

Mandarin speakers undergoing attrition produce more explicit referring expressions¹

Yajun Liu^{a,1}, Antonella Sorace^{a,2}, Kenny Smith^{a,3}

^a School of Philosophy, Psychology and Language Sciences, University of Edinburgh

¹ yajun.liu@ed.ac.uk; ² a.sorace@ed.ac.uk; ³ kenny.smith@ed.ac.uk

Please address correspondence to:

Yajun Liu
School of Philosophy, Psychology and Language Sciences
University of Edinburgh
Dugald Stewart Building
3 Charles Street
Edinburgh EH8 9AD
yajun.liu@ed.ac.uk

¹ Acknowledgements: Following the policies of the University of Edinburgh, for the purpose of open access we have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

Abstract: Continuous immersion in a second language causes speakers' first language to change, a phenomenon known as L1 attrition. We explored (1) whether bilingual native Mandarin speakers display attrition-related changes in their use of referring expressions in Mandarin after exposure to English, and (2) whether the severity of attrition is affected by the amount of exposure to both Mandarin (L1) and English (L2) and English proficiency. All participants completed a questionnaire to assess their language experience, and a picture description task in spoken Mandarin. The results show that where more-monolingual Mandarin speakers preferred null pronouns, bilingual speakers tended to use overt pronouns, suggesting attrition-related changes in their native language which favoured explicitness. Our study also shows that decreased use of L1 coupled with increased use of L2 and higher L2 proficiency are likely to result in a greater degree of attrition, although such an association is statistically unreliable in some models.

Keywords: L1 attrition, Mandarin Chinese, Referring expressions, Production

1. Introduction

L1 attrition refers to a process where the native language (L1) of bilinguals undergoes gradual changes due to continuous immersion in a second language (L2) environment and reduced use of their L1. Traditionally, attrition has been viewed as a form of language loss or gradual erosion over time (Seliger & Vago, 1991). However, more recent research suggests that attrition may not be a mere loss but rather an adaptive restructuring in response to a changed linguistic environment (Chamorro et al., 2016; Hicks & Domínguez, 2020; Laméris et al., 2024; Sorace, 2019). Under this view, syntactic attrition is understood as change in linguistic preferences and processing strategies, rather than language loss; for instance, at the syntax-pragmatics interface in pronominal use (our focus in this paper), bilingual speakers of pro-drop languages (e.g., Italian, Spanish) undergoing L1 attrition do not lose the ability to use null pronouns but rather develop a stronger preference for overt pronouns in contexts where monolingual speakers would typically omit them (Fernando, 2023; Tsimpli et al., 2004).

Previous research has generally assumed that the effects of attrition emerge gradually over time and, as a result, has primarily focused on bilingual speakers who immigrated after puberty to a country where their L2 is the dominant language and lived there for many years. However, studies on lexical attrition, characterized by e.g. slower retrieval times, reduced vocabulary diversity, and a tendency to use simpler words (Jarvis, 2019; Schmid & Jarvis, 2014), challenge this assumption, showing that even a short period of immersion in an L2 environment can lead to adjustments in L1 lexical retrieval patterns (six months in Baus, et al., 2013; three months in Linck et al., 2009). These studies suggest that L1 attrition-related changes may begin sooner than previously thought. While attrition at the syntax-pragmatics interface appears to emerge more slowly than lexical attrition, it remains unclear how early such changes can begin.

In this paper we explore referent production in L1 Mandarin L2 English speakers. A growing number of Chinese people study, work, and live abroad, and are likely to be experiencing attrition. Despite this, the effects of L1 attrition in Mandarin Chinese have not been adequately explored, specifically regarding referring expressions. Mandarin is a radical pro-drop language that permits both subject and object drop, and is distinct from often closely related Indo-European languages like Spanish and Italian that have been the focus of many studies on attrition. Investigating L1 attrition in Mandarin can reveal whether attrition phenomena reported in the literature are specific to certain languages, or reflects a wider cross-linguistic phenomena. The speakers in our two experimental groups had resided in the UK for up to 12 months and at least 33 months, respectively. Furthermore, our participants' duration of residence in the L2 environment is notably shorter than in previous research (our two experimental groups had resided in the UK for less than 12 months and more than 33 months, respectively) allowing us to observe how early syntactic attrition might begin to emerge.

In the sections that follow we review the literature on reference comprehension and production in monolinguals (section 1.1) and individuals undergoing attrition (section 1.2); since there is no literature on reference in Mandarin Chinese speakers undergoing attrition, in section 1.3 we review the most relevant available evidence, from L2 learners of Mandarin. The general picture from this literature review is that speakers undergoing attrition in their L1 (and L2 learners) have a preference for more explicit forms relative to monolingual speakers. In section 1.4 we review the various theories proposed to account for attrition effects, before turning to our own study of attrition effects in L1 Mandarin L2 English individuals in sections 2-3.

1.1 Reference comprehension and production across languages

This study investigates potential attrition effects at the syntax-discourse/pragmatics interface in Mandarin Chinese, specifically regarding reference. Before addressing L1 attrition at this interface, we will examine the cross-linguistic similarities and differences in reference comprehension and production among Mandarin Chinese, English, Italian, and Spanish.

Languages can be broadly classified as pro-drop and non-pro-drop based on whether they permit the omission of subjects (and, to a limited extent, objects). English, for instance, is a non-pro-drop language that generally requires overt subjects (as shown in 1b). Mandarin, Italian, and Spanish are pro-drop languages that share a similar inventory of pronouns (i.e., null and overt pronouns). However, the licensing conditions and distributions of these forms vary across languages.

Pro-drop languages like Italian and Spanish primarily allow subject omission (as shown in sentences 1c and 1d), while object omission remains generally more restricted and context-dependent. Due to their rich morphological systems, omitted referents can be tracked through verb inflection, which provides information such as person and number. In example (1) c, the Italian verb form *partito* includes a gender marker *o* (M = masculine), indicating that the omitted subject is male. In contrast, Mandarin Chinese allows both subject and object to be omitted flexibly (as shown in 1a) and due to its minimal verbal morphology, it relies more on discourse context for recovering omitted subjects and objects. As such, Mandarin is classified as a radical pro-drop language, distinguishing it from agreement-based pro-drop language like Italian and Spanish.

(1) a. Null arguments in Mandarin

A: Zhangsan kanjian Lisi le ma?

Zhangsan see Lisi LE¹ Q

‘Did Zhangsan see Lisi?’

B: *e* kaijian *e* le.

[He] saw [him]. (Huang, 1984, p. 533)

b. Overt subject in English

Susan gave Betsy a pet hamster.

Susan/She reminded her such hamsters were quite shy. (Gorden et al. 1993, p.313)

c. Null subject in Italian

E’ partito

Is-3s gone-M

‘He left.’ (Tsimpli et al., 2004, p259)

d. Null subject in Spanish

Pedro/*pro* salió del restaurante.

Pedro/*pro* left of+the restaurant.

‘(Peter) left the restaurant.’ (Chamorro, 2018, p2)

Comprehension

From a pragmatic perspective, Accessibility Theory (Ariel, 1990) posits that the appropriateness of a referring expression is associated with the accessibility of the referent under discussion. The accessibility of the referent in the speaker’s mental representation is influenced by multiple factors, with topicality being one of them. In many cases, a discourse topic is typically the grammatical subject of a sentence. As such, an antecedent in the subject/topic position has higher accessibility than an antecedent in a non-subject/non-topic position, and is thus more likely to be referred to with a more reduced expression (e.g., overt or null pronouns). Furthermore, NPs and (in pro-drop languages) overt pronouns can be used to refer to a referent of lower accessibility, signalling topic shift (Chamorro, 2018).

¹ LE: the marker for perfective or inchoative aspect (Huang, 1984).

Generally, the interpretational biases of Mandarin, English, Italian, and Spanish pronouns are compatible with the Accessibility Theory. However, these languages differ in when and how they use pronouns, each following its own syntactic and pragmatic rules. English pronouns often refer to prominent discourse topics, usually in the subject position, indicating a strong subject bias (Garrod & Sanford, 1982; Gordon et al., 1993). Italian null and overt pronouns exhibit a clear division of labour, with null pronouns being interpreted as referring to the subject referent and overt pronouns to the object referent (as shown in 2), known as the Position of Antecedent Strategy (Carminati, 2002). Similar to Italian, Spanish null pronouns are consistently interpreted as referring to the subject antecedent; however, the overt pronoun in Spanish behaves differently from the overt pronoun in Italian, and is roughly equally likely to be interpreted as referring to subject and object antecedents (Filiaci et al., 2014).

- (2). Marta_i scriveva frequentemente a Piera_j quando \emptyset_i /lei_j era negli Stati Uniti.
'Marta_i wrote frequently to Piera_j when \emptyset_i /she_j was in the United States.'

In Mandarin Chinese, both null and overt pronouns are more likely to refer to subject referents, i.e. both display a subject bias similar to English (Yang et al., 1999)². However, recent findings by Zhang and Kwon (2022) reveals a subtle distinction in Mandarin sentences like that shown in (3); while both null and overt pronouns favour subject references, null pronouns exhibit an even stronger subject bias than overt pronouns.

- (3). Li Gang_{male} gei Wang Qiang_{male} da dianhua desh hou, \emptyset /ta_{male} haizai bangongshi.
'When Li Gang_{male} called Wang Qiang_{male}, (he) was in the office.'

² As pointed out in Zhang and Kwon (2022), Figure 3 in the results of Experiment 4 in Yang et al. (1999) shows that in the shift condition, null pronouns elicited longer reaction times than overt pronouns when coreferential with the object referent of the preceding sentence. This suggests that null pronouns are less likely to co-refer with object referents than overt pronouns. However, Yang et al. (1999) do not discuss this finding in detail.

Production

Existing research on reference production has examined speakers of Mandarin, Italian, and Spanish using various experimental paradigms, such as picture description and storytelling. These studies differ in task design and analysis, making direct comparisons challenging. The use of reference follows a general pattern: null pronouns are consistently used to refer to highly prominent antecedent in the subject/topic position, NPs are consistently preferred when referring to referents with a relatively low degree of prominence, such as an antecedent in the non-subject/topic position. The distribution of overt pronouns, however, may vary depending on specific tasks and focus of the study. Some studies, such as Contemori et al. (2023), Montrul (2004), Wu (2020), and Belletti et al. (2007), show that overt pronouns are the least used in their control groups of monolingual speakers of Mandarin, Italian, and Spanish, leading to an alternation of null pronouns and NPs. This seems to contrast with the comprehension preferences reviewed above, where overt pronouns are more likely to be understood as referring to these less prominent antecedents. This asymmetry between comprehension and production in reference suggests that different mechanisms underlie these two processes.

Many studies have explored the underlying mechanisms of reference production. Two of these, conducted in Mandarin Chinese and English, respectively, are particularly relevant to the current research, as the production experiment in this study is adapted from their designs. Hwang (2021) explored the production of referring expressions in written and spoken Mandarin, using methods adapted from similar studies conducted in English (Arnold & Griffin, 2007; Fukumura & van Gompel, 2010). Mandarin Chinese uses distinct written forms for third-person singular pronouns "他 (he)" and "她 (she)" to indicate gender, whereas in spoken Mandarin, they share the same pronunciation and are therefore gender-neutral.

Hwang used a story-continuation task in which participants were asked to read a sentence and then to continue the story. Participants were presented with one-character prompts (example 5a) or two-character prompts (examples 5b-c) where the two characters were of the same (example 5b) or different (example 5c) genders. In the two-character conditions, participants were instructed to continue the story with either the subject or the non-subject antecedent. Participants were instructed to avoid the use of null pronouns as they can lead to ungrammatical sentences in written Mandarin but not in spoken Mandarin.

