Acquisition of a transparent gender system:

A comparison of Italian and Croatian

1 **Keywords: gender acquisition, gender agreement, gender transparency, Italian, Croatian**

2 **Abstract**

3 For a long time, transparency has been seen as a key facilitator for early acquisition of gender.
4 This study compares the acquisition of the gender system of Italian and Croatian, both of which
5 are considered transparent systems. The study focuses on the different degrees of transparency
6 between the two languages by taking into account their extended nominal paradigms.
7   We have conducted an adjective elicitation task on a total of 60 monolingual Italian and
8 Croatian children divided in two age groups (Italian=3;0 and 3;10, Croatian= 2;10 and 4;2).
9   The results reveal that the Italian gender system is mastered already by the youngest child
10 (age=2;6) and that the two gender values are acquired simultaneously. However, the Croatian
11 children show a significant difference in the error ratio between the two age groups, which
12 indicates that the gender system is not yet acquired in the younger group (average age=3;0).
13 Additionally, the results suggest that feminine is the first gender to be mastered in Croatian due
14 to the regularity of its paradigm, and that neuter is the most problematic gender for the children,
15 likely due to its lower frequency and syncretism with masculine throughout the case paradigm.
16   This paper adds to the body of research indicating that transparency of the gender system
17 is not merely a binary feature, it underlines the relevance of placing the languages on a
18 continuum with respect to transparency in order to make predictions related to the acquisition
19 of the gender system.
20

21 **1. Introduction**

22
23 The aim set forth in this article is to compare how the level of transparency of a gender system
24 influences its acquisition. We will do so by comparing how Italian and Croatian monolingual
25 children acquire their gender system and the individual gender values therein.
26   Grammatical gender is an inherent property of the noun reflected in agreement with other
27 elements (i.e. determiners and adjectives) (Corbett, 1991). A transparent gender system entails
28 that the gender of the noun is evident from its phonological form, as nouns have a gender
29 assigned as it is a part of their lexical entry (Kupisch et al., 2002:109). Both Italian and Croatian
30 have transparent gender systems and rich nominal morphology (Kovačević et al., 2009;
31 Gudmundson, 2010). Nevertheless, all languages diverge from transparency to some extent
32 (Audring, 2014:6) which means that we cannot consider transparency a categorical property,
33 i.e. a binary distinction between a transparent and opaque system. Rather, it should be placed
34 on a continuum.
35   The different degree of transparency between the two languages will be defined by taking
36 into consideration (i) the complexity of the system, expressed through syncretism and the
37 number of gender values, and (ii) gender agreement. In relation to our first point, the declension
38 classes of the nouns in each language (i.e. number for Italian, number and case for Croatian)
39 have been taken into consideration in order to account for syncretism. Syncretism is defined as
40 the relation between two or more words in a paradigm that have different morphosyntactic
41 features but are identical in form (Matthews, 2007) and this adds to the complexity of the system

1 as having identical forms reduces the likelihood that the markers will be unambiguous
2 (Audring, 2014:12). The number of gender values is also a factor related to the complexity of
3 a gender system (Audring, 2014:14), and may also contribute to the transparency of the system.
4 Croatian has three gender values (masculine, feminine, and neuter) while Italian has two
5 (masculine and feminine). In Croatian the three gender values are not equally well represented
6 in child directed speech (CDS) with the neuter amounting to only 6% of nouns (Kovačević et
7 al., 2009:161), whereas no such discrepancy is noted for Italian. Agreement will be accounted
8 for as gender is rarely marked on the noun itself, but it is present overwhelmingly elsewhere
9 (Audring, 2014:7). Italian has obligatory gender-marked articles that occur with the noun. Chini
10 (1995) defines the Italian articles as the first and most frequent syntactic cue for gender, thus
11 enhancing the transparency of the language. Croatian does not have articles and thus nouns are
12 often bare, even if the noun may be accompanied by gender-marked demonstratives,
13 possessives, or adjectives. We considered these factors to compare the relative transparency of
14 the languages.
15 　　　We tested the acquisition of gender through agreement by testing children on an adjective
16 elicitation task. Adjectives were chosen as these are optional and gender-marked in both
17 languages, and were therefore considered a comparable testing ground. The results have shown
18 considerable differences between the two language groups, and we discuss this based on the
19 different degrees of transparency.

