# THE ACQUISITION OF GENDER IN A ROMANIAN-HUNGARIAN BILINGUAL SETTING

**Veronica Tomescu**[*]

**Abstract:** This paper analyses the acquisition of Romanian gender agreement in a Romanian-Hungarian bilingual setting, based on two longitudinal corpora and a corpus of narratives, with a view to identifying the causes that lead to the vulnerability of the gender feature in this particular language combination. The fact that Hungarian is a genderless language causes some delay in the acquisition of Romanian gender with bilinguals. While phonological and semantic transparency do not seem to have had much influence, it has been found that agreement at a distance represents an obstacle, probably due to processing difficulties inherent to bilingual language acquisition. The two main facilitating factors have been found to be adjacency to the noun and the presence of the definite and indefinite articles.

**Keywords:** gender, Romanian-Hungarian bilinguals, phonological and semantic transparency, processing difficulties.

## 1. Introduction

The aim of the paper is to determine what makes gender a vulnerable feature in bilingual language acquisition, in a Romanian-Hungarian bilingual setting. Previous research has stressed the importance of the quantity of input, of the combination of languages involved, as well as of the degree of morpho-phonological, semantic and syntactic transparency of the language considered. This study analyses the acquisition of the Romanian gender feature in a Romanian-Hungarian bilingual setting, on the basis of two longitudinal corpora of spontaneous utterances, as well as a collection of narratives produced by a group of kindergarten age Romanian-Hungarian bilinguals. In view of the fact that Hungarian has no grammatical gender, not even with personal pronouns, it might be expected that this could cause some delay in the acquisition of gender agreement in the other language, especially with unbalanced bilinguals. Since Romanian nouns are phonologically transparent to a certain extent as regards gender, and [+animate] nouns also exhibit some degree of transparency, the corpora were examined with a view to establishing whether either phonological or semantic transparency might have had a facilitating effect. Two other important factors are adjacency to the noun of the category agreeing with it, as well as the presence of the definite or indefinite article which might prompt correct gender agreement.

The organization of the paper is as follows: section 2 presents an overview of the literature on gender acquisition, as well as a brief description of Romanian gender; section 3 describes the corpora, section 4 contains the analysis proper, followed by a discussion section.

[*] [1] University of Bucharest, Department of English, veronica.tomescu@lls.unibuc.ro

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

## 2. Previous research on the acquisition of gender

### 2.1 The acquisition of gender

According to Carstens (2010), the gender feature is an anomaly with respect to Chomsky's (2001) classification of features: interpretable and valued vs. uninterpretable and unvalued. Uninterpretable features have to be eliminated before the Conceptual-Intentional Interface; they are valued by means of an Agree relation and subsequently deleted. This is what happens in the case of person and number. But gender is semantically arbitrary, hence a valued feature, part of the lexical entry of the noun. Carstens proposes that it need not be deleted by valuation, it is recognized as uninterpretable and ignored. This allows it to be available in successive Agree relations.

Hawkins and Franceschina (2004) consider gender features parametrized features that are made available by UG and may or may not be activated in a specific grammar. Thus, English (or Hungarian) does not select gender features, while Romance languages do. Learners need positive evidence in order to incorporate the gender feature into their grammars. According to Caramazza et al. (2001), the selection of grammatical features is the automatic consequence of the selection of a lexical node; meaning that the gender feature becomes available as soon as the noun has been selected. Other categories are not lexically marked for gender and therefore enter the derivation unmarked. Gender with categories other than nouns is a syntactic feature resolved at a later point in the derivation, like number (Cantone and Müller 2007).

Acquisition of gender in L1 is by all acounts fast and errorless in Romance, but poses some difficulty in Germanic, especially Dutch (van der Linden and Hulk 2009, Cornips et al. 2006 etc.). As regards bilinguals, different results have been reported by various authors. Research in bilingual language acquisition (Costa et al. 2003, Cantone and Müller 2008) has shown that, at least with highly proficient bilinguals, the gender systems of the two languages are autonomous. That being said, it is also evident from the data that bilinguals perform differently from monolinguals as regards the acquisition of gender.

Both delay and acceleration have been reported regarding the acquisition of gender by bilinguals; the factors leading to the differences are the importance of input – unbalanced bilinguals might do worse in the weaker language, the influence of the other language – the language combination is not irrelevant, morpho-phonological, semantic or syntactic transparency. Also, bilinguals may encounter computational, processing difficulties.

Difficulties have been recorded in a Basque-Spanish and Spanish-English setting in Montrul (2004); in a French-Swedish bilingual context in Granfeldt et al. (2007); with the acquisition of the Dutch gender feature in Cornips et al. (2006), Unsworth (2007); Schwartz et al. (2015) investigate delay in the acquisition of Russian gender with various groups of bilinguals, whose other language is English, Finnish, German or Hebrew; French gender is acquired later in a Dutch-French bilingual context according to Hulk and van der Linden (2009) (although Dutch gender is acquired sooner). Kuchenbrandt (2005) reports that agreement within the DP comes at a later stage in a Spanish-German setting than in a monolingual setting, although gender itself does not seem to be directly

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

affected. Kupisch et al. (2002) find that Italian was acquired faster in an Italian-German context than French in a French-German context, but it might have been due to individual differences; additionally, the number of children studied was very small.

On the other hand, French seems to have had a positive influence on the acquisition on Dutch gender in the case of Dutch-French, Dutch-Spanish bilinguals (Hulk and van der Linden 2009). Similarly, no difficulties were encountered by the Spanish-English bilinguals in Silva-Corvalan (2014).

There has been some debate in the literature whether the masculine could not be considered a default gender in acquisition (Hawkins and Franceschina 2004, Alarcon 2011) – default being defined as the option used in the absence of agreement (Tsimpli and Hulk 2013). Bruhn de Garavito and White (2002) find that usually it is the masculine which is overgeneralized, but some speakers show a preference for the feminine. According to Greenberg (1966 in Arias-Trejo et al. 2103), the masculine is unmarked, hence easier to acquire; Montrul (2004) also reports that the masculine gender in Spanish is acquired sooner than the feminine by Spanish-English bilinguals. By contrast, López-Ornat (1997 in Arias-Trejo et al. 2013) and Pizutto and Caselli (1992 in Arias-Trejo et al. 2013) find that there are fewer errors with the feminine due to its phonological simplicity.

### 2.2 Factors influencing delay and acceleration in a bilingual setting

The role of input and language dominance has been emphasised in Ter Avest and Mulder (2009 in van der Linden and Hulk 2009), Cornips and Hulk (2008), van der Linden (2009), Unsworth et al. (2014), Rodina and Westergaard (2013a, 2013b), Montrul (2004). As regards the nature of the two languages that are paired, van der Linden and Hulk (2009), Hulk and Cornips (2006) report acceleration with Heerlen-Dutch bidialectals in contrast with delay with bilinguals; van der Linden (2009) affirms that Romance-Germanic combinations are helpful, for instance Dutch gender is more easily acquired in the context of French-Dutch bilingualism; Eichler et al. (2013) on the other hand consider that a combination between a three- and a two-gender system will result in delay in the three-gender system. Schwartz et al. (2015) show that bilinguals whose other language does have a gender feature outperformed bilinguals with no gender in their other language.

Another factor is morpho-phonological transparency. Several authors discuss the facilitating effect of the transparency of the gender system in languages such as Italian (Kupisch et al. 2002), Russian (Rodina and Westergaard 2013a, 2013b), Greek (Unsworth et al. 2014) or Spanish (Arias-Trejo 2013). Brehmer and Rothweiler (2012) show that children rely heavily on morpho-phonetic clues; with non-transparent nouns they either manipulate the form of the noun or guess wrong. Tsimpli and Hulk (2013) give the example of Dutch nouns derived by the diminutive suffix in Dutch, which are always neuter.

Against this, however, Bates et al. (1995 in Tsimpli and Hulk 2013) argue that phonological cues are only helpful at the post-lexical level, with unknown words or borrowings. Tsimpli and Hulk (2013) also quote Miozzo and Caramazza (1997) in stating that gender is activated pre-lexically, before the complete phonological representation of the word is accessed. Hawkins and Franceschina (2004) stress the importance of

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

phonological cues in the early stages of acquisition, but older children, when unsure of the gender of the noun, use a default masculine or employ other means (analogy). Müller (1994) also considers that phonological cues are of some relevance with very young children, before the age of 2;0, when the gender feature is in fact not yet available; she reports the case of a child who assigns gender on the basis of surface rhyming. Rhyming agreement is also mentioned as being helpful in Russian in Voeikova (2013).

