

Prediction on the basis of gender and number in Mandarin-Italian bilingual children

Jasmijn E. Bosch*¹, Mathilde Chailleux*¹, Jia'en Yee², Maria Teresa Guasti¹ & Fabrizio Arosio¹

1 University of Milan-Bicocca

2 University Putra Malaysia

Abstract

The present study used a visual world eye-tracking paradigm to investigate online processing of grammatical gender and number in Mandarin-Italian bilingual children, in comparison to monolingual Italian children. We examined how children anticipated upcoming nouns on the basis of grammatical gender and number information on the preceding article. While monolingual speakers are able to employ such predictive mechanisms from a very young age, to our knowledge, this is the first study that compares gender and number processing in bilingual children.

The results show that, overall, participants made linguistic predictions on the basis of articles, although a post-hoc analysis focusing on a subset of our bilingual participants did not confirm the prediction effect in Mandarin-Italian bilingual children. We found a greater difference in the bilingual group than in the monolingual group, in that bilinguals tended to be slower when processing gender than number. Finally, we found that L2 proficiency had a significant effect on gender processing in the bilingual group. One interpretation of these findings is that the discrepancy between gender and number may be due to transfer, since Mandarin does not have grammatical gender while it does have a conceptual notion of number. Another factor may be L2 proficiency, and especially lexical knowledge, since gender is an arbitrary property stored in the lexicon, while number is concretely linked to the referential context.

Keywords: visual world paradigm, gender, number, child bilingualism, predictive processing

Jasmijn E. Bosch, Mathilde Chailleux, Jia'en Yee, Maria Teresa Guasti & Fabrizio Arosio. 2021. Prediction on the basis of gender and number in Mandarin-Italian bilingual children. In Dalila Ayoun (ed.), *The acquisition of gender*. John Benjamin Publishing Company, 243-272. DOI 10.1075/sibil.63.

1. Introduction

Humans are able to make predictions while processing speech, making language comprehension fast and efficient. In a classic language processing study, Altmann and Kamide (1999) found that listeners who heard the sentence *The boy eats the cake* inferred when hearing the word *eats* that the upcoming noun must denote an edible object. In other words, listeners anticipated how the sentence would continue based on the lexical meaning of the verb. Similarly, a native speaker of Italian who hears the sentence *Prendi la banana* ‘Take the.FEM.SING banana’ could anticipate certain features of the upcoming noun on the basis of morphosyntactic information on the article. Firstly, the gender feature indicates that the following noun must be feminine. Secondly, the number marking of the article indicates that the noun should refer to a single object rather than two or more. In contrast, when hearing the sentence *Prendi i mandarini* ‘Take the-MASC-PLU mandarins-MASC-PLU’, Italian listeners can predict, as soon as they hear the article, that the following noun will refer to more than one object with masculine gender.

While previous research has found that monolinguals make rapid use of determiner features during noun recognition from an early age (e.g. Lew-Williams & Fernald 2007; Robertson, Shi & Melançon 2012), adult second language (L2) learners appear to be less efficient when processing gender (e.g. Foucart & Frenck-Mestre 2011; Lew-Williams & Fernald 2010). We turned our attention to bilingual children who started acquiring their L2 at an early age, by investigating linguistic prediction based on gender and number in Mandarin-Italian bilinguals, in comparison to monolingual Italian children. Mandarin is relevant for this purpose since it has neither articles nor grammatical gender, while it does express number, albeit in a different way than in Italian. This provides us with an interesting test case for investigating the effect of linguistic differences on language processing in early L2 learners.

Italian has a binary gender system, in which nouns can be classified as feminine or masculine. Gender assignment is largely arbitrary, although the gender of animate nouns tends to match the biological sex of their referents (e.g. *uomo*-MASC ‘man’ and *donna*-FEM ‘woman’). Furthermore, there are some morphophonological regularities associated with the gender classification of Italian nouns. Most feminine singular nouns end with the phoneme /a/ (e.g. *matita*-FEM ‘pencil’) while most masculine singular nouns end with the phoneme [o] (e.g. *libro*-MASC ‘book’), but there are also exceptions to this rule (e.g. *moto*-FEM ‘motorbike’; *poema*-MASC ‘poem’). Additionally, Italian has a subclass of nouns ending in [e], which may be either masculine or feminine (e.g. *noce*-FEM, ‘nut’; *sole*-MASC, ‘sun’), as well as loan words ending in consonants that tend to be masculine (e.g. *computer*-MASC ‘computer’) (Ayoun & Maranzana 2020; Chini 1995;

Padovani & Cacciari 2003; Panzini 2017). The present study only includes nouns with prototypical, morphologically transparent endings, i.e. masculine nouns ending in $-[o]$ or feminine nouns ending in $-[a]$.

In Italian, number is expressed on nouns through suffixes (syncretic with gender marking), such that the vocalic endings of nouns reveal information about both gender and number. Typically, the plural form of regular masculine nouns ends in $-i$ (e.g. *libro*-MASC-SING 'book' vs *libri*-MASC-PLU 'books') and the plural form of regular feminine nouns ends in $-e$ (e.g. *matita*-FEM-SING 'pencil' vs *matite*-FEM-PLU 'T-shirts'). The plural form of nouns ending in $-e$ typically ends in $-i$, regardless of grammatical gender (e.g. *noci*-FEM, 'nuts'; *solì*-MASC, 'suns'). It should be noted, however, that the system is also characterized by exceptions; some nouns do not mark plurality through suffixation (e.g. *computer*- MASC-SING 'computer' vs *computer*- MASC-PLU 'computers'; *città*- FEM-SING 'city' vs *città*- FEM-PLU 'cities'), and other nouns change their grammatical gender according to number (e.g., *dito*-MASC-SING 'finger', *dita*- FEM-PLU 'fingers') (Ferrari & Zampese 2016; Panzini 2017).

The gender and number of nouns can typically be determined by looking at the preceding article. Italian has definite and indefinite articles that are unambiguously marked for gender and number; singular masculine nouns are typically preceded by *il* (definite) or *un* (indefinite), singular feminine nouns by *la* (definite) or *una* (indefinite), plural masculine nouns by *i*, and plural feminine nouns by *le*. Nouns that starts with a vowel are an exception, since they require the use of the elided article *l'*, which can either be followed by a feminine or a masculine noun (e.g. *l'amico*- MASC and *l'amica*- FEM 'the friend'). For a small group of masculine nouns, the articles *lo*, *uno* or *gli* are used, based on phonological rules (e.g. nouns starting with *z*; *lo zaino* 'the backpack'; *uno zaino* 'a backpack'; *gli zaini* 'the backpacks') (Ayoun & Maranzana 2020; Ferrari & Zampese 2016; Panzini 2017).

In addition to determiners, gender and number of nouns are also marked through agreement with personal pronouns, adjectives, and past participles (Panzini 2017). The example in (1) shows the typical agreement pattern with a masculine singular noun (here *vaso* 'vase'), which includes the masculine singular article *il* and the vocalic ending $[o]$ for the personal pronoun *mio* 'my', the adjective *rosso* 'red' and the past participle *stato rotto* 'is broken'. In the case of a feminine singular noun like *tazza* 'cup' in (2), the corresponding article *la* and the vocalic ending $[a]$ are typically used. The agreement with masculine plural nouns, such as *vasi* 'vases' in (3), typically involves the article *i*, as well as the vocalic ending $[i]$ for the other elements of the sentence (*miei* 'my', *rossi* 'red', *stati rotti* 'are broken'). Finally, the typical agreement pattern for feminine plural nouns such as *tazze* 'cups' is shown in (4), in which the article *le* and the vocalic ending $[e]$ are used.

1. *Il mio vaso rosso è stato rotto*
the-MASC-SING my-MASC-SING vase-MASC-SING red-MASC-SING is been-MASC-SING
broken-MASC-SING
‘My red vase has been broken’
2. *La mia tazza rossa è stata rotta*
the-FEM-SING my-FEM-SING cup-FEM-SING red-FEM-SING is been-FEM-SING broken-
FEM-SING ‘My red cup has been broken’
3. *I miei vasi rossi sono stati rotti*
the-MASC-PLU my-MASC-PLU vase-MASC-PLU red-MASC-PLU are been-MASC-PLU
broken-MASC-PLU
‘My red vases have been broken’
4. *Le mie tazze rosse sono state rotte*
the-FEM-PLU my-FEM-PLU cup-FEM-PLU red-FEM-PLU are been-FEM-PLU broken-FEM-
PLU
‘My red cups have been broken’

Mandarin presents several contrasts with Italian. First, it lacks articles, as bare nouns may either have a generic or existential meaning, or a definite interpretation. Second, Mandarin does not have grammatical gender, but it uses nominal classifiers. These lexical items, which indicate to which class a noun belongs, appear obligatorily with numerals or quantifiers (Frankowsky & Ke 2016). It has been argued that in some aspects classifier systems may not be qualitatively different from grammatical gender systems, as both systems entail nominal classification (Fedden & Corbett 2017; Haspelmath 2018). Unlike grammatical gender in Italian, however, classifiers in Mandarin do not involve morphosyntactic agreement. Nominal classification in Mandarin is based on semantic properties of the noun, such as its material and physical shape (Huang & Ahrens 2003).