20 　　　The paper is structured as follows: In the next section we provide a general overview of
21 the acquisition of gender systems, in section 3 we move on to describing the genders systems
22 of the two languages we are investigating, with a subsection on previous research on gender
23 acquisition in these Italian and Croatian. In section 4, we describe the aims of the current study
24 and lay out the research questions, following that the methodology is described (section 5).
25 Section 6 focuses on the results, followed by a discussion (section 7) and conclusions (section
26 8).

27 **2.　The acquisition of gender: a general overview**

28
29 From an acquisition perspective, there are two types of cues that contribute the acquisition of
30 gender: formal and semantic cues. The former are morpho-phonological cues that appear on the
31 noun, the latter correlate to the natural gender of the referent. Research on various languages,
32 such as Hebrew (Levy, 1983), French (Karmiloff-Smith, 1981), Russian (Rodina, 2008), and
33 Spanish (Pérez-Pereira, 1991) has shown that children rely more on formal than semantic cues.
34 Hence, these will not be discussed further. Thus, if the formal cues are clear, i.e. transparent,
35 children should acquire them more easily. In this study we consider a gender assignment system
36 to be transparent when the phonological or morphological assignment of gender allows for an
37 accurate inference of the gender of the noun without having to rely on agreement on other
38 arguments; as a high complexity of the nature and number of assignment rules may lead to a
39 greater difficulty for the acquiring child (Audring, 2014:16). An example of a transparent
40 system is Spanish with the final vowel of the noun signaling its gender: -*o* for M and -*a* for F,
41 i.e. *armario*-M 'closet', *mesa*-F 'table' (Pérez-Pereira, 1991), whereas languages like
42 Norwegian are deemed to be an opaque system as the noun itself offers no, or very little,
43 indication of the gender, i.e. *stol*-M 'chair, *seng*-F 'bed', and *skap*-N 'closet' (Rodina and
44 Westergaard, 2015).
45 　　　Other arguments receive gender by agreement. As Corbett (1991) states, in order to
46 establish the gender of a noun we have to use agreement as a test, since agreement is the way
47 in which gender is realized in language use. So, nouns have gender by assignment (i.e. they are
48 assigned a gender value in the lexicon), while the elements that agree with the noun receive

1 gender through agreement (Kupisch et al., 2013:153). Unlike assignment, a rich agreement
2 system facilitates acquisition as languages in which a lot of elements agree with the noun (e.g.
3 Spanish, French) is mastered earlier than a system with few agreement markers (e.g. Dutch)
4 (Audring, 2014:16).
5   Previous studies have correlated the time course in the acquisition of the gender system
6 with its transparency: if the gender assignment system is transparent it will be acquired more
7 easily; as this property facilitates early use of correct agreement (Kupisch et al., 2002:141;
8 Rodina and Westergaard, 2015:197). On the other end of the spectrum, in languages with
9 opaque gender systems, this property is acquired late, such as in Norwegian (Rodina and
10 Westergaard, 2015:200).
11   Here the acquisition of a gender system is observed through agreement (of the adjective)
12 as has been done in previous studies (Karmiloff-Smith, 1981; Kupisch et al., 2002; Rodina,
13 2008). Pérez-Pereira (1991) conducted a study on 160 Spanish children of a wide age range (4-
14 11-year-olds) that tested the semantic, morphological, syntactic cues in conditions where two
15 of the factors either indicated the same gender or were in conflict. Converging cues made it
16 easier to determine the correct gender, but in conflict conditions, children payed more attention
17 to agreement that to assignment. Pérez-Pereira (1991:584) also found that children tended to
18 attribute M to F, which might be because of the unmarked status of M.