As regards semantic transparency, Cornips et al. (2006) find that the gender of [+animate] nouns appears to pose less of a problem for bilinguals. Semantic transparency is also mentioned in Brehmer and Rothweiler (2012): semantic cues appear to overrule morpho-phonetic ones.

Of greater importance seem to be syntactic cues, such as the definite article in French (Taft and Meunier 1998 in Tsimpli and Hulk 2013), the indefinite article in Spanish (Arias-Trejo et al. 2013). Arias-Trejo et al. (2013) also stress the importance of reiterative marking of gender in Romance.

Certain asymmetries have been recorded in the acquisition of gender agreement with various categories: for instance, Arias-Trejo et al. (2013) report more difficulties with the definite rather than the indefinite article in Spanish. By contrast, in both German and Spanish, more errors were reported with the indefinite rather than the definite article in Müller (1994), Bruhn de Garavito and White (2002). In a study by Cornips et al. (2006), with both L1 and 2L1 Dutch children, a difference was found between agreement with common and neuter nouns across morphological categories: in the case of common nouns it is agreement with adjectives which exhibits the greatest degree of accuracy, outperforming relative pronouns and definite determiners, while with neuter nouns, it is determiners which are used more accurately than adjectives.

Alarcón (2011) holds that gender errors with bilinguals are sometimes the result of computational, processing difficulties; in this study bilinguals perform more accurately in the comprehension than in the production task. It is argued here that as regards competence there is no difference between L2, 2L1 or L1: it is performance which differs. Gender agreement is acquired earlier with attributive rather than predicative adjectives, and there is greater accuracy with determiners than with adjectives as regards gender agreement (Bruhn de Garavito and White 2002, Alarcón 2011). By contrast, no difference was found between article and adjective gender agreement in Kuchenbrandt (2005).

### 2.3 Romanian gender

Romanian nominally has a three-gender system (masculine, feminine, neuter), but in fact only two gender forms are available for functional categories and adjectives, since what is called the "neuter" gender has a masculine singular and a feminine plural. In Romanian, it is adjectives, determiners and quantifiers that have to agree with the noun, as well as the participle of the passive verb.

As regards semantic transparency, it exists to a certain extent in the case of [+animate] nouns since the gender of the noun will usually match the gender of the entity denoted (e.g. *băiat* M 'boy', *fată* F 'girl', *iapă* F 'mare', *motan* M 'tomcat'). Nouns denoting various animal species however are arbitrarily masculine (e.g. *cal* 'horse') or

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

feminine (e.g. *pisică* ‘cat’). Nouns denoting professions also usually come in pairs, according to the gender of the individual (*profesor* M – *profesoară* F ‘teacher’), but with some the gender is not transparent and the word may apply to both men or women (părinte M ‘parent’, rudă F ‘relative’, star N ‘star, e.g. singer or actor’).

There is also phonological transparency, since nouns ending in *–ă* are always feminine (e.g. *apă* ‘water’). Nouns ending in consonants always have a masculine singular, but may be neuter and require a feminine plural (compare the masculine *stup* ‘hive’ and the neuter *borcan* ‘jar’). Nouns ending in *-e* are not transparent (e.g. *perete* M ‘wall’; but *urechе* F ‘ear’), and neither are plurals, where some homonymy is noticeable between plural endings across gender forms (e.g. *pereți* M PL ‘walls’, *urechi* F PL ‘ears’). The Romanian definite article is a clitic on the noun; this means that any gender errors committed in the case of the definite article will lead to a distortion of the noun. Additionally, the input can be confusing in the sense that two of the forms in the paradigm are homonymous, as is apparent from Table 1.

**Table 1.**
**Romanian definite articles.**

| Definite article | Masculine | Feminine |
|---|---|---|
| Singular | -l, -le | -a |
| Plural | -i | -le |

### 3. Description of the data

The study uses two longitudinal corpora (ages 1;10-3;0 and 1;10-2;8), and a corpus of narratives produced by kindergarten age children; both the two children in the longitudinal corpora and the group of kindergarten children are simultaneous Romanian-Hungarian bilinguals, living in Bucharest.

#### 3.1 The longitudinal corpora

The longitudinal corpora consist in recordings of naturalistic, non-structured conversations in a Romanian-Hungarian bilingual context (Tomescu 2013, 2017b) . The children were recorded 30 to 60 minutes per week between the ages 1;10 – 3;0 for the eldest, Toma, and 1;10 – 2;8 for the youngest, Petru. The two children are brothers living in a Romanian community; they speak Hungarian to their mother, older brother and maternal grandmother; they were also exposed to Hungarian by means of stories read aloud to them by their mother. Their father and the other family members are Romanian monolinguals. They are unbalanced bilinguals with Hungarian their weaker language, as shown by the graphs in the Appendix (Figures 1-4), representing the MLU for both children, as well as the number of Romanian/Hungarian utterances per recording. The number of Hungarian utterances as well as the MLU is lower for Hungarian with both children.

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

An overview of the corpora (see also Tomescu 2017a) has found that gender errors occur with adjectives and with all functional categories that have gender agreement, except – oddly enough – the indefinite article.

After an initial period in which the masculine definite article is overgeneralized – which is a natural stage in evolution with Romanian monolinguals as well and for other languages (Müller 1994, Avram 2001) – there are no more errors with the indefinite article. However, the early overgeneralization of the masculine is not a gender error: the child uses a non-adult form, possibly a numeral, as replacement (Avram 2001). More importantly, there are no overextensions of the feminine indefinite article in the two bilingual longitudinal corpora. The feminine indefinite article is first attested at 2;2 with both children and by the age of 2;4 we have seen the last of the overextensions of the masculine.

**Table 2.**
**Overextensions of the masculine indefinite article. Age 2;2-2;3**

| Toma | Petru |
|---|---|
| 23% (29/126) | 6% (5/88) |

Adjectives are correctly used by 3;0 and 2;4 respectively. For a while longer, however, gender errors can still occur on occasion with functional categories: pronouns, clitics (e.g. 1).

(1) *mama face cu **a lui** (Toma 5;2)
mother makes with **his**
Intended: 'Mother uses hers.'

A summary of the gender errors (both feminine and masculine) in the longitudinal corpus is given in Table 3. The percentage of gender errors with various categories out of the total contexts containing the respective category is given. For example, out of the 556 adjectives that are variable for gender in the Toma corpus, 152 (27%) bear the incorrect gender agreement (masculine or feminine). The numbers for the 3rd person possessives (in fact the genitive form of the personal pronoun paradigm) are given separately from the 1st and 2nd person possessives (part of the possessive pronoun paradigm). With the 3rd person genitive personal pronoun, gender agreement is with the possessor (2a), while with the possessive pronoun[2] gender agreement is with the object possessed (2b).

(2) a. era mizeria aia din patul **lui** (Petru 2;3)
was mess-the that of bed-the his
'it was the mess in his bed'

[2] The 3rd person of the possessive pronoun paradigm is rather formal and does not appear in child utterances or child directed speech; gender agreement is with the object possessed: e.g. *patul M său M* bed-the his/her; *barca F sa F* his/her boat

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

b. asta să fie barca **mea** (Toma 2;5)
this SĂ be boat-the.F my.F
'this will be my boat'

**Table 3.**
**Summary of gender errors, longitudinal corpora**

| | **Toma (1;10-3;0)** | **Petru (1;10-2;8)** |
|---|---|---|
| | % (gender errors / total contexts) | |
| Adjectives (variable) | 27% (152/556) | 23% (73/316) |
| Acc Clitics (3rd) | 23% (127/547) | 11% (14/133) |
| Possessives (1st, 2nd) | 10% (5/48) | 4% (3/85) |
| Possessives (3rd) | 0/20 | 15% (2/13) |
| Quantifiers | 27% (32/120) | 4% (4/101) |
| Personal pronouns (3rd) | 44% (8/18) | 0 |

The difference between the children might be due to the fact that Toma is a more balanced bilingual than his younger brother. Petru's stronger language is Romanian, hence – possibly – fewer gender errors.

### 3.2 The corpus of narratives

In addition, the study uses the narratives produced by 19 kindergarten age Romanian-Hungarian bilinguals (age range: 3;3 – 5;10, mean age 53 months) from Bucharest. The narratives are based on Mercer Mayer's *Frog Where Are You*? A few of the children produced additional utterances, not related to the storybook, before or after narrating the story of the boy and the frog. Some engaged in conversation with the investigator on other topics, while some were reluctant to narrate the story because they did not seem to like the pictures; after a brief session of warm-up, however, they condescended to carry out the task as well. These additional utterances were also included in the analysis. 14 (15[3]) out of the 19 children, up to the age of 5;6, committed a total of 74 gender errors, with adjectives, possessives, indefinite articles, personal pronouns, accusative clitics, quantifiers, definite articles, examples (3). A summary is given in Table 4.

