be explained by the different OR magnitudes, lower for students than L1 controls and residents in the first comparison, and greater for students than L1 controls and residents in the latter. Between groups, no differences were found within conditions (all ps > .05). Predicted values are illustrated in Figure 5.

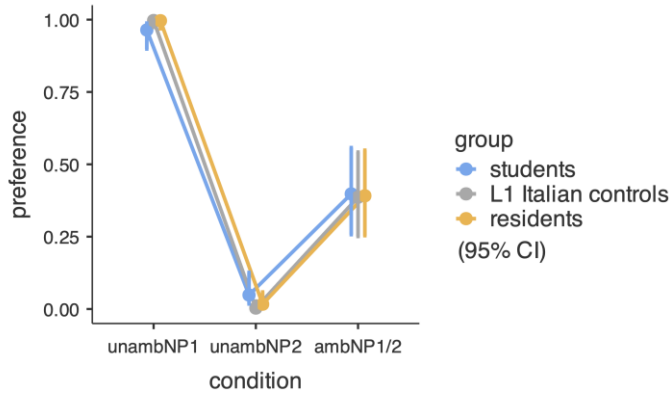

Figure 5. Predicted probability of preference (1 = NP1, 0 = NP2) in different conditions of the Italian self-paced reading task for university students of Italian in the UK, long-term residents in Italy and L1 (Italian) controls. UnambNP1 = unambiguous NP1 trials; UnambNP2 = unambiguous NP2 trials; AmbNP1/2 = ambiguous NP1/2 trials. Error bars = 95% confidence interval (CI).

The linear mixed model fitted to predict RT in English included group (L1-English controls, residents, students), region (5, 6, 7), condition (unambiguous NP1, unambiguous NP2, ambiguous NP1/2), the three-way interaction group × region × condition, age, and reading-span score. All main effects and interactions involving region and condition were significant (ps < .05), except for group, which was not (p = .88). The three-way interaction between group, region, and condition was also significant (p = .006). Within groups, residents were the only group whose RTs in ambiguous NP1/2 trials were significantly longer than RTs in both unambiguous NP1 trials (estimated difference = 104.51 ms, SE = 23.8, t(74.4) = 4.39, p = .001) and unambiguous NP2 trials (estimated difference = 119.59 ms, SE = 24.2, t(79.5) = 4.94, p < .001). Moreover, in the ambiguous NP1/2 trials, residents were the only group whose region 5 RTs were significantly longer than region 6 RTs (estimated difference = 165.74, SE = 30.2, t(3829) = 30.2, p < .001), and whose region 7 RTs were significantly longer than region 5 RTs (estimated difference = 206.536, SE = 29.8, t(3824) = 6.924, p < .001). In the same trials, region 7 RTs were also significantly longer than region 6 RTs for both residents (estimated difference = 372.28 ms, SE = 30.7, t(3830) = 12.13, p < .001) and L1-English controls (estimated difference = 197.69 ms, SE = 32.8, t(3831) = 6.04, p < .001), though not for students (p > .05). Between-group comparisons in ambiguous NP1/2 trials showed that, when compared to residents, region 7 RTs were significantly shorter for both L1-English controls (estimated difference = −233.73 ms, SE = 57.2, t(141) = −4.08, p = .026) and students (estimated difference = −271.37 ms, SE = 59.2, t(145) = −4.58, p = .003). In the same trials, RTs in regions 5 and 6 did not differ significantly between groups (all ps > .05). Additionally, neither age nor reading-span score significantly predicted RT (both ps > .05). Residualised RTs by group, condition, and region are illustrated in Figure 6.

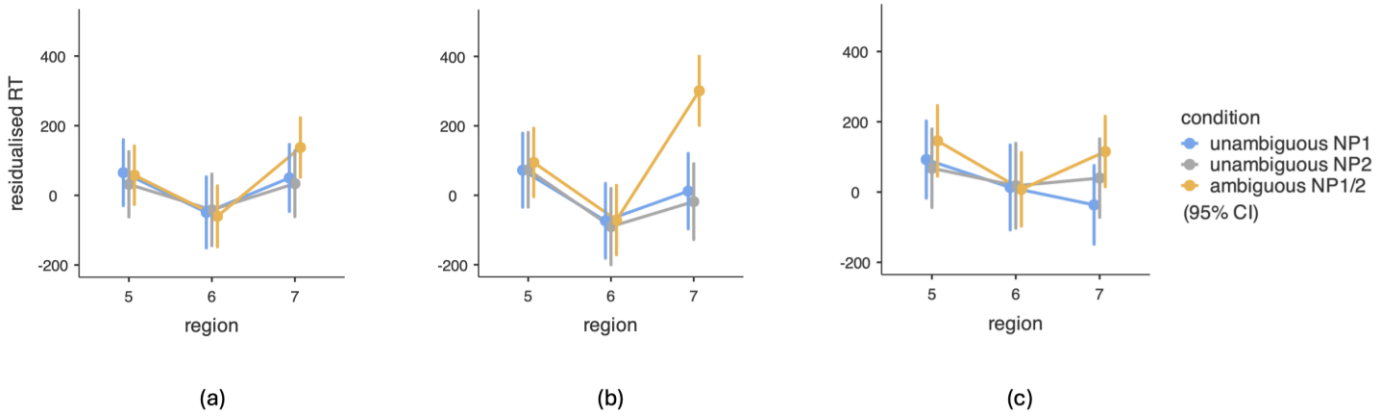

Figure 6. Residualised RT in critical region (5) and post-critical regions (6–7) for different conditions of the English self-paced reading task for (**a**) L1-English controls, (**b**) long-term residents in Italy, and (**c**) university students of Italian in the UK. E.g., "She is very polite": "she" = region 5; "is" = region 6; "very polite" = region 7. Error bars = 95% confidence interval (CI).

The linear mixed model fitted to predict RT in Italian included group (L1-Italian controls, residents, students), region (5, 6, 7), condition (unambiguous NP1, unambiguous NP2, ambiguous NP1/2), the three-way interaction group × region × condition, age, and reading-span score. The three-way interaction was not significant (p = .323), but all two-way interactions involving group, region, and condition were (ps < .05). The main effects of region and condition were also significant, while group was not (p = .66). To unpack the significant group × condition interaction, within groups, RTs in ambiguous NP1/2 trials were longer than in unambiguous NP1 trials for students (estimated difference = 111.36 ms, SE = 23.5, t(179.7) = 4.74, p < .001) and residents (estimated difference = 132.24 ms, SE = 21.1, t(118.0) = 6.26, p < .001). RTs in ambiguous NP1/2 trials were also longer than in unambiguous NP2 trials for students (estimated difference = 109.86 ms, SE = 23.2, t(169.2) = 4.74,

$p < .001$) and residents (estimated difference = 122.28 ms, $SE = 21.1$, $t(118.3) = 5.79$, $p < .001$). No such differences were found for L1-Italian controls (all ps > .05). Between-group comparisons did not reveal any significant differences within conditions (all ps > .05). Finally, neither age nor working memory score was a significant predictor in the analysis (both ps > .05). Residualised RTs by group, condition, and region are illustrated in Figure 7.