Number is unspecified in bare nouns in Mandarin, that is, they are not singular nor plural, but they have general number (Corbett 2000; Rullmann & You 2006). There are, however, several ways of expressing the concept of number or quantity, including using numerals such as *yī* ‘one’ or *liǎng* ‘two’, or quantifiers such as *duì* (denoting a pair) or *tào* (denoting a set). Two frequently discussed plural-encoding forms are *xīe* and *-mén* (Iljic 1994). While *xīe* ‘some’ marks collectiveness, it is not obligatory, and it can also be used with mass nouns. The suffix *-mén* is obligatorily used with pronouns to denote number, but it is restricted to human and definite referents. Moreover, it has been argued that plurality may be expressed by the reduplication of

classifiers, and singularity may be expressed by using a simple classifier in the absence of a numeral (Zhang 2014).

To sum up, while the grammatical notion of gender is absent in Mandarin, number can be expressed in various ways, albeit differently than in Italian. This allows us to test the role of L1 characteristics in L2 processing in bilingual children.

This chapter will first discuss previous research on processing of grammatical gender and number, in both monolingual and bilingual adults and children. Then we will turn to our hypotheses, methods and results, before concluding with a discussion of our findings.

2. Literature review

2.1 Processing of grammatical gender

Gender processing occurs rapidly in native speakers. Evidence for this comes from studies using the visual world paradigm, in which participants' eye movements are recorded when presented with a visual scenario, while they are listening to sentences (Lew-Williams & Fernald 2010; Dahan et al. 2000). For example, Lew-Williams and Fernald (2010) presented participants with two pictures of objects that either matched or mismatched in grammatical gender, instructing them to look for an object displayed on the screen (i.e. *Encuentra el/la ...* 'Find the-MASC/the-FEM'). Monolingual Spanish speakers showed faster identification of the target referent when there was a mismatch between target and competitor, allowing the target noun to be identified based on the gender-marked article (Lew-Williams & Fernald 2010). Using a similar paradigm, Dahan et al. (2000) found that French speakers were able to pre-activate the upcoming noun based on the gender-marked article. The results of these eye-tracking studies are supported by neurophysiological evidence showing that native speakers attend to gender cues when processing sentences, and that they are able to use this information to anticipate upcoming nouns (Wicha, Moreno & Kutas 2004; Barber & Carreiras 2005).

Monolingual children also make use of grammatical gender cues during online processing. Recent eye-tracking studies have demonstrated that, similarly to adults, German eight- and nine-year-old children can predict upcoming nouns on the basis of grammatical gender marking on determiners and adjectives (Cholewa et al. 2019; Lemmerth & Hopp 2019). Furthermore, three-year-old monolingual Spanish-speaking children have been shown to effectively exploit gender information on articles, leading to faster recognition of subsequent nouns (Lew-Williams & Fernald 2007). Using a preferential looking paradigm, Van Heugten & Shi (2009) reported a similar effect in even younger children; French-learning two-year-olds were faster to turn their gaze to a target

referent when it was preceded by a disambiguating gender-marked article. Moreover, when a noun was preceded by an incongruent article, this disagreement in grammatical gender subsequently hindered comprehension.

In contrast to native speakers, adult L2 learners tend to be less efficient when processing grammatical gender. For example, using an auditory naming task, Guillelmon and Grosjean (2001) tested whether native English late L2 learners of French, early English-French bilinguals and French monolinguals experienced a facilitation effect when repeating nouns that were preceded by a grammatical gender-marked article, and an inhibition effect when there was a mismatch in grammatical gender. Their results showed clear facilitation and inhibition effects for both monolinguals and early bilinguals, resulting in faster reaction times when the appropriate gender was used and slower reaction times when there was a gender mismatch, while late bilinguals were completely insensitive to gender marking.

Furthermore, predictive processing may be slower in the L2, even in highly proficient speakers whose offline comprehension is on target. For example, Grüter, Lew-Williams & Fernald (2012) investigated offline comprehension, elicited production, and online processing of gender in advanced L2 speakers of Spanish (L1 English) as well as L1 Spanish speakers. Even though the L2 participants performed at-ceiling in a picture selection task, they made some errors in production, and they made significantly less use of gender cues when processing familiar nouns.

Focusing on L1 English advanced to near-native L2 speakers of German, Hopp (2012) found that predictive processing based on gender-marked articles was significantly slower in L2 speakers with variable performance in a gender assignment production task than in native speakers and L2 speakers who consistently marked grammatical gender correctly. When processing gender-marked adjectives, L2 speakers with consistent high accuracy were significantly slower than native speakers, while L2 speakers with variable accuracy did not show any predictive processing. As for L2 learners with lower levels of proficiency, Lew-Williams and Fernald (2010) found that in contrast to monolingual adults and children, intermediate L2 learners of Spanish failed to make use of gender cues on the article during online processing. In order to test whether this may have been due to differences in frequency of exposure, follow-up experiments were conducted with novel words, which were taught to the participants before the experimental phase. The results showed that after a training, L2 participants were able to make use of the gender-marked article, albeit to a lesser extent than L1 participants. However, when participants were trained with indefinite articles and tested with definite articles, L2 speakers were no longer able to do so, indicating that they could not exploit gender cues on articles when the task required them to generalize across forms.

In addition to factors related to the language proficiency, the typological distance between two languages and the degree of phonological overlap between translation equivalents may contribute to the observed differences between monolingual and bilingual processing of grammatical gender. Dussias et al. (2013) explored the effect of language distance on Spanish L2 processing, by comparing native speakers of English (a language lacking grammatical gender) to native speakers of Italian (a language which has a grammatical gender system comparable to Spanish), while taking into account different levels of proficiency. The results showed that, like Spanish monolingual speakers, native English high-proficiency L2 learners exhibited predictive processing on the basis of gender-marked articles, while native English low-proficiency L2 learners failed to do so. Interestingly, native Italian low-proficiency L2 learners of Spanish showed anticipatory eye movements, but only with the feminine article *la*, which is identical to its Italian counterpart. This suggests that low-proficiency participants experienced positive transfer when there was overlap between their L1 and their L2. Another example of transfer in bilingual gender processing is provided by the ‘gender congruency effect’, which refers to the often-replicated finding that bilinguals who speak two gender-marked languages are faster when processing nouns which have the same grammatical gender in both languages, as compared to nouns with incongruent gender (Morales, Paolieri, Dussias, Kroff, Gerfen & Bajo 2016; Sá-Leite, Fraga & Comesaña 2019).

At the neurophysiological level, some studies have found that L2 speakers respond differently to violations in grammatical gender agreement than native speakers, depending on the morphosyntactic realization and the degree of syntactic overlap between the L1 and the L2 (Foucart & Frenck-Mestre 2011, 2012). For example, Foucart & Frenck-Mestre (2011) found that native German advanced L2 learners of French showed a native-like P600 effect in response to grammatical gender violations when agreement rules were similar in both languages (i.e. between determiners and nouns), while there was no such effect when agreement rules differed across languages (i.e. between adjectives and nouns). However, Tokowicz & MacWhinney (2005) found that native English low-proficiency L2 learners of Spanish were sensitive to agreement violations with grammatical gender, despite the fact that this feature was absent in the L1.

Whilst there is extensive (yet inconclusive) literature on the differences between online processing of grammatical gender in native speakers versus adult L2 learners, online processing of grammatical gender in bilingual children remains largely unexplored. A recent study investigated linguistic prediction based on gender-marked determiners and adjectives in German, by comparing eight- and nine-year-old German-Russian bilingual children to a group of age-matched monolingual controls in a visual world experiment (Lemmerth & Hopp 2019). The bilingual children were

divided in two groups; simultaneous bilinguals, who were exposed to both languages from birth, and early sequential bilinguals, who were exposed to German before the age of four. The results showed that, similarly to monolingual children, simultaneous bilinguals anticipated upcoming nouns on the basis of the grammatical gender of the article or prenominal adjective. In contrast, sequential bilinguals only did so when the grammatical gender of the target noun was the same as that of its Russian translation equivalent. This means that native Russian early L2 learners of German experienced a gender congruency effect, i.e. their processing was delayed when there was a mismatch in grammatical gender between their two languages, due to cross-linguistic interference during the pre-activation of nouns based on articles and adjectives.