19   Transparency is not a binary feature and comes in degrees, therefore previous studies
20 have found pivotal differences in the acquisition of two languages with fairly similar gender
21 systems. Smoczyńska (1985) compared Polish and Russian, and found that the Polish gender
22 system is more easily acquired. The reason for this is the different degrees of transparency
23 between the two languages. Rodina and Westergaard (2017) define two classes of opaque nouns
24 in Russian: nouns ending in a palatalized consonant which may belong either to the M or the F
25 gender, and N nouns ending in an unstressed -*o* are indistinguishable from the typical -*a* ending
26 of F nouns making the nouns in question ambiguous between F and N. The latter types of nouns
27 also remain ambiguous with agreement as adjectival endings are also unstressed and thus have
28 the same ambiguity. The ambiguity can be resolved through prenominal modifiers with stress
29 on the final syllable. Russian children were found to overgeneralize these nouns to the F gender.

30   An effect of the degrees of transparency was also found between French and Italian as
31 gender in Italian is acquired earlier than in French (Kupisch et al., 2002:116) because the formal
32 regularities are more reliable in Italian. The French system is not considered opaque, as in both
33 languages gender is unambiguously marked only on determiners, but Italian has a more reliable
34 morphology (Kupisch et al., 2002:108). Thus, also in the current study, we will look at degrees
35 of transparency. Kupisch et al. (2018) have already used this kind of approach to study the
36 acquisition of gender in Russian children acquiring German and have developed a continuum
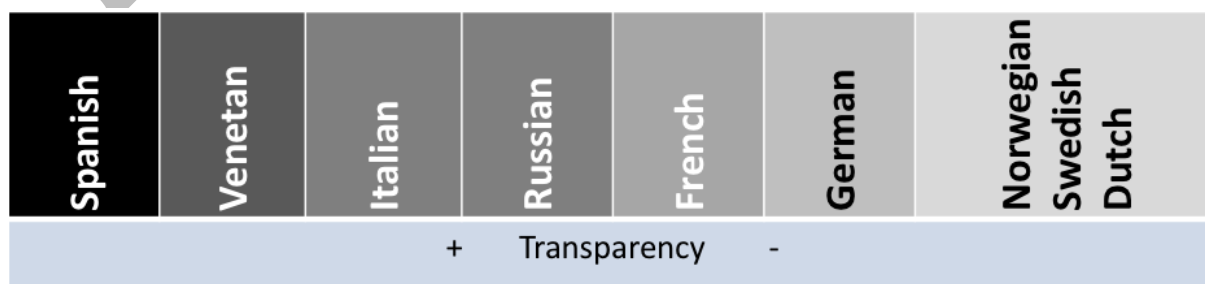37 of gender transparency that places these languages in the context of others (figure 1).

38
39 Figure 1: Transparency scale of languages from Kupisch et al. (2018)

40 In the following section the gender systems of Italian and Croatian will be thoroughly

1 explained, and we will propose where Croatian might be placed on this scale.

2 **3.The gender systems of Croatian and Italian**
3
4 In the following sections we will describe the gender system of Italian and Croatian in order to
5 outline the level of transparency of each language.
6

7 **3.1 The Italian gender system**

8 The Italian gender system has two values, masculine (M) and feminine (F), and the gender is
9 expressed through morphophonological properties of the noun ending (Chini, 1995). The
10 distribution of the two genders in the language is 60% of M nouns and 40% of F nouns (Costa
11 et al., 2003:186). According to Noccetti (2002:, M is the unmarked gender, and deriving F
12 nouns is possible via suffixations when the nouns are animate. Italian nouns can be divided in
13 various declension classes based on the noun endings in the singular and plural. The full
14 categorization of Italian nouns and their respective declension classes along with the
15 frequencies from LIP is displayed in table 1. The table is a summary from Gudmundson (2010),
16 we use the same terminology as the author.