(5) a. Xiaohong_{female} xiangqu chaoshi.

‘Xiaohong_{female} wants to go to a supermarket.’

b. Xiaohong_{female} xianggen Xiaoli_{female} quchaoshi.

‘Xiaohong_{female} wants to go to a supermarket with Xiaoli_{female}.’

c. Xiaohong_{female} xianggen Xiaogang_{male} quchaoshi.

‘Xiaohong_{female} wants to go to a supermarket with Xiaogang_{male}.’

Consistent with findings for a similar task in English (Arnold & Griffin, 2007), Hwang’s participants used more NPs (i.e., proper names) in the two-character condition compared to the one-character condition. In the written task, speakers produced more pronouns in different-gender contexts than in same-gender contexts, while no significant difference between the two gender contexts in pronoun use was observed in the spoken task. Mandarin speakers’ pronoun usage therefore seems to be influenced by referential ambiguity, indicating a preference for avoiding potentially ambiguous expressions. However, this strategy does not fully explain the notable reduction in pronoun usage (and preference for NPs) observed in the different-gender context of the two-character condition in the written Mandarin task, where pronoun usage is not inherently ambiguous. Arnold and Griffin (2007) proposed that the

decreased usage of pronouns in the two-character condition in English is driven by semantic competition, where two referents compete for attention. This implies that reference production is constrained by speaker-internal cognitive pressures such as attention, as well as partner-directed factors such as potential ambiguity; overall, the process of reference production might involve an interplay of different strategies.

1.2 L1 attrition at the syntax-pragmatics interface

Attrition is not a deterministic process, and its effect may vary across individuals: not all bilinguals exhibit the same level or overt signs of attrition, making it a dynamic process, with individual differences in the degree and rate of attrition (Opitz, 2019). However, consistent patterns emerge at the group level, such as the general preference toward over-explicitness in reference reviewed below.

Notably, syntactic attrition is selective, insofar as certain linguistic structures are more easily affected by attrition than others (Chamorro & Sorace, 2019; Gürel, 2004). The Interface Hypothesis (Sorace & Filiaci, 2006) predicts that structures involving syntax and other cognitive domains, such as pragmatics, are more susceptible to change than those that do not involve such an interface. This susceptibility arises from the need for speakers to have both (1) “knowledge of the structure and of the mapping conditions that operate within interface components” and (2) “the processing principles that apply in the real-time integration of information from different domains” (Sorace, 2011, p.12). In cases where speakers are inefficient in integrating these interface properties, the use of grammar becomes more vulnerable, manifesting as “emerging optionality” in L1 attrition (Sorace, 2011, p.5). Therefore, the effects of attrition stem from reduced efficiency in integrating contextual

information from multiple domains, impacting bilinguals' real-time language processing rather than their underlying representational knowledge.

Effects of attrition on comprehension

The pronominal structure in null-subject languages serves as a testbed for the investigation of bilingual language development, due to its reliance on both syntactic knowledge and pragmatic constraints. The Interface Hypothesis predicts that while null pronouns are expected to remain relatively stable in bilinguals, overt pronouns are likely to undergo change. This is partly because the antecedent preferences of overt pronouns are more flexible than those of null subjects. For instance, adult monolingual speakers in Italian are more likely to use an overt pronoun to refer to the subject referent in sentence 6 (a), where there is no ambiguity, than in sentence 6 (b), where two equally possible antecedents can cause ambiguity (Sorace, 2011). Additionally, as previously discussed, variation among pro-drop languages is primarily limited to the constraints of overt pronouns, whereas null pronouns consistently refer to the subject or topic referents.

6 (a) Paolo ha detto che pro/lui andra al matrimonio si Maria.

‘Paolo has said that \emptyset /he will go to the wedding of Maria.

6 (b) Marta scriveva spesso ad Anna quando pro / lei era in vacanza.

‘Marta wrote frequently to Anna when \emptyset /she was on holiday.’

This prediction is supported by studies exploring the interpretation of null and overt pronouns in potentially attrited speakers. Tsimpli et al. (2004) found that while monolingual Italian speakers tended to coreference the overt pronoun with the object referent in sentences similar to 6(b) (as in Carminati, 2002), their English-proficient Italian speakers (presumably

undergoing attrition) tended to interpret the overt pronoun as referring to the subject referent, suggesting that L1 attrition affects pronoun interpretation, and that attrited speakers have a preference for more explicit forms of reference (i.e. using an overt pronoun for a subject referent is more explicit than the alternative null pronoun).

Chamorro et al. (2016) conducted a similar experiment exploring attrition in Spanish.

The experiment consisted of two tasks, covering both online processing (via eye-tracking) and offline interpretation (via acceptability judgements) of Spanish null and overt pronouns in forward anaphoric sentences similar to Tsimpli et al. (2004). While the offline acceptability judgement task showed no difference between groups, the eye-tracking data revealed, in the attrited group only, a lack of sensitivity to a pronoun mismatch condition where an overt pronoun is coreferential with a subject referent. These findings are consistent with a greater acceptance of explicit reference (i.e. use of overt rather than null pronouns) in speakers undergoing attrition. Chamorro et al. (2016) also found that re-immersion to the L1 can partially reverse the process of L1 attrition, supporting the view that L1 attrition reflects dynamic changes in processing strategies as an adaptation to different linguistic environments rather than permanent language loss or irreversible changes in mental representations.

Effects of attrition on production

The literature on attrition effects in production is relatively limited and presents mixed results. Fernando (2023) explored the distribution of referring expressions in Spanish with a group of monolinguals and two groups of L1 Spanish L2 English bilinguals (advanced instructed bilinguals in Spain, and immersed bilinguals in the UK) on two corpus-based video-retelling tasks. The results show that both groups of bilingual speakers produced more overt pronouns than L1 Spanish monolinguals, and that immersed bilinguals produced

significantly more NPs than both monolinguals and instructed bilinguals in contexts of topic continuity. This again suggests a tendency toward over-explicitness in reference among bilingual speakers. While the Interface Hypothesis predicts pronominal preference in comprehension, it does not account for NP overuse in production. However, its principles may still provide insight into this pattern, as attrition-related processing difficulties could lead to a preference for more explicit referential forms, such as NPs. We will also explore alternative explanations for the overuse of NPs later, such as ambiguity avoidance. In contrast, Giannakou (2018) found no significant differences between Greek monolinguals and L1 Greek-L2 Spanish speakers (potential L1 attriters) in their use of lexical subjects, null subjects, and overt subject pronouns in a storytelling task.

1.3 The syntax-pragmatics interface in Mandarin Chinese in L2 Mandarin speakers

No research to date has looked at L1 attrition effects on pronoun comprehension and production in Mandarin Chinese. The most relevant studies we are aware of are those examining the comprehension and production of referring expressions among L2 learners (Slabakova et al., 2024; Zhao, 2014), child bilinguals (Zhou et al., 2022), and heritage speakers (Jia & Paradis, 2015; Wu, 2020) of Mandarin Chinese. As reviewed below, these studies suggest that referential over-explicitness is not consistently observed in all bilingual populations or linguistic contexts, likely because the use of reference is highly influenced by contextual factors.

Comprehension

To our knowledge, studies on comprehension of Mandarin anaphoric expressions in L2 Mandarin speakers do not use sentence stimuli comparable to those in typical attrition studies (e.g., Tsimpli et al., 2004: two possible referents were mentioned in the preceding context,

causing ambiguity or competition for attention). For instance, Zhao (2014) used forward anaphora sentences, as shown in (6), where only one referent was mentioned in the preceding clause, reducing potential ambiguity in reference. Their L1 English L2 Mandarin learners did not differ from native Mandarin speakers, consistently interpreting null pronouns as referring to subject referents and also allowing overt pronouns to be coreferential with subject referents (albeit with indeterminacy). This is perhaps not surprising given that both null and overt pronouns in Mandarin are strongly subject-biased (similar to English pronouns) (also see footnote five for a discussion on when-clauses in Mandarin Chinese).

(6) Xiao Zhang_i chi fan de shihou, e_i/ta_{i/j} dai zhe yi tiao haokan de xianglian.

‘When Xiao Zhang is eating, he is wearing a pretty necklace.’

Slabakova et al. (2024) investigated the interpretation of null and overt pronouns in Mandarin using resultative constructions, as illustrated in (7). In this structure, the null and overt pronouns have a clear division of labour: the null element refers to the matrix subject, whereas the overt pronoun has to refer to someone else in the discourse. This constraint arises because the coreferential reading of *ta* in (7) is argued to be ruled out by Binding Principle B, which states that a pronoun cannot refer to an antecedent within the same local domain (Huang J., 1992; Huang Y., 1994) They found that L2 learners more frequently corefer overt pronouns with subject referents than native speakers.

(7) Daxiang_i change-de Ø_{i/*j}/ta_{i/j}/Little Monday_j ku qilai le.

‘Big Elephant sang, and as a result he began to cry.’

Production

Studies comparing reference production in Mandarin between L1 Mandarin and L2 Mandarin speakers have presented mixed findings. For instance, Wu (2020) investigated the acquisition of Mandarin pronouns among two groups of L1 English L2 Mandarin learners with low and high Mandarin proficiency and two corresponding groups of heritage speakers in an oral narration task of picture sequences. Production patterns of highly proficient L2 learners and heritage speakers were in line with native speakers, having comparable use of NPs and both overt and null pronouns when maintaining previously mentioned referents. Conversely, L2 learners and heritage speakers of low Mandarin proficiency diverged from native speakers in two ways. L2 learners used a higher percentage of overt pronouns compared to native speakers, while heritage speakers used more NPs than native speakers. This study did not explicitly differentiate the referring expressions used for subject referents and those for non-subject referents; further in-depth analysis may therefore reveal variations in the distribution of the three types of reference.

1.4 The causes of bilingual over-explicitness in reference

The role of language exposure and proficiency

What are the causes of L1 attrition? Researchers often consider factors such as exposure to and use of L1 as key predictors of attrition effects. The Activation Threshold Hypothesis (Paradis, 1993) proposes that each language has an activation threshold and a linguistic element in one language may fall below this threshold if it is not frequently used; attrition occurs when an element in the L1 is not used frequently enough and competes with a corresponding element in the L2 that is used more often. This hypothesis is supported by Chamorro et al. (2016) suggesting an inverse relationship between L1 exposure and the severity of attrition. Other studies have also observed a higher attrition level in participants

experiencing diminishing exposure to their L1 (Bergmann et al., 2016; Flores, 2012; Kasparian et al., 2017; Opitz, 2013; Schmid & Yılmaz, 2018). However, some studies did not find a reliable relationship between L1 use and attrition effects (Jarvis, 2003; Schmid & Jarvis, 2014).

These divergent outcomes may arise from the inherent difficulty in quantifying the amount of language use and exposure (Schmid, 2007). For instance, the decrease in L1 exposure typically aligns with an increase in L2 exposure among bilinguals, thus making it difficult to disentangle these two intertwined processes (Schmid & Yılmaz, 2018). The intricate nature of these challenges pinpoints the importance of considering the dynamic relationship between L1 and L2 exposure in understanding the process of L1 attrition.

Speakers who use L1 less and L2 more tend to achieve higher L2 proficiency, thus making L2 proficiency another key predictor of L1 attrition (Yazawa et al., 2024). Accordingly, early research in L1 attrition often focused on individuals with high L2 proficiency, particularly those at a near-native level, and suggested that attrition is more pronounced in this group (e.g., De Leeuw et al., 2010; Flege, 1987; Mayr et al., 2012; Tsimpli et al., 2004).