Russian and German both have a tripartite grammatical gender system (distinguishing between feminine, masculine and neuter gender), but they differ in the way in which gender is expressed. While both languages mark grammatical gender on prenominal adjectives, German also marks gender on articles, which are absent in Russian. This allowed the authors to explore the effect of linguistic differences, by testing whether participants would be more likely to show prediction based on adjectives than on articles.

A previous study on native Russian adult L2 learners of German (Hopp & Lemmerth 2018) found that high-intermediate speakers showed predictive processing in German based on gender-marked adjectives but not on gender-marked articles, suggesting that cross-linguistic overlap in the morphosyntactic realization of grammatical gender facilitates processing in L2 learners. However, Lemmerth and Hopp (2019) observed identical patterns for both types of stimuli, which means that the delayed processing of grammatical gender in sequential bilingual children cannot be attributed to the absence of articles in Russian.

Using a similar paradigm, Bosch and Foppolo (submitted) investigated predictive processing in Italian-German bilingual children between the ages of six and nine. Participants were resident in Italy or in Germany, and the sample included both simultaneous and sequential bilinguals with varying levels of language dominance. The results showed rapid prediction based on gender-marked articles in an Italian experiment (which was conducted in both countries) and in a German experiment (which was only conducted in Germany). Although in the Italian experiment prediction was significantly delayed in the incongruent gender condition (in which there was a mismatch between German and Italian gender), participants were still able to anticipate the target referent after hearing the disambiguating article. In other words, while children were influenced by their knowledge of German when processing sentences in Italian, they exhibited prediction regardless of gender incongruency. Crucially, the efficiency of linguistic prediction was related to relative language proficiency, i.e. Italian-dominant children showed more anticipation and less interference

from German in an Italian task than German-dominant children. On the other hand, no gender congruency effect was found in the German experiment, in which mostly German-dominant children participated.

To summarize, the limited empirical evidence that is available suggests that bilingual children, like monolingual children and adults, are able to process grammatical gender in a fast and efficient manner (at least if the two languages have a gender system). Although language dominance appears to be an important factor, it still remains unclear how child L2 learners with lower proficiency may differ from simultaneous bilinguals, and to what extent bilingual children may be influenced by the linguistic properties of their L1.

2.2. Processing of number

The concept of number appears very early in life, as shown by studies reporting pre-linguistic understanding of number using manual-search tasks (Feigenson & Carey 2005; Li, Ogura, Barner, Yang & Carey 2009). In Li et al. (2009) infants were able to discriminate a “one” condition, in which a single object was presented, from a “more-than-one” condition, in which infants saw several objects of the same type. From the age of two, children are sensitive to the morphological marking of number (Corrêa, Augusto, Ferrari-Neto 2005) and they are able to distinguish singular from plural marking. Using a preferential looking paradigm, Kouider et al. (2006) found that English-speaking children were sensitive to double number marking on the verb and the article at the age of 24 months. They showed participants two pictures, one depicting a single novel object and one depicting several identical novel objects. When hearing ‘There are some blinkets’, 24-month-olds looked significantly longer to the target picture depicting several objects. In a second experiment, they found that only 36-month-olds (but not 24-month-olds) were sensitive to the singular/plural distinction when only the phonological marking of the noun was presented (‘Look at the blinkets’). With the same type of paradigm, Robertson and colleagues (2012) investigated processing of number information carried by the determiner in French-speaking two-year-olds. To do so, they manipulated the degree of information given by the determiner. In the uninformative condition, target and competitor pictures were matched in number so that the number marking of the determiner was not disambiguating. In the informative condition, target and competitor pictures mismatched in number so that children could rely on the number features of the determiner to anticipate the noun. Their results showed that toddlers looked longer at the target in an informative condition (in which the target could be identified by relying on the number marking on the determiner) than in an uninformative condition (in which the determiner did not reveal any

identifying information about the target). Although participants did not preselect the noun based on the article, they did show a facilitation effect of number agreement in processing the following noun.

Previous studies investigating agreement effects in monolingual adults consistently found a facilitation effect of number on subsequent language processing, both within the noun phrase and across phrases (i.e. between subjects and verbs). Within the nominal phrase, the number facilitation effect has been evidenced in lexical decision tasks using a grammatical priming paradigm, in which participants are presented with two dependent words (typically Determiner-Noun or Noun-Adjective) that either match or mismatch in number, while they have to perform a lexical decision task on the second element. Lukatela, Carello and Turvey (1987) reported an effect of number congruency between the possessive adjective and the noun in Serbo-Croatian speakers. Noun recognition was faster in congruent trials than in incongruent ones, suggesting that participants may have pre-activated the number feature of the noun on the basis of the preceding article, or that they experienced a facilitation effect of number agreement during subsequent processing. Similar results were found for French and Spanish (Cole & Segui 1994; Faussart, Jakubowicz, Costes 1999). Number agreement effects have also been evidenced in neuroimaging and electrophysiological studies, reporting a sensitivity to number agreement violations (Barber & Carreiras 2003, 2005; Carreiras, Carr, Barber & Hernandez 2010). ERP data have revealed that number is integrated both at the syntactic level and at the semantic level. Barber and Carreiras (2003, 2005) found that number agreement violations elicit a Left Anterior Negativity (LAN) as well as a P600, which respectively correspond to early morphosyntactic integration and syntactic reanalysis (Friederici 2002). They also found an N400, which is generally associated with semantic integration (*ibid*).

Considering L2 learners, the acquisition of number agreement generally appears to be rather unproblematic. Looking at production data, White and colleagues (2004) investigated number accuracy in late L2 learners of Spanish (L1 French and English) with varying proficiency levels. All groups performed at ceiling when computing number agreement between the determiner and the noun, with a minimum accuracy of 94.4% for the French low-proficiency group. When producing the Determiner-Noun-Adjective sequence, minimum accuracy was still 83.5%, showing that number agreement is acquired even at low proficiency levels.

Moreover, different studies have shown that adult L2 learners of Spanish (Lew-Williams & Fernald, 2009) and German (Hopp, 2012) are able to use number-marking on articles to rapidly distinguish between single versus multiple referents. In other words, similarly to native speakers, late L2 learners seem to be able to anticipate nouns based on the number feature of the determiner. Studies using electrophysiological methods also provide evidence for native-like processing of

number in L2 speakers, at least at the higher levels of proficiency (Gabriele, Fiorentino, Bañón 2013; Gillon-Dowens et al. 2010; Rossi et al. 2006). For example, Gabriele et al. (2013) tested English-speaking late learners of Spanish with a grammaticality judgment task addressing number violations between the noun and the adjective (*El banco es un edificio muy *siguros* ‘The bank is a building-MASC-SING very safe-MASC-PLUR’) while ERP responses were recorded. Results showed that participants who were proficient in Spanish displayed a native-like pattern both in the behavioural data (97% accuracy) and in the elicitation of the P600. Using Barber & Carreiras (2005)’s material, Gillon-Dowens et al. (2010) found a native-like biphasic LAN-P600 pattern in native English advanced L2 learners of Spanish. Consistent with this finding, Rossi et al. (2006) found a similar pattern for high-proficiency L2 learners of Italian and German.

While late learners can reach native-like processing of number, there seem to be two major modulating factors. First, effects of L1/L2 similarity have been reported. To investigate the effect of L1 characteristics on L2 processing, Gillon-Dowens et al. (2011) replicated their earlier study with native English learners of Spanish with Chinese learners of Spanish. This time, number violations elicited only the P600, showing no LAN. The authors argue that the different ERP patterns between the two groups may be due to L1 effects on the L2: contrary to English and Spanish, Chinese does not mark number through morphosyntactic agreement. Thus, the absence of the number feature in the L1 may affect the early syntactic processing of number violations.

A second factor that may modulate number processing is L2 proficiency, as studies focusing on beginner L2 learners have reported non-nativelike ERP patterns (Gabriele et al. 2013; Tokowicz & MacWhinney 2005). For example, Gabriele et al. (2013) investigated number processing in native English late learners of Spanish with varying levels of proficiency (low, intermediate and advanced). While advanced and intermediate learners exhibited a P600 response, only a marginally significant P600 effect was found in low-proficiency learners. Another study testing native English low-proficiency learners of Spanish reported no P600 response to number violations at all (Tokowicz & MacWhinney 2005).

From very early in life, number marking is used as a cue for predicting upcoming speech, thus facilitating language processing. Late L2 learners can reach native-like abilities, as shown in both behavioural and electrophysiological data. Yet, performance might be modulated by the degree of similarity between the L1 and the L2, as well as by L2 proficiency. Until now, the L2 literature has mostly focused on late L2 learners and their ability to acquire new linguistic features. However, it remains unknown how early bilingual children process number, and whether the L1 affects number processing in early L2 learners.