Noun class	Gender	Ending sg/pl	Example	Translation	Frequency in LIP (%)
A	M	-o/-i	Libro/i	Book/s	39,80
B	F	-a/-e	Padella/e	Pot/s	31,40
C1	F	-e/-i	Volpe/i	Fox/s	21,49
C2	M	-e/-i	Cane/i	Dog/s	
C3	M&F	-e/-i	Insegnate/i	Teacher/s	
D	F	-á/-á	Abilità	Ability	2,96
E	M	-a/-i	Problema/i	Problems/s	1,46
F	M	-cons/-cons	Camion	Truck	0,70
G1	M	-a/-i	Artista/i	Artist/s	0,44
G2	F	-a/-e	Artista/e	Artist/s	
H1	F	-i/-i	Analisi	Analysis	0,40
H2	M	-i/-i	Domani	Tomorrow	
I	F	-o/-o	Foto	Photograph	0,24
J	M	-í/-í	Lunedì	Monday	0,29
K	M.sg -F.pl	-o/-a	Uovo/a	Egg/s	0,25
L	M	-è/-è	Caffè	Coffee/s	0,09
M	M	-a/-a	Cinema	Cinema	0,07
N	F	-ù/-ù	Virtù	Virtue	0,02
Other					0,41

17 Table 1: The Italian declension classes

1   From table 1 we can see that there is much variation when it comes to the noun endings in
2   Italian. Classes A and B are the transparent classes as the *-o/-i* and *-a/-e* endings unequivocally
3   signal M and F respectively. But with so much variation, how can this system be considered
4   transparent? The corpus study on the LIP corpus[1] (Voghera et al., 2014) conducted by
5   Gudmundson (2010: and summarized in table 1 revealed that the two groups of transparent
6   nouns constitute 71,2% of tokens used. The third most frequent noun class amounting to 21%
7   of the corpus consisted of the *–e/-i* ending and is ambiguous with respect to gender as it contains
8   M nouns (cane/i- "dog/s"), F nouns (volpe/i- "fox/es"), or both depending on the natural gender
9   of the referent (insegnate/i- "teacher/s"). However, this class can be broken down into groups
10 based on derivational morphemes. Some of these morphemes have been found to
11 unambiguously signal the gender of a noun, such as *-ione* (F) (stagione- "season"), *-tore* (M)
12 and *-trice* (F) respectively (attore/attrice- "actor/actress"), *-iere* (M) (giocolliere- "juggler"), -
13 *ame* (M) (fogliame- "foliage"), and *-udine* (F) (solitudine- "loneliness"); which according to
14 Gudmundson (2010: account for 52% of nouns class C. Thus, even if the majority of noun
15 classes in Italian can be classified as opaque, when the size of the groups and the nouns that
16 appear there in is considered together with frequency, it is evident that the majority of the nouns
17 (tot=92,68%) are transparent (Gudmundson, 2010:14). Thus, the Italian gender system is
18 transparent because the majority of Italian nouns ends in *-o/-i* (sg./pl.) which signals M, or in -
19 *a/-e* which signals F.

20   The gender of the opaque nouns becomes explicit through agreement, thus the rich
21 inflectional system of the nominal domain in Italian should be a valuable resource for the
22 acquiring child. Elements that have gender agreement with the noun are: articles, determiners,
23 adjectives, quantifiers, possessives, *wh*-words, relative clauses and the past participle (Chini,
24 1995; Gudmundson, 2010). Here, we will focus on the description of the article and adjectival
25 systems: the paradigm of the former is crucial for our hypotheses while the latter is relevant
26 because we used adjectives as test items in our task. We start by describing the article system
27 because it is an obligatory marking of the noun.

28   Italian articles are marked for definiteness, number, and gender, there is no plural form
29 for the indefinite article. All the Italian articles are presented in table 2.
30

Gender	Singular	Example	Plural	Example
	Definite			
M	il	Il libro the-M.SG book-M	i	I libri The-M.PL books-M.PL
	lo	Lo zaino the-M.SG backpack-M	gli	Gli zaini the-M.PL backpacks-M.PL
	l'	L' albero the-M.SG tree-M		Gli alberi the-M.PL trees-M.PL
F	la	La padella the-F.SG pot	le	Le padelle the-F.PL pots
	l'	L' ape the-F.SG bee		Le api the-F.PL bees
	Indefinite			
M	un	Un libro, a-M.SG book	NA	

---

[1] Corpus of spoken Italian with diaphasic, diatopic and diamesic varieties.