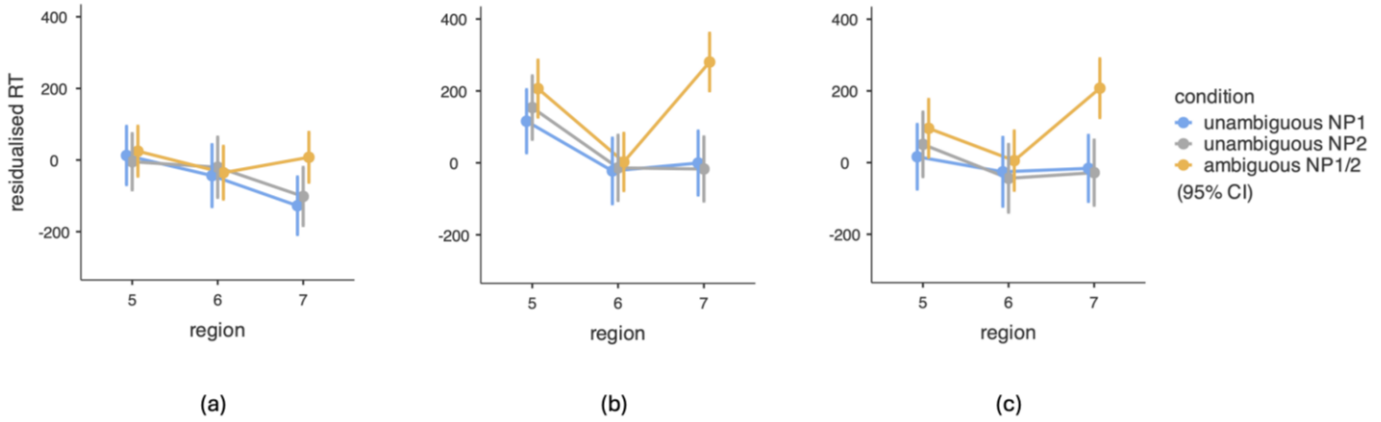

Figure 7. Residualised RT in critical region 5 and post-critical regions (6–7) for different conditions of the Italian self-paced reading task for (**a**) L1-Italian controls, (**b**) long-term residents in Italy, and (**c**) university students of Italian in the UK. E.g., "Lei è molto educata": "lei" = region 5; "è" = region 6; "molto educata" = region 7. Error bars = 95% confidence interval (CI).

3.3. PNT Results

Table 5 below reports the mean percentage of trials where relative prominence was judged as adjusted (PA) or not adjusted (PNA) for different conditions in each language.

Table 5. Mean percentage of trials where relative prominence was judged as adjusted/not adjusted in the English and Italian picture-naming tasks.

	University Students in UK		**Long-Term Residents in Italy**		**L1 Controls**	
	PA	**PNA**	**PA**	**PNA**	**PA**	**PNA**
English						
CC	19.2%	80.8%	15.4%	84.6%	19.6%	80.4%
CG	87.3%	12.7%	89.6%	10.4%	92.2%	7.8%
GC	17.9%	80.1%	14%	86%	17.3%	82.7%
Italian						
CC	2.4%	97.6%	0.4%	99.6%	0.7%	99.3%
CG	35.4%	64.6%	10.1%	89.9%	0.5%	99.5%
GC	0.3%	99.7%	0.2%	99.8%	0%	100%

Note. PA = prominence (judged as) adjusted; PNA = prominence (judged as) not adjusted. CC = contrastive/contrastive; CG = contrastive/given; GC = given/contrastive. L1 controls are English monolinguals in the English picture-naming task and Italian monolinguals in the Italian task, respectively.

The results in Table 5 clearly show that, in English, lower percentages of prominence judged as adjusted were reported (to a similar degree) on CC and GC items than on CG items for L1 controls (respectively, 19.6% and 17.3% versus 92.2%). This was expected given that the difference between CC and GC items (e.g., "GREEN TIGER" vs. "green TIGER") is both phonetically more subtle and possibly of a different phonological status than the difference involved in the CG trials (e.g., "GREEN tiger"), where the possible prominence patterns are arguably categorically distinct. In Italian, on the other hand, low percentages of prominence judged as adjusted were reported on all items for L1 controls (respectively, 0.7%, 0.5%, and 0%, for CC, CG, and GC items). This was also expected

given that prominence patterns are not commonly adjusted for repeated material in Italian. However, the percentages of prominence being judged as adjusted on CC and GC English trials were unexpectedly higher than those on the same trials in Italian, for all groups alike. Further investigations revealed that in most of these trials, the adjustment of prominence was triggered by an unforeseen complication arising due to the experimental design (as later discussed in Section 3.3.2).

3.3.1. Confirmatory Analyses for Prominence Adjustment

The logistic mixed model fitted to predict the judgment of relative prominence adjustment in English included group (students, L1-English controls, residents), condition (CC, CG, GC), and the group × condition interaction. The results showed a main effect of condition ($p < .001$), no main effect of group ($p = .39$), and no significant interaction between group and condition ($p = .35$). To unpack the main effect of condition, across groups, relative prominence was significantly more likely to be judged as having been adjusted in CG trials compared to CC trials ($OR = 55.14$, $SE = 7.29$, $z = 30.33$, $p < .001$) and GC trials ($OR = 62.52$, $SE = 9.18$, $z = 28.16$, $p < .001$). In contrast, the likelihood of relative prominence adjustment did not differ significantly between CC and GC trials across groups ($p > .05$). Predicted values are illustrated in Figure 8.

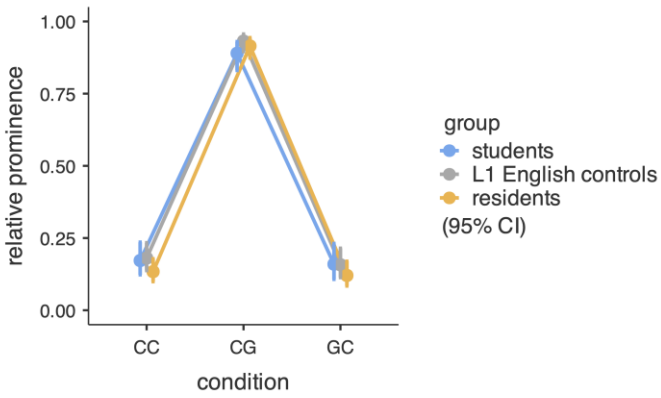

Figure 8. Predicted probability of relative prominence (1 = judged as adjusted, 0 = judged as not adjusted) in the contrastive/contrastive (CC), contrastive/given (CG), and given/contrastive (GC) conditions of the English picture-naming task for university students of Italian in the UK, long-term residents in Italy, and L1 (English) controls. Error bars = 95% confidence interval (CI).

The logistic mixed model fitted to predict the judgment of relative prominence adjustment in Italian included group (students, L1-Italian controls, residents), condition (CC, CG, GC), and the group × condition interaction. The results showed a main effect of condition ($p < .001$), no main effect of group ($p = .108$), and a significant interaction between group and condition ($p < .001$). Within-group comparisons revealed that relative prominence was significantly more likely to be judged as having been adjusted in CG trials than in GC trials for students ($OR = 437.83$, $SE = 456.12$, $z = 5.84$, $p < .001$) and residents ($OR = 64.83$, $SE = 66.56$, $z = 4.06$, $p = .002$), but not for L1-Italian controls ($p > .05$). Relative prominence was also significantly more likely to be judged as having been adjusted in CG trials than CC trials for students ($OR = 56.20$, $SE = 20.90$, $z = 10.83$, $p < .001$) and residents ($OR = 42.64$, $SE = 26.33$, $z = 6.08$, $p < .001$), but not for L1-Italian controls ($p > .05$). In contrast, the likelihood of relative prominence adjustment did not differ significantly between CC and GC trials for any group (all ps $> .05$). Between-group comparisons revealed that, in CG trials, students were significantly more likely to adjust prominence than both L1-Italian controls ($OR = 327.09$, $SE = 345.44$, $z = 5.48$, $p < .001$) and residents ($OR = 12.78$, $SE = 9.07$, $z = 3.59$, $p = .012$). L1-Italian controls and residents did not differ significantly in CG trials

($p > .05$). No other between-group comparison within CC or GC trials reached significance (all $ps > .05$). Predicted values are illustrated in Figure 9.

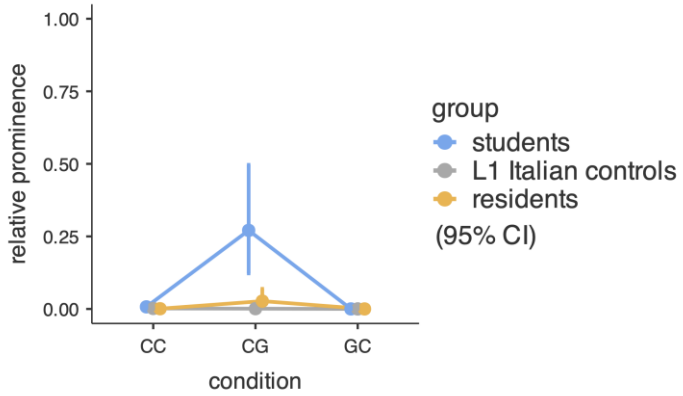

Figure 9. Predicted probability of relative prominence (1 = judged as adjusted, 0 = judged as not adjusted) in the contrastive/contrastive (CC), contrastive/given (CG), and given/contrastive (GC) conditions of the Italian picture-naming task for university students of Italian in the UK, long-term residents in Italy, and L1 (Italian) controls. Error bars = 95% confidence interval (CI).