2.3. The discrepancy between gender and number

Although advanced L2 learners are able to reach native-like proficiency in the production, comprehension and online processing of both grammatical gender and number (White, Valenzuela, Kozłowska-Macgregor & Leung 2014; Gabriele et al. 2013; Gillon-Dowens, Vergara, Barber & Carreiras 2010; Gillon-Dowens, Guo, Guo, Barber & Carreiras 2011), grammatical gender appears to be more challenging to acquire than number.

Several studies have directly compared number processing to gender processing in L2 populations, but the results are mixed. For example, Tokowicz and MacWhinney (2005) investigated the production and online processing of grammatical gender and number agreement in native English late L2 learners of Spanish. They found that even though participants were less accurate with grammatical gender than with number in a grammaticality judgement task, their ERP patterns showed sensitivity to agreement violations with gender but not with number. The authors argued that their participants were implicitly sensitive to morphosyntactic agreement with a feature that is unique to their L2 (i.e. gender), but not with a feature that exists in both the L1 and the L2 but which is expressed differently in the two languages (i.e. number).

In contrast, other studies found an advantage for number over gender in L2 learners, both in behavioral tasks (White et al. 2004) and in online language processing (Gabriele et al. 2013; Gillon-Dowens et al. 2010; Lew-Williams & Fernald 2009; Hopp 2012). For example, capitalizing on the fact that articles in Spanish are marked for both grammatical gender and number, Lew-Williams & Fernald (2009) found that native English late L2 learners of Spanish made rapid use of number cues when identifying subsequent nouns, while they were not able to rely on grammatical gender cues unless the same pairs of articles and nouns were used throughout the experiment. Similarly, studying native English high-proficient L2 speakers of German, Hopp (2012) found more variable performance with gender processing than with number processing. While L2 speakers reliably anticipated nouns on the basis of the number marking on articles and adjectives, participants with sub-optimal performance in an offline gender assignment task were less likely to anticipate nouns on the basis of grammatical gender.

Moreover, several studies have compared ERP responses to gender and number agreement violations in bilingual adults. Testing native English late learners of Spanish, Gabriele et al. (2013) found no difference between gender and number for low-proficiency speakers, who showed only a marginal P600 effect for both gender and number agreement violations. The dissociation between number and gender appeared in the intermediate learners, who showed native-like ERP patterns for number but not for gender. In advanced learners, the advantage for number over gender was only

quantitative, in that the P600 was significantly weaker when processing gender as compared to number agreement violations. A similar pattern was observed in the native English high-proficiency L2 learners of Spanish tested by Gillon-Dowens and colleagues (2010), who showed a delayed LAN and a weaker P600 when processing agreement violations with grammatical gender but not with number.

Thus, although there is some conflicting evidence in the literature, most studies suggest that grammatical gender develops more slowly than number in L2 learners (White et al. 2004), which may lead to more non-nativelike processing of gender as compared to number, in particular at the lower levels of language proficiency (Gabriele et al. 2013).

This dissociation between gender and number has also been evidenced in the monolingual population, both in adults (Antón-Méndez, Nicol & Garrett 2002; Barber & Carreiras 2003, 2005; Carreiras 2010; Igoa, García-Albea & Sánchez-Casas 1999; Lukatela, Carello & Turvey 1987) and children (Dispaldro, Ruggiero & Scali 2015). Monolinguals tend to produce more gender errors than number errors (Antón-Méndez et al. 2002) and are slower to process gender than number (Lukatela et al. 1987). Barber & Carreiras's (2003, 2005) findings suggest that although early processes seem to be similar, differences arise in later processing. More specifically, they found that the second phase of the P600 response to agreement violations was larger for gender than for number. They argue that this difference in the ERP pattern reflects more complex reanalysis processes for gender than for number. The dissociation between the two features is also supported by neuroimaging data that showed that the processing of gender and number activate different areas of the brain (Carreiras et al. 2010).

It has been claimed that this dissociation between number and gender is related to different representations at the conceptual level (Antón-Méndez, Nicol, Garrett 2002; Igoa, García-Albea, Sánchez-Casas 1999). Grammatical gender (feminine vs masculine in Italian) is an abstract feature usually described as an intrinsic characteristic of the noun. Thus, grammatical gender is assumed to be stored at the level of the lexicon (Harris 1991). On the other hand, in many languages number (singular vs plural) is a morphosyntactic feature that is attached to the lexical stem and that refers to the quantity of the referent. Contrary to grammatical gender, number is not a property of a noun and can be retrieved from the referential context (Ritter 1993), making it a more meaningful feature for L2 learners.

In sum, although there are some mixed results in the literature, grammatical gender tends to be more challenging than number. This difference has been explained at the representational level, where gender is an arbitrary part of the lexicon and number is a morphosyntactic feature expressing

a semantic property of the referent. The current study aims to contribute to this debate by contrasting gender and number processing in early L2 learners of Italian.

3. Methods

3.1 Hypotheses and predictions

The present study addresses how Mandarin-Italian bilingual children use grammatical gender and number marking on articles to anticipate upcoming nouns, in comparison with monolingual Italian children. Whereas monolingual children have been shown to effectively rely on gender and number cues in spoken word recognition, to date very few studies have focused on bilingual children. The combination of Mandarin and Italian provides an interesting test case: in Italian prenominal articles are marked for gender and number, while Mandarin lacks both articles and grammatical gender (but not number).

Firstly, we hypothesize that predictive processing is less efficient for child L2 learners than for monolingual children, especially if they have to rely on a grammatical category that is absent in their L1. Since Mandarin does not have articles, we predict that bilinguals will be less likely to anticipate upcoming nouns on the basis of determiner features. In other words, in a condition in which there is a mismatch in grammatical gender and number between the target picture and its competitor, we expect that participants will start directing their gaze toward the target picture after the onset of the article and before the onset of the noun, but bilingual participants may be delayed in comparison with monolingual participants.

Secondly, we hypothesize that grammatical gender is a more challenging feature to acquire for L2 learners than number, due to greater reliance on arbitrary lexical knowledge. We therefore predict that bilingual children will be slower when anticipating nouns based on a gender as compared to a number cue, and we predict that the discrepancy between gender and number processing will be greater for bilinguals than for monolinguals. The fact that Mandarin lacks grammatical gender while it does have a conceptual notion and grammatical expression of number may also contribute to such an asymmetry.

Thirdly, hypothesizing that gender processing becomes more native-like as L2 proficiency increases, we predict that for bilingual children, prediction of nouns based on gender-marked articles will be influenced by their proficiency in Italian in terms of vocabulary knowledge.

3.2. Participants

We tested a total of 63 children ranging from eight to eleven years old. After excluding two monolinguals and four bilinguals (see the Analysis section for more details), our final sample included 32 Italian monolinguals ($M_{\text{Age}} = 9;7$, $SD = 1;1$, Range = 8;0 – 10;8) and 25 L2 learners of Italian who spoke Mandarin as their L1 ($M_{\text{Age}} = 9;9$, $SD = 1;2$, Range = 8;6 – 11;9).

Bilingual children were from an immigration background and were raised in Mandarin-speaking families living within the Chinese community in Milan. In this community, families typically speak only Mandarin at home, while children are first exposed to Italian when they enter school (between the age of three and six). Based on the school's information, all children had at least three years of exposure to Italian.

All participants attended an Italian school in Milan. Bilingual children also attended a Chinese heritage language school where they received literacy training in Mandarin Chinese on Saturdays during the scholastic year and on weekdays during two months in the summer break. Bilingual children took the Peabody Picture Vocabulary Test (PPVT-R; Stella, Pizzoli & Tressoldi 2000) in order to obtain a measure of their Italian proficiency. The mean standard score on this task was 85 ($SD = 15$)¹, indicating that our bilingual participants had substantially smaller vocabulary sizes in Italian in comparison to monolingual norms ($M = 100$, $SD = 15$).

3.3. Task and Materials

Participants took part in a visual world eye-tracking task to measure their ability to predict the upcoming noun on the basis of the gender and number marking of the article.

The auditory stimuli consisted of Italian sentences divided into three parts: introduction (*Adesso trova* 'Now find...'), definite article (*il* 'the-MASC-SING', *la* 'the-FEM-SING', *le* 'the-FEM-PLU'), and target noun. The sentences were recorded by a female native speaker of Italian. The audios were manipulated with the Audacity® software (Audacity Team 2020). We spliced the audio fragments in order to ensure that the introduction of the sentence (*Adesso trova il/la/le* 'Now find the') was always the same. The determiners *la*, *le* and *il* had a duration of 315 ms, 350 ms and 370 ms respectively. The noun always started exactly 750 ms after the determiner onset, so that in each

¹ Note that we had missing data on the vocabulary test for four children.

condition anticipatory eye movements could be observed before the onset of the noun. Background noise was removed.