[3] One of the girls speaks three languages – she is Russian-Hungarian, but her nanny is Romanian and she is rather fluent; she produced a feminine agreement on a participle (i) (agreeing with a feminine noun actually), which is not at all expected in Romanian. She is also one of the children who produced a distorted noun with the feminine instead of masculine plural definite article (**papucile* instead of the masc.pl. *papucii*). She had no other gender errors. Her being trilingual however led to her being eliminated from all the percentages and totals in the paper – the number of children who committed gender errors is 15 if we include her.

i. a găsite (Vera 3;6)
has found F
Intended: 'He found them.' (the frogs = broaştele F)

BDD-A28524 © 2017 Universitatea din Bucureşti
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

(3) a. şi asta e foarte *roşu (maşina) (Eva 3;11)
also this.F is very red.M (car-the.F)
'this one is also very red'

b. (albina) …să îl ţeapă cu acul *lui (Norbi 4;6)
(bee-the.F) SĂ him sting with needle-the his
Intended: 'to sting him with its sting'

c. are *un pălărie (Maria 3;4)
has a.M hat F

d. s -au speriat că *el nu mai face nimic (musca) (Mark 5;6)
REFL have scared that he not more does nothing (the fly F)
'they were scared that it wouldn't do anything'

e. şi a prins*-o băiatu(l) (căţelul) (Benedek 5;6)
and has caught her boy-the (doggie-the.M)
Intended: 'and the boy caught him'

f. *doi maşini (Eva 3;11)
two M cars F

g. furnici*-i (Oli 4;6)
ants.F -the.M

**Table 4.**
**Summary of gender errors, kindergarten corpus.**

| Age | A | possessives | personal pronouns | clitics | Q | indefinite articles | definite articles | Total |
|---|---|---|---|---|---|---|---|---|
| Otto 3;3 | 2 | 2 | | | | 4 | | **8** |
| Maria 3;4 | | | | | | 1 | | **1** |
| Dominik 3;6 | | | 1 | | | 2 | | **3** |
| Arpi 3;7 | | | 1 | 6 | | | | **7** |
| Eva 3;11 | 5 | | | | 1 | | 2 | **8** |
| Evelin 4;0 | | | | | 1 | | | **1** |
| Oli 4;6 | 5 | | 1 | 1 | 1 | 7 | 1 | **16** |
| Norbi 4;6 | | 1 | | 3 | | | | **4** |
| Bence 5;0 | 2 | 4 | 4 | | | | | **10** |
| Agripina5;1 | | | | 1 | | | | **1** |
| Benedek 5;6 | | | | 1 | | | | **1** |
| Mark 5;6 | 2 | | 2 | 1 | | | | **5** |
| Sasa 5;6 | 1 | | 1 | 3 | | 3 | | **8** |
| Alexa 5;6 | | | | 1 | | | | **1** |
| **Total** | **17** | **7** | **10** | **17** | **3** | **17** | **3** | **74** |

In order to establish that the pervasiveness of gender errors is a result of the bilingual setting, I compared the narratives produced by the bilingual children with the

BDD-A28524 © 2017 Universitatea din Bucureşti
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

narratives produced by the Romanian monolingual children in a study by Buja (2008); the narratives were based on the same story (*Frog Where Are You?*). It is striking that there are close to no gender errors in the monolingual narratives, even with the youngest children; a look at the transcriptions available at the end of Buja's study shows that only 3 out of 7 3-year-old children produced a total of 4 gender errors (with Acc clitics only). The older (age 4-5) monolingual children no longer produce any gender errors.

## 4. Analysis

### 4.1 Masculine vs. feminine gender errors

Both masculine and feminine incorrect forms are attested, as exemplified in (4) for the longitudinal corpora and (5) for the narrative corpus. In (4a) and (5b) the adjective *albastru* and the accusative clitic should be feminine to agree with the feminine nouns *carte* and *broască*. In (4b) and (5a) on the other hand, the possessive and the adjective should be masculine, to agree with the neuter (masculine singular form) *avion* (note that the demonstrative pronoun coindexed with it is indeed masculine) and the masculine *cățel*.

(4) a. acuma să strângem cartea *albastru (Toma 2;2)
now SĂ put-away book-the.F blue.M
'Now let's put away the blue book.'

b. ba *a mea ăla cu balon (avionul) (Petru 2;4)
no mine.F that.M with balloon (plane-the.M[4])
'No, the one with the balloon is mine.'

(5) a. şi a zis că a fost *obraznică căţelu(l). (Sasa 5;6)
and has said that has been cheeky.F doggie-the.M
Intended: 'And he told the doggie off for being cheeky.'

b. *îl caută sub farfurie (broasca) (Arpi 3;7)
him search under plate (frog-the.F)
Intended: 'And (the dog) is looking for (the frog) in the jar.'

It is also noteworthy that masculine and feminine forms alternate seemingly randomly, sometimes even in the same utterance, or in successive utterances. Similar findings were recorded in Müller (1994). This is true both in the case of the longitudinal corpora (examples under 6) and in the case of the kindergarten group (examples under 7). It looks as though gender were not firmly set as a feature exclusive to each noun; gender agreement is at first sight a flexible phenomenon. There are recordings where the child speaks at length about a certain object and alternately modifies it with masculine and feminine adjectives or functional categories. Example (6a) is taken from a recording where Toma alternates between singular masculine and feminine forms for the same (neuter) noun *tren*. The same for the neuter *elicopter* in (6b). In (6c) masculine and

[4] For simplification, neuter singular nouns will be marked masculine in the glosses.

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

feminine forms alternate freely in the same utterance, though the antecedent is unspecified. In utterance (6d), the masculine proper name (the male character in *Cars* which the boys have as a toy) is accompanied by a masculine adjective (*vechi*) and a feminine predicative (*nouă*). A similar juxtaposition of a feminine and masculine adjective modifying the same entity is visible in (7a). In (7b), the feminine noun *bufniță* is preceded by a masculine indefinite article (*un*), then by a feminine demonstrative (*altă*), and finally is coindexed with a masculine personal pronoun (*el*).

(6) a. trenu(l) *roşie … roşu trenu(l) (Toma 1;11)
train-the.M red.F red.M train-the.M

b. că e *lipită... (e)licopteru(l) negru ... ăla roşu...
because is glued.F helicopter-the black.M that.M red.M
ăsta e prea *grea
this.M is too heavy.F (Toma 2;2)
'because it is glued... the black helicopter... the red one... this one is too heavy'

c. să mănânc unu mică mică mic (Toma 2;3)
SĂ eat one.M small.F small.F small M
'let me eat a small small small one'

d. am făcut pe McQueen vechi *nouă. (Petru 2;3)
have made PE McQueen.M old.M new F
Intended: 'I have made the old McQueen new.'

(7) a. e *micuţă… e foarte mic… e cam atâta. (Mark 5;6)
is small F is very small M is about this-big
Intended: 'It is small, it is very small, it is about this big.'

b. încă *un bufniţă ... altă bufniţă ... şi ciugulea
yet a.M owl.F another.F owl F and pecked
şi *el
also too (Dominik 3;6)
Intended: 'another owl… and it pecked (him) too.'

As regards the longitudinal corpora, there appears to be a preference for the feminine. Most erroneous forms are feminine, as illustrated in Table 5. The percentages show the incorrect feminine forms out of all adjectives/accusative clitics/possessives/quantifiers incorrectly marked for gender. For example, in the Toma corpus, there are 152 adjectives that do not agree in gender with the noun they modify (e.g. (8)): out of this total of 152, 99 (65%) are feminine forms modifying a masculine or neuter singular noun, the rest are masculine forms with a feminine or neuter plural noun.

(8) un poloboc ca ăla *galbenă (Toma 2;3)
a level.M like that.M yellow.F
'a bubble level like the yellow one'

An exception seems to be the third person genitive personal pronoun, which agrees with the possessor, as noted above. Importantly, in the great majority of contexts (19 out

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

of 20 with Toma and 11 out of 13 with Petru) the possessor is masculine (unsurprisingly, usually one of the brothers). Toma has no gender errors at all with this morphological category, and Petru overgeneralizes the masculine. See also Table 3 above.

Additionally, the younger brother uses fewer incorrect feminine clitics, but the total percentage of feminine gender errors is higher with him.