Crosslinguistic interference

In studies exploring L1 attrition at the pronominal interface structure, where participants' L2 is a non-null-subject language, such as English, the preference for a more explicit form of reference is often attributed to crosslinguistic interference – the transfer effect of L2 on L1 (Tsimpli et al., 2004). For instance, the extension/overuse of overt pronouns is understood to be influenced by English, which has only one pronominal option in the specific discourse context examined in those studies.

The Attrition via Acquisition model proposed by Hicks and Domínguez (2020), adapted from the L1 acquisition model in Lidz and Gagliardi (2015), conceptualises grammatical attrition as an active restructuring process driven by resolving conflicts between L1 grammar and extensive L2 input and, crucially, successful internalization of new input (intake), rather than a passive loss. For adults with a stable L1 grammar, the inference engine, a mechanism for grammar adjustment based on new input, typically becomes dormant. However, in an environment where L2 is the dominant language, substantial L2 input may create mismatches with the existing L1 grammar in adult speakers, reactivating the inference engine to resolve these mismatches. As a result, grammatical attrition may occur, leading to modifications in the L1 that may cause it to resemble the L2.

However, these accounts may over-predict attrition effects. For instance, the model attributes grammatical attrition to the active resolution of mismatches between existing L1 knowledge and new L2 input; if this were the case, any L1 structures that conflict with L2 could undergo change. However, numerous studies suggest that attrition is selective and not always directly tied to specific L1-L2 structural contrasts (Sorace, 2019).

Additionally, crosslinguistic interference alone may not fully account for attrition effects. For instance, studies on child bilinguals (Sorace et al., 2009) and L2 learners (e.g., Spanish and Italian: Belletti et al., 2007; Margaza & Bel, 2006; Spanish and Greek: Lozano, 2018) of two null-subject languages also reported the same over-extension of overt pronouns. However, interference between two null-subject languages would not be expected to lead to a preference for overt pronouns under these accounts emphasising mismatches between the grammars of L2 and L1.

A preference for redundancy

Lozano (2016, 2018) proposed the Pragmatic Principles Violation Hypothesis, which suggests that the overuse of overt pronouns may be driven by general principles of pragmatic economy, where learners simply prefer being redundant to being ambiguous, often using more explicit forms to avoid ambiguity in referential contexts. This hypothesis is also applied to explain the overuse of NPs observed in L1 Spanish attriters (Fernando, 2023) and L2 learners (Ryan, 2015). While there is little direct evidence supporting this hypothesis in the literature, the contrast between the one-character and two-character conditions in our experiment speaks to this issue, which we return to in the discussion.

2. The current study

Here we report an experiment investigating the use of three referential forms (null and overt pronouns, and NPs) in spoken Mandarin through a picture description task. In order to explore the effects of reduced L1 exposure / increased L2 exposure, we tested three groups of speakers: a control group consisting of more monolingual speakers based in China, and two experimental groups of more-bilingual speakers were resident in the UK. We address two research questions: (1) Do L1 Mandarin-L2 English speakers undergoing attrition in their L1 Mandarin tend to be more explicit in reference, as seen in L2 Mandarin and in attrition in L1 Italian and Spanish? (2) If so, is there a correlation between their attrition-related changes and their usage of L1 and L2 in their daily lives or their L2 proficiency?

2.1 Methods

Participants completed a questionnaire adapted from the Language and Social Background questionnaire (Anderson et al., 2017), to assess their use of and exposure to Chinese and

English, then completed an experiment consisting of a picture description task in spoken Mandarin, adapted from Arnold and Griffin (2007) and Hwang (2021).

Participants

We recruited participants for two experimental groups and a control group. The control group consisted of 31 mainland Chinese PhD researchers who had not travelled abroad, with an age range from 22 to 31 years (Mean: 26.87, SD: 2.23). The first experimental group consisted of 35 UK-based Masters students aged 20 to 29 years (Mean: 23.43, SD: 1.79) who had resided in the UK for 1 to 12 months (Mean 7.31 months, SD: 3.12). This group is referred to as the Short-Term English Exposure group. The second experimental group consisted of 35 UK-based PhD researchers who were either studying or had completed their PhD program in the UK. They ranged in age from 22 to 37 years (Mean: 28.6, SD: 3.34) and had spent between 33 and 131 months in the UK (Mean: 60.57 months, SD: 23.39). This group is referred to as the Long-Term English Exposure group. The three groups did not differ substantially in their age of English acquisition, with most participants beginning to learn English in primary school: the age range was 5-13 years for the Control group (Mean: 8.35, SD: 2.30), and 3-12 for both the Short-Term (Mean: 6.63, SD: 6.63) and Long-Term English Exposure groups (Mean: 6.94, SD: 2.17; see also Section 3.1 and 3.2 for a more detailed analysis of language proficiency and use). Each participant was paid £5 for their participation.

Stimuli

In the picture description task³, we used a sentence structure similar to Hwang (2021)⁴. The task was built using JsPsych (De Leeuw et al., 2023) to elicit spoken production from

³ To access the full list of sentence and image stimuli, please go to the Stimuli folder via the link: https://osf.io/q2ev3/?view_only=d83cee3363ec4c1a91dfb2298dd402be

⁴ We used a sentence structure similar to Hwang (2021) instead of the extensively-examined *when*-clause in L1 attrition studies (e.g., Chamorro et al., 2016; Tsimpli et al., 2004) and in Zhang and Kwon (2022). This is because

participants in response to images, audio descriptions, and text. The images were based on scenarios adapted from those in Zhang and Kwon (2022) and Hwang (2021). We designed images depicting a total of 224 scenarios, comprising 128 scenarios for the critical two-character condition, 64 for the one-character condition and 32 filler scenarios. These scenarios were constructed around four referents, comprised of two female characters (Xiaozi “Little Purple” and Xiaohong “Little Red”) and two male characters (Xiaolan “Little Blue” and Xiaohuang “Little Yellow”). The color of their hair and clothes matches their name, making it easier for participants to remember these character names.

Each scenario consisted of two actions, illustrated in two separate images – the context image and the target image (see Figure 1). The critical condition is the two-character condition, where the context image featured two animate referents of either the same or different genders, and the target image featured one of those two referents (either the subject or non-subject referent from the context image). We constructed critical two-character trials around 16 pairs of verbs (one verb for the context image, one verb for the target image; e.g. the context verb in the scenario in Figure 1 is “greeted”, the target verb is “picked up”); for each of these verb pairs we created 4 same-gender and 4 different-gender combinations, featuring different assignments of characters to the various roles in the context and target events, yielding an inventory of 128 possible two-character trials.

Scenarios in the one-character condition were adapted from the two-character condition, but the context image featured only one referent who reappeared in the target image. Each of the

in the “when” temporal clause in Mandarin, where the pronoun is null, the sentence entails an adverbial clause inserted inside the matrix one; whereas in the case of the overt pronoun, this is the structure in which the adverbial clause precedes the matrix one (Yan, 2022). So, in principle, these two sentences are structurally different, thus naturally leading to different results.

16 verb pairs used for the two-character condition provided four possible one-character scenarios which differed only in the character involved, providing an inventory of 64 items for the one-character condition.

In filler scenarios, two animate referents, either two human characters (e.g., “Xiaozi and Xiaohong”) or one human and one animal character (e.g., “Xiaozi and the little bunny”), are depicted performing actions together in both context and target image, thereby forming a compound subject within a coordinative structure. Target image descriptions on these filler trials featured neither the NPs nor the pronoun forms we were interested in in the critical trials (e.g., a target description in a filler trial might be “Xiaozi and the little bunny took a nap together. *They* felt hungry afterwards.”).

We generated 16 experimental lists using the Latin Square method. Each list has 64 trials, including one variation of each of the 16 verb pairs from both the two-character and one-character conditions, alongside 32 fillers. Each referent appears an equal number of times within each list and across lists. Each variation of the 16 verb pairs in both two-character and one-character conditions was equally distributed across the lists. The presentation order of trials within each list was randomised, starting with two filler trials, and then following the pattern of filler trial – one-character trial – filler trial – two-character trial throughout the task. Additionally, the 64 trials of each experimental list were divided into two blocks, each comprising 32 trials; to avoid potential priming effects between two-character and one-character trials, if a two-character scenario featuring a particular verb pair appeared in the first block, then the corresponding one-character scenario adapted from it would not be presented in the same block.

Continuity Context: Subject



Shift Context: Non-subject



Figure 1: An example of image stimuli used in the picture description task. In this given example, the audio description says “Xiaozi [Little Purple] greeted Xiaohong [Little Red]”. Participants are prompted to “Please repeat what you heard and then complete the story”, and provided with a prompt word, in this example meaning “picked up”. In this scenario, in the subject continuity condition, we expected the participant to say (in Chinese) something like “Xiaozi greeted Xiaohong, then null/she/ Xiaozi [Little Purple] picked up the backpack”; in the shift condition we expected something like “Xiaozi greeted Xiaohong, then null/she/Xiaohong [Little Red] picked up the backpack”.

Procedure

All control participants in mainland China took part in the experiment online via a Zoom or Tencent (a widely used online meeting platform in mainland China) session with the researcher. During the online meeting, participants shared their screens with the researcher while completing the tasks – we adopted this procedure after finding in a pilot experiment that unsupervised remote participation from mainland China resulted in very low-quality production data. UK-based participants were offered the flexibility to take part in the

experiment either online (via Zoom or Teams, following the same screen-sharing procedure) or in person in a lab at the University.

During the experiment, participants first completed the questionnaire and then proceeded to the picture description task. During the picture description task, the context image and an audio description of the image were presented first in the centre of the screen for five seconds. Then, the target image, a prompt word positioned above the image, and a microphone icon below were displayed on the screen. Participants were instructed to repeat what they heard and then proceed to describe the target image using the provided prompt word. Recording of verbal responses was initiated by clicking the microphone icon, and participants could stop recording and move on to the next trial by clicking the icon again.

2.2 Predictions

We aim to address two primary research questions. First, we attempted to probe whether L1 Mandarin-L2 English bilinguals exhibit attrition effects in their L1, particularly concerning the choice of referential forms. As reviewed above, individuals undergoing attrition generally show a preference for the more explicit referential choice (e.g., Tsimpli et al., 2004); we therefore expect our experimental groups to use more explicit referring expressions than our control group. However, exactly how this preference will manifest itself in Mandarin Chinese is unclear a priori. Increased explicitness could be achieved by using overt pronouns rather than null pronouns (as seen in L1 Italian or Spanish speakers undergoing attrition), or full NPs rather than pronouns. Various language-specific factors in Mandarin suggest the latter solution might be preferred. Overt pronouns in spoken Mandarin are gender-neutral, and therefore inherently ambiguous in the two-character condition; indeed they might be more ambiguous in practice than null pronouns, which have a stronger subject bias (Zhang &

Kwon, 2022). Additionally, the Interface Hypothesis prediction regarding overt pronouns is based on pro-drop languages like Italian, where null and overt pronouns exhibit a clear division of labour in the two-character condition. In contrast, both null and overt pronouns in Mandarin Chinese are generally strongly subject-biased. This prediction of more explicit referential choices is clearest for our Long-Term English Exposure group; the potential differences between the control group and the Short-Term English Exposure group remains uncertain due to that group's shorter stay in the UK.

We expected minimal distinctions among the three groups in the one-character condition (as shown in Zhao, 2014). This expectation arises from the fact that there is only one referent in the context, eliminating any ambiguity/competition introduced by an additional character. Additionally, since both overt and null pronouns exhibit a strong subject bias, the use of an overt pronoun is expected to be as effective as the use of a null pronoun.