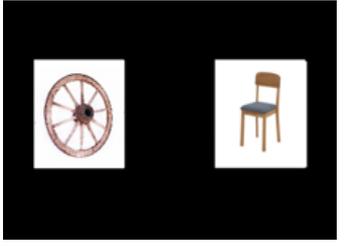
Sixty high-frequency nouns were selected, and their corresponding pictures were created. All nouns referred to inanimate concrete objects and were two or three syllables long. Pictures depicted simple objects on a white background, with either one single object or two identical objects. Targets and competitors were matched in syllable length.

Three conditions manipulated predictability and type of cue: Early Gender, Early Number, and Late (see Table 2). In both Early conditions, there was a mismatch between the target and the competitor, allowing for prediction on the basis of the article. In the Early Gender condition, pictures showed one feminine object and one masculine object. Thus, there was a gender mismatch so that the gender marking of the article (*il* vs *la*) could lead to anticipatory eye-movements towards the target. In the Early Number condition, pictures displayed one single object on one side and two identical objects on the other side. Thus, there was a number mismatch (with matched feminine gender) so that anticipatory eye-movements to the target could be generated from the number marking of the article (*la* vs *le*). In the Late condition, the target and the competitor were matched in both gender and number in three different types of pairs : *la* vs *la*, *il* vs *il* and *le* vs *le*. In this condition, the target could not be identified before the onset of the noun.

There were 10 items in each Early condition, and 15 items in the Late condition. Each noun was used only once in the experiment. Gender and number of the target were balanced within the corresponding condition so that five items targeted a feminine noun in the Gender condition, and five items targeted a plural noun in the Number condition. In the Late condition, five targets were masculine singular, five were feminine singular and five were feminine plural. The target appeared on the right side in 17 trials and on the left side in 18 trials. To avoid any preference bias towards one of the pictures, we created a second list of 35 items in which targets and competitors were swapped. The two lists were balanced across participants. The presentation order of the items was randomised.

Table 2. An overview of experimental conditions.

Condition	Picture	Audio
-----------	---------	-------

<p>Early Gender (predictable)</p>		<p>Adesso trova la barca Now find the-FEM-SING boat-FEM-SING 'Now, find the boat'</p> <p>Competitor: <i>il letto</i> the-MASC-SING bed-MASC-SING</p>
<p>Early Number (predictable)</p>		<p>Adesso trova la foglia Now find the-FEM-SING leaf-FEM-SING 'Now, find the leaf'</p> <p>Competitor: <i>le torte</i> the-FEM-PLUR cakes-FEM-PLUR</p>
<p>Late (unpredictable)</p>		<p>Adesso trova la ruota Now find the-FEM-SING ruota-FEM-SING 'Now, find the wheel'</p> <p>Competitor: <i>la sedia</i> the-FEM-SING chair-FEM-SING</p>

3.4. Procedure

Participants were tested individually in a quiet room with a portable computer. Data were collected with a Tobii Pro X3-120 eye-tracker that captured participants' gaze at 120Hz. Participants sat in front of the computer screen, with their eyes 60 to 70 cm away from the display. Participants underwent a short familiarisation phase consisting of three items, after which calibration occurred. Calibration consisted of 9 red dots that participants had to fixate.

Each trial was divided into two phases. The first one was a passive phase in which gaze data was collected. Participants faced a visual setting showing a pair of pictures while listening to an audio stimulus. The pair of pictures appeared one second before the sentence started, in order to give time for visual exploration and lexical retrieval. The trial ended one second after the end of the audio stimulus, so as to capture the whole gaze pattern. In the second phase, offline accuracy was collected. At the end of the trial, a question mark appeared, indicating to the participants that they

had to select the right picture by clicking on the mouse. Before moving on to the next trial, children saw a fixation cross; the experiment would only continue to the next item once they fixated on the middle of the screen. No feedback was given during the experimental phase.

3.5. Analysis

Data from four bilingual children were removed from the original sample because of poor calibration of the eye-tracker. We also excluded the data of two monolingual children who had more than 50% track-loss during the experimental phase. Thus, the analysis included 32 monolinguals and 25 bilinguals. The analysis was based on trials for which the offline response was accurate, excluding trials in which there was more than 35% track-loss. As a result, 161 trials were removed for the monolinguals (32 in the Early Gender condition, 60 in the Early Number condition and 69 in the Late condition), and 160 for the bilinguals (57 in the Early Gender condition, 42 in the Early Number condition and 61 in the Late condition), leaving respectively 959 and 711 trials for the statistical analysis.

For the purpose of our study, we created three time windows corresponding to the introduction (starting 500 ms before the onset of the article), the determiner (750 ms starting from the article onset until the noun onset) and the noun (ending 500 ms after the onset of the noun). The boundaries of each time window were shifted by 200 ms, as this is the estimated time required for the planning and execution of saccadic eye movements (Altmann 2011). We used the *eyetrackingR* (Dink & Ferguson 2015) and *ggplot2* (Wickham 2016) packages to visualize the eye gaze patterns in *R* (R Core Team 2019).

In the statistical analyses we compared looks toward the target versus competitor during the introduction and the determiner. Data were analyzed with a generalized linear mixed-effects model in which the odds of looking to the target served as the binary dependent variable (yes vs no), using the *glmer* function of the *lme4* package (Bates, Maechler, Bolker & Walker 2015) in *R*. In the first analysis, we compared bilinguals to monolinguals, to examine (1) whether they differ in the extent to which they use determiner features to predict upcoming nouns and (2) whether they differ with respect to gender versus number processing. We considered various predictors such as Time Window (Intro vs Determiner), Condition (Early Gender vs Early Number vs Late), Group (Monolingual vs Bilingual), Age, List and Item Order, as well as random intercepts for Item and

Subject, and random slopes for Time window and Condition. We set the contrasts as follows: for Time Window, Intro was coded as 0 and Determiner as 1, and for Group, Monolingual was coded as +1/2 and Bilingual was coded as -1/2. For Condition, we used a ternary contrast. When comparing the two early conditions against the late condition, Late was coded as -2/3, Early Gender was coded as +1/3 and Early Number was coded as +1/3. When comparing gender versus number, Late was coded as 0, Early Gender was coded as -1/2 and Early Number was coded as +1/2. All the numerical variables were rescaled and centered around the mean. The best model was selected through a bottom-up stepwise model comparison using the Bayesian Information Criterion (BIC).

In a follow-up analysis, we focused on the subset of bilinguals who completed the vocabulary test, to test whether they were able to anticipate nouns on the basis of the gender and number of articles, and to investigate the role of L2 proficiency in their ability to do so. This time, we considered main effects and interactions of Time Window, Condition (with Late coded as the baseline) and Italian Vocabulary scores, a main effect of Item Order, and random intercepts for Subject and Item with random slopes for Time Window and Condition. Again, numeric predictors were rescaled and centered around the mean.

4. Results

The mean offline accuracy on the task was 98.9% for bilingual participants ($SD = 10\%$) and 99.9% for monolingual participants ($SD = 3.7\%$). The analysis of the eye-tracking data includes data from accurate trials only.

The time course of target fixations during the sentence in the three different conditions is displayed in Figure 1 for monolinguals and in Figure 2 for bilinguals.

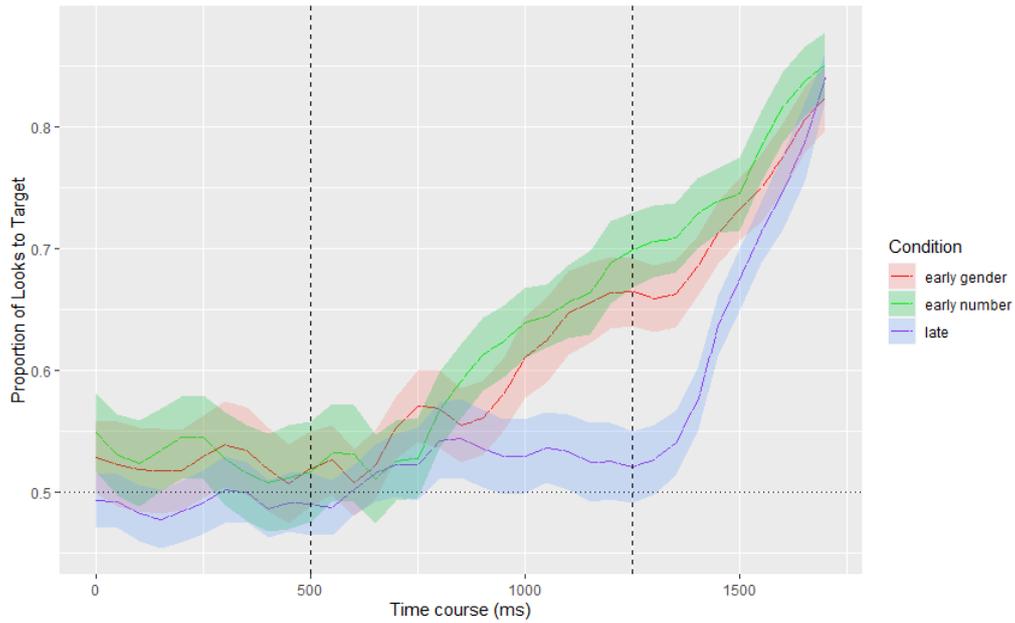


Figure 1. Time course of the proportions of looks toward the target (versus competitor) in the three conditions for monolingual participants. The first vertical line represents determiner onset and the second vertical line represents noun onset, shifted 200 ms to account for saccade planning.