3.3.2. Exploratory Analyses for Cross-Item Prominence Adjustment

As can be seen from Table 5, the percentages of prominence being judged as adjusted on CC and GC English trials were unexpectedly higher than those on the same trials in Italian, for all groups alike. This is due to an unforeseen complication that attests to the very strong tendency to deaccent repeated material in English; we refer to this complication as "cross-item adjustment". It sometimes happened that the target NP (the description of the second image) in one test item was prosodically adjusted because of a relation to the second image in the preceding test item. For example, in a sequence of two test items [*red dragon, green flower*] and [*blue pumpkin, yellow flower*], the target NP *yellow flower* in the second test item might be prosodically adjusted to reflect the contrast with *green flower* in the previous test item, even though, within the context of its own item, *flower* contrasts with *pumpkin*. Unfortunately, as a consequence of the full randomisation of the experimental items, cases which may trigger cross-item prominence adjustment were not controlled for; in fact, the randomisation meant that different participants had a different number of opportunities in which cross-item prominence adjustment might be triggered. Although this may be regarded as a methodological flaw (as discussed later in Section 4), such a context allowed for further exploration of the results by comparing the proportion of judgments also in the uncontrolled cases that potentially elicited cross-item prominence adjustment across items. The percentages for cross-item prominence adjustment in English are reported in Table 6 below—where CC, CG, and GC refer to the second of the two test items, and where only the second noun of the second test item is relevant for the analysis.

Table 6. Mean percentage of trials where relative prominence was judged as adjusted/not adjusted in cases with potential for cross-item adjustment in the English picture-naming task.

	University Students in UK			**Long-Term Residents in Italy**			**L1 Controls**		
	N	**PA**	**PNA**	***N***	**PA**	**PNA**	***N***	**PA**	**PNA**
CC	109	62%	38%	164	56%	44%	175	65%	35%
CG	56	86%	14%	90	96%	4%	90	92%	8%
GC	55	69%	31%	82	44%	56%	89	61%	39%

Note. *N* = total number of cases with potential for cross-item adjustment per group in each condition. PA = prominence (judged as) adjusted; PNA = prominence (judged as) not adjusted. CC = contrastive/contrastive; CG = contrastive/given; GC = given/contrastive. CC, CG, and GC refer to the second of the two test items, and only the second noun of the second test item is relevant for the analysis.

As can be seen from Table 6, in CG trials (where the adjustment of prominence is expected by definition), prominence was more often judged as having been adjusted than not for all groups. By contrast, interesting differences can be noted across groups in the other two conditions. Of specific interest to the aim of this L1 attrition study is the fact that, on both CC and GC items, the percentage of relative prominence judged as adjusted was reportedly lower for residents than for L1-English controls. On the other hand, similar percentages of prominence judged as being adjusted were reported for students and L1-English controls. Specifically, the percentage of prominence judged as adjusted was 56% for residents, 62% for students, and 65% for L1 controls on CC items with potential for cross-item prominence adjustment. On GC items with the same potential, the percentage was only 44% for residents, but 69% for students and 61% for L1 controls.

4. Discussion

For ease of reference, Table 7 below summarises whether evidence of L1 attrition and L2 acquisition was found across the three domains (i.e., the lexicon, syntax–pragmatics interface, and prosody) in the two bilingual groups examined (i.e., university students of Italian in the UK and long-term residents in Italy).

Table 7. Summary of evidence of L1 attrition and L2 acquisition found in the lexicon, syntax–pragmatics, and prosody for university students of Italian in the UK and long-term residents in Italy.

	Lexicon		**Syntax–Pragmatics Interface**		**Prosody**	
	L1 Attrition	**L2 Acquisition**	**L1 Attrition**	**L2 Acquisition**	**L1 Attrition**	**L2 Acquisition**
Students in UK	✓	✓	✗	✓	✗	✓
Residents in Italy	✓	✓	✓	✓	?	✓

Note. ✓ = some evidence; ✗ = no evidence; ? = supplementary evidence.

4.1. L1 Attrition in Different Bilingual Speakers and Language Domains

The present study asked to what extent L1 attrition would be found in the two groups of late bilingual speakers (RQ1), and whether attrition would affect the lexicon, syntax–pragmatics interface, and prosody within the same individuals (RQ2). Based on the assumption that all bilinguals are (potential) attriters (cf. Schmid & Köpke, 2017a, 2017b; see Section 1.1), we expected attrition in both groups, though to varying extents due to differences in input (cf. see Section 1.3)—with more pronounced attrition effects (e.g., slower processing, lower accuracy, and greater divergence in preferences from monolingual controls) predicted for long-term residents in Italy compared to university students in the UK. Further, we expected certain domains to be more susceptible to attrition: although lexical access was considered vulnerable in both groups, the syntax–pragmatics interface and prosody were more likely to be affected in residents (see Sections 1.2.1–1.2.3).

The data partly confirmed our predictions. In line with our expectations, students exhibited more limited attrition, affecting only lexical access, while residents experienced attrition in all three domains, including the syntax–pragmatics interface and prosody. However, contrary to our expectations, the extent to which each domain was affected by attrition was not always different between groups, as discussed below.

With specific regard to the lexicon, residents exhibited significantly lower accuracy and longer RTs in semantic distractor trials, consistent with our expectations of increased semantic interference linked to L2 immersion (cf. Linck et al., 2009). However, contrary to expectations, students also showed significantly reduced accuracy and longer RTs in semantic distractor trials, rather than in trials targeting lexical interference. While such effects were expected primarily for residents due to L2 immersion and differences in input,

their presence in students as well suggests that even less proficient bilinguals in classroom settings can exhibit early signs of semantic interference, aligning with recent claims about the onset of semantic sensitivity in lower-proficiency learners (Kroll & Ma, 2018; Ma et al., 2017). These patterns are discussed further in Section 4.2 in light of the RHM (Kroll & Stewart, 1994). As a methodological consideration for future studies, since the translation-recognition task involves the concurrent activation of both languages and may thus facilitate the occurrence of attrition effects, it might be useful to investigate lexical attrition through the use of a verbal fluency task, which involves the activation of only one language (see note 5).

Turning to the syntax–pragmatics interface, our expectation, based on Sorace and Filiaci's (2006) Interface Hypothesis, was that this domain would be particularly vulnerable to attrition. However, following Hulk and Müller's (Hulk & Müller, 2000; Müller & Hulk, 2001) account of CLI directionality, no attrition driven by influence from Italian was predicted in the present study on L1-English L2-Italian bilinguals. This is because CLI is expected only when partial overlap exists, from the language that presents only one option (e.g., English) to the one that presents more than one (e.g., Italian), but not vice versa. The results aligned with these expectations. Although all groups showed a similar mean preference for NP1 resolution in ambiguous NP1/2 trials (ranging between 53% and 57%), it was only for residents that RTs in these trials were longer than RTs in both unambiguous conditions. Residents' RTs in post-critical region 7 were also significantly longer than those of students and L1-English controls in the same region, thus pointing to delays in processing affecting residents specifically. These effects suggest that L1 attrition can manifest at the syntax–pragmatics interface due to increased processing difficulty rather than CLI—a point we return to in detail in Section 4.3. Moreover, this evidence seems to suggest that, even for highly proficient/immersed L2 speakers, attrition does not necessarily affect preference alongside RTs but may, in some cases, manifest only in terms of slowdown effects. These results are consistent with the view that attrition may primarily affect access to linguistic knowledge in real-time (cf. Sorace, 2011). However, whether representational changes in the L1 lexicon/grammar may also be at play cannot be determined from the present study and may thus be investigated in future studies through the use of offline measures (e.g., grammaticality judgment tasks). Moreover, given that in this study the pronoun was presented in the subject position of an independent sentence, thus limiting its structural relationship with potential antecedents (NP1 or NP2), it would be valuable to explore whether similar results occur with anaphora ambiguity within the same sentence and with more than two antecedents. Additionally, since in the present design each ambiguous NP1/2 item had only one disambiguated version (unambiguous NP1 or unambiguous NP2), future studies may want to fully manipulate within items to increase statistical power.