		un albero a-M.SG tree	
	uno	Uno zaino a-M.SG backpack	
F	una	Una padella a-F.SG pot	NA
	un'	Un' ape a-F.SG bee	

Table 2: The Italian article system

Italian articles are sensitive to the phonology of the initial syllable of the noun: il/un (M) and la, una (F) are the unmarked forms; lo, gli, and uno are used when the M noun starts with a */s/+consonant* (e.g. lo stallone/gli stalloni – "stallion/s") or with a */z/, /x/, /j/, /pn/, /ps/, /gn/,* and */sc/* (lo zaino/gli zaini- "backpack/s") whereas *l'* (both for M and F nouns) and *un'* (for F) are used with nouns beginning with a vowel (e.g. l'albero-M "tree", l'anatra-F "duck") (Pizzuto and Caselli, 1992:513; Gudmundson, 2010:25). Chini (1995) defines the paradigm of M article as more complex than that of the F article, because it contains more allomorphs, and also because the unmarked form for il does not end in *-o* as the typical M nouns do. The gender cue is however neutralized with *l'* especially if the noun also belongs to the C-class (l'ape, "bee").

The article is the first and most frequent agreement cue for gender, moreover, it is the only cue when the noun does not have transparent assignment (Chini, 1995:93).

Adjectives also agree in gender with the noun and their ending morphemes are equivalent to A and B declension classes in table 1, taking *-o/-i* for M and *-a/-e* for F. However, there are exceptions as some adjectives agree only in number and have the *-e/-i*, while some do not have variation at all. The types of adjectival agreement are presented in table 3.

Adjective	Gender	Ending sg./pl.	Example
Class 1	M	-o/-i	Piccolo libro[2]/ piccoli libri Small-M.SG book / small-M.PL books
	F	-a/-e	Piccola padella/ piccole padelle Small pot-F.SG /small pots-F.PL
Class 2	/	-e/-i	Grande libro/ grandi libri Big book-SG / big-PL books Grande padella/ grandi padelle Big-SG pot / big-PL pots
Class 3	/	/	Libro blu[3] / libri blu Book blue / books blue Padella rosa/ padelle rosa Pot pink / pots pink

Table 3: The Italian adjectival classes

---

[2] Note that the Italian adjective can be placed in pre- and postnominal position (i.e. *il libro piccolo/il piccolo libro*) and some variation to meaning may apply. However, this is not the focus of the current study as the adjective is gender marked in either position.

[3] Some adjectives can be placed pre- and post- nominally while others cannot, but this is beyond the scope of this paper, so this is not a differentiating criterion for the current study. Adjective also agree with the noun both in attributive (as above) and predicative position, but again, gender is marked in both cases so we will not differentiate between the two.

1   The adjective and the article co-occur with the noun as in Il libro bianco (The-M book
2 white-M) or La padella bianca (the-F pot white-F). Summarizing, the majority of Italian nouns
3 have a transparent gender assignment, but because of obligatory article inflected for gender
4 even these nouns are disambiguated.