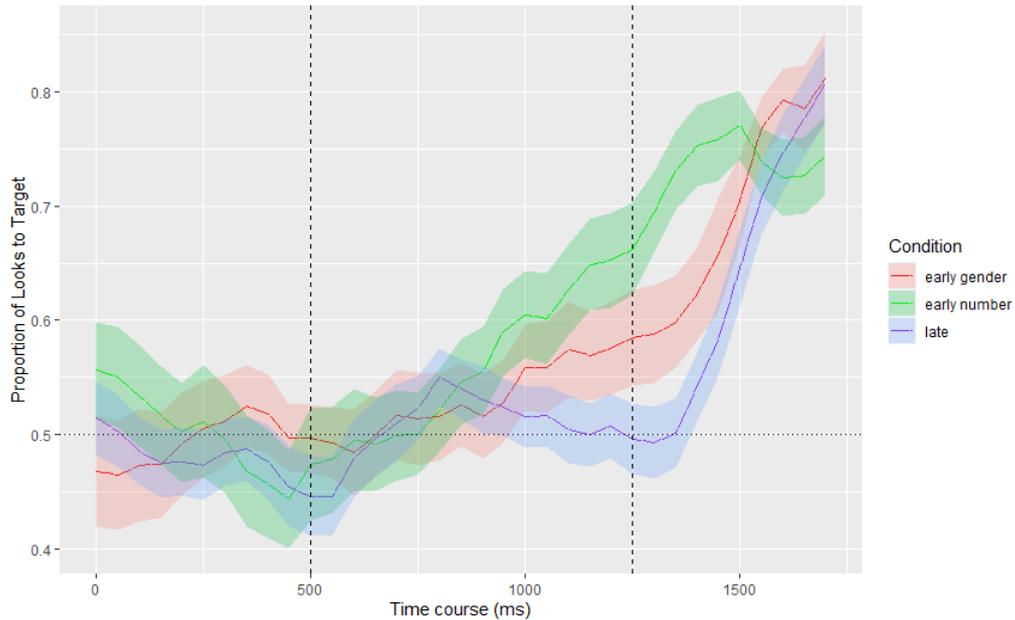


Figure 2. Time course of the proportions of looks toward the target (versus competitor) in the three conditions for bilingual participants. The first vertical line represents determiner onset and the second vertical line represents noun onset, shifted 200 ms to account for saccade planning.

As can be seen in Figure 1, for monolingual children the proportion of looks to the target picture started increasing during the article, before the onset of the noun, in the early gender and the early number condition, but not in the late condition. This suggests that monolinguals anticipated nouns on the basis of articles, with nearly identical patterns for gender and number. In contrast, as can be seen in Figure 2, bilinguals also showed anticipation of the target noun, but they seemed to be considerably slower to direct their gaze to the target picture in the gender condition than in the number condition.

The statistical analysis confirmed these patterns. The most predictive model that converged included a three-way interaction between Time Window, Condition and Group as well as their main effects, a main effect of Item Order, and random intercepts for Item and Subject with random slopes for Time window and Condition. Table 3 provides a summary of the model output.

Table 3. Estimated odds ratios, 95% confidence intervals and associated p-values of main and interaction effects for the comparison between monolingual and bilingual participants.

Generalized linear mixed model			
Looks to target (yes or no) ~ Time window (intro vs determiner) * Condition (early gender vs early number vs late) * Group (monolingual vs bilingual) + Item order + (1 Item) + (1 + Time window + Condition Subject)			
Fixed factor	Est. odds ratio	95% CI	<i>p</i>
Time window	1.20	1.1 .. 1.4	.002
Condition (early vs late)	1.17	.92 .. 1.5	.209
Condition (gender vs number)	1.08	.79 .. 1.5	.622
Group	1.09	.92 .. 1.3	.304
Item order	.933	.92 .. .94	<.001
Time window : Condition (early gender and early number vs late)	1.05	1.0 .. 1.1	.010
Time window : Condition (early gender vs early number)	.983	.94 .. 1.0	.481
Time window: Group			
Condition (early gender and early number vs late) : Group	1.08	.86 .. 1.4	.514
Condition (early gender vs early number): Group	1.15	.88 .. 1.5	.310
Time window : Condition (early gender and early number vs late) : Group	.858	.58 .. 1.3	.446
Time window: Condition (early gender vs early number): Group	1.02	.95 .. 1.1	.570
Group	1.31	1.1 .. 1.4	<.001

Overall, participants were more likely to look at the target picture during the determiner than during the introduction, as shown by the significant main effect of Time window. Item order also had a significant effect; participants became less likely to look at the target during the course of the experiment.

The significant interaction between Time window and Condition (Early vs Late) shows a prediction effect; the difference between the (predictable) Early conditions and the (unpredictable) Late condition was significantly greater during the article than during the introduction, indicating that participants anticipated the target noun on the basis of determiner features. We found no significant difference in prediction abilities between monolinguals and bilinguals, since the interaction between Time Window, Condition (early gender and early number vs Late) and Group

was not significant. When comparing early number to early gender, however, there was a significant three-way interaction between Time Window, Condition (early gender vs early number) and Group, indicating that the difference between gender and number processing was significantly greater for bilinguals than for monolinguals. In other words, overall we found a prediction effect, but bilingual participants were less efficient when processing gender as compared to number.

We then conducted a follow-up analysis, focusing on bilingual participants only, to test whether they were able to anticipate nouns based on gender- and number-marked articles, taking their Italian vocabulary knowledge into account. This analysis included only 21 bilingual participants, since we had missing data on the vocabulary test for four children. The model included main and interaction effects of Time Window, Condition and Vocabulary, a main effect of Item order, random intercepts for Subject and Item and random slopes for Condition and Time Window. The model output is provided in Table 4.

Table 4. Estimated odds ratios, 95% confidence intervals and associated p-values of main and interaction effects for bilingual participants only.

Generalized linear mixed model			
Looks to target (yes or no) ~ Time window (intro vs determiner) * Condition (early gender vs early number vs late) * Vocabulary + Item order + (1 Item) + (1 + Condition + Time window Subject)			
Fixed factor	Est. odds ratio	95% CI	<i>p</i>
Time window	1.19	.97 .. 1.5	.094
Condition (late vs early gender)	1.09	.70 .. 1.7	.707
Condition (late vs early number)	1.15	.73 .. 1.8	.549
Vocabulary	.977	.18 .. 1.2	.808
Item order	.887	.87 .. .90	<.001
Time window : Condition (late vs early gender)	.994	.92 .. 1.1	.872
Time window : Condition (late vs early number)	.961	.89 .. 1.0	.285
Time window: Vocabulary	.877	.72 .. 1.1	.200
Condition (late vs early number) : Vocabulary	.844	.62 .. 1.2	.302
Condition (late vs early gender): Vocabulary	1.06	.76 .. 1.5	.735
Time window : Condition (late vs early gender) : Vocabulary	1.33	1.2 .. 1.4	<.001
Time window: Condition (late vs early number): Vocabulary	.931	.87 .. 1.0	.054

In the second analysis we found no significant prediction effect for bilingual participants as a group (neither for gender nor number), since there was no significant interaction between Time Window and Condition. However, the degree to which bilinguals used gender-marked articles to anticipate nouns was influenced by their Italian vocabulary knowledge, as shown by the three-way interaction between Time Window, Condition (late vs early gender) and Vocabulary. Note that there was also a marginally significant three-way interaction between Time Window, Condition (late vs early number) and Vocabulary, suggesting that there may be a relation between L2 proficiency and prediction based on number, too.