Lastly, looking at prosody, residents were expected to be less likely to adjust prominence in L1-English compared to students and L1-English controls in contrastive/given trials, due to potential L2-to-L1 intonational transfer (cf. van Maastricht et al., 2016). The results of the confirmatory analyses did not support this prediction, with all three groups reported to be significantly more likely to be judged to have adjusted prominence in contrastive/given trials than in both other types of trials. However, exploratory analyses examining cross-item prominence adjustment (triggered by repeated nouns across adjacent test items) revealed lower rates of adjustment for residents compared to L1-English controls in both contrastive/contrastive and given/contrastive trials with potential for cross-item prominence adjustment. This seems to be indicative of the fact that some form of L2-to-L1 transfer is taking place for long-term residents, since prominence adjustment is triggered to a lesser extent for them than for L1-English controls across test items. The

reason why some evidence of intonational attrition may be found in these contexts, but not on contrastive/given trials, is possibly due to participants being less aware of the aim of the task across trials than within trials—resulting in more natural prominence production on their behalf. Hence, it seems reasonable to suggest that future studies explicitly control for instances where prominence adjustment may be triggered across test items. Moreover, although using expert judgments in the analysis allowed for a broader examination of prominence adjustment, the phonetic details of prominence remain under-researched to date.

Taken together, these results contribute to the relatively small body of research on attrition in L1-English (e.g., Porte, 1999; Dostert, 2004; Sučková, 2020) by showing that all aspects tested in the present study can indeed undergo attrition in L1-English. Moreover, the findings lend support to the view that attrition emerges along a continuum (Schmid & Köpke, 2017a, 2017b) and that neither full L1 disuse nor extended immersion in an L2 environment is a necessary precondition for L1 changes to arise (cf. Yılmaz & Schmid, 2018). Even instructed bilinguals who remain in an L1-dominant environment may experience subtle processing difficulties or shifts in access to L1 forms as a result of sustained L2 use. However, the degree to which different bilingual speakers experience attrition in different domains is likely influenced by varying factors—such as proficiency, input, dominance, and individual differences (as expanded on in Section 4.3).

4.2. The Relationship Between L1 Attrition and L2 Acquisition

The novel design of this study allows us to assess L1 attrition in relation to L2 acquisition. Following an "integrated view" of L2 acquisition and L1 attrition (Cherciov, 2011), whereby higher degrees of L1 attrition are associated with the increased presence and influence of the L2, we would expect performance in the L1 to be inversely related to performance in the L2. Within each domain, higher L2 accuracy/more similar preferences to Italian controls and faster processing speeds would be likely to correspond to lower L1 accuracy/more dissimilar preferences from English controls and slower processing speeds. In this study, support for this view is mixed and appears to depend on the domain.

Starting with the lexicon, the translation-recognition task does not allow for a direct L1 vs. L2 performance comparison due to the nature of the task engaging both languages. Yet, it still offers a useful window into how bilingual lexical access may be shaped by language proficiency and use. Specifically, both bilingual groups showed accuracy above chance on *yes* trials (78.2% for students, 94.1% for residents), providing evidence of L2 lexical acquisition in both cases, though with a clear performance difference that reflects the groups' differing L2 proficiency levels. Interestingly, on the more critical *no* trials, where interference was expected, both groups exhibited the same type and degree of difficulty: semantic distractor trials significantly reduced accuracy and increased RTs relative to unrelated trials for both groups. As explained in Section 4.1, this was contrary to initial predictions that residents would show semantic interference while students would show lexical interference. Although unexpected, these findings remain compatible with Kroll and Stewart's (1994) Revised Hierarchical Model, which posits that lexical access becomes increasingly conceptually mediated with growing L2 proficiency (see also Kroll & Tokowicz, 2005). Given that the students were in their final year of Italian studies, they were likely beyond the early stages of L2 learning. This interpretation aligns with Kroll and Ma's (2018, p. 30) observation that semantic sensitivity may emerge even in lower-intermediate bilinguals (see also Ma et al., 2017), suggesting that the students may have passed a critical stage of learning and developed conceptual links between L2 words and concepts similar to those of more advanced bilinguals. Future studies could further investigate the developmental trajectory of bilingual lexical access by recruiting learners at

earlier stages of L2 acquisition (e.g., first- or second-year students), whilst also having to pay particular attention to the design of the items employed in the task to avoid confounds due to unfamiliarity with words rather than interference. Additionally, examining lexical performance in the L1 vs. L2 using separate verbal fluency tasks may offer a clearer picture of the relationships between L1 attrition and L2 acquisition in the bilingual lexicon.

With regard to the syntax–pragmatics interface, in L2-Italian, both bilingual groups—and especially students—were predicted to show a preference for NP1 resolution and report longer RTs in ambiguous NP1/2 trials, under the influence of L1-English, in contrast to the expected NP2 preference and shorter RTs of L1-Italian controls. These predictions were only partially confirmed. In terms of preferences, contrary to expectations, no differences were reported on ambiguous trials between groups. Importantly, all groups showed a similar mean preference for NP2 resolution (ranging between 57% and 59%), suggesting that for both bilingual groups, L1-English influence on L2-Italian resolution preferences may be weaker than anticipated. In terms of RTs, and in line with expectations, both bilingual groups showed longer RTs in ambiguous NP1/2 trials compared to unambiguous trials. Interestingly, this difference was not observed in the L1-Italian control group, suggesting that the increased processing cost for ambiguous trials is specific to bilinguals and not simply a result of trial type. Therefore, these findings seem notable: although both bilingual groups displayed target-like preferences in Italian, they required greater processing effort to do so—a pattern consistent with the notion of "residual optionality" in L2 acquisition and increased processing demands under interface conditions (cf. Sorace, 2005). Together with the evidence of L1 attrition in real-time processing found in the case of residents but not students, these findings challenge the view of a simple inverse relationship between L1 attrition and L2 acquisition. Rather than improved L2 processing corresponding to reduced L1 performance, highly proficient bilinguals (like the residents) appear to experience increased processing demands in both languages. This aligns with accounts of bilingual language use that highlight the ongoing cognitive costs of managing two languages, especially at the syntax–pragmatics interface (cf. Sorace, 2011). Thus, the findings point to a bidirectional vulnerability in high-proficiency bilinguals under complex interface conditions, rather than a unidirectional trade-off. Future studies could probe into this further by also investigating whether L2 resolution preferences for null pronouns in Italian, which were not examined here, align with those of Italian controls or are influenced by English preferences.

Lastly, turning to prosody, in L2-Italian, it was expected that university students, due to their lower L2 proficiency, would show greater reliance on L1 prosodic patterns, leading to more frequent prominence adjustment in CG trials compared to both residents and L1-Italian controls. The results partly confirmed this prediction. Both students and residents were more likely to adjust prominence in CG trials than in CC or GC trials, with students showing a stronger effect than residents (e.g., CG vs. GC: OR = 437.83 for students, OR = 64.83 for residents; CG vs. CC: OR = 56.20 for students, OR = 42.64 for residents). In contrast, no significant effects were observed for L1-Italian controls. This suggests that students may still be transferring prominence adjustment patterns from their L1 (English), where deaccentuation is normal, to their L2 (Italian), where it is not. Residents also showed evidence of prominence adjustment to a lesser degree than students, indicating a shift toward more target-like prosodic patterns. These findings are consistent with previous research on the modulating role of L2 proficiency in L1-to-L2 transfer (e.g., Swerts & Zerbian, 2010; Reichle & Birdsong, 2014). However, it is also important to point out that the percentage of CG trials where students were judged as having adjusted prominence was only 35.4%; in other words, in 64.6% of the cases, their prominence was judged as not adjusted. By contrast, the percentages of prominence judged as not adjusted were much lower for long-term residents and L1-Italian controls on CG

trials (i.e., 89.9% and 99.5%, respectively). Still, the fact that for students the percentage was above chance level (i.e., 50%) seems to indicate that they are in the process of acquiring L2 prominence patterns, although they may not be performing as L1 speakers of the language, or as highly proficient/immersed L2 speakers. This is an interesting result, as it seems to be indicative of the fact that L2 prominence patterns can be learned even in classroom settings, where they are not usually taught explicitly. Importantly, these findings are of theoretical significance in relation to the properties of the two languages examined. From the perspective of a transitional stage in the shift from L1-English to L2-Italian, the differences between the two bilingual groups likely reflect differing stages of prosodic development. Long-term residents (i.e., more proficient, immersed L2 speakers) seem to have acquired Italian prominence patterns to a greater extent than university students (i.e., less proficient, classroom-based learners), who rely more heavily on L1 prosodic patterns. Together with the findings of reduced cross-item prominence adjustment in L1-English for residents, but not in the case of students, the results lend tentative support to the idea of an inverse relationship between the two languages—at least for prosody.