The second research question concerns the role of exposure to L1 and L2 as well as L2 proficiency in attrition effects. Previous studies indicate that increased exposure to L2, coupled with reduced use of L1, may lead to more severe attrition effects. Additionally, attrition appears to be more pronounced in bilinguals with higher L2 proficiency.

Accordingly, we hypothesized that speakers who report more English exposure (less Mandarin exposure) and/or higher English proficiency would show more severe attrition in their reference production, showing a stronger preference for the more explicit form. This difference should show up in a coarse-grained fashion by comparing across our three groups but can also be assessed in a more fine-grained way by correlating our questionnaire data with participants' explicitness in the production task.

2.3. Data Analysis

We focused on participants' production of three referential forms, namely NPs (which were always proper names in the context of our experiment), overt pronouns and null pronouns, in the two conditions of the picture description task. We analysed the referential forms that speakers used in their first complete sentence when mentioning the target referent. Empty responses and responses containing plural forms (such as "Xiaozi and Xiaohong", or "They") or possessive forms (such as "Her hands") were excluded from data analysis. We also omitted responses in instances where participants altered the order of the context and target images and provided descriptions accordingly, for example, "Xiaohong picked up a backpack and greeted Xiaolan". Responses were also excluded in cases where there was a topic shift before speakers described the target referent in the topic-continuity context of the two-character condition, for example, "Little Red met Little Blue on the campus. Little Blue is Little Red's enemy, so Little Red was particularly unhappy."

Consequently, a total of 1515 trials in the two-character condition and a total of 1551 trials in the one-character condition were analysed using Bayesian ordinal logistic regression, specifically the adjacent category model, with the *brms* package (Bürkner, 2017) in R (R Core Team, 2023). This approach was chosen because the three referential forms produced by speakers can be conceptualized along a continuum of increasing explicitness (null, overt, NP). The adjacent category model allows for a comparison of differences between adjacent categories (i.e. null to overt pronouns, overt pronouns to NPs) across groups. The probability of direction (*pd*) was obtained accordingly using the function *pd()* from the *bayestestR* package (Makowski et al., 2019). For each model, we used very weakly informative priors with mean 0 and standard deviation 1.5 (log-odds) for both the intercept and the other effects (which corresponds to a 95% Credible Interval between -3 and + 3 log-odds, equal to almost

0 to 100% probability). Four MCMC chains of 4000 iterations each were executed and the first 1000 iterations were warmup.

3. Results

We begin this section by summarising participants' questionnaire responses on their language proficiency and use. Next, we present the statistical results on reference production in the picture description task, comparing across groups. Finally, we analyse how participants' language proficiency and use relate to the explicitness of their referential choices.

3.1 Language proficiency and use

Figure 2⁵ illustrates the mean proficiency of English and Mandarin in listening, speaking, reading, and writing across groups, self-reported on a scale of 0-10. Speakers from the Long-Term English Exposure group reported the highest English proficiency in all four skills. All speakers reported the highest proficiency in Mandarin Chinese, with the Long-Term English Exposure and Short-Term English Exposure groups slightly surpassing the control group.

Figure 3 depicts the mean proportion of language use in English and Mandarin in the four skills by group. Figure 4 shows the proportion of English use in 12 specific daily situations by group. We only included 12 contexts out of 20 included in the questionnaire (“at school”, “with roommates”, “with neighbours”, “with friends”, “social events”, “activities”, “shopping”, “reading”, “emails”, “texting”, “on social media”, “watching shows”), because speakers in all three groups reported using over 90% of Chinese in six contexts related to

⁵ We also asked participants about their use of regional varieties of Chinese, including e.g., Cantonese and Wu. 16 out of 35 speakers in the Long-Term English Exposure group, 19 out of 35 in the Short-Term English Exposure group, and 10 out of 31 in the control group indicated a lack of daily use or proficiency in these regional varieties. Speakers in the control group reported higher proficiency in listening and speaking compared to their counterparts in the other groups, and report greater use of them than the other groups, particularly in listening and speaking.

communication with family members at home, and two contexts related to communication with colleagues at work were not applicable to speakers in the Short-Term English Exposure group. As expected, speakers in the Long-Term English Exposure group reported the highest percentage of English use in the four skills and in daily contexts, whereas the Control group reported the lowest English use.

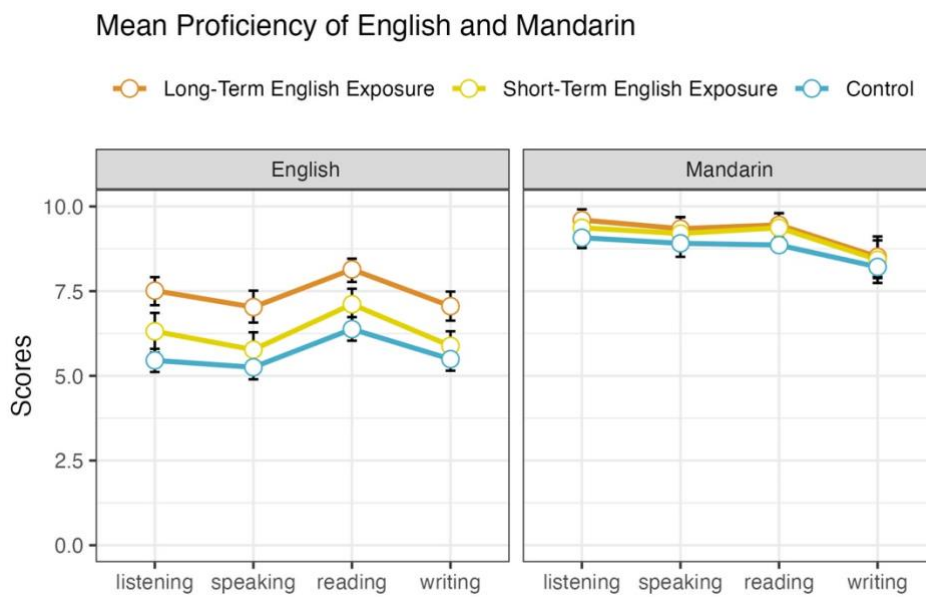


Figure 2: The mean proficiency scores in English and Mandarin across our three groups. Error bars show bootstrapped 95% confidence intervals of the mean.

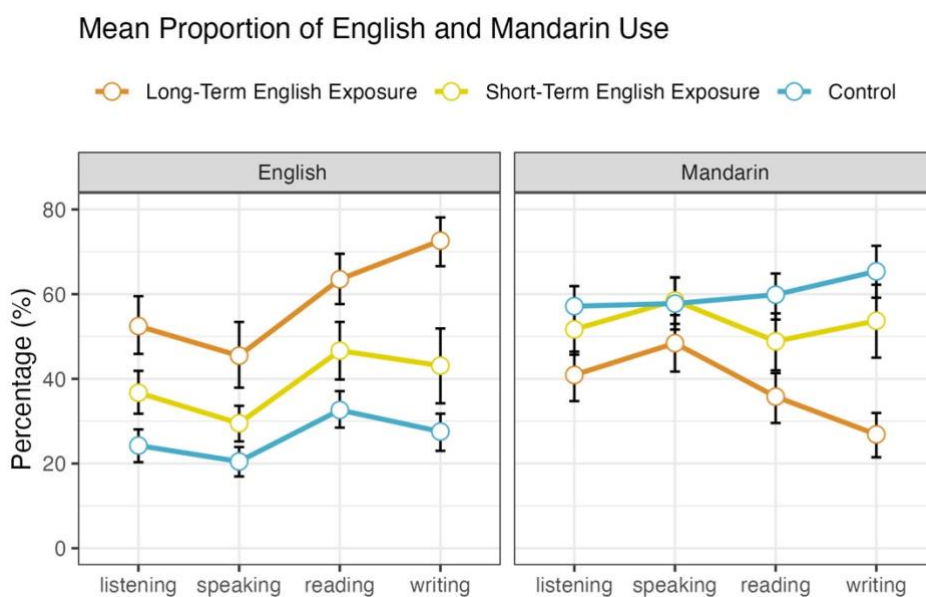


Figure 3: The mean percentage of language use in English and Mandarin in the respective four skills across groups. Plotting conventions as in Figure 2. Use of Chinese dialects are not shown and make up the remaining percentages.

Mean Proportion of English Use in Specific Contexts against Chinese

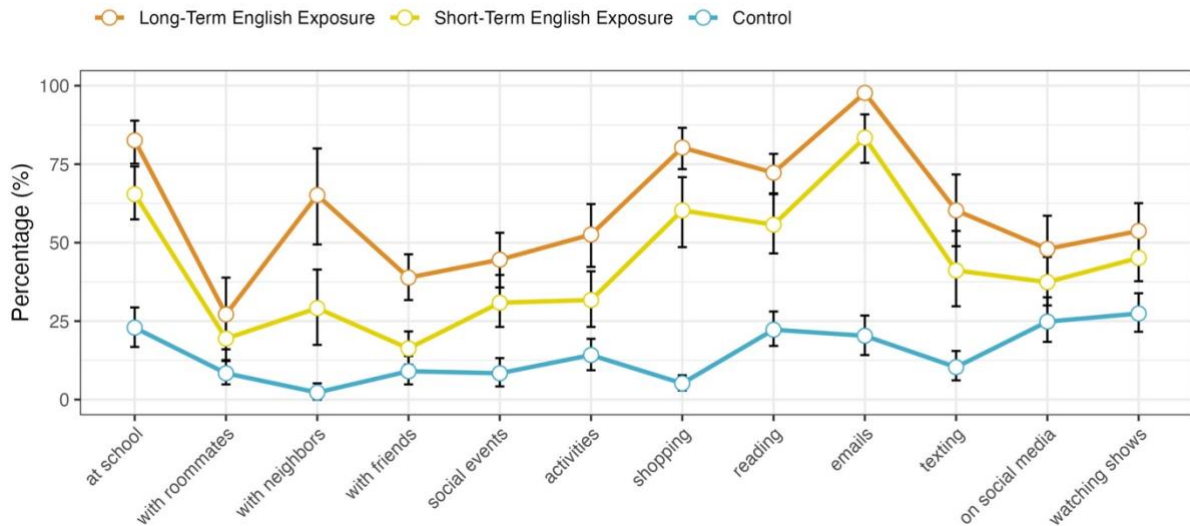


Figure 4: The mean percentage of language use in English in 12 specific daily situations. Plotting conventions as in Figure 2 and 3.

3.2 Reference production

The questionnaire responses reveal the expected differences among the three groups in their self-reported English proficiency and language use. In this subsection, we analyse whether the three groups show different preferences in reference production. Figure 5 and Figure 6 show the production distribution of the three referential forms in the two-character condition and one-character condition, respectively.