As in the previous model, there was also a significant main effect of Item order, indicating that participants became less likely to look at the target during the course of the experiment.

5. Discussion and conclusion

This study addressed linguistic prediction on the basis of determiner features in Mandarin-Italian bilingual children. We aimed to contrast the processing of grammatical gender with number, since gender has been hypothesized to be particularly challenging for L2 speakers, specifically when this feature is absent in the L1. The bilingual participants were compared to a group of monolingual Italian children.

The first aim of this study was to investigate to what extent native Mandarin-speaking children who are learning Italian as their L2 use articles to anticipate upcoming nouns in Italian, in comparison with monolingual controls. We hypothesized that bilingual children use less predictive processing than monolingual children, especially when they need to attend to a grammatical category that is absent in their L1. Therefore, we predicted that Mandarin-Italian bilingual participants would be less likely to anticipate target nouns on the basis of articles than monolingual Italian participants. This prediction was not borne out by the data: we found no significant difference between bilinguals and monolinguals in the extent to which they were able to anticipate, providing no evidence for a general difference in predictive processing.

Our results suggest that, overall, participants were able to anticipate nouns based on morphosyntactic agreement with the article, complementing previous research showing that children process speech predictively (Cholewa et al 2019; Lew-Williams & Fernald 2007; Mani & Huettig 2012). However, in a post-hoc analysis focusing on the effect of language proficiency in a subset of bilingual children, we found no significant prediction effects. This result should be interpreted with caution since the post-hoc analysis only focused on a smaller sample, which may have led to insufficient statistical power. Previous research suggests that bilingual children are able

to process speech predictively (Lemmerth & Hopp 2019; Bosch & Foppolo submitted), and the gaze pattern of bilingual participants in the present study suggests a similar tendency, at least for number processing. Therefore, future studies that include a larger number of participants are needed to determine the extent to which child L2 learners are able to anticipate upcoming words when they have to attend to a grammatical category that is absent in their L2.

Our second aim was to compare the processing of grammatical gender and number in Mandarin-Italian bilingual children with that of their monolingual Italian peers. We hypothesized that there would be a discrepancy between gender and number processing for bilingual participants, since gender processing requires greater reliance on arbitrary lexical knowledge, and Mandarin lacks grammatical gender despite having a conceptual notion and grammatical expression of number. Our results lend support to this hypothesis. As predicted, the difference between gender and number processing was significantly greater for bilinguals than for monolinguals. Although this pattern was not statistically confirmed by our post-hoc analysis focusing on bilinguals only, bilingual children showed a tendency to be slower with grammatical gender than with number processing.

Previous studies have found a dissociation between gender and number in late L2 learners; while one study found faster processing of gender in comparison with number (Tokowicz & MacWhinney, 2005), other studies found that gender is more challenging for L2 speakers than number (Gabriele et al. 2013; White et al. 2004; Gillon-Dowens et al. 2010; Lew-Williams & Fernald 2009; Hopp 2012). Such an advantage for number over gender processing may be due to different representations at the conceptual level. Assuming that L2 learners with poorer language proficiency have more difficulties with lexical retrieval, they may experience more difficulties in gender processing than in number processing, since gender processing requires reliance on arbitrary lexical knowledge, whereas number is conceptually more tangible and strongly linked to the referential context. Moreover, L1 characteristics may play a role since most studies reporting a discrepancy between gender and number processing focus on L2 learners whose L1 does not have grammatical gender (typically, English).

Our third aim was to test the effect of L2 proficiency in terms of vocabulary knowledge. The results support our hypothesis that grammatical gender processing becomes more native-like as L2 proficiency increases. As predicted, the extent to which bilingual participants used the grammatical gender of the article to anticipate subsequent nouns was significantly influenced by their Italian vocabulary knowledge. This complements previous research showing that grammatical gender processing is related to language proficiency (Lew-Williams & Fernald 2010; Dussias et al. 2013; Gabriele et al. 2013; Bosch & Foppolo submitted).

The effect of proficiency also provides a potential explanation for the observation that the Mandarin-Italian bilinguals in our study appeared to be considerably slower when processing grammatical gender as compared to bilingual children in previous research. Other studies have found that as long as there is no interference due to cross-linguistic gender incongruency, bilingual children efficiently anticipate nouns based on grammatical gender (Lemmerth & Hopp 2019 in Russian-German bilinguals; Bosch & Foppolo submitted in Italian-German bilinguals). However, the German-Italian children tested by Bosch and Foppolo were attending bilingual schools and living in bilingual communities with strong ties to Italy, while the Mandarin-Italian children in our study were growing up in a relatively closed Chinese community, with considerably less exposure to Italian outside of school. As a result, our participants had lower L2 proficiency, as indicated by weaker vocabulary skills in Italian. Bosch and Foppolo's participants were tested with the same standardized vocabulary test, and the results showed that their Italian language proficiency ($M = 98$, $SD = 20$) was considerably higher than that of the Mandarin-Italian bilinguals in the present study ($M = 85$, $SD = 15$). Moreover, age of acquisition may play a role. While participants in Lemmerth and Hopp (2019) were either simultaneous or very early sequential bilinguals (age of onset before age three), like the majority of participants in Bosch and Foppolo (submitted), participants in the current study only started learning Italian when they entered kindergarten after the age of three.

In addition to child-level factors such as proficiency, age of acquisition and language use, the difference between our results and previous findings about gender processing in bilingual children may be due to the different characteristics of the L1. German, Russian and Italian all have grammatical gender, thus enabling positive transfer from the L1 to the L2. Mandarin, on the other hand, does not have grammatical gender, which means that no such transfer is possible. Consequently, grammatical gender in the L2 may be more difficult for Mandarin-speaking children than for children with a gender-marked L1.

At least three limitations should be acknowledged. Firstly, our sample size was relatively small, and follow-up studies are needed to confirm the patterns observed in the present study. Secondly, we are not able to tease apart the two possible explanations for a potential discrepancy between gender and number processing for bilingual participants. In other words, bilinguals showed a larger difference between gender and number processing than monolinguals, but we do not know whether this is due to L1 effects or to child-level factors such as proficiency. More research is needed to disentangle this by systematically examining the effect of individual differences in input and language proficiency in both the L1 and the L2, or by testing different language pairs. One possibility would be to repeat the current study with bilingual children who speak two languages that have both grammatical gender and number, as to test whether bilinguals differ from

monolinguals independently of L1 effects. Thirdly, even though we only included trials that were answered correctly, we cannot be certain if participants actually knew all the nouns and their grammatical gender. Future studies may therefore want to include a production task, in order to test participants' knowledge of the nouns used in the experiment.

Acknowledgements

We thank Francesca Foppolo for her contribution to the experimental design and analysis, Silvia Silleresi for her help in creating the experimental stimuli, Giulia Mornati for her involvement in implementing the task, Abby Deng and Antonella Guarascio for helping with data collection and data entry, and Carlo Toneatto for technical support. Finally, we are very grateful to the children, their families and the school for their participation.

Funding

This project has received funding from the European Union's Horizon2020 research and innovation programme under the Marie Skłodowska Curie grant agreement No 765556, as a part of the MultiMind project.

Author contributions

Jasmijn Bosch and Mathilde Chailleux designed and implemented the experiment, collected and analyzed the data and wrote the first draft of the chapter. Jia'en Yee advised about Mandarin Chinese and helped with the communication with the parents, and Fabrizio Arosio was involved in the design of the study. All authors commented on the final draft.

References

- Altmann, G. T. (2011). The mediation of eye movements by spoken language. In *The Oxford Handbook of Eye Movements*, S. Liversedge, I. Gilchrist & S. Everling (eds), pp. 979-1004. Oxford University Press.
- Altmann, G. T. & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition* 73(3): 247-264.
- Antón-Méndez, I., Nicol, J. L. & Garrett, M. F. (2002). The relation between gender and number agreement processing. *Syntax* 5(1): 1-25.
- Audacity Team (2020). Audacity(R): Free Audio Editor and Recorder [Computer application].
- Ayoun, D. & Maranzana, S. (2020). Italian grammatical gender: a corpus study from a second language acquisition perspective. *Italica* 97(2): 306-333.
- Barber, H. & Carreiras, M. (2003). Integrating gender and number information in Spanish word pairs: An ERP study. *Cortex* 39(3): 465-482.
- Barber, H. & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: an ERP comparison. *Journal of cognitive neuroscience* 17(1): 137-153.
- Bates, D., Maechler, M., Bolker, B. & Walker S. (2015). Fitting Linear Mixed Effects Models Using lme4. *Journal of Statistical Software* 67(1): 1-48.
- Bosch, J.E. & Foppolo, F. (submitted). Predictive processing of grammatical gender in bilingual children: The effect of cross-linguistic incongruence and language dominance.
- Carreiras, M., Carr, L., Barber, H. A. & Hernandez, A. (2010). Where syntax meets math: right intraparietal sulcus activation in response to grammatical number agreement violations. *NeuroImage* 49(2): 1741-1749.
- Chini, M. (1995). Grammatiche a confronto: la categoria grammaticale del genere nella competenza di nativi italofofoni e nelle interlingue di apprendenti dell'italiano come L2. In *L'universo delle lingue. Confrontare lingue e grammatiche nella scuola*, P. Desideri (ed.), pp. 277-294. Firenze: Quaderni del Giscel, La Nuova Italia.
- Cholewa, J., Neitzel, I., Bürsgens, A. & Günther, T. (2019). Online-processing of grammatical gender in noun-phrase decoding: An eye-tracking study with monolingual German 3rd and 4th graders. *Frontiers in Psychology* 10: 2586.
- Cole, P. & Segui, J. (1994). Grammatical incongruency and vocabulary types. *Memory & Cognition* 22(4): 387-394.
- Corbett, G. G. (2000). *Number*. Cambridge textbooks in linguistics.
- Corrêa, L. S., Augusto, M. R. & Ferrari-Neto, J. (2005). The early processing of number agreement