5

3.1.1 The acquisition of the Italian gender system

7 It has generally been reported that the Italian gender system is acquired in a short time and with
8 very few errors in the article system (Kupisch et al., 2002:117). When it comes to the
9 distribution of declension classes in child speech and child directed speech (CDS), De Marco
10 (2005) found that the two transparent classes are the most represented ones, with nouns of what
11 we referred to as the C-class being less frequent. No reports on the other, much less frequent
12 declension classes, are made in that study. Thus De Marco (2005) found that A and B-class
13 nouns (singular) amount to 63% in CDS (average across the corpus), and to 66% in child
14 speech; whereas the C-class nouns are produced at 11% and 15%, respectively, by each group.
15 Noccetti (2002:62) also reports that the A and B declension classes (to which she refers to as
16 productive microclasses) are the most frequent ones both in CDS, in the child's utterances, and
17 also in adult to adult speech. She also provides an analysis which shows how the proportion of
18 noun types and tokens divided per gender is fairly similar in the child's productions and in
19 CDS.
20   When it comes to the opaque nouns, various studies (Caselli et al., 1993; Kupisch et al.,
21 2002; Noccetti, 2002; Belletti and Guasti, 2015) report that also these are acquired easily. A
22 corpus investigation conducted by Kupisch et al. (2002:115) revealed that 90% of all noun
23 tokens belong to the two noun classes defined as transparent, and additionally, in CDS,
24 diminutive suffixes are used rather often which makes the gender of the noun available also on
25 C class nouns (i.e. il cane/il canino- "the dog/doggie", la volpe/la volpicina- "the fox/foxy").
26 The Italian monolingual corpus investigated by Kupisch et al. (2002:117) contained very few
27 gender errors and they conclude that gender in Italian is mastered early without significant
28 deviations from the target gender. According to Kupisch et al. (2002:116) articles have a
29 relevant contribution for the early acquisition of gender in Italian, because transparency
30 reinforces the early acquisition of articles, and once they are acquired it is an additional
31 facilitation for the acquisition of gender. This is also found in Caselli et al. (1993) as children
32 produced articles in the same proportion with all noun classes. However, Noccetti (2002:
33 provides a more detailed analysis of a corpus of a single child and discusses a more sequential
34 acquisition of Italian gender. In the *premorphology phase* (2;00-2;3) of gender acquisition the
35 child makes some assignment errors (i.e. gender mistakes in the plural) and the nouns rarely
36 have agreement, i.e. the article is omitted. But by the end of this phase, there are no longer
37 assignment errors in the number paradigms. In the *protomorphology phase* (2;3-2;8) mistakes
38 in agreement of C-class nouns are present as they are generalized to M; however, the B-class
39 nouns (F gender) are productive, for example the child pluralized a proper name. In the final
40 phase, *modularized morphology* (2;8-3;4), the child uses the paradigms and agreement of all
41 genders correctly. Thus, even though there are obviously stages in the acquisition of Italian
42 gender, we can see form Noccetti's (2002) data that these phases are quite brief which probably
43 why there were not observed by other, perhaps less detailed, studies.
44   Thus, the presence of the article is thus crucial as it can be regarded as a kind of gender
45 morpheme (Chini, 1998:43). Accordingly, Pizzuto and Caselli (1992) found based on corpus
46 data of three children that the F article *la* is the first one to be attested (age=1;6). The errors of
47 commission observed in the study are related to the phonological choice of M article (Pizzuto
48 and Caselli, 1992:540). Pizzuto & Caselli's (1992) analysis shows that even if the articles are

1   attested early, only *la* can be considered acquired by all three children in the corpus. Moreover,
2   what they conclude is that none of the major inflectional paradigms that they investigated
3   (articles, pronouns, clitics, and verbs) is fully mastered by the age of 3;0. Thus, article as a
4   gender cue certainly contributes to acquiring the gender system, it does not entail that the system
5   is immediately mastered. Therefore, investigating whether gender is productive on agreeing
6   elements other than articles can give us insight into when the system is acquired. For example,
7   Caselli et al. (1993) conducted an elicitation task on three age groups and found that *la* has a
8   better accuracy than *il* in the youngest group, but agreement on adjectives is at ceiling in all
9   three age groups (Caselli et al., 1993:384). According to Caselli et al. (1993) Italian
10   grammatical morphemes reach full mastery at ages 3;6-4;0.
11   Bottari et al. (1993) analyzed the production of monosyllabic place holders that emerge
12   before lexical items and that the correct morpheme *la* form the F article is realized before its M
13   counterpart. They explain this in relation to *la* having vowel features analogous the noun, which
14   the M article does not have. This is strengthened by the fact that the plural article *le* is also
15   learned before *il*, in spite of the lower frequency of the plural in the input. They report that the
16   full version of *il* and la take over the place holders at 3;1 and 3;8 and thus claim that there are
17   two stages that lead to a correct production of articles: in the *linear agreement strategy* the
18   placeholder agrees with the noun ending, therefore *-a* is used for F, while an approximate vowel
19   is used for M. Thus, *il* is acquired when *morphological insertion strategy* takes place (Bottari
20   et al., 1993:363) when the right morpheme has to be inserted.
21   Previous studies on Italian cited in this section have reported an early mastery of the
22   gender system, even for the classes that are deemed less transparent, because of a frequent
23   gender cue in the form of an article.
24