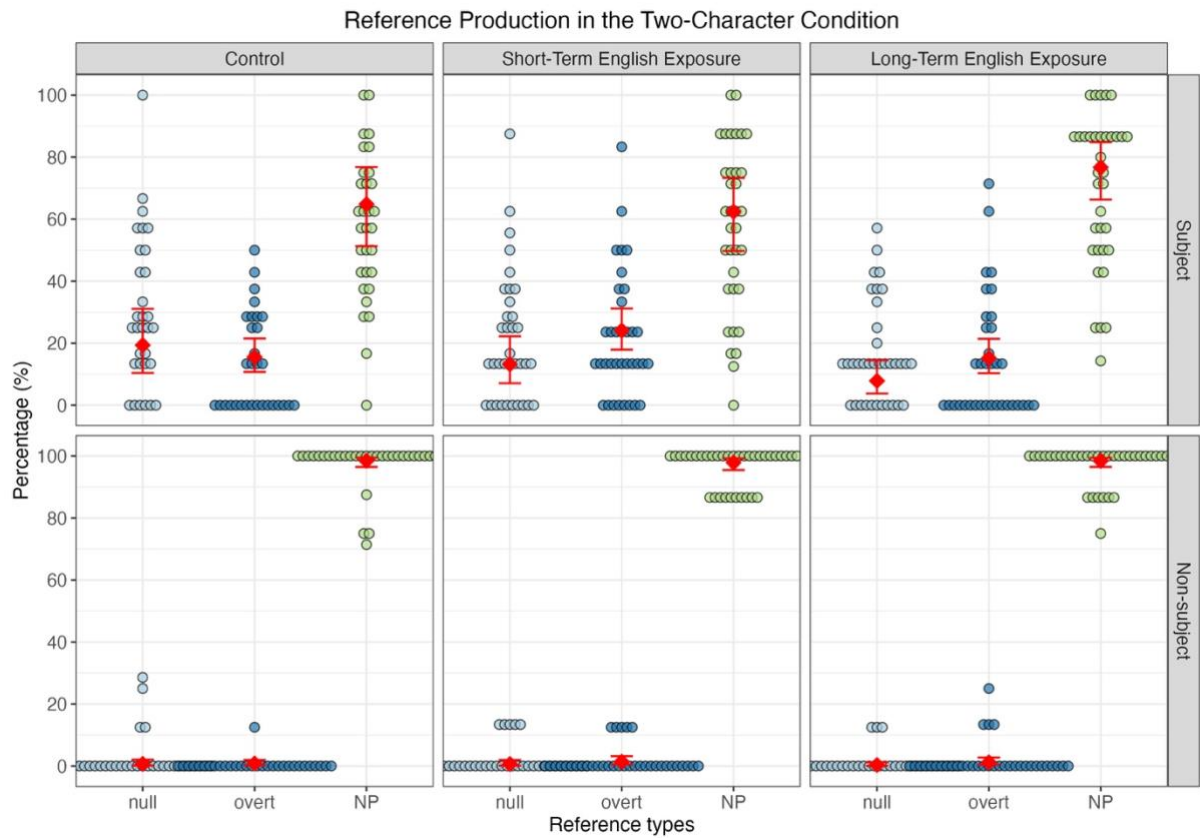


Figure 5: The production of three referential forms in the two-character condition across the three groups, where the target referent is either the subject or non-subject of the previous context. Each dot corresponds to the results of one participant. The diamond shape represents the estimated mean, with error bars showing 95% credible intervals, both derived from the Bayesian model.

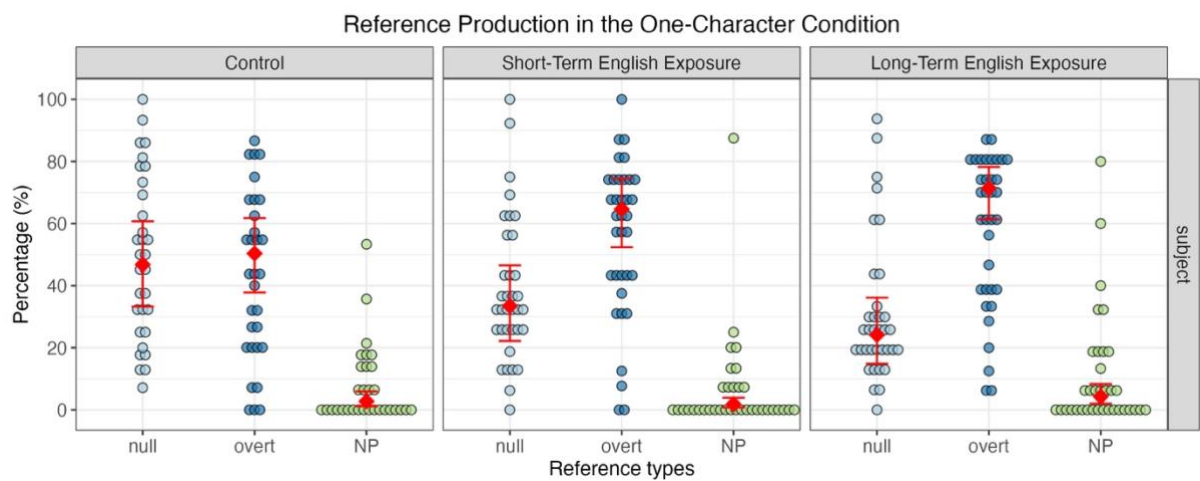


Figure 6: The production data of three referential forms in the one-character condition, where there is only one referent in the context. Plotting conventions as in Figure 5.

Our analysis⁶ of referring expressions for two-character scenes included fixed effects of the referent role (subject or non-subject of the context sentence), group (Control, Short-Term English Exposure, Long-Term English Exposure), and their interaction. The model also included by-participant and by-item random intercepts and slopes for referent role. We used the default treatment contrast for both Role and Group, with the Subject role and the Control group set as reference levels for Role and Group, respectively. This produces two Group fixed effects: one which compares the Short-Term English Exposure group against the reference level Control group, and a second which compares the Long-Term English Exposure group against the Control group. The adjacent category model with category-specific effects allows us to specifically compare group differences across two referential contrasts, i.e. overt versus null pronouns; NPs versus overt pronouns. Table 1 in the Appendix provides the results of the Bayesian model in the two-character condition.

Looking first at the Control group, the analysis reveals no clear preference between null and overt pronouns, for both subject and non-subject referents, as indicated by the wide range of credible intervals that include both negative and positive values (intercept, indicating subject role: $b = 0.21$, $CrI = [-0.36, 0.75]$, $pd = 77\%$; effect of non-subject role: $b = 0.42$, $CrI = [-0.86, 1.79]$, $pd = 73\%$). The control speakers used more NPs than overt pronouns for subject referents ($b = -1.43$, $CrI = [-1.94, -0.93]$, $pd = 100\%$), and this preference for NPs was even stronger for non-subject referents ($b = 3.40$, $CrI = [2.42, 4.48]$, $pd = 100\%$).

In terms of group comparison, both Short-Term and Long-Term English Exposure groups preferred overt pronouns over null pronouns more than the Control group (Short-Term vs

⁶ The full analysis can be accessed in the Acat Analysis folder via this link: https://osf.io/q2ev3/?view_only=d83cee3363ec4c1a91dfb2298dd402be

Control: $b = 0.82$, $\text{CrI} = [0.15, 1.48]$, $\text{pd} = 99\%$; Long-Term vs Control: $b = 0.88$, $\text{CrI} = [0.17, 1.61]$, $\text{pd} = 99\%$); although the wide range of credible intervals suggests some uncertainty about the magnitude of these effects (that is, this effect could be very small or large).

Neither of the bilingual groups showed a stronger preference for NPs over overt pronouns compared to the Control group, as suggested by the credible intervals spanning both negative and positive values, indicating substantial uncertainty regarding the directionality and magnitude of these effects (Short-Term English Exposure relative to Control: $b = -0.48$, $\text{CrI} = [-1.08, 0.13]$, $\text{pd} = 94\%$; Long-Term English Exposure relative to Control: $b = 0.19$, $\text{CrI} = [-0.43, 0.81]$, $\text{pd} = 73\%$). All interaction terms have credible intervals that encompass a wide range of both negative and positive values, suggesting considerable uncertainty about the direction and magnitude of these effects; as can be seen from Figure 5, reference to non-subjects in all 3 groups is dominated by the use of NPs.

We also analysed reference production in the one-character condition. Since this condition features a single animate referent consistently assuming the role of the subject in the context sentence, the model contained only a fixed effect of Group (Control, Short-Term English Exposure, Long-Term English Exposure, coded as before), with by-participant and by-item random intercepts. The results of the Bayesian model in the one-character condition are presented in Table 2 in the Appendix.

The model indicates that speakers in the Control group had no clear preference between null and overt pronouns ($b = -0.07$, $\text{CrI} = [-0.62, 0.47]$, $\text{pd} = 61\%$) but they strongly favoured overt pronouns over NPs ($b = 2.91$, $\text{CrI} = [2.27, 3.57]$, $\text{pd} = 100\%$). The Short-Term English Exposure group showed a tendency to prefer overt pronouns over null pronouns more than

the Control group; however, the directionality and magnitude of the effect remain uncertain, as the credible intervals include both negative and positive values ($b = 0.58$, $\text{CrI} = [-0.10, 1.25]$, $\text{pd} = 96\%$). The Long-Term English Exposure group showed a stronger and clearer preference for overt over null pronouns than the Control group ($b = 1.01$, $\text{CrI} = [0.34, 1.70]$, $\text{pd} = 100\%$), although again the wide credible interval indicates this difference could be quite small or large. No differences were observed between either of the bilingual groups and the Control group in the comparison of NPs and overt pronouns (Short-Term vs Control: $b = -0.66$, $\text{CrI} = [-1.51, 0.15]$, $\text{pd} = 94\%$; Long-Term vs Control: $b = 0.09$, $\text{CrI} = [-0.71, 0.77]$, $\text{pd} = 59\%$).

Separate analyses were conducted to directly compare the Short-Term and Long-Term English Exposure groups (the same statistical model with the Short-Term English Exposure set as the reference level). Results are provided in Table 3 and 4 in Appendix. In the two-character condition, the Long-Term English Exposure group showed a stronger preference for NPs over overt pronouns compared to the Short-Term English Exposure group ($b = 0.63$, $\text{CrI} = [0.05, 1.21]$, $\text{pd} = 98\%$), whereas there was no difference between the groups in their use of overt relative to null pronouns ($b = 0.10$, $\text{CrI} = [-0.56, 1.21]$, $\text{pd} = 61\%$). In the one-character condition, no robust difference emerged for either overt versus null pronouns ($b = 0.38$, $\text{CrI} = [-0.28, 1.03]$, $\text{pd} = 88\%$) or NP versus overt pronouns ($b = 0.59$, $\text{CrI} = [-0.19, 1.36]$, $\text{pd} = 93\%$).

In summary, in both the one-character and two-character conditions, we found that our more-bilingual groups showed a preference for more explicit forms (specifically, overt rather than null pronouns), relative to the control group, consistent with attrition effects shown for Italian

and Spanish. Additionally, the Long-Term English Exposure group used more NPs than the Short-Term English Exposure group in the two-character condition.

3.3 Over-explicitness and L2 proficiency and use

The analyses in the previous section look at the effect of group (Control, Short-Term English Exposure, Long-Term English Exposure) on referential choices. Here we conduct three additional sets of analyses, each using a different continuous predictor from participants' questionnaire responses to predict referential choices: (1) English proficiency, (2) English use in four skills (listening, speaking, reading, and writing), and (3) English use in specific contexts. These analyses address our second question in a fine-grained manner: whether increased L2 (English) proficiency or use correlates with a preference for more explicit forms of reference. For each participant, we computed (1) the average score of English proficiency in the 4 skills of listening, speaking, reading, and writing; (2) the average percentage of English use for each participant in the 4 skills; and (3) the average self-reported percentage of English use in 12 specific contexts. To ensure consistency, we scaled and centred all continuous predictors in the models. Results are summarised in Table 5 to 10 in the Appendix.