The potential relationship between L1 attrition and L2 acquisition, found here in one domain for highly proficient bilingual speakers under different input conditions, warrants further investigation. Understanding the linguistic and cognitive reconfigurations that allow some adult language learners to go further in L2 acquisition may provide valuable insights. This would involve examining L1–L2 relationships within the same individuals, considering individual variation methodologically, and exploring L1–L2 interactions across different linguistic domains. In this regard, very recent research has highlighted a very complex relationship, for instance, between L2 acquisition and L1 grammatical attrition within individual bilingual speakers (Baker, 2024).

4.3. L1 Attrition: CLI and/or Processing Demands?

This study also contributes to ongoing debates surrounding the mechanisms underlying L1 attrition by providing evidence that CLI is not the only pathway through which changes in the L1 can occur.

With specific reference to the syntax–pragmatics interface, as discussed in Section 4.1, some evidence of L1 attrition emerged in participants' RTs, especially among long-term residents. These findings cannot be explained in terms of CLI since, following the assumptions of directionality in Hulk and Müller's (Hulk & Müller, 2000; Müller & Hulk, 2001) account, influence from a null-subject language (Italian) onto a non-null subject language (English) is not generally expected. Instead, the observed effects suggest that L1 attrition may stem from the cognitive load involved in managing competing structures from both languages in real-time, leading to longer processing times. This is more in line with claims by Sorace (2011, 2012, 2016), who emphasises the role of computational complexity and processing demands in the overextension of the scope of overt pronouns for bilinguals of two non-null subject languages (where this would not be expected; cf. Section 1.2.2). Ultimately, this constitutes support for the idea that attrition is not necessarily due to CLI (i.e., the direct influence of the L2 on the L1) and, therefore, that the terms "attrition" and "CLI" are not synonymous: namely, the former indicates a broad phenomenon, whilst the latter indicates only one of its possible causes.

Such processing demands appear particularly pronounced for long-term residents in an L2 environment, raising important questions about the role of input, proficiency, and dominance. While the present study did not include proficiency, input, or dominance measures in the modelling, contextual differences between the groups allow for plausible inferences. Specifically, the residents were more proficient, immersed in their L2 environment, and plausibly more L2-dominant, while students were less proficient,

learning the L2 in an L1 environment, and plausibly more L1-dominant. Prior research has shown that L2 dominance influences the degree of L1 attrition (Martín-Villena, 2023), and differences in both quality and quantity of input (see Section 1.3) are also known to increase the likelihood of attrition. In addition, increased L1 inhibition resulting from L2 immersion (e.g., Linck et al., 2009) may further compound these effects by reducing the availability of cognitive resources for L1 processing. Ultimately, this is particularly likely among bilinguals whose L2 is more active and less easily inhibited, such as the long-term residents tested here (cf. Green, 1986, 1998).

As a final point, it is also worth noting that this study did not account for the role of individual differences in language experience and general cognitive abilities, such as executive functions, which may shape how bilinguals integrate discourse-level information. Prior research in child bilingualism (e.g., Torregrossa et al., 2021) has shown that both language experience and executive functions can affect referential choices. Future research should thus explore whether similar mechanisms operate in adult bilinguals and, if so, how these factors may modulate the extent and nature of attrition effects.

5. Conclusions

This study presented a novel experimental design that allowed the examination of L1 attrition (i.e., selective changes in real-time L1 comprehension and production) alongside L2 acquisition in three language areas (i.e., the lexicon, syntax–pragmatics interface, and prosody) among two types of late bilingual speakers: university students in the UK and long-term residents in Italy.

Across tasks, the results showed that L1 attrition can manifest across multiple domains, though the extent to which different types of bilinguals are affected can vary. On the one hand, in the lexicon, both groups displayed similar types and degrees of semantic interference. At the syntax–pragmatics interface and in prosody, on the other hand, more selective effects were revealed: residents showed attrition in both domains, while students did not. These findings highlight the potential role of a combination of factors (i.e., L2 proficiency, immersion, quality and quantity of input, and individual differences) in shaping the extent to which attrition manifests.

Importantly, this study also sheds light on the complex and domain-specific relationship between L1 attrition and L2 acquisition. At the syntax–pragmatics interface, slower processing was observed in both languages for residents, but only in L2-Italian for students. By contrast, within prosody, an inverse relationship was revealed: residents performed similarly to L1-Italian controls but diverged from L1-English controls in our exploratory analyses. The relationship between the L1 and the L2 thus appears to be complex and potentially inverse at least with regard to some domains—with more research needed in this direction.

Finally, the attrition reported at the syntax–pragmatics interface in the present study is unlikely to be driven by CLI. Rather, it appears to reflect the cognitive demands of navigating competing options in real-time processing for bilinguals. These findings seem to suggest that L1 attrition can represent a natural outcome of bilingualism itself, even in the absence of direct influence from the L2: after all, a bilingual is "not the sum of two monolinguals" (Grosjean, 1989, p. 6).

Author Contributions: Conceptualisation, M.Z., V.C., D.R.L., and A.S.; Methodology, M.Z., V.C., D.R.L., and A.S.; Formal Analysis, M.Z.; Investigation, M.Z., V.C., D.R.L., and A.S.; Writing—Original Draft Preparation, M.Z.; Writing—Review and Editing, M.Z., V.C., D.R.L., and A.S.; Visualisation, M.Z.; Supervision, V.C., D.R.L., and A.S.; Funding Acquisition, M.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Arts and Humanities Research Council (grant number 2287831), the School of Philosophy, Psychology, and Language Sciences at the University of Edinburgh (research support grants), and the School of Education, Language, and Psychology at York St John University (quality-related funding).

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of The University of Edinburgh (protocol code 167-2021/7, accepted on 25 January 2021).

Informed Consent Statement: Informed consent was obtained from all subjects involved in this study.

Data Availability Statement: The original data presented in this study are openly available from the Open Science Framework (OSF) at https://osf.io/bm3tj/ (accessed on 28 August 2025).

Acknowledgments: We would like to extend our sincerest thanks to Judith Kroll, Gretchen Sunderman, Fengyang Ma, Dalia Garcia, Carlos Gussenhoven, Hannah Rhode, Alex Lorson, Patrick Sturt, and Ianthi Maria Tsimpli for their precious help in various phases of this project. We would also like to express our gratitude to all the participants of our study.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of this study, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

Notes

1 In the original PhD work that this article is based on, RQ2 also examined whether the resolution of different structures within the same domain (i.e., pronouns and relative clauses) is affected to the same extent. However, to keep the discussion focused and manageable in scope, this article focuses on pronoun resolution only. Full coverage of relative clause resolution is available in Zingaretti (2022) and forthcoming work.

2 Individual predictions were also formulated and preregistered on the AsPredicted platform for each section of this study. For lexical access: https://aspredicted.org/g2cp-db3b.pdf (accessed on 28 August 2025); for syntactic interfaces: https://aspredicted.org/vs69-2pfh.pdf (accessed on 28 August 2025); for prosody: https://aspredicted.org/wkxw-kbsc.pdf (accessed on 28 August 2025).

3 As explained in Section 2, the participants chosen as English and Italian controls are "monolingual" in the sense that they are not fluent in any language other than their L1. However, they may have studied other languages in school, highlighting an increasing difficulty in finding fully monolingual speakers nowadays—thus supporting recent proposals to replace monolingual controls in second language research (see Rothman et al., 2023).

4 In the original prediction 2 (see note 1 above), all domains were expected to be affected for long-term residents in Italy, given the inclusion of relative clauses, which were expected to undergo attrition (cf. Zingaretti, 2022).