3.2 The gender system of Croatian

26   Croatian has a rich inflection system and nouns are marked for gender, number and case. The
27   gender of the noun is an inherent property, while number and case are considered morphological
28   properties (Mihajlović, 2014:121). Croatian has three gender values: M, F, and N. An
29   investigation of the 4000 most spoken (noun) lemmas in adult spoken language revealed that
30   M nouns amount to 43,3%, F to 42,9% are F, and N takes up the remaining 13,7% (Vuletić,
31   1991) in (Kovačević et al., 2009:157).

32   The morpho-phonological properties of the noun act as a reliable cue for gender: M nouns
33   end in a consonant (jelen "deer", prozor "window"); F nouns end in –*a*, (kuća "house"; and N
34   nouns end in –*o* or –*e*, (nebo "sky", more "sea"). However, some exceptions apply to M and F
35   as there are M nouns that end in an *-o* (pepeo "ash", ugao "corner"), but also in an *-a* (gazda
36   "boss/owner"); and some F nouns end in a consonant (kost "bone") (Barić et al., 2005).

37   Croatian has three declension classes, but these do not match the three gender values.
38   The criteria for grouping the nouns in declension classes are based on the ending that the noun
39   takes in the genitive singular: class *-a*, class *-e*, and class *-i*. The first class includes M and N
40   nouns, the second class includes mostly F nouns with some M exceptions such as gazda
41   "boss/owner", while the *-i* class contains only F nouns, the ones ending in a consonant (Barić
42   et al., 2005:103). Thus, F nouns span over two classes, whereas M and N share a declension
43   class. Nouns in the *-i* declension class (F nouns ending in a consonant) are opaque and were not
44   included the study.
45   In table 4 we present the full paradigm of Croatian declension. This includes the three
46   declension classes divided into their gender values. The M gender additionally has an animacy
47   sub-gender which entails a separation between M animate and inanimate nouns. These differ

1 only in accusative singular[4] where the inanimate noun takes a -Ø ending (prozor-acc "window")
2 unlike the animate noun that has an -a ending (jelena-acc "deer").
3

Singular	Class -a		Class -e	Class -i
	M	N	F	F
NOM	Jelen (deer)	nebo (sky)	Kuća (house)	Kost (bone)
ACC	jelen-a	nebo	kuć-u	kost
GEN	jelen-a	neb-a	kuć-e	kost-i
DAT	jelen-u	neb-u	kuć-i	kost-i
VOC	jelen-e	nebo	kuć-o	kost-i
LOC	jelen-u	neb-u	kuć-i	kost-i
INS	jelen-om	neb-om	kuć-om	kost-i
Plural				
NOM	jelen-i	neb-a	kuć-e	kost-i
ACC	jelen-e	neb-a	kuć-e	kost-i
GEN	jelen-a	neb-a	kuć-a	kost-iju
DAT	jelen-ima	neb-ima	kuć-ama	kost-ima
VOC	jelen-i	neb-a	kuć-e	kost-i
LOC	jelen-ima	neb-ima	kuć-ama	kost-ima
INS	jelen-ima	neb-ima	kuć-ama	kost-ima

4 Table 4: The Croatian declension system
5

6 In table 4 we can see the overwhelming syncretism between M and N (shaded in gray). This is
7 not surprising considering that these two genders are part of the same declension class. In this
8 kind of system, the declension class is not really informative for gender and this affects the
9 overall transparency of the system as the gender of a noun is opaque when expressed in these
10 cases. The syncretism of the -ima (pl) ending also extends to the F -i stem declension class in
11 DAT, LOC, and INS case. The F nouns from class -e, included in this study, are only syncretic
12 (same form and function of the suffix) with the other two gender values in INS.SG and GEN.PL
13 We can see from table 4 that the F nouns in class -e have also some other shared suffixes as the
14 suffix -a is shared by NOM-F and ACC-M, the suffix -u is shared between dat.m and ACC.F, -e is
15 shared between VOC.M and GEN.F. But since words are rarely produced in isolation, it will be
16 obvious from the context, for example, whether a word is in DAT or in ACC case. Thus, when
17 case is disambiguated, the gender should not be ambiguous either. Thus, even if there are
18 syncretism across the paradigm including all three genders, it is crucial to note that M and N
19 are part of the same declension class and that are syncretic within the same case, whereas the
20 syncretism between M and F is across case. Taking this into account, the F (-e class) is the most
21 transparent of the three genders in Croatian as it has a clearly distinguishable gender marker
22 across the paradigm. Thus, the Croatian system is transparent when the base form (i.e.
23 Nominative) is considered, but when we look at the full paradigm ambiguities between M and
24 N arise.