English proficiency in the four skills

Recall that in the two-character condition, the group-based analysis found that our bilingual groups had a stronger preference for overt over null pronouns than the less English-proficient Control group, and the Long-Term English Exposure group used more NPs than the Short-Term English Exposure group. In the analysis using self-reported English proficiency, the association between increased English proficiency and the preference for overt pronouns over null pronouns is in the expected direction, but not reliable ($b = 0.20$, $CrI = [-0.11, 0.51]$, $pd =$

90%); for NPs over overt pronouns the directionality of the effect is unclear ($b = -0.06$, $\text{CrI} = [-0.32, 0.20]$, $\text{pd} = 67\%$). In the one-character condition, our group-based analysis again showed that our bilingual groups had a stronger preference for overt over null pronouns relative to the Control group; as expected, increased self-reported English proficiency was positively associated with the preference for overt pronouns over null pronouns ($b = 0.38$, $\text{CrI} = [0.10, 0.68]$, $\text{pd} = 99\%$), although the wide range of credible intervals indicate that this effect might be small. There was no clear effect of English proficiency on the use of NPs over overt pronouns ($b = 0.26$, $\text{CrI} = [-0.09, 0.61]$, $\text{pd} = 92\%$), consistent with the absence of those effects in the group-based analysis.

English use in the four skills

In the two-character condition, the association between increased English use in the four skills and the preference for overt pronouns over null pronouns is in the expected direction, but not reliable ($b = 0.20$, $\text{CrI} = [-0.10, 0.52]$, $\text{pd} = 91\%$); similarly for NPs over overt pronouns ($b = 0.15$, $\text{CrI} = [-0.10, 0.41]$, $\text{pd} = 89\%$). The first interaction term, which compares overt to null pronouns, indicates a potential (albeit weak) positive association between the English use in four skills and the preference of the more explicit form, i.e. overt pronouns, for the non-subject referents ($b = 0.88$, $\text{CrI} = [-0.04, 1.87]$, $\text{pd} = 97\%$), but the direction and magnitude of this effect remain uncertain as the credible intervals include zero. The credible intervals for the second interaction term, comparing NPs to overt pronouns, indicates substantial uncertainty of directionality and magnitude ($b = -0.29$, $\text{CrI} = [-0.87, 0.31]$, $\text{pd} = 83\%$).

In the one-character condition, consistent with the group-based analysis, more English use in the four skills is likely to be positively associated with the use of overt pronouns rather than

null pronouns, but there is some chance this association could be very small or negative ($b = 0.28$, $CrI = [-0.02, 0.57]$, $pd = 97\%$). Increased English use in the four skills does result in a preference for NPs over overt pronouns ($b=0.36$, $CrI = [0.03, 0.71]$, $pd = 98\%$), an effect not seen in the group-based analysis, although this effect could be very small.

English use in the daily contexts

In the two-character condition, this analysis reveals a positive association between increased English use in daily contexts and a preference for overt pronouns over null pronouns ($b = 0.40$, $CrI = [0.11, 0.71]$, $pd = 100\%$), consistent with the group-based analysis. This positive association is not seen in the contrast of NPs versus overt pronouns ($b = 0.05$, $CrI = [-0.19, 0.31]$, $pd = 66\%$). Similarly, in the one-character condition, the increased English use in daily contexts is positively associated with the use of overt rather than null pronouns ($b = 0.46$, $CrI = [0.18, 0.75]$, $pd = 100\%$), but no such association is suggested for the use of NPs over overt pronouns ($b = 0.18$, $CrI = [-0.18, 0.53]$, $pd = 85\%$).

In general, the relationship between continuous measures of English use/proficiency and referential explicitness is more pronounced in the use of overt pronouns over null pronouns in the one-character condition. However, in the two-character condition, these associations are less clear than in the group-based analysis, except for a reliable link between English use in daily contexts and a preference for overt pronouns over null pronouns.

Comparison across the three continuous predictors

We conducted two model comparisons across the three continuous predictors for the two-character and one-character conditions, respectively, to determine which factor is a stronger predictor of referential overexplicitness. We used the `loo_compare` (criterion = "LOO")

function (Vehtari et al., 2024). The results indicate no reliable difference among the three models in either condition, indicating that they perform similarly in predicting overexplicitness. Detailed results are presented in Table 11 in Appendix.

4. Discussion

We examined reference production in L1 Mandarin speakers with varying degrees of exposure to L2 English, in order to probe for attrition effects in their native language. Speakers were asked to complete stories in two discourse conditions, a two-character condition involving two animate referents, and a one-character condition.

Our findings partially align with our predictions. In line with previous studies on reference production by Arnold and Griffin (2007) and Hwang (2021), all speakers, irrespective of their proficiency in their L2, used more NPs than pronouns in the two-character condition. Our more bilingual speakers prefer more explicit forms of reference compared to their more monolingual peers. This indicates potential attrition occurring in their L1, manifesting as changes in referential preferences. Our data also indicates attrition-related change even after a relatively short period of L2 immersion, i.e., no more than 12 months, as seen in the Short-Term English Exposure group. Our findings align with the Interface Hypothesis and its specific predictions for L1 attrition (Sorace, 2011) and the results reported by Fernando (2023) with L1 Spanish L2 English speakers. In the two-character condition we found that both the Short-Term and Long-Term English Exposure groups tended to prefer to use more overt pronouns than Control speakers. Additionally, we observe a preference for NPs over overt pronouns in the Long-Term English Exposure group relative to the Short-Term English Exposure group. This suggests that while the Short-Term English Exposure group made a strong shift toward overt pronouns, the Long-Term English Exposure group did not further

reinforce this tendency; instead, they appeared to rely more on NPs than the Short-Term English Exposure group, possibly to mitigate the potential ambiguity in the two-character condition. In the one-character condition, most participants used pronouns rather than NPs, as expected, but again we see a preference for more explicit forms of reference (more overt than null pronouns) in our Long-Term and Short-Term English Exposure groups relative to the Control group, with this effect being clearest in the Long-Term English Exposure group. While our findings suggest a general preference for the more explicit form of reference among our bilingual groups, the magnitude of this effect remains variable. This variability likely reflects individual differences in language use, suggesting that attrition occurs along a continuum, with general patterns emerging at the group level.

Our experiment also involved self-reports regarding speakers' language proficiency and use in both Mandarin and English. When asking speakers to assess their language exposure, we took into consideration the dynamic change between L1 and L2, reflecting decreased L1 coupled with increased L2. Speakers were also asked to assess their language exposure through two lenses: (1) general language skills encompassing listening, speaking, reading, and writing, and (2) language use in more specific contexts such as communicating with friends, shopping, and participating in social events. This approach aims to capture a more comprehensive picture of speakers' language use across various aspects of their daily routines and interactions. Our results show that using more L2 (English) and less L1 (Mandarin), as well as higher L2 proficiency, are likely to deepen the attrition process, specifically a stronger preference for overt pronouns over null pronouns, particularly pronounced in the one-character condition. This association is consistent with previous studies on the role of language exposure (e.g., Chamorro et al., 2016) and the role of L2 proficiency (e.g., Tsimpli et al., 2004) in the attrition process. In contrast, the two-character condition, while these

associations are broadly in line with the group-based analysis, the direction and strength of the associations are inconsistent across models, with a reliable effect only observed for English use in the four general skills. The extent to which language proficiency and use drives this shift therefore appears to depend on contexts. As suggested by Arnold and Griffin (2007) and Hwang (2021), referential choice in the two-character condition is much more complex and can be influenced by an interplay of multiple factors (e.g., speaker-internal cognitive pressure and partner-directed factors such as ambiguity avoidance), compared to the much simpler, less ambiguous one-character condition.

Why does attrition result in the particular pattern of reference use that we see here? We reviewed two theories from the literature in the introduction: crosslinguistic interference, and a more general preference for redundancy/clarity/ambiguity avoidance.

Our results are hard to reconcile with the latter explanation, because the clearest attrition effects we see (i.e. the strongest preference for more explicit referential forms) tend to be in the one-character condition, where there is little referential ambiguity introduced by using less explicit referential forms. While this does not rule out some preference for redundancy or ambiguity avoidance being the cause of these effects, we think it complicates this account.

Our data do not rule out a crosslinguistic interference account: one possibility is that the pronominal system in Mandarin for our bilingual speakers might simply be influenced by the English pronominal system, potentially resulting in increased usage of overt pronouns over null pronouns throughout. To delve deeper into the impact of cross-linguistic influence in the attrition process, we next plan to study late bilingual speakers whose two languages both permit subject drop, by testing native Mandarin speakers currently residing in Italy. Given

that Italian and Mandarin are pro-drop languages, but with different distributions of null and overt pronouns, we are intrigued to see how Mandarin-Italian speakers make referential choices in those discourse contexts.

5. Conclusion

We used a picture description task in spoken Mandarin to explore referential over-explicitness, a phenomenon often regarded as a sign of L1 attrition, among L1 Mandarin speakers with English as their L2. Speakers' referential choices were examined in both the two-character and one-character conditions. Our bilingual speakers show a preference for more explicit forms in both conditions compared to their more monolingual peers, suggesting attrition-related changes in their native language, even after a relatively short period of L2 immersion (as seen in the Short-Term English Exposure group). Specifically, bilingual speakers tended to use overt pronouns over null pronouns to refer to the subject referent of the preceding context sentence, compared to more-monolingual speakers.

Furthermore, this tendency toward over-explicitness is likely positively associated with language exposure to L1 and L2 and L2 proficiency, with this association being clearer in the one-character condition. However, the underlying mechanisms that motivate bilingual over-explicitness of reference in attrited speakers still require further investigation, in particular to differentiate between cross-linguistic interference and other cognitive mechanisms.