**Table 5.**
**Incorrect feminine forms. Longitudinal corpora**

| *F / gender errors | Toma | Petru |
|---|---|---|
| Adjectives | 65% (99/152) | 78% (57/73) |
| Accusative clitics | 69% (87/127) | 29% (4/14) |
| Possessives: 1st, 2nd | 64% (18/28) | 100% (24/24) |
| Possessives: 3rd | - (0/0) | 0/2 |
| Quantifiers: | 63% (20/32) | 75% (3/4) |
| **Total** | **66% (224/339)** | **75% (88/117)** |

As for the kindergarten corpus, there is some individual variation (see Table 6). Some children use exclusively masculine forms irrespective of the context, some use both feminine and masculine incorrect forms. Four of the children have more feminine gender errors than masculine gender errors.

**Table 6.**
**Feminine vs. masculine gender errors. Narratives.**

| Child (age) | feminine gender errors | masculine gender errors |
|---|---|---|
| Otto 3;3 | 63% (5/8) | 37% (3/8) |
| Maria 3;4 | 0/1 | 100% (1/1) |
| Dominik 3;6 | 0/3 | 100% (3/3) |
| Arpi 3;7 | 0/7 | 100% (7/7) |
| Eva 3;11 | 0/8 | 100% (8/8) |
| Evelin 4;0 | 0/1 | 100% (1/1) |
| Oli 4;6 | 13% (2/16) | 87% (14/16) |
| Norbi 4;6 | 75% (3/4) | 25% (1/4) |
| Bence 5;0 | 80% (8/10) | 20% (2/10) |
| Agripina 5;1 | 100% (1/1) | 0/1 |
| Benedek 5;6 | 100% (1/1) | 0/1 |
| Mark 5;6 | 0/5 | 100% (5/5) |
| Alexa 5;6 | 0/1 | 100% (1/1) |
| Sasa 5;6 | 75% (6/8) | 25% (2/8) |
| **Total** | **34% (28/74)** | **66% (49/74)** |

An interesting case is one of the girls aged 5;6 who uses exclusively feminine postverbal clitics, irrespective of the gender of the noun. This might as a matter of fact

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

arise from a preference for the postverbal rather than the preverbal form, without considering its gender feature, since with the Romanian *perfect compus* form the feminine clitic is always postverbal, while the masculine is always preverbal. She has several omissions (9b), but only in contexts requiring preverbal clitics, and she also commits one word order error which seems to confirm the preference for the postverbal clitic – the clitic should in fact be preverbal in (9a). She has only one preverbal clitic (9d) – which happens to be an overextension. Moreover, she has several overextensions of the feminine clitic *–o,* with intransitives (both unaccusatives and unergatives: 9c, d). All in all, the errors she makes do not appear to be gender related but point to the choice of the postverbal over the preverbal form, even though it does indeed seem as if she were completely disregarding the gender feature of the clitic.

(9) a. a venit şi a vrut să ciupit ***-o**. (Sasa 5;6)
has come and has wanted SĂ pinched her
Intended: 'It came and wanted to sting him.'

b. a venit şi a vrut să *_ ciupeşte. (Sasa 5;6)
has come and has wanted SĂ pinche
Intended: 'It came and wanted to sting him.'

c. şi a zis omul că a dispărut ***-o** broasca (Sasa 5;6)
and has said man-the that has disappeared her frog-the
Intended: 'And the boy said the frog had disappeared.'

d. şi un căţel care ***o** doarme. (Sasa 5;6)
and a doggie which her sleeps
Intended: 'and a doggie that is sleeping.'

This preference for the postverbal clitic might arise from an analogy[5] with the Hungarian definite agreement object marker (E.Kiss 2004), which is a suffix on verbs whose direct object is definite or null but recoverable from the context (see 10a). Contexts requiring the Romanian accusative clitic more or less overlap with contexts requiring this marker in Hungarian[6].

(10) a. meg -csíp -t **-e**
PERF.P sting-PAST-DEF

b. a ciupit**-o**
has stung her

To conclude this section, there appears to be no consistent explanation across the board for why some of the children choose predominantly masculine or feminine forms. There seems to have been some individual preference, one of the children in the longitudinal study (Toma) prefers the feminine, to a lesser extent his brother as well,

[5] For the facilitating effect of this analogy on the acquisition of Romanian Accusative clitics by Romanian-Hungarian bilinguals see Tomescu and Avram (2016).

[6] An anonymous reviewer suggested that this preference for the feminine form might be due to its phonological prominence over the masculine *îl*/*-l*.

BDD-A28524 © 2017 Universitatea din Bucureşti
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

while some of the children in the kindergarten group overgeneralize masculine forms. Nevertheless, masculine and feminine forms alternate quite freely and gender agreement seems rather flexible.

The next subsections will examine the importance of phonological and semantic transparency, as well as syntactic cues and agreement at a distance.

### 4.2. Phonological transparency

The section on phonological transparency looks at three categories of contexts.

First, nouns that are not accompanied by a definite or indefinite article but enter into a gender agreement relation with other functional categories or adjectives. If the noun is also accompanied by an article, it becomes difficult to decide whether it was this syntactic cue that prompted (or failed to prompt) correct agreement or whether it was rather the phonological shape of the noun. However, the number of bare nouns is quite small.

In the longitudinal corpora, the number of phonologically transparent bare nouns is 8 and 6 and the number of erroneous adjectives accompanying them 1 and 2 respectively (e.g. 11). The noun *bluză* ends in the vowel specific for feminine forms, but there is no article present that might have facilitated agreement. The adjective modifying it is however masculine.

(11) că are bluză *roşu (Toma 2;9)
because has shirt.F red M
'Because she has a red shirt.'

In the kindergarten corpus of narratives, there are two quantifiers (e.g. 12) and three adjectives with erroneous gender agreement, all of which adjacent to a phonologically transparent noun. The noun *broască* is also transparent phonologically, yet the quantifier preceding it is masculine.

(12) *aicea a văzut câinele şi băieţelul *doi broască[7] (Evelin 4;0)
here has seen dog-the and boy-the two.M frog F
'Here the dog and the little boy saw two frogs.'

Table 7 offers a summary.

[7] The lack of plural agreement on the noun is not surprising given that in Hungarian number agreement need only appear once; with plural quantifiers the noun must be singular.

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

**Table 7.**
**Phonological transparency. Bare nouns**

| Child | gender errors / total transparent contexts |
|---|---|
| Evelin 4;0 | 100% (1/1) |
| Bence 5;0 | 100% (2/2) |
| Toma | 13% (1/8) |
| Petru | 33% (2/6) |

Second, at least in the corpus of narratives, there are gender errors with indefinite articles that precede phonologically transparent nouns. 5 of the children in the kindergarten group did use indefinites articles which failed to match the gender of the noun they selected (Table 7). All the errors occurred in phonologically transparent contexts, such as (13), with a masculine noun ending in a consonant.

(13) este doar ***o** animal. (Otto 3;3)
is just a.F animal.M
'It is just an animal.'

As highlighted in section 3.1.1, in the longitudinal corpora there are no gender errors with the indefinite article.

**Table 7.**
**Overextensions – masculine/feminine indefinite article, kindergarten corpus.**

| Child Age | overextensions masculine | overextensions feminine | total indefinite articles transparent | non-transparent |
|---|---|---|---|---|
| Otto 3;3 | 8% (1/13) | 23% (3/13) | 13 | |
| Maria 3;4 | 33% (1/3) | 0 | 3 | |
| Dominik 3;6 | 40% (2/5) | 0 | 5 | 1 |
| Oli 4;6 | 50% (7/14) | 0 | 14 | 2 |
| Sasa 5;6 | 22% (2/9) | 11% (1/9) | 9 | |

Third, gender errors with the clitic definite article also belong in this section on phonological transparency, attested both in the longitudinal corpora and in the kindergarten corpus, since such errors by necessity presuppose the phonological distortion of the noun. While interesting, though, these errors are very rare and the data have no quantitative relevance regarding the facilitating effect of phonological transparency on gender agreement. However, these few individual errors merit some comments.

In some cases, the wrong clitic definite article is forced onto a clearly phonologically transparent noun: **spum-u(l)* (Toma 2;6). The singular *spumă* 'foam' is evidently feminine, the end vowel should have been an obvious clue as to its gender,

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

nevertheless it is distorted by means of a masculine definite article. Further, probably under the influence of his initial error, the child produces the masculine sounding **spum*, which he will use for weeks (as noted by personal observation): this backformation might in fact be prompted by the masculine article he initially mistakenly chose. Similarly, he distorts the feminine *floare-a* 'flower-the': **floru(l)* (Toma 2;6), and the neuter singular *ghiveci-ul* 'flowerpot-the': **ghiveci-a* (Toma 2;6).

In other cases, the error might in fact be a direct result of the child taking the phonological shape of the noun as cue: **tractor-i-i* (Eva 3;11) (cf. N PL *tractoare-le* 'tractors-the'). The neuter noun *tractor* was treated as a masculine noun and attached a masculine plural (*-i*) and definite article (*-i*), which, while of course unacceptable, sounds marginally less odd than the examples above.