- in the DP: evidence from the acquisition of Brazilian Portuguese. Talk given at the 30th Boston University Conference on Language Development, Boston.
- Dahan, D., Swingle, D., Tanenhaus, M. K. & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of memory and Language* 42(4): 465-480.
- Dink, J. W. & Ferguson, B. (2015). EyetrackingR: An R library for eye-tracking data analysis. Available at www.eyetracking-r.com
- Dispaldro, M., Ruggiero, A. & Scali, F. (2015). Real-time comprehension of gender and number in four-to seven-year-old children: a study of the relationship between Italian clitic pronouns and visual picture referents. *Journal of Child Language* 42(4): 917-931.
- Dussias, P. E., Kroff, J. R. V., Tamargo, R. E. G. & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition* 35(2): 353-387.
- Faussart, C., Jakubowicz, C. & Costes, M. (1999). Gender and number processing in spoken French and Spanish. *Rivista di Linguistica* 11(1): 75-101.
- Fedden, S., & Corbett, G. G. (2017). Gender and classifiers in concurrent systems: refining the typology of nominal classification. *Glossa* 2(1): 1-47.
- Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition* 97(3): 295-313.
- Ferrari, A. & Zampese, L. (2016). *Grammatica: parole, frasi, testi dell'italiano* (Vol. 173). Carocci.
- Foucart, A. & Frenck-Mestre, C. (2011). Grammatical gender processing in L2: electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism: Language and Cognition* 14(3): 379-399.
- Foucart, A. & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, 66(1), 226-248.
- Frankowsky, M. & Ke, D. (2016). Humanness and classifiers in Mandarin Chinese. *Language and Cognitive Science* 2(1): 55-67.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences* 6(2): 78-84.
- Gabriele, A., Fiorentino, R. & Bañón, J. A. (2013). Examining second language development using event-related potentials: a cross-sectional study on the processing of gender and number agreement. *Linguistic Approaches to Bilingualism* 3(2): 213-232.
- Gillon-Dowens, M., Vergara, M., Barber, H. A. & Carreiras, M. (2010). Morphosyntactic processing in late second-language learners. *Journal of Cognitive Neuroscience* 22(8): 1870-1887.

- Gillon-Dowens, M., Guo, T., Guo, J., Barber, H. & Carreiras, M. (2011). Gender and number processing in Chinese learners of Spanish—Evidence from Event Related Potentials. *Neuropsychologia* 49(7): 1651-1659.
- Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: a production or a real-time processing problem?. *Second Language Research* 28(2): 191-215.
- Guillelmon, D., & Grosjean, F. (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory & cognition* 29(3): 503-511.
- Haspelmath, M. (2018). Toward a new conceptual framework for comparing gender systems and some so-called classifier systems. Talk presented at Stockholm University, Department of Linguistics on April, 13, 2018.
- Harris, J. W. (1991). The exponence of gender in Spanish. *Linguistic inquiry* 22(1): 27-62.
- Hopp, H. (2012). The On-line Integration of Inflection in L2 Processing: Predictive Processing of German Gender. In Proceedings of the 36th annual Boston University Conference on *Language Development*, A.K. Biller, E. Y. Chung & A. E. Kimball (eds.), pp. 226-245. Somerville: Cascadilla Press.
- Hopp, H. & Lemmerth, N. (2018). Lexical and syntactic congruency in L2 predictive gender processing. *Studies in Second Language Acquisition* 40(1): 171-189.
- Huang, C. R. & Ahrens, K. (2003). Individuals, kinds and events: classifier coercion of nouns. *Language Science* 25(4): 353–373.
- Igoa, J. M., García-Albea, J. E. & Sánchez-Casas, R. (1999). Gender-number dissociations in sentence production in Spanish. *Rivista di Linguistica* 11(1): 163-196.
- Iljic, R. (1994). *Quantification in Mandarin Chinese: Two markers of plurality*. Berlin: Mouton de Gruyter.
- Kouider, S., Halberda, J., Wood, J. & Carey, S. (2006). Acquisition of English number marking: the singular-plural distinction. *Language Learning and Development* 2(1): 1-25.
- Lemmerth, N. & Hopp, H. (2019). Gender processing in simultaneous and successive bilingual children: cross-linguistic lexical and syntactic influences. *Language Acquisition* 26(1): 21-45.
- Lew-Williams, C. & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science* 18(3): 193-198.
- Lew-Williams, C. & Fernald, A. (2009). Fluency in using morphosyntactic cues to establish reference: How do native and non-native speakers differ? In *Proceedings of the 33rd annual Boston University Conference on Language Development*, J. Chandlee, M. Franchini, S. Lord & G-M. Rheiner (eds), pp. 290-301. Somerville: Cascadilla Press.

- Lew-Williams, C. & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language* 63(4): 447-464.
- Li, P., Ogura, T., Barner, D., Yang, S. J. & Carey, S. (2009). Does the conceptual distinction between singular and plural sets depend on language? *Developmental Psychology* 45(6): 1644-1653.
- Lukatela, G., Carello, C. & Turvey, M. T. (1987). Lexical representation of regular and irregular inflected nouns. *Language and Cognitive Processes* 2(1): 1-17.
- Mani, N. & Huettig, F. (2012). Prediction during language processing is a piece of cake – but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance* 38(4): 843-847.
- Morales, L., Paolieri, D., Dussias, P. E., Kroff, J. R. V., Gerfen, C. & Bajo, M. T. (2016). The gender congruency effect during bilingual spoken-word recognition. *Bilingualism: Language and Cognition* 19(2): 294-310.
- Padovani, R. & Cacciari, C. (2003). Il ruolo della trasparenza morfologica nel riconoscimento di parole in Italiano. *Giornale italiano di psicologia* 30(4): 749-772.
- Panzini, A. (2017). *Grammatica italiana*. Edizioni Lindau.
- R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Ritter, E. (1993). Where's gender? *Linguistic inquiry* 24(4): 795-803.
- Robertson, E. K., Shi, R. & Melançon, A. (2012). Toddlers use the number feature in determiners during online noun comprehension. *Child Development* 83(6): 2007-2018.
- Rossi, S., Gugler, M. F., Friederici, A. D. & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: evidence from event related potentials. *Journal of cognitive neuroscience* 18(12): 2030-2048.
- Rullmann, H. & You, A. (2006). General number and the semantics and pragmatics of indefinite bare nouns in Mandarin Chinese. *Where semantics meets pragmatics*, 175-196.
- Sá-Leite, A. R., Fraga, I. & Comesaña, M. (2019). Grammatical gender processing in bilinguals: An analytic review. *Psychonomic Bulletin & Review* 26(4): 1148-1173.
- Stella, G., Pizzoli, C. & Tressoldi, P. E. (2000). *Il Peabody Test. Test di vocabolario recettivo*. Torino: Omega.
- Tokowicz, N. & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in Second Language Acquisition* 27(2): 173-204.
- Van Heugten, M. & Shi, R. (2009). French-learning toddlers use gender information on

- determiners during word recognition. *Developmental Science* 12(3): 419-425.
- White, L., Valenzuela, E., Kozłowska–Macgregor, M. & Leung, Y. K. I. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics* 25(1): 105-133.
- Wicha, N. Y., Moreno, E. M. & Kutas, M. (2004). Anticipating words and their gender: an event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of cognitive neuroscience* 16(7): 1272-1288.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.
- Zhang, N. N. (2014). Expressing number productively in Mandarin Chinese. *Linguistics* 52(1): 1-34.