5 We also aimed to investigate accuracy and speed in L1/L2 word production through the use of two verbal fluency tasks (with semantic and phonemic categories, respectively). However, due to technical issues, participants' verbal fluency recordings did not save properly (i.e., recordings either did not save at all, or the initial parts were missing from the saved files) and could thus not be analysed.

6 The higher number of fillers in Italian was to account for twelve additional items with ambiguous *null* pronouns initially included in the language. However, due to the potential resolution bias introduced by the lack of unambiguous null pronoun items, ambiguous null pronoun items were later excluded from the final analysis.

7 Only a subset (n = 92) of the participants completed the PNT.

8 Contrary to the preregistration, separate models were fitted for English and Italian in both the SPRT and PNT, as different participant groups completed each language version (e.g., L1-Italian controls completed the tasks in Italian, while L1-English controls completed them in English).

9 While not part of the preregistered analysis, it was necessary to add the variable region in the analyses in order to detect at what point, critically or post-critically, processing costs would show if they occurred.

10 Contrary to the preregistered analyses, *yes* trials were added to the regression model by removing the independent variable relatedness and including all conditions in one model.

References

Amenta, S., Badan, L., & Brysbaert, M. (2021). LexITA: A quick and reliable assessment tool for Italian L2 receptive vocabulary size. *Applied Linguistics*, *42*(2), 292–314. [CrossRef]

Andersen, R. W. (1982). Determining the linguistic attributes of language attrition. In R. D. Lambert, & B. F. Freed (Eds.), *The loss of language skills* (pp. 83–118). Newbury House.

Baker, L. (2024). *The role of linguistic input in adult grammars: Modelling L1 morphosyntactic attrition* [Ph.D. thesis, University of Southampton]. [CrossRef]

Baroni, M., Bernardini, S., Ferraresi, A., & Zanchetta, E. (2009). The WaCky wide web: A collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation*, *43*(3), 209–226. [CrossRef]

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*, *68*(3), 255–278. [CrossRef] [PubMed]

Baus, C., Costa, A., & Carreiras, M. (2013). On the effects of second language immersion on first language production. *Acta Psychologica*, *142*(3), 402–409. [CrossRef]

Bini, M. (1993). La adquisición del italiano: Más allá de las propiedades sintácticas del parámetro pro-drop. In J. Liceras (Ed.), *La lingüística y el análisis de los sistemas no nativos* (pp. 129–139). Dovehouse.

Bolinger, D. (1972). *Degree words* (Vol. 53). Janua Linguarum. Series Maior. De Gruyter Mouton. [CrossRef]

Bonfieni, M. (2018). *Bilingual continuum: Mutual effects of language and cognition* [Ph.D. thesis, University of Edinburgh]. Available online: https://era.ed.ac.uk/handle/1842/31365 (accessed on 28 August 2025).

Bonfieni, M., Branigan, H. P., Pickering, M. J., & Sorace, A. (2019). Language experience modulates bilingual language control: The effect of proficiency, age of acquisition, and exposure on language switching. *Acta Psychologica*, *193*, 160–170. [CrossRef]

Cenoz, J. (2003). The intercultural style hypothesis: L1 and L2 interaction in requesting behaviour. In V. Cook (Ed.), *Effects of the second language on the first* (pp. 62–80). Multilingual Matters.

Chamorro, G., & Sorace, A. (2019). The interface hypothesis as a framework for studying L1 attrition. In M. S. Schmid, & B. Köpke (Eds.), *The oxford handbook of language attrition* (pp. 24–35). Oxford University Press. [CrossRef]

Chamorro, G., Sorace, A., & Sturt, P. (2016a). What is the source of L1 attrition? The effect of recent L1 re-exposure on Spanish speakers under L1 attrition. *Bilingualism: Language and Cognition*, *19*(3), 520–532. [CrossRef]

Chamorro, G., Sturt, P., & Sorace, A. (2016b). Selectivity in L1 attrition: Differential object marking in Spanish near-native speakers of English. *Journal of Psycholinguistic Research*, *45*(3), 697–715. [CrossRef]

Cherciov, M. (2011). *Between attrition and acquisition: The dynamics between two languages in adult migrants* [Ph.D. thesis, University of Toronto].

Chomsky, N. (2000). Minimalist inquiries. In R. Martin, D. Michaels, & J. Uriagereka (Eds.), *Step by step: Essays on minimalist syntax in honor of howard lasnik* (pp. 89–155). MIT Press.

Cook, V. (Ed.). (2003). *Effects of the second language on the first*. Multilingual Matters. [CrossRef]

Cruttenden, A. (1993). The de-accenting and re-accenting of repeated lexical items. In *Proceedings of the ESCA workshop on prosody* (pp. 16–19). Working Papers 41. Department of Linguistics and Phonetics, University of Lund.

Crystal, D. (2012). *English as a global language* (2nd ed.). Cambridge University Press.

Cuza, A. (2010). On the L1 attrition of the Spanish present tense. *Hispania*, *93*(2), 256–272. [CrossRef]

de Leeuw, E. (2007). Hesitation markers in English, German, and Dutch. *Journal of Germanic Linguistics*, *19*(02), 85–114. [CrossRef]

de Leeuw, E. (2017). How phonetics and phonology inform L1 attrition (narrowly defined) research. *Linguistic Approaches to Bilingualism*, *7*(6), 725–729. [CrossRef]

de Leeuw, E., Tusha, A., & Schmid, M. S. (2018). Individual phonological attrition in Albanian–English late bilinguals. *Bilingualism: Language and Cognition*, *21*(2), 278–295. [CrossRef]

De Prada Pérez, A. (2009). *Subject expression in Minorcan Spanish: Consequences of contact with Catalan* [Ph.D. thesis, The Pennsylvania State University].

Długosz, K. (2021). L2 Effects on L1 in foreign language learners: An exploratory study on object pronouns and verb placement in WH-questions in Polish. *Prace Językoznawcze*, *23*(4), 153–172. [CrossRef]

Domínguez, L., & Hicks, G. (2016). Synchronic change in a multidialectal community: Evidence from Spanish null and postverbal subjects. In A. Cuza, L. Czerwionka, & D. Olson (Eds.), *Inquiries in Hispanic linguistics: From theory to empirical evidence* (pp. 53–72). John Benjamins. [CrossRef]

Dostert, S. (2004). Sometimes I feel as if there's a big hole in my head where English used to be!: Attrition of L1 English. *International Journal of Bilingualism*, *8*(3), 189–209. [CrossRef]

Duñabeitia, J. A., Baciero, A., Antoniou, K., Antoniou, M., Ataman, E., Baus, C., Ben-Shachar, M., Çağlar, O. C., Chromý, J., Comesaña, M., Filip, M., Đurđević, D. F., Dowens, M. G., Hatzidaki, A., Januška, J., Jusoh, Z., Kanj, R., Kim, S. Y., Kırkıcı, B., . . . Pliatsikas, C. (2022). The multilingual picture database. *Scientific Data*, *9*(1), 431. [CrossRef] [PubMed]

Gallucci, M. (2019). *GAMLj Suite for jamovi*. Available online: https://gamlj.github.io (accessed on 30 October 2022).

Gargiulo, C. (2020). *On L1 attrition and prosody in pronominal anaphora resolution* [Ph.D. thesis, Lund University]. Available online: https://www.ht.lu.se/en/series/9855440/ (accessed on 28 August 2025).

Green, D. W. (1986). Control, activation, and resource: A framework and a model for the control of Speech in bilinguals. *Brain and Language*, *27*(2), 210–223. [CrossRef]

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67–81. [CrossRef]

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*(1), 3–15. [CrossRef]

Guido Mendes, C., & Iribarren, C. (2007). Fixação do parâmetro do sujeito nulo na aquisição do português europeu por hispanofalantes. In *XXII Encontro nacional da associação portuguesa de linguística: Textos seleccionados* (pp. 483–497). Associação Portuguesa de Linguística.