Other words are perhaps more ambiguous: the plural ending *-i* attaches to both feminine and masculine nouns, therefore the children can be forgiven for chosing the incorrect definite article for these ambiguous plurals: **furnic-i-i* (Oli 4;6) (cf. F PL *furnic-i-le* ants-the*),* **ochi-le* (Petru 2;3) *(*cf. M PL *ochi-i* 'eyes-the'), **papuc-i-le* (Vera 3;4) (cf. M PL *papuc-i-i* slippers-the).

Another ambiguity is the clitic *–le* which is homonymous between the masculine singular and the feminine plural: **covrigelele* (Toma 2;6) (cf. M PL *covrigei-i* 'pretzels-the'). Toma also has difficulties with the masculine noun *burete*, requiring the definite article *-le*, whose gender (and number) he is clearly confused about. Note the following: **toate buretele* all F PL sponge-the M 'all the sponges'; **buretul* - with the other masculine definite article -l; **o burete* - with the feminine indefinite article, *un *buret* - where he snips off the end vowel (Toma 2;3). Similar difficulties are encountered by Petru with the masculine noun *bebe*:

(14) a. *unde sunt bebe-le meu? (Petru 2;4)
where are baby-the.M my.M.SG
b. *unde sunt bebe-le mele? (Petru 2;4)
where are baby-the.M my F PL
Intended: 'Where are my babies?'/'Where is my baby?' (?)

### 4.3 Semantic transparency

One other aspect investigated was whether semantic transparency was at all helpful in assigning the correct gender. I looked at all nouns and pronouns with a [+human] referent separately. The data is summarized in Tables 8 and 9. If grammatical gender is an arbitrary feature, semantic gender would not necessarily contribute to correct gender agreement. Crucially, Hungarian has no grammatical gender, even personal pronouns lack the gender feature. The expectation would be that semantic transparency would not facilitate gender agreement.

Indeed, in the case of the longitudinal data, especially Toma seems to disregard the gender of the individual concerned and uses masculine or feminine forms indiscriminately. As much as 45% (17/38) of all adjectives modifying nouns denoting persons (15a), and more than a third (37%, 7/19) of [+human] clitics (15 c) are incorrect with respect to the gender feature.

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

However, with the younger child there is some evidence that semantic transparency may have helped. Petru produced only 8 incorrect adjectives (11%) out of a total of 70 adjectives accompanying a noun with a [+human] referent (e.g. 15b) and no incorrect clitics with [+human] referent. The data is summarized in Table 8 (see also Tomescu 2017 a, b).

(15) a. sunt *tăiată (Petru 2;2)
am sliced.F
'I am sliced.'

b. Ioana *mic (Toma 2;6)
Ioana.F little M
'little Ioana'

c. *l -am lovit pe Ioana (Toma 2;10)
him have hit PE Ioana F
'I hit Ioana.'

d. vreau s- *o văd pe Petru (Toma 2;6)
want SĂ her see PE Petru
'I want to see Petru.'

**Table 8.**
**Gender errors in [+human] contexts. Longitudinal corpus.**

| Child | A adjacent to[+human] N | [+human] clitic |
|---|---|---|
| Toma | 45% (17/38) | 37% (7/19) |
| Petru | 11% (8/70) | 0% |

With respect to the narrative corpus, there is great individual variation. It appears that the children with the highest number of gender errors disregard semantic transparency to a greater extent than the children whose total number of gender errors is low. The three least proficient children get wrong as many as 40% and even 100% of [+human] contexts (Bence, Oli, Sasa). On the other hand, with some children there are no gender errors with [+human] referents. True, some of these children have few [+human] contexts (e.g. Otto, Alexa). But with some of the children (e.g. Mark, Maia, Arpi, Eva), it is not implausible to assume that semantic transparency might have boosted correct gender agreement.

(16) a. cerbu(l) a căzut pe *ea (Bence 5;0)
stag-the has fallen on her
Intended: 'And the stag fell on him.'

b. şi a dat *-o jos bufniţa (Norbi 4;6)
and has given her down owl-the
Intended: 'And the owl made him fall.'

BDD-A28524 © 2017 Universitatea din Bucureşti
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

c. tu *singur? (Oli 4;6)
you alone.M
'By yourself?' (to female investigator)

**Table 9.**
**Gender errors in [+human] contexts. Kindergarten corpus.**

| Age | incorrect gender/total +human contexts | total gender errors |
|---|---|---|
| Otto 3;3 | 0/1 | 8 |
| Maia 3;6 | 0/8 | 1 |
| Arpi 3;7 | 0/3 | 7 |
| Eva 3;11 | 0/3 | 8 |
| Oli 4;6 | 40% (2/5) | 16 |
| Norbi 4;6 | 20% (1/5) | 4 |
| Bence 5;0 | 100% (4/4) | 10 |
| Agripina 5;1 | 20% (1/5) | 1 |
| Mark 5;6 | 0/5 | 5 |
| Alexa 5;6 | 0/2 | 1 |
| Sasa 5;6 | 100% (4/4) | 8 |
| **Total** | **26% (12/47)** | **74** |

All in all, the data is not uniform across the board: we may state that semantic transparency is not a factor in some cases, but it is not to be overlooked in others.

### 4.4 Agreement at a distance

The most important factor prompting correct gender agreement was found to be the adjacency of the adjective/functional category to the noun. Agreement at a distance appears to have been much more difficult for the bilinguals both in the longitudinal corpora and in the kindergarten group. Possibly, therefore, it is a question of processing difficulties, which has been found to be a problem with bilinguals before (Unsworth et al. 2014, Patuto et al. 2011, Serratrice 2013, Sorace 2011, Alarcón 2011).

As shown above in Table 3, out of the total number of variable adjectives in the two longitudinal corpora, 27% for Toma and 23% for Petru are incorrectly inflected for gender. But if contexts where agreement takes place at a distance (e.g. 17 c, d) are eliminated, and only those contexts are counted where the adjective is adjacent to the noun (e.g. 17a), the percentage of errors decreases by half (see also Tomescu 2017 a, b). The difference is statistically significant ($p < 0.0001$, $p = 0.0081$). Table 10 offers a summary: the first row shows the gender errors in the totality of contexts containing variable adjectives, while the second row only shows those contexts where the adjective directly follows the noun it modifies.

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

(17) a. asta e oc(hi)u(l) *ascunsă (Petru 2;2)
this is eye-the.M hidden.F
'this is the hidden eye.'

b. oliţa *albastru (Toma 2;2)
potty-the.F blue.M

c. *s- arăt la mama că e galben (piatra) (Toma 2;2)
SĂ show to mother that is yellow.M (stone.M)
'let me show mother that it is yellow'

d. *ruptă ăsta (tren) (Petru 2;0)
broken.F this M (train.M)
'this one is broken'

**Table 10.**
**Adjacency to nouns. Adjectives.**

| | **Toma** | **Petru** |
|---|---|---|
| all A | 27% (152/556) | 23% (73/316) |
| A adjacent to N | 11% (22/202) | 11% (11/104) |

As regards the Toma corpus, early on, the child produces some prenominal adjectives, which is rather a formal choice for Romanian, if not downright odd in certain contexts. There might have been some cross-linguistic influence at play here, since in Hungarian adjectives are prenominal. Most prenominal adjectives (13/15) are incorrect as regards gender agreement – which might also be due to computational difficulties (see also Tomescu 2017 b). These 15 adjectives were not included in the total of 202 in Table 10. Two of the more unbalanced bilinguals in the kindergarten corpus also produced two prenominal adjectives (one correct, one incorrect: 18b) – note that the number of adjectives in the narratives is rather small, due possibly to the nature of the task: there was no need for extensive descriptions, the story is quite dynamic.

(18) a. punem *nouă scutecu(l)[8]. (Toma 2;0)
put new.F diaper-the.M
'We put on a new diaper.'

b. şi *mic broască *_ mergat[9] (Bence 5;0)
also small.M frog M gone
Intended: 'the small frog also went.'

[8] Note that when the adjective is used prenominally, the definite article cliticizes on the adjective rather than the noun; example (12a) would not be correct even with a masculine adjective. The correct version would have been: *noul scutec* (new-the diaper), and it would have sounded too formal.

[9] Incorrect participle: cf. correct *mers*. The child distorts most participles in his narrative: **durat* (*a durea*, part. *durut* 'hurt'), **ştiat* (*a şti*, part. *ştiut* 'know'), **fugeat* (*a fugi*, part. *fugit* 'run'). He also omits on occasion the auxiliary required to form the *perfect compus* form. The correct version for the verb (18b) would have been: *a mers* has gone. He is the least proficient of the group.