References

- Anderson, J. A. E., Mak, L., Keyvani Chahi, A., & Bialystok, E. (2017). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods*, *50*(1), 250–263.
<https://doi.org/10.3758/s13428-017-0867-9>
- Ariel, M. (1990). *Assessing Noun-Phrase Antecedents (RLE Linguistics B: Grammar)*. Routledge. <https://doi.org/10.4324/9781315857473>
- Arnold, J. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition*, *76*(1), B13–B26.
[https://doi.org/10.1016/S0010-0277\(00\)00073-1](https://doi.org/10.1016/S0010-0277(00)00073-1)
- Arnold, J. E., & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, *56*(4), 521–536.
<https://doi.org/10.1016/j.jml.2006.09.007>
- Baus, C., Costa, A., & Carreiras. (2013). On the effects of second language immersion on first language production. *Acta Psychologica*. *142*(3), 402-409.
<https://doi.org/10.1016/j.actpsy.2013.01.010>
- Belletti, A., Bennati, E., & Sorace, A. (2007). Theoretical and developmental issues in the syntax of subjects: Evidence from near-native Italian. *Natural Language & Linguistic Theory*, *25*(4), 657–689. <https://doi.org/10.1007/s11049-007-9026-9>
- Bergmann, C., Nota, A., Sprenger, S. A., & Schmid, M. S. (2016). L2 immersion causes non-native-like L1 pronunciation in German attriters. *Journal of Phonetics*, *58*, 71–86.
- Bürkner, P.-C. (2017). **brms**: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1). <https://doi.org/10.18637/jss.v080.i01>
- Carminati, M. N. (2002). The processing of Italian subject pronouns. [Doctoral dissertation, University of Massachusetts at Amherst]. ScholarWorks@UMassAmherst.
<https://scholarworks.umass.edu/dissertations/AAI3039345/>
- Chamorro, G. (2018). Offline interpretation of subject pronouns by native speakers of Spanish. *Glossa: A Journal of General Linguistics*, *3*(1). <https://doi.org/10.5334/gjgl.256>
- 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.
<https://doi.org/10.1093/oxfordhb/9780198793595.013.42>
- Chamorro, G., Sorace, A., & Sturt, P. (2016). 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. <https://doi.org/10.1017/S1366728915000152>
- Contemori, C., Tsuboi, N., & Armendariz Galaviz, A. L. (2023). Referential choice in two languages: The role of language dominance. *International Journal of Bilingualism*, 0(0). <https://doi.org/10.1177/13670069231170545>
- De Leeuw, J. R., Gilbert, R. A., & Luchterhandt, B. (2023). jsPsych: Enabling an Open-Source Collaborative Ecosystem of Behavioral Experiments. *Journal of Open Source Software*, 8(85), 5351. <https://doi.org/10.21105/joss.05351>
- De Leeuw, E, Schmid M. S., & Mennen I. (2010). The effects of contact on native language pronunciation in an L2 migrant setting. *Bilingualism: Language and Cognition*, 13(1): 33–40. <https://doi.org/10.1017/S1366728909990289>
- Fernando, M. V. (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. Doctoral dissertation, Universidad De Granada. <https://digibug.ugr.es/handle/10481/81920>
- Filiaci, F., Sorace, A., & Carreiras, M. (2014). Anaphoric biases of null and overt subjects in Italian and Spanish: A cross-linguistic comparison. *Language, Cognition and Neuroscience*, 29(7), 825–843. <https://doi.org/10.1080/01690965.2013.801502>
- Flege, J. E. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15(1), 47–65. [https://doi.org/10.1016/S0095-4470\(19\)30537-6](https://doi.org/10.1016/S0095-4470(19)30537-6)
- Flores, C. (2012). Differential effects of language attrition in the domains of verb placement and object expression. *Bilingualism: Language and Cognition*, 15(3), 550–567. <https://doi.org/10.1017/S1366728911000666>
- Fukumura, K., & van Gompel, R. P. G. (2010). Choosing anaphoric expressions: Do people take into account likelihood of reference? *Journal of Memory and Language*, 62(1), 52–66. <https://doi.org/10.1016/j.jml.2009.09.001>
- Garrod, S. C., & Sanford, A. J. (1982). The mental representation of discourse in a focused memory system: Implications for interpretation for the anaphoric noun-phrases. *Journal of Semantics* 1, 21–41.
- Giannakou, A. (2018). Spanish and Greek subjects in contact: Greek as a heritage language in Chile. [Doctoral dissertation, University of Cambridge]. <https://doi.org/10.17863/CAM.30352>

- Gorden, P. C., Grosz, B. J., & Gilliom, L.A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 17(3), 311-347.
- Gürel, A. (2004). Selectivity in L2-induced L1 attrition: A psycholinguistic account. *Journal of Neurolinguistics*, 17(1), 53–78. [https://doi.org/10.1016/S0911-6044\(03\)00054-X](https://doi.org/10.1016/S0911-6044(03)00054-X)
- Hicks, G., & Domínguez, L. (2020). A model for L1 grammatical attrition. *Second Language Research*. 36(2), 143-165. <https://doi.org/10.1177/0267658319862011>
- Huang, C.-T. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry*, 15(4), 531-574. <https://www.jstor.org/stable/4178404>
- Huang, Y. (1994). *The Syntax and Pragmatics of Anaphora: A Study with Special Reference to Chinese*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511554292>
- Hwang, H. (2021). Avoidance of gender-ambiguous pronouns as a consequence of ambiguity-avoidance strategy. *Discourse Processes*, 58(3), 251–259. <https://doi.org/10.1080/0163853X.2020.1844965>
- Huang, J. C.-T. (1992). Complex Predicates in Control. In R. K. Larson, S. Iatridou, U. Lahiri, & J. Higginbotham (Eds.), *Control and Grammar* (pp. 109–147). Springer. https://doi.org/10.1007/978-94-015-7959-9_4
- Jarvis, S. (2003). Probing the Effects of the L2 on the L1: A Case Study. In V. Cook (Ed.), *Effects of the Second Language on the First* (pp. 81–102). Multilingual Matters. <https://doi.org/10.21832/9781853596346-007>
- Jarvis, S. (2019). Lexical Attrition. In M. S. Schmid & B. Köpcke (Eds.), *The Oxford Handbook of Language Attrition* (pp. 241–250). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198793595.013.20>
- Jia, R., & Paradis, J. (2015). The use of referring expressions in narratives by Mandarin heritage language children and the role of language environment factors in predicting individual differences. *Bilingualism: Language and Cognition*, 18(4), 737–752. <https://doi.org/10.1017/S1366728914000728>
- Kasparian, K., Vespignani, F., & Steinhauer, K. (2017). First Language Attrition Induces Changes in Online Morphosyntactic Processing and Re-Analysis: An ERP Study of Number Agreement in Complex Italian Sentences. *Cognitive Science*, 41(7), 1760–1803. <https://doi.org/10.1111/cogs.12450>
- Laméris, T. J, Kubota, M., Kupisch, T., Cabrelli, J., Snape, N., & Rothman, J. (2024). Language change in Japanese-English bilingual returnee children over the course of five years: Evidence from accent-rating. *Second Language Research*, 41(1), 191-216. <https://doi.org/10.1177/02676583241230854>

- Lidz, J., & Gagliardi, A. (2015). How nature meets nurture: Universal Grammar and statistical learning. *Annual Review of Linguistics 1*: 333–52.
<https://doi.org/10.1146/annurev-linguist-030514-125236>
- 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.
<https://doi.org/10.1111/j.1467-9280.2009.02480.x>
- Lozano, C. (2016). Pragmatic principles in anaphora resolution at the syntax-discourse interface: advanced English learners of Spanish in the CEDEL2 corpus. In M. Alonso Ramos (Ed.), *Spanish learner corpus Research: Current trends and future perspectives* (pp. 236-265). Amsterdam: John Benjamins. <https://doi.org/10.1075/scl.78.09loz>
- Lozano, C. (2018). The Development of Anaphora Resolution at the Syntax-Discourse Interface: Pronominal Subjects in Greek Learners of Spanish. *Journal of Psycholinguist Research* 47, 411–430. <https://doi.org/10.1007/s10936-017-9541-8>
- Makowski, D., Ben-Shachar, M., & Lüdecke, D. (2019). bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *Journal of Open Source Software*, 4(40), 1541. <https://doi.org/10.21105/joss.01541>
- 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* (pp. 88–97). Cascadilla Press.
<https://www.lingref.com/cpp/gasla/8/paper1491.pdf>
- Mayr, R., Price, S., & Mennen, I. (2012). First language attrition in the speech of Dutch–English bilinguals: The case of monozygotic twin sisters. *Bilingualism: Language and Cognition*, 15(4), 687-700. <https://doi.org/10.1017/S136672891100071X>
- Montrul, S. (2004). Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition*, 7(2), 125–142.
<https://doi.org/10.1017/S1366728904001464>
- Opitz, C. (2013). A dynamic perspective on late bilinguals' linguistic development in an L2 environment. *International Journal of Bilingualism*, 17(6), 701–715.
<https://doi.org/10.1177/1367006912454621>
- Opitz, C. (2019). A complex dynamic systems perspective on personal background variables in L1 attrition. In M. S. Schmid & B. Köpcke (Eds.), *The Oxford Handbook of Language Attrition* (pp. 49–60). Oxford University Press.

<https://doi.org/10.1093/oxfordhb/9780198793595.013.5>

- Paradis, M. (1993). Linguistic, psycholinguistic, and neurolinguistic aspects of ‘interference’ in bilingual speakers: The activation threshold hypothesis. *International Journal of Psycholinguistics*, 9(2), 133–145.
- R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing [Computer software]. <https://www.r-project.org/>
- Ryan, J. (2015). Overexplicit Referent Tracking in L2 English: Strategy, Avoidance, or Myth? *Language Learning*, 65(4), 824–859. <https://doi.org/10.1111/lang.12139>
- Schmid, M. S. (2007). The role of L1 use for L1 attrition. In B. Köpcke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language Attrition: Theoretical perspectives* (pp. 135–153). John Benjamins Publishing Company. <https://doi.org/10.1075/sibil.33.10sch>
- Schmid, M. S., & Jarvis, S. (2014). Lexical access and lexical diversity in first language attrition. *Bilingualism: Language and Cognition*, 17(4), 729–748. <https://doi.org/10.1017/S1366728913000771>
- Schmid, M. S., & Yilmaz, G. (2018). Predictors of Language Dominance: An Integrated Analysis of First Language Attrition and Second Language Acquisition in Late Bilinguals. *Frontiers in Psychology*, 9, Article1306. <https://doi.org/10.3389/fpsyg.2018.01306>
- Seliger, H. W., & Vago, R. M. (1991). The study of first language attrition: an overview. In: H. W. Seliger and R. M. Vago (Eds). *First Language Attrition* (pp.3-16). Cambridge University Press. <https://doi.org/10.1017/CBO9780511620720.001>
- Slabakova, R., Zhao, L., Baker, L., Turner, J., & Tuniyani, E. (2024). Null and overt pronoun interpretation in L2 Mandarin resultative constructions. *Linguistic Approaches to Bilingualism*, 14(2), 178–217. <https://doi.org/10.1075/lab.20120.sla>
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism*, 1(1), 1–33. <https://doi.org/10.1075/lab.1.1.01sor>
- Sorace, A. (2019). L1 attrition in a wider perspective. *Second Language Research*, 36(2), pp. 203-206. <https://doi.org/10.1177/0267658319895571>
- Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, 22(3), 339–368. <https://doi.org/10.1191/0267658306sr271oa>
- 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. <https://doi.org/10.1016/j.lingua.2008.09.008>
- Tsimpli, I., 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.

<https://doi.org/10.1177/13670069040080030601>

- Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P., Paananen, T., & Gelman, A. (2024). “loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models.” R package version 2.8.0. <https://mc-stan.org/loo/>
- Wu, S.-L. (2020). Crosslinguistic influence in development of reference realization: A comparison of foreign language learners and heritage language learners. *Chinese as a Second Language Research*, 9(2), 227–257. <https://doi.org/10.1515/caslar-2020-2009>
- Yan, Q. C. (2022). Deriving Temporal Adverbial Clauses in Mandarin Chinese. [Unpublished Manuscript].
- Yang, C. L., Gordon, P. C., Hendrick, R., & Wu, J. T. (1999). Comprehension of Referring Expressions in Chinese. *Language and Cognitive Processes*, 14(5–6), 715–743. <https://doi.org/10.1080/016909699386248>
- Yazawa, K., Konishi, T., Pérez-Ramón, P., & Kondo, Mariko. (2024). More foreign-accented but more comprehensible: Attrition and amelioration of L1 speech in proficient L2 learners. *F1000Res*, 13:867. <https://doi.org/10.12688/f1000research.148193.1>
- Zhang, A., & Kwon, N. (2022). The interpretational preferences of null and overt pronouns in Chinese. *Journal of Linguistics*, 58(3), 649–676. <https://doi.org/10.1017/S0022226721000402>
- Zhao, L. X. (2014). Ultimate attainment of anaphora resolution in L2 Chinese. *Second Language Research*, 30(3), 381–407. <https://doi.org/10.1177/0267658314521107>
- Zhou, J., Mai, Z., Cai, Q., Liang, Y., & Yip, V. (2022). Reference production in Mandarin–English bilingual preschoolers: Linguistic, input, and cognitive factors. *Frontiers in Psychology*, 13, Article 897031. <https://doi.org/10.3389/fpsyg.2022.897031>

Appendix

Reference Production: Reference Level = Control Group

Table 1: The outputs of the Bayesian model for effects involving Role and Group in the two-character condition. For Intercept, positive values indicate a preference for the less explicit form (listed on the right), whereas negative values indicate a preference for the more explicit form (listed on the left). For the remaining effects, it is the opposite (positive values indicate a preference for the more explicit form, listed on the right). Role = Subject and Group = Control were set as reference levels, respectively; group is dummy coded.