Gussenhoven, C. (2004). *The phonology of tone and intonation* (1st ed). Cambridge University Press. [CrossRef]

Gürel, A. (2004). Selectivity in L2-induced L1 attrition: A psycholinguistic account. *Journal of Neurolinguistics*, *17*(1), 53–78. [CrossRef]

Gürel, A. (2007). (Psycho)linguistic Determinants of L1 Attrition. In B. Köpke, M. S. Schmid, M. Keijzer, & S. C. Dostert (Eds.), *Language attrition: Theoretical perspectives* (pp. 99–121). John Benjamins. Available online: https://benjamins.com/catalog/sibil.33.08gur (accessed on 28 August 2025).

Hicks, G., & Domínguez, L. (2020). A model for L1 grammatical attrition. *Second Language Research*, *36*(2), 143–165. [CrossRef]

Hirschberg, J. (1993). Pitch accent in context: Predicting intonational prominence from text. *Artificial Intelligence*, *63*(1–2), 305–340. [CrossRef]

Horne, M. (1990). *Accentual patterning in 'new' vs 'given' subjects in English* (Vol. 36, pp. 81–97). Working papers. Lund University, Department of Linguistics and Phonetics.

Hulk, A., & Müller, N. (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition*, *3*(3), 227–244. [CrossRef]

Iverson, M. B. (2012). *Advanced language attrition of Spanish in contact with Brazilian Portuguese* [Ph.D. thesis, University of Iowa]. Available online: http://rgdoi.net/10.13140/RG.2.2.27341.84962 (accessed on 28 August 2025).

Jarvis, S. (2019). Lexical Attrition. In M. S. Schmid, & B. Köpke (Eds.), *The oxford handbook of language attrition* (pp. 240–250). Oxford University Press. [CrossRef]

Jenkins, J., Baker, W., & Dewey, M. (2018). *The Routledge handbook of English as a lingua franca*. Routledge. [CrossRef]

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*(2), 228–238. [CrossRef]

Kaltsa, M., Tsimpli, I. M., & Rothman, J. (2015). Exploring the source of differences and similarities in L1 attrition and heritage speaker competence: Evidence from pronominal resolution. *Lingua*, *164*(B), 266–288. [CrossRef]

Kecskes, I. (1998). The state of L1 knowledge in foreign language learners. *WORD*, *49*(3), 321–340. [CrossRef]

Kecskes, I., & Papp, T. (2000). *Foreign language and mother tongue*. Lawrence Erlbaum Associates.

Kecskes, I., & Papp, T. (2003). How to demonstrate the conceptual effect of the L2 on L1? Methods and techniques. In V. Cook (Ed.), *The effect of the second language on the first* (pp. 247–267). Multilingual Matters.

Kim, S. H. O., & Starks, D. (2008). The role of emotions in L1 attrition: The case of Korean-English late bilinguals in New Zealand. *International Journal of Bilingualism*, *12*(4), 303–319. [CrossRef]

Köpke, B. (2002). Activation thresholds and non-pathological first language attrition. In F. Fabbro (Ed.), *Advances in the neurolinguistics of bilingualism* (pp. 119–142). Forum.

Kroll, J. F., & Ma, F. (2018). The bilingual lexicon. In E. M. Fernández, & H. S. Cairns (Eds.), *The handbook of psycholinguistics* (pp. 294–319). Wiley Blackwell.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*(2), 149–174. [CrossRef]

Kroll, J. F., & Tokowicz, N. (2005). Models of bilingual representation and processing: Looking back and to the future. In J. F. Kroll, & A. M. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 531–553). Oxford University Press.

Kupisch, T., Bayram, F., & Rothman, J. (2017). Terminology matters II: Early bilinguals show cross-linguistic influence but are not attriters. *Linguistic Approaches to Bilingualism*, *7*(6), 719–724. [CrossRef]

Ladd, D. R. (1980). *The structure of intonational meaning: Evidence from English*. Indiana University Press.

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge University Press. [CrossRef]

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, *10*(8), 707–710.

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychological Science*, *20*(12), 1507–1515. [CrossRef]

Lozano, C. (2006). Focus and split-intransitivity: The acquisition of word order alternations in non-native Spanish. *Second Language Research*, *22*(2), 145–187. [CrossRef]

Lozano, C. (2009). Selective deficits at the syntax-discourse interface: Evidence from the CEDEL2 corpus. In Y. Leung, N. Snape, & M. Sharwood-Smith (Eds.), *Representational deficits in second language acquisition* (pp. 127–166). John Benjamins. Available online: http://hdl.handle.net/10481/22165 (accessed on 28 August 2025).

Ma, F., Chen, P., Guo, T., & Kroll, J. F. (2017). When late second language learners access the meaning of L2 words: Using ERPs to investigate the role of the L1 translation equivalent. *Journal of Neurolinguistics*, *41*, 50–69. [CrossRef]

Margaza, P., & Bel, A. (2006). Null subjects at the syntax-pragmatics interface: Evidence from Spanish interlanguage of Greek speakers. In M. G. O'Brien, C. Shea, & J. Archibald (Eds.), *Proceedings of the 8th generative approaches to second language acquisition conference (GASLA 2006)* (pp. 88–97). Cascadilla Proceedings Project.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. [CrossRef]

Marinis, T. (2010). Using on-line processing methods in language acquisition research. In E. Blom, & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 139–162). John Benjamins Publishing Company.

Martín-Villena, F. (2023). *L1 Morphosyntactic attrition at the early stages: Evidence from production, interpretation, and processing of subject referring expressions in L1 Spanish-L2 English instructed and immersed bilinguals* [Ph.D. thesis, Universidad de Granada]. Available online: https://hdl.handle.net/10481/81920 (accessed on 28 August 2025).

Martín-Villena, F., & Lozano, C. (2020). Anaphora resolution in topic continuity: Evidence from L1 English–L2 Spanish data in the CEDEL2 corpus. In J. Ryan, & P. Crosthwaite (Eds.), *Referring in a second language: Studies on reference to person in a multilingual world* (pp. 119–141). Routledge. [CrossRef]

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics*, *32*(4), 543–563. [CrossRef]

Michnowicz, J. (2015). Maya-spanish contact in yucatan, mexico: Context and sociolinguistic implications. In S. Sessarego, & M. G. Rivera (Eds.), *New perspectives on Hispanic contact linguistics in the Americas* (pp. 21–42). Iberoamericana/Vervuert.

Montrul, S., & Yoon, J. (2019). Morphology and language attrition. In *Oxford research encyclopedia of linguistics*. Oxford University Press. [CrossRef]

Morrill, T. H., Dilley, L. C., & McAuley, J. D. (2014). Prosodic patterning in distal speech context: Effects of list intonation and F_0 downtrend on perception of proximal prosodic structure. *Journal of Phonetics*, *46*, 68–85. [CrossRef]

Müller, N., & Hulk, A. (2001). Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition*, *4*(1), 1–21. [CrossRef]

Nooteboom, S. G., & Terken, J. M. B. (1982). What makes speakers omit pitch accents? An experiment. *Phonetica*, *39*(4–5), 317–336. [CrossRef]

Paradis, M. (1993). Linguistic, psycholinguistic, and neurolinguistic aspects of 'interference' in bilingual speakers: The activation threshold hypothesis. *International Journal of Psycholinguistics*, *9*, 133–145.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. John Benjamins.

Paradis, M. (2007). L1 attrition features predicted by a neurolinguistic theory of bilingualism. In B. Köpke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Studies in bilingualism* (pp. 121–133). John Benjamins Publishing Company.

Payne, B. R., Grison, S., Gao, X., Christianson, K., Morrow, D. G., & Stine-Morrow, E. A. L. (2014). Aging and individual differences in binding during sentence understanding: Evidence from temporary and global syntactic attachment ambiguities. *Cognition*, *130*, 157–173. [CrossRef]

Porte, G. (1999). English as a forgotten language: The perceived effects of language attrition. *ELT Journal*, *53*(1), 28–36. [CrossRef]

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510–532. [CrossRef] [PubMed]

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Available online: https://www.R-project.org (accessed on 28 August 2025).