BDD-A28524 © 2017 Universitatea din Bucureşti
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

An examination of 1st and 2nd person possessives – that agree in gender with the object possessed - in the longitudinal corpora shows a tendency similar to adjectives (see also Tomescu 2017b). Adjacency to the noun (to the object possessed that is, such as example (19a) will result in a much lower number of gender errors compared to contexts where the object possessed is not overtly expressed (19b). Table 11 summarizes the data. The first row presents all contexts containing a possessive, while the second row shows the percentages only for those contexts where the possessive is directly adjacent to the noun it agrees with.

(19) a. piratul meu (Petru 1;10)
pirate-the.M my.M
b. *a mea e cu motoare (avion) (Petru 2;4)
mine.F is with engines (airplane.M)
'Mine has engines.'

**Table 11.**
**Adjacency to nouns. Possessives: 1st, 2nd person.**

| Gender errors out of total contexts: | Toma | Petru |
|---|---|---|
| all contexts | 23% (28/144) | 15% (26/172) |
| adjacent to noun | 10% (5/48) | 4% (3/85) |

Third person possessives are not considered here: the third person of the possessive pronoun paradigm is absent in the corpora, while the genitive form of the personal pronoun, which is preferred in colloquial language, enters into gender agreement relation with the possessor rather than the object possessed (see section 3.1 and footnote 3).

If we take a look at the frog story corpus, we find that 88% (65/74) of gender errors occur when agreement is at a distance (predicative adjectives (20a), clitics (20b), 3rd person possessives (20c) and personal pronouns (20d)); in the rest of the cases the erroneous category is adjacent to the noun – of these 17 are indefinite articles.

(20) a. e atât de *mică (scaun). (Otto 3;3)
is so small.F (chair.M)
'It is so small.'
b. aici *îl strigă (broasca) (Alexa 5;6)
here him calls (frog.F)
'Here he is calling it.'
c. albina merge la casa *lui. (Bence 5;0)
bee-the.F goes to house-the his
Intended: 'The bees are going home.'
d. *încearcă să vadă ce se întâmplă cu **el** (broasca). (Arpi 3;7)
tries SĂ see what REFL happens with him (frog.F)
'He is trying to see what happened to it.'

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

Table 12 offers a summary for each child: the left-hand column shows the total number of gender errors, while the right-hand column only shows the percentages of gender errors where there is no adjacency to the noun.

**Table 12.**
**Adjacency. Kindergarten corpus.**

| Child (age) | total number of gender errors | non-adjacency to N |
|---|---|---|
| Otto 3;3 | 8 | 50% (4/8) |
| Maria 3;4 | 1 | 0/1 |
| Dominik 3;6 | 3 | 33% (1/3) |
| Arpi 3;7 | 7 | 100% (7/7) |
| Eva 3;11 | 8 | 63% (5/8) |
| Evelin 4;0 | 1 | 100% (1/1) |
| Oli 4;6 | 16 | 38% (6/16) |
| Norbi 4;6 | 4 | 100% (4/4) |
| Bence 5;0 | 10 | 80% (8/10) |
| Agripina 5;1 | 1 | 100% (1/1) |
| Benedek 5;6 | 1 | 100% (1/1) |
| Mark 5;6 | 5 | 100% (5/5) |
| Alexa 5;6 | 1 | 100% (1/1) |
| Sasa 5;6 | 8 | 63% (5/8) |
| Total | 74 | 88% (65/74) |

In order to have an analysis comparable with the longitudinal study, it was attempted to calculate the percentage of gender errors out of all contexts with noun adjacency across categories. However, the small number of contexts, as well as the brevity of the individual utterance was a hindrance in this respect. In most cases the noun is absent from the utterance. Five of the children used no adjectives at all, and only two of the others used adjectives adjacent to the noun it modifies. One of these two children committed 5 out of 6 gender errors with adjectives, and one of these happened to be adjacent to the noun. The numbers for the other child resemble the findings for the longitudinal study: 16% (5/32) of all adjectives are incorrect, but none of the 5 adjectives adjacent to the noun is incorrect. Note that I eliminated from the count the invariable adjective *mare* 'big' which – ironically - appeared quite frequently. There are no 1$^{st}$ or second person possessives in the narratives. Table 13 summarizes the data:

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

**Table 13.**
**Adjacency to nouns. Adjectives. Kindergarten corpus.**

| Child (age) | all adjectives | A adjacent to N |
|---|---|---|
| Otto 3;3 | 50% (2/4) | - |
| Maria 3;4 | 0/9 | - |
| Dominik 3;6 | 0/1 | - |
| Arpi 3;7 | 0/1(3) | - |
| Eva 3;11 | 16% (5/32) | 0/5 |
| Evelin 4;0 | - | - |
| Oli 4;6 | 83% (5/6) | 1/1 |
| Norbi 4;6 | 0/3 | - |
| Bence 5;0 | 100% (2/2) | 1/1 |
| Agripina 5;1 | 0/3 (9) | - |
| Benedek 5;6 | - | - |
| Mark 5;6 | 67% (2/3) | - |
| Alexa 5;6 | - | - |
| Sasa 5;6 | - | - |

### 4.5 Syntactic cues

It was found that correct agreement was also helped along by the presence of the article, since after all the overwhelming majority of nouns in the longitudinal corpora bear (especially the definite) article. I counted all the contexts with the indefinite article *un/o* (these were not very numerous) (e.g. (21a), the feminine definite article *-a* (21b) or the masculine definite article *-l* (21c) and an adjective modifying the noun. Plural contexts and singular contexts with the definite article *-le* were eliminated, because of the confusion that might have been caused by the homonymy between the plural feminine or singular masculine *-le*. The masculine definite article *-i* might also have been misleading because of its similarity with the masculine/feminine plural marker *-i*. The results show that the percentage of gender errors in contexts of this type was much lower than the percentage of gender errors out of the totality of adjectives to be found in the corpora: 16 and 14% compared to 27% and 23% respectively. The difference is statistically significant ($p = 0.0031$) in Toma's case; in Petru's case the second sample size may be too small for comparison. Table 14 summarizes the results.

(21) a. Toma vrea **o** gumă mestecată. (Toma 2;0)
Toma wants a.F gum.F chewed F
Intended: 'Toma wants chewing gum.'

b. nu ştiu unde e telecomand**-a** *roşu (Toma 2;2)
not know where is remote-the.F red M
Intended: 'The old brush.'

BDD-A28524 © 2017 Universitatea din Bucureşti
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

c. vagon-**u(l)** ăla *roşie (Petru 2;1)
car-the.M that.M red.F
'the red car.'

**Table 14.**
**Relevance of syntactic cues.**

| Gender errors out of total contexts: | Toma | Petru |
|---|---|---|
| all A | 27% (152/556) | 23% (73/316) |
| A + N -a/-l/un/o | 16% (29/176) | 14% (8/58) |

In the case of the kindergarten corpus, there are only two adjectives cooccurring with a DP: in one case the adjective is in fact correct, while the article is not, in the other, the adjective is incorrect.

(22) a. *un mică broască (Sasa 5;6)
a.M small.F frog.F
b. scaun-ul este aşa atât de *mică (Otto 3;3)
chair-the.M is so so small.F
'the chair is so small'

### 4.6 Summary

This section looked at whether phonological transparency, semantic transparency, adjacency to the noun and the presence of definite/indefinite articles are helpful with respect to correct gender agreement.

As regards phonological transparency, in the longitudinal corpora, transparent bare nouns modified by adjectives enter into a correct gender agreement relation in the majority of the cases (although the number of contexts is quite small). Furthermore, the indefinite article, which is quite close to the noun, was never incorrectly used by the brothers in the longitudinal study. The odd error with the definite article cannot be quantitatively relevant. Conversely, in the kindergarten group, indefinite articles often fail to match the gender of the noun. The three examples containing a bare noun modified by a variable morphological category are all incorrect. In the narrative corpus, phonological transparency does not seem to have had any facilitating effect.

Semantic transparency was a factor with some children but not at all with others. In the case of the longitudinal Toma corpus, and also with some of the children in the corpus of narratives, the natural gender of the human referent did not greatly contribute to the correct grammatical gender agreement, with error percentages of up to 100% in two instances. In the case of some of the other children in the kindergarten corpus, however, there is evidence to the effect that semantic transparency was helpful. Also, the younger brother in the longitudinal study has no gender errors with [+human] clitics and fewer with adjectives in [+human] contexts than his brother.

BDD-A28524 © 2017 Universitatea din Bucureşti
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

Adjacency to the noun was very helpful for correct gender agreement both in the case of the longitudinal corpora and in the narratives. Agreement at a distance was more of a hurdle, because of processing difficulties inherent to bilingual language acquisition.