Predictors	Estimates	95% CrI	PD (%)
Intercept, Overt vs Null (Control group, Role = Subject)	0.21	[-0.36, 0.75]	77
Intercept, NP vs Overt (Control group, Role = Subject)	-1.43	[-1.94, -0.93]	100
Role = Non-subject, Null vs Overt (Control group)	0.42	[-0.86, 1.79]	73
Role = Non-subject, Overt vs NP (Control group)	3.40	[2.42, 4.48]	100
Short-Term English Exposure vs Control, Null vs Overt (Role=Subject)	0.82	[0.15, 1.48]	99
Short-Term English Exposure vs Control, Overt vs NP (Role=Subject)	-0.48	[-1.08, 0.13]	94
Long-Term English Exposure vs Control, Null vs Overt (Role=Subject)	0.88	[0.17, 1.61]	99
Long-Term English Exposure vs Control, Overt vs NP (Role=Subject)	0.19	[-0.43, 0.81]	73
Role = Non-subject * Short-Term English Exposure, Null vs Overt	-0.18	[-1.71, 1.37]	59
Role = Non-subject * Short-Term English Exposure, Overt vs NP	-0.09	[-1.32, 1.17]	56
Role = Non-subject * Long-Term English Exposure, Null vs Overt	0.25	[-1.38, 1.95]	61
Role = Non-subject * Short-Term English Exposure, Overt vs NP	-0.59	[-1.84, 0.71]	82

Table 2: The outputs of the Bayesian model for effects involving Group in the one-character condition. For Intercept, positive values indicate a preference for the less explicit form (listed on the right), whereas negative values indicate a preference for the more explicit form (listed on the left). For the remaining effects, it is the opposite (positive values indicate a preference for the more explicit form, listed on the right). Group = Control was set as the reference levels and was dummy coded.

Predictors	Estimates	95% CrI	PD (%)
Intercept Null vs Overt (Control group)	-0.07	[-0.62, 0.47]	61
Intercept Overt vs NP (Control group)	2.91	[2.27, 3.57]	100
Short-Term English Exposure vs Control, Overt vs Null	0.58	[-0.10, 1.25]	96
Short-Term English Exposure vs Control, NP vs Overt	-0.66	[-1.51, 0.15]	94
Long-Term English Exposure vs Control, Overt vs Null	1.01	[0.34, 1.70]	100
Long-Term English Exposure vs Control, NP vs Overt	0.09	[-0.71, 0.88]	59

Reference Production: Reference Level = Short-Term English Exposure

Table 3: The outputs of the Bayesian model for effects involving Role and Group in the two-character condition. For Intercept, positive values indicate a preference for the less explicit form (listed on the right), whereas negative values indicate a preference for the more explicit form (listed on the left). For the remaining effects, it is the opposite (positive values indicate a preference for the more explicit form, listed on the right). Role = Subject and **Group = Short-Term English Exposure group** were set as reference levels, respectively; group is dummy coded.

Predictors	Estimates	95% CrI	PD (%)
Intercept, Overt vs Null (Short-Term English Exposure group, Role = Subject)	-0.59	[-1.12, -0.09]	99
Intercept, NP vs Overt (Short-Term English Exposure group, Role = Subject)	-0.97	[-1.42, -0.53]	100
Role = Non-subject, Null vs Overt (Short-Term English Exposure)	0.36	[-0.82, 1.64]	72
Role = Non-subject, Overt vs NP (Short-Term English Exposure)	3.10	[2.30, 4.00]	100
Control vs Short-Term English Exposure, Null vs Overt (Role=Subject)	-0.84	[-1.54, -0.16]	99
Control vs Short-Term English Exposure, Overt vs NP (Role=Subject)	0.43	[-0.18, 1.04]	91
Long-Term English Exposure vs Short-Term English Exposure, Null vs Overt (Role=Subject)	0.10	[-0.56, 0.79]	61
Long-Term English Exposure vs Short-Term English Exposure, Overt vs NP (Role=Subject)	0.63	[0.05, 1.21]	98
Role = Non-subject * Short-Term English Exposure, Null vs Overt	-0.39	[-2.13, 1.32]	67
Role = Non-subject * Short-Term English Exposure, Overt vs NP	0.99	[-0.43, 2.55]	91
Role = Non-subject * Long-Term English Exposure, Null vs Overt	0.31	[-1.25, 1.98]	64
Role = Non-subject * Short-Term English Exposure, Overt vs NP	-0.28	[-1.41, 0.88]	69

Table 4: The outputs of the Bayesian model for effects involving Group in the one-character condition. For Intercept, positive values indicate a preference for the less explicit form (listed on the right), whereas negative values indicate a preference for the more explicit form (listed on the left). For the remaining effects, it is the opposite (positive values indicate a preference for the more explicit form, listed on the right). **Group = Short-Term English Exposure group** was set as the reference levels and was dummy coded.

Predictors	Estimates	95% CrI	PD (%)
Intercept, Overt vs Null (Short-Term English Exposure)	-0.71	[-1.26, -0.20]	100
Intercept NP vs Overt (Short-Term English Exposure)	3.40	[2.68, 4.06]	100
Control vs Short-Term English Exposure, Null vs Overt	-0.73	[-1.40, -0.06]	98
Control vs Short-Term English Exposure, Overt vs NP	0.35	[-0.45, 1.18]	80
Long-Term English Exposure vs Short-Term English Exposure, Null vs Overt	0.38	[-0.28, 1.03]	88
Long-Term English Exposure vs Short-Term English Exposure, Overt vs NP	0.59	[-0.19, 1.36]	93

Reference Production and English proficiency

Table 5: The outputs of the Bayesian model for effects of role and the average English proficiency in the four skills across two referential contrasts in the two-character condition. For intercept, positive values indicate a preference for the less explicit form (listed on the right), whereas negative values indicate a preference for the more explicit form (listed on the left). For the remaining effects, it is the opposite (positive values indicate a preference for the more explicit form, listed on the right). The subject role was the reference level.

Effects	Estimate	95% CrI	PD (%)
Intercept, Overt vs Null (Role = Subject)	-0.38	[-0.79, 0.01]	97
Intercept, NP vs OVERT (ROLE = SUBJECT)	-1.30	[-1.62, -0.98]	100
Role = Non-subject, Null vs Overt	0.31	[-0.72, 1.42]	71
Role = Non-subject, Overt vs NP	3.31	[2.60, 4.07]	100
English proficiency, Null vs Overt	0.20	[-0.11, 0.51]	90
English proficiency, Overt vs NP	-0.06	[-0.32, 0.20]	67
Role = Non-subject * English proficiency Null vs Overt	0.61	[-0.31, 1.54]	91
Role = Non-subject * English proficiency Overt vs NP	-0.52	[-1.21, 0.14]	94

Table 6: The outputs of the Bayesian model for the effect of the average English proficiency in the four skills across two referential contrasts in the one-character condition. For intercept, positive values indicate a preference for the less explicit form (listed on the right), whereas negative values indicate a preference for the more explicit form (listed on the left). For the remaining effects, positive values indicate a preference for the more explicit form (listed on the right), whereas negative values indicate a preference for the less explicit form (listed on the left).

Effects	Estimate	95% CrI	PD (%)
Intercept, Overt vs Null	-0.62	[-1.01, -0.24]	100
Intercept, NPs vs Overt	3.11	[2.66, 3.57]	100
English proficiency, Null vs Overt	0.38	[0.10, 0.68]	99
English proficiency, Overt vs NP	0.26	[-0.09, 0.61]	92

Reference Production and English Use in the Four Skills

Table 7¹: The outputs of the Bayesian model for effects of role and English use in four skills across two referential contrasts in the two-character condition. Positive values indicate a preference for the more explicit form (listed on the right), whereas negative values indicate a preference for the less explicit form (listed on the left). The subject role was the reference level.

Effects	Estimate	95% CrI	PD (%)
English use in the four skills, Null vs Overt (Role = Subject)	0.21	[-0.10, 0.52]	91
English use in the four skills, Overt vs NP (Role = Subject)	0.15	[-0.10, 0.41]	89
Role = Non-subject * English use in the four skills, Null vs Overt	0.88	[-0.04, 1.87]	97
Role = Non-subject * English use in the four skills, Overt vs NP	-0.29	[-0.87, 0.31]	83

Table 8²: The outputs of the Bayesian model for the effects of English use in four skills across two referential contrasts in the one-character condition. Positive values indicate a preference for the more explicit form (listed on the right), whereas negative values indicate a preference for the less explicit form (listed on the left).

Effects	Estimate	95% CrI	PD (%)
English use in the four skills, Null vs Overt	0.28	[-0.02, 0.57]	97
English use in the four skills, Overt vs NP	0.37	[0.03, 0.71]	98

¹ We do not repeat coefficients for the Intercept (Role = Subject) and Role (= Non-subject) in the two-character condition in the following tables as they have been presented in Table 5.

² We do not repeat coefficients for the Intercept (Role = Subject) in the one-character condition in the following tables as they have been presented in Table 6.

Reference Production and English Use in the Daily Contexts

Table 9: The outputs of the Bayesian model for effects of role and English use in the 12 daily contexts across two referential contrasts in the two-character condition. Positive values indicate a preference for the more explicit form (listed on the right), whereas negative values indicate a preference for the less explicit form (listed on the left). The subject role was the reference level.

Effects	Estimate	95% CrI	PD (%)
English use in 12 daily contexts, Null vs Overt (ROLE = SUBJECT)	0.40	[0.11, 0.71]	100
English use in 12 daily contexts, Overt vs NP (Role = Subject)	0.05	[-0.19, 0.31]	66
Role = Non-subject * English use in 12 daily contexts, Null vs Overt	0.21	[-0.66, 1.10]	68
Role = Non-subject * English use in 12 daily contexts, Overt vs NP	-0.32	[-0.94, 0.29]	84

Table 10: The outputs of the Bayesian model for the effect of English use in the 12 daily contexts across two referential contrasts in the one-character condition. Positive values indicate a preference for the more explicit form (listed on the right), whereas negative values indicate a preference for the less explicit form (listed on the left).

Effects	Estimate	95% CrI	PD (%)
English use in 12 daily contexts, Null vs Overt	0.46	[0.18, 0.75]	100
English use in 12 daily contexts, Overt vs NP	0.18	[-0.15, 0.53]	85

Model Comparison across the Three Continuous Predictors

Table 11: Two model comparisons were conducted across the three continuous predictors for the two-character and one-character conditions, respectively. The `loo_compare` function ranks the models from best to worst based on their Expected Log Predictive Density (ELPD). The top-ranked model serves as the baseline, with its ELPD difference and standard error (SE) difference set to zero. A difference in ELPD between models is considered reliable when its absolute value is at least 2 times (or more conservatively, 4 times) larger than the corresponding SE difference³. ELPD represents the model's predictive accuracy, whereas SE difference quantifies uncertainty in the ELPD difference between models.

Condition	Models	ELPD difference	SE difference
two-character	English use in four skills	0.0	0.0
	English use in 12 daily contexts	-0.3	2.3
	English proficiency	-0.7	2.5
one-character	English use in 12 daily contexts	0.0	0.0
	English proficiency	-1.9	1.9
	English use in four skills	-3.4	2.6

³ Vehtari, A. (2018, September 26). *Interpreting output from compare of LOO*. The Stan Forums. <https://discourse.mc-stan.org/t/interpreting-output-from-compare-of-loo/3380/2>