Reichle, R. V., & Birdsong, D. (2014). Processing focus structure L1 and L2 French: L2 proficiency effects on ERPs. *Studies in Second Language Acquisition*, *36*(3), 535–564. [CrossRef]

Requena, P. E., & Berry, G. M. (2021). Cross-linguistic influence in L1 processing of morphosyntactic variation: Evidence from L2 learners. *Applied Psycholinguistics*, *42*(1), 153–180. [CrossRef]

Rezlescu, C., Danaila, I., Miron, A., & Amariei, C. (2020). More time for science: Using testable to create and share behavioral experiments faster, recruit better participants, and engage students in hands-on research. *Progress in Brain Research*, *253*, 243–262. [CrossRef]

Rodríguez-Ordóñez, I., & Sainzmaza-Lecanda, L. (2018). Bilingualism effects in Basque subject pronoun expression: Evidence from L2 basque. *Linguistic Approaches to Bilingualism*, *8*(5), 523–560. [CrossRef]

Rothman, J., Bayram, F., DeLuca, V., Di Pisa, G., Duñabeitia, J. A., Gharibi, K., Hao, J., Kolb, N., Kubota, M., Kupisch, T., & Laméris, T. (2023). Monolingual comparative normativity in bilingualism research is out of 'control': Arguments and alternatives. *Applied Psycholinguistics*, *44*(3), 316–329. [CrossRef]

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. [CrossRef]

Schmid, M. S. (2016). First language attrition. *Language Teaching*, *49*(2), 186–212. [CrossRef]
Schmid, M. S. (2019). The impact of frequency of use and length of residence on L1 attrition. In M. S. Schmid, & B. Köpke (Eds.), *The oxford handbook of language attrition*. Oxford University Press. [CrossRef]
Schmid, M. S., & Cherciov, M. (2019). Introduction to extralinguistic factors in language attrition. In M. S. Schmid, & B. Köpke (Eds.), *The oxford handbook of language attrition* (pp. 505–522). Oxford University Press. [CrossRef]
Schmid, M. S., & Jarvis, S. (2014). Lexical access and lexical diversity in first language attrition. *Bilingualism: Language and Cognition*, *17*(4), 729–748. [CrossRef]
Schmid, M. S., & Köpke, B. (2009). L1 attrition and the mental lexicon. In A. Pavlenko (Ed.), *The bilingual mental lexicon* (pp. 209–238). Multilingual Matters. [CrossRef]
Schmid, M. S., & Köpke, B. (2017a). The relevance of first language attrition to theories of bilingual development. *Linguistic Approaches to Bilingualism*, *7*(6), 637–642. [CrossRef]
Schmid, M. S., & Köpke, B. (2017b). When is a bilingual an attriter? Response to the commentaries. *Linguistic Approaches to Bilingualism*, *7*(6), 721–755. [CrossRef]
Schmid, M. S., & Köpke, B. (Eds.). (2019). *The oxford handbook of language attrition* (1st ed.). Oxford University Press. [CrossRef]
Seliger, H. W., & Vago, R. M. (Eds.). (1991). *First language attrition*. Cambridge University Press. [CrossRef]
Sorace, A. (2005). Selective optionality in language development. In L. M. E. A. Cornips, & K. P. Corrigan (Eds.), *Current issues in linguistic theory* (Vol. 265, pp. 55–80). John Benjamins Publishing Company. [CrossRef]
Sorace, A. (2011). Pinning down the concept of 'interface' in bilingualism. *Linguistic Approaches to Bilingualism*, *1*(1), 1–33. [CrossRef]
Sorace, A. (2012). Pinning down the concept of interface in bilingual development: A reply to peer commentaries. *Linguistic Approaches to Bilingualism*, *2*(2), 209–217. [CrossRef]
Sorace, A. (2016). Referring expressions and executive functions in bilingualism. *Linguistic Approaches to Bilingualism*, *6*(5), 669–684. [CrossRef]
Sorace, A. (2020). L1 attrition in a wider perspective. *Second Language Research*, *36*(2), 203–206. [CrossRef]
Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, *22*(3), 339–368. [CrossRef]
Sorace, A., Serratrice, L., Filiaci, F., & Baldo, M. (2009). Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua*, *119*(3), 460–477. [CrossRef]
Sučková, M. (2020). Acquisition of a foreign accent by native speakers of English living in the Czech Republic. *ELOPE: English Language Overseas Perspectives and Enquiries*, *17*(1), 71–85. [CrossRef]
Sunderman, G. (2013). Translation recognition tasks. In J. Jegerski, & B. VanPatten (Eds.), *Research methods in second language psycholinguistics* (pp. 185–211). Routledge.
Sunderman, G., & Kroll, J. F. (2006). First language activation during second language lexical processing: An Investigation of lexical form, meaning, and grammatical class. *Studies in Second Language Acquisition*, *28*(3), 387–422. [CrossRef]
Swerts, M. (2007). Contrast and accent in Dutch and Romanian. *Journal of Phonetics*, *35*(3), 380–397. [CrossRef]
Swerts, M., Krahmer, E., & Avesani, C. (2002). Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of Phonetics*, *30*(4), 629–654. [CrossRef]
Swerts, M., & Zerbian, S. (2010). Intonational differences between L1 and L2 English in South Africa. *Phonetica*, *67*(3), 127–146. [CrossRef]
Talamas, A., Kroll, J. F., & Dufour, R. (1999). From form to meaning: Stages in the acquisition of second-language vocabulary. *Bilingualism: Language and Cognition*, *2*(1), 45–58. [CrossRef]
The jamovi project. (2022). *jamovi (version 2.3)*. Available online: https://www.jamovi.org (accessed on 30 October 2022).
Torregrossa, J., Andreou, M., Bongartz, C., & Tsimpli, I. M. (2021). Bilingual acquisition of reference: The role of language experience, executive functions and cross-linguistic effects. *Bilingualism: Language and Cognition*, *24*(4), 694–706. [CrossRef]
Tsimpli, I. M., & Sorace, A. (2006). Differentiating Interfaces: L2 performance in syntax-semantics and syntax-discourse phenomena. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *Proceedings of Boston University Conference on Language Development 30* (pp. 653–664). Cascadilla Press.
Tsimpli, I. M., Sorace, A., Heycock, C., & Filiaci, F. (2004). First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism*, *8*(3), 257–277. [CrossRef]
Van Dijk, C., Van Wonderen, E., Koutamanis, E., Jan Kootstra, G., Dijkstra, T., & Unsworth, S. (2022). Cross-linguistic influence in simultaneous and early sequential bilingual children: A meta-analysis. *Journal of Child Language*, *49*(5), 897–929. [CrossRef] [PubMed]
van Maastricht, L., Krahmer, E., & Swerts, M. (2016). Prominence patterns in a second language: Intonational transfer from Dutch to Spanish and vice versa. *Language Learning*, *66*(1), 124–158. [CrossRef]
Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *The Quarterly Journal of Experimental Psychology*, *49*(1), 51–79. [CrossRef]

Weinreich, U. (1953). *Languages in contact: Findings and problems*. Linguistic Circle of New York.

Weltens, B., & Grendel, M. (1993). Attrition of vocabulary knowledge. In R. Schreuder, & B. Weltens (Eds.), *The Bilingual lexicon* (pp. 135–156). John Benjamins Publishing Company. [CrossRef]

Yılmaz, G., & Schmid, M. S. (2018). First language attrition and bilingualism. In D. Miller, F. Bayram, J. Rothman, & L. Serratrice (Eds.), *Bilingual cognition and language: The state of the science across its subfields* (pp. 225–250). John Benjamins.

Zingaretti, M. (2022). *First language attrition in late bilingualism: Lexical, syntactic and prosodic changes in English-Italian bilinguals* [Ph.D. thesis, University of Edinburgh]. [CrossRef]

Zingaretti, M., Garraffa, M., & Sorace, A. (2024). The impact of bilingualism on hate speech perception and slur appropriation: An initial study of Italian UK residents. In S. Cruschina, & C. Gianollo (Eds.), *An investigation of hate speech: Use, identification, and perception of aggressive language in Italian* (pp. 193–230). Helsinki University Press. [CrossRef]

Zingaretti, M., & Spelorzi, R. (2021). Bridging the gap between native and non-native English-speaking teachers: Insights from bilingualism research. *Spark: Stirling International Journal of Postgraduate Research*, *7*(1), 1–14. Available online: https://spark.stir.ac.uk/wp-content/uploads/2021/10/Zingaretti-and-Spelorzi_final.pdf (accessed on 28 August 2025).