Another positive factor in the case of the longitudinal corpora was the presence of the definite articles which are not ambiguous as regards gender: the feminine *-a* and the masculine *-l*, as well as the indefinite articles. In the corpus of narratives on the other hand, the indefinite article itself was often incorrect. As regards the definite article, there was only one utterance that contained a definite DP and an adjective modifying it.

## 5. Discussion

While both masculine and feminine incorrect forms are attested both in the longitudinal data and in the narratives, there is individual variation and variation across categories regarding any preference for one particular gender. The longitudinal corpora do contain a higher number of feminine gender errors, although there are differences between the two brothers. In the kindergarten group, some of the children had a preference for the masculine gender. The fact that the same noun appears accompanied by masculine and feminine adjectives or determiners in the same sentence or in successive utterances however does seem to prove that gender agreement (or gender agreement errors) may be inconsistent and random.

Most gender errors appear to have been committed at a distance, when the noun is not adjacent to the category that agrees with it. It would appear therefore that agreement errors are caused by processing difficulties, which is not unexpected in bilingual language acquisition (see Unsworth et al. 2014, Patuto et al. 2011, Serratrice 2013, Sorace 2011, Alarcón 2011).

As regards the longitudinal corpora, another helpful factor seems to have been the presence of the definite or indefinite article, prompting correct gender agreement with other morphological categories. As additional evidence in favour of the hypothesis that it is agreement at a distance which is problematic for bilinguals, in the longitudinal corpora at least, indefinite articles, never separated from the noun, are never incorrectly used. However, the same thing cannot be said for the corpus of narratives, where the indefinite article was often incorrect.

While nouns have gender in their lexical entry and it is activated automatically at the moment of lexical selection (Caramazza et al. 2001), adjectives and pronouns are not lexically marked for gender, which is a syntactic feature resolved later in the derivation, in a way similar to number agreement (Cantone and Müller 2007). Whereas number agreement does not appear to be more problematic for bilinguals that for monolinguals (Hungarian does have number after all), Romanian-Hungarian bilinguals are somewhat hampered by the fact that in Hungarian the gender feature is not activated. Serratrice (2013) also considers the possibility of underspecification as manifestation of cross-linguistic influence: when bilinguals are faced with conflicting evidence they tend towards a more flexible interpretation of the phenomenon. Possibly then, confused by the lack of gender feature in Hungarian, bilinguals choose freely between the

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

masculine/feminine form available to them and fail to strictly match the gender of the e.g. clitic to that of the noun it is coindexed with.

It is worth noting that number errors occur in a different way. We do not find in the data plural adjectives with singular nouns. The selection of the singular/plural adjective is not in fact random. The only error type attested is that a singular adjective may be selected alongside a plural noun[10]. Even with clitics, number errors are only attested in the longitudinal corpora, and their rarity is striking compared with the percentage of gender errors: 2% vs. 26% and 23% percent respectively. Of these, there is only one plural clitic in one of the longitudinal corpora corresponding to a singular noun. Note how hesitant and lacking in fluency the sentence is.

(23) *să **le** căutăm pe **ăla** … **ăla** **mic** să mă joc cu … (Toma 2;3)
SĂ them.F search PE that.M that.M small.M SĂ REFL play with
Intended: 'Let's look for the small one, I want to play with it.'

To conclude, therefore, two main factors have been identified as causes for gender errors: processing difficulties inherent to bilingual language acquisition, translated into a higher number of gender errors when agreement must occur at a distance, and the influence of the other language, which in this case has no gender features. Additionally, in the case of the longitudinal data, evidence has been found for the facilitating effect of the presence of the article on the noun.

## 6. Conclusion

The study has analysed the gender errors committed in two longitudinal corpora and a collection of narratives produced by kindergarten age children, all Romanian-Hungarian bilinguals living in Bucharest. The data disproved the importance of phonological and semantic transparency in gender agreement and highlighted the relevance of adjacency to the noun and possibly of the presence of the definite article. Since lack of adjacency to the noun leads to a higher number of gender errors, it is not implausible to assume that gender errors are mainly committed as a result of processing difficulties that bilinguals are often confronted with.

[10] Which actually is the correct choice in Hungarian, since Hungarian adjectives are only plural in the absence of the noun (iii). Number agreement in Hungarian must only occur once, whether on the noun or on the numeral:

i. két piros labda
two red ball
ii. piros labdák
red balls
iii. a pirosak
the reds
'the red ones'

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)

## References


Alarcón, I.V. 2011. Spanish gender agreement under complete and incomplete acquisition: Early and late bilinguals' linguistic behavior within the noun phrase. *Bilingualism: Language and Cognition* 14: 332-350.

Arias-Trejo, N., Falcon, A. Alva-Canto, E. A. 2013. The gender puzzle: Toddlers' use of articles to access noun information. *Psicologica* 34: 1-23.

Avram, L. 2001. Early omission of articles in child Romanian and the emergence of DP. *Revue Roumaine de Linguistique* XLVI (1-4): 105-123.

Brehmer, B., Rothweiler, M. 2012. The acquisition of gender agreement marking in Polish. In K. Braunmüller and C. Gabriel (eds.), *A Study of Bilingual Polish-German-Speaking Children, Multilingual Individuals and Multilingual Societies*, 81-100. Amsterdam/Philadelphia: John Benjamins.

Bruhn de Garavito, J., White, L. 2002. The L2 acquisition of Spanish DPs. The status of grammatical features. In A. T. Perez-Leroux and J. Liceras (eds.), *The Acquisition of Spanish Morphosyntax: The L1/L2 Connection*, 151-176. Dordrecht: Kluwer.

Buja, E. 2008. *Relating Events in Narrative: A Case Study of Romanian*, Braşov: Editura Universităţii Transilvania.

Cantone, K. F., Müller, N. 2007. Un nase or una nase? What gender marking within switched DPs reveals about the architecture of the bilingual language faculty. *Lingua* 118: 810-826.

Caramazza, A., Miozzo, M., Costa, A., Schiller, N., Alario, F. X., 2001. Lexical selection: A cross-language investigation of determiner production. In E. Dupoux (ed.), *Language, Brain, and Cognitive Development: Essays in Honor of Jacques Mehler*, 209-226. Cambridge, MA: MIT Press.

Carstens, V. 2010. Implications of grammatical gender for the theory of uninterpretable features. In M. Putnam (ed.), *Exploring Crash-Proof Grammars*, 31-57. Amsterdam/Philadelphia: John Benjamins.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed), *Ken Hale: A Life in Language*, 1-52. Cambridge, MA: MIT Press.

Cornips, L., van der Hoek, M., Verwer, R. 2006. The acquisition of grammatical gender in bilingual child acquisition of Dutch (by older Moroccan and Turkish children). The definite determiner, attributive adjective and relative pronoun. *Linguistics in the Netherlands 2006*: 40-51.

Cornips, L., Hulk, A. 2008. Factors of success and failure in the acquisition of grammatical gender in Dutch. *Second Language Research* 24 (3): 267-296.

Costa, A., Kovacic, D., Franck, J., Caramazza, A. 2003. On the autonomy of the grammatical gender systems of the two languages in a bilingual. *Bilingualism: Language and Cognition* 6 (3): 181-200.

Eichler, N., Jansen, V., Müller, N. 2013. Gender acquisition in bilingual children: French–German, Italian–German, Spanish–German and Italian–French. *International Journal of Bilingualism* 17 (5): 550-572.

E. Kiss, K. 2004. *The Syntax of Hungarian*. Cambridge: Cambridge University Press.

Granfeldt, J, Schlyter, S., Kihlstedt, M. 2007. French as L2, 2L1 and L1 in pre-school children. *PERLES Lund University* 24: 7-42.

Hawkins, R., Franceschina, F. 2004. Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish, In P. Prevost, J. Paradis (eds.) *The Acquisition of French in Different Contexts: Focus on Functional Categories*, 175-206. Amsterdam/Philadelphia: John Benjamins.

Hulk, A., Cornips. L. 2006. Neuter gender and interface vulnerability in child L2/2L1 Dutch. In S. Unsworth, T. Parodi, A. Sorace and M. Young-Scholten (eds.), *Paths of Development in L1 and L2 Acquisition* 107-134. Amsterdam/Phuiladelphia: John Benjamins.

Hulk, A., van der Linden, E. 2009. L'acquisition du genre chez les enfants bilingues et le rôle de l'autre langue. In P. Bernardini, V. Egerland, J. Granfeldt (eds.), *Mélanges plurilingues offerts à Suzanne Schlyter à l'occasion de son 65ème anniversaire*, 147-160. Lunds Universitet. Språk- och litteraturcentrum.

Kuchenbrandt, I. 2005. Gender acquisition in bilingual Spanish. In J. Cohen, K.T. McAlister, K. Rolstad and J. MacSwan (eds.), *Proceedings of the 4th International Symposium on Bilingualism*, 1252-1263. Somerville, MA: Cascadilla Proceedings Project.

Kupisch, T., Müller, N., Cantone, K. F. 2002. Gender in monolingual and bilingual first language acquisition: Comparing Italian and French. *Lingue e linguaggio* 1 (1): 107-151.



**BDD-A28524 © 2017 Universitatea din Bucureşti**

**Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)**

van der Linden, E. 2009. Est-ce la ou le? Un apprentissage difficile. *De la linguistique à la didactique des langues: le probleme des modèles linguistiques. Actes du Colloque international du 7 juin 2009, organisé par l'école doctorale "Langues et identités culturelles" de l'Université de Bucarest*, 141-154. Bucharest: Paralela 45.

van der Linden, E., Hulk, A. 2009. The vulnerability of gender on determiners in L1, 2L1 and L2 acquisition. *Bucharest Working Papers in Linguistics* 2: 97-107.

Montrul, S. 2004. *The Acquisition of Spanish. Morphosyntactic Development in Monolingual and Bilingual L1 Acquisition and Adult L2 Acquisition*. Amsterdam/Philadelphia: John Benjamins

Müller, N. 1994 Gender and number agreement within DP. In J. M. Meisel (ed.), *Bilingual First Language Acquisition: French and German Grammatical Development*, 53-88. Amsterdam/Philadelphia: John Benjamins.

Patuto, M., Repetto, V., Müller, N. 2011. Delay and acceleration in bilingual first language acquisition. In E. Rinke, T. Kupisch (eds.), *The Development of Grammar: Language Acquisition and Diachronic Change. In honour of Jürgen, M. Meisel*, 231-261. Amsterdam/Philadelphia: John Benjamins.

Rodina, Y., Westergaard, M. 2013a.The acquisition of gender: A bilingual case study. In S. Stavrakaki, M. Lalioti, P. Konstantinopoulou (eds.), *Advances in Language Acquisition*, 355-365. Newcastle upon Tyne: Cambridge Scholars Publishing.

Rodina, Y., Westergaard, M. 2013b. The acquisition of gender and declension class in a non-transparent system: Monolinguals and bilinguals. *Studia Linguistica* 67 (1): 47-67.

Schwartz, M., Minkov, M., Dieser, E., Protassova, E., Moin, V., Polinsky, M. 2015. Acquisition of Russian gender agreement by monolingual and bilingual children. *International Journal of Bilingualism* 19 (6): 726-752.

Serratrice, L. 2013. Cross-linguistic influence in bilingual development. Determinants and mechanisms. *Linguistic Approaches to Bilingualism* 3 (1): 3-25.

Sorace, A. 2011. Pinning down the concept of 'interface' in bilingualism. *Linguistic Approaches to Bilingualism* 1 (1): 1 -33.

Silva-Corvalan, C. 2014. *Bilingual Language Acquisition, Spanish and English in the First Six Years*. Cambridge: Cambridge University Press.

Tomescu, V. 2013. The acquisition of directional and locative particles in a Romanian-Hungarian bilingual context. In L. Avram, A. Sevcenco (eds.), *Topics in Language Acquisition and Language Learning in a Romanian Context*, 107-128. Bucharest: Editura Universităţii din Bucureşti.

Tomescu, V. 2017a. Gender errors in the acquisition of Romanian in a Romanian-Hungarian bilingual context. In M. Burada, O. Tatu, R. Sinu (eds.), *Language Diversity in a Globalized World. Proceedings of the 13th Conference on British and American Studies, Transilvania University of Brasov*, 104-121. Newcastle upon Tyne: Cambridge Scholars Publishing.

Tomescu, V. 2017b. *Acquisition in a Romanian-Hungarian Bilingual Setting. A Case Study*. Bucharest: Editura Universităţii din Bucureşti.

Tomescu, V., Avram, L. 2016 Peripheral cross-linguistic interference in the acquisition of Accusative clitics by Romanian-Hungarian simultaneous bilinguals. *Probus: International Journal of Latin and Romance Linguistics*. 10.1515/probus-2016-0014.

Tsimpli, I. M., Hulk, A. 2013. Grammatical gender and the notion of default: Insights from language acquisition. *Lingua* 137: 128-144.

Unsworth, S. 2007. Age and input in early child bilingualism: The acquisition of grammatical gender in Dutch. In A. Belikova, L. Meroni and M. Umeda (eds.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*, 448-458. Somerville, MA: Cascadilla Proceedings Project.

Unsworth, S., Argyri, F., Cornips, L., Hulk, A., Sorace, A., Tsimpli, I. 2014. The role of age of onset and input in early child bilingualism in Greek and Dutch. *Applied Psycholinguistics* 35 (04): 765-805.

Voeikova, M. D. 2013. Russian agreement patterns: General strategy and individual differences in their acquisition. Workshop on the Acquisition of Adjectives Across Languages, November 28-29 2013, Utrecht.

**BDD-A28524 © 2017 Universitatea din Bucureşti**
**Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)**

## Appendix

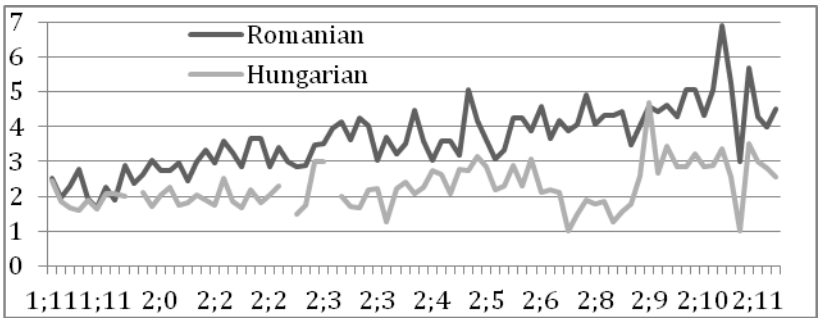


Figure 1. MLU Toma.

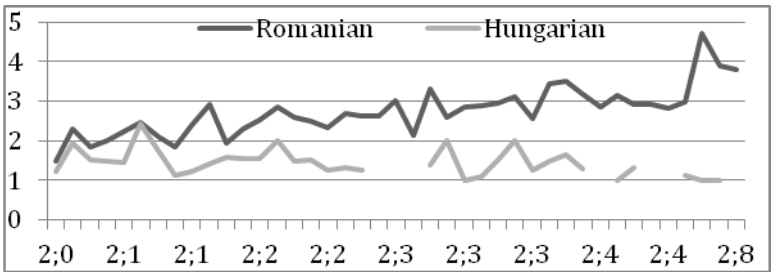


Figure 2. MLU Petru.

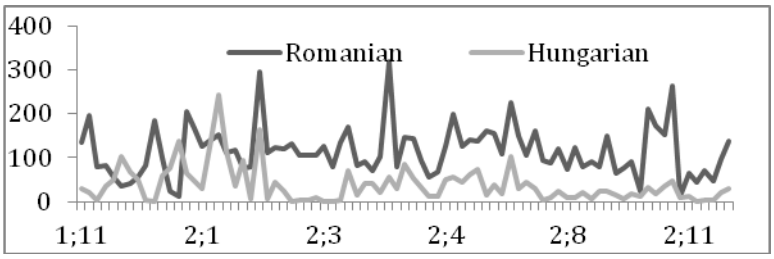


Figure 3. Toma: number of utterances per recording.

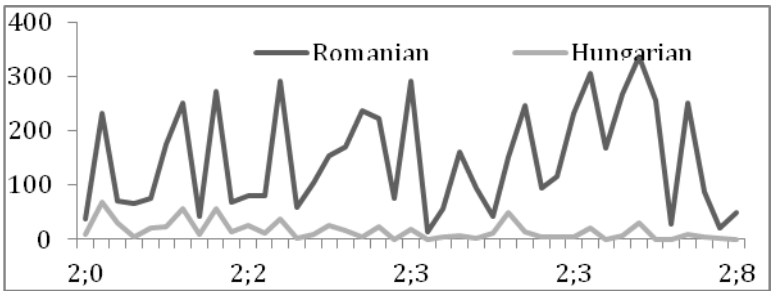


Figure 4. Petru: number of utterances per recording.

BDD-A28524 © 2017 Universitatea din București
Provided by Diacronia.ro for IP 216.73.216.169 (2026-11-07 03:08:20 UTC)