25 When it comes to agreement, demonstratives, possessives, adjectives and the periphrastic
26 past tense show gender agreement. In (1) we can see the demonstrative agreeing with the M,
27 the possessive with the F, and the adjective with the N noun. The verb is in agreement with the
28 subject, which in (1) is boy-M. However, since these are not obligatory elements, the Croatian

---

[4] This sub-division of the M gender value is not expected to interfere with the current study because (i) the target of the elicutation task is in Nominative, (ii) the declesion of adjectives does not differ based on the animacy criteria

1 noun is often bare, with only the morphology of the noun itself serving as a gender cue; like in
2 (2): consider the three nouns that are present- boy, house, and tree- and how none of the nouns
3 co-occurs with a determiner.

4 (1) Onaj   dječak   je sagradio   svoju   kučicu
5   That-NOM.M boy- NOM.M is-aux built-M his-REFL.ACC.f house-DIM.ACC.F

6   na   velikome   stablu.
7   on big-LOC.N tree- LOC.N

8 (2) Dječak   gradi kučicu   na   stablu.
9   Boy-NOM.M builds house- DIM.ACC.F on tree-LOC.N

10

11 Regarding the agreement paradigm, we will outline only on the adjectival one because
12 adjectives were elicited in our task. In table 5, the declension of the adjective beautiful is
13 displayed per gender. Adjectives agree with gender and are not dependent on the declension
14 class of the noun. Notice that M has a subdivision of adjectives in *short* and *long*: the main
15 distributional difference is that the short adjective can be in predicative position and the long
16 one cannot, while both can be in attributive position and were thus considered valid responses
17 in the elicitation task. The meaning and the use of the two variants is beyond the scope of this
18 paper, for an overview check (Aljović, 2002; Trenkić, 2004; Velnić, 2015)

19

	Singular				Plural		
	M		N	F	M	N	F
	short	long					
NOM	lijep-ø/	lijep-i	lijep-o	lijep-a	lijep-i	lijep-a	lijep-e
ACC	lijep-a/	lijep-og	lijep-o	lijep-u	lijep-e	lijep-a	lijep-e
GEN	lijep-a/	lijep-og	lijep-og	lijep-e	lijep-ih	lijep-ih	lijep-ih
DAT	lijep-u	lijep-om	lijep-om	lijep-oj	lijep-im	lijep-im	lijep-im
VOC	/	lijep-i	lijep-o	lijep-a	lijep-i	lijep-a	lijep-e
LOC	lijep-u	lijep-om	lijep-om	lijep-oj	lijep-im	lijep-im	lijep-im
INS	lijep-im	lijep-im	lijep-im	lijep-om	lijep-im	lijep-im	lijep-im

20 Table 5: The Croatian adjectival paradigm

21 From table 5 we can see once more how M (long) and N are neutralized in some places of the
22 paradigm, but also how in the plural adjectives, gender distinction is barely made. However,
23 we are currently not testing plurals.

24   Overall, the gender system of Croatian is transparent as the noun endings in the
25 nominative singular quite unambiguously indicate the gender value of the noun; also, the
26 agreeing elements display a transparent gender marker. However, when the whole nominal
27 paradigm is taken into consideration, the initially straightforward transparency becomes more
28 limited: the M and N genders are revealed as much closer than initially shown by their
29 declension in the nominative case, as syncretism of the declension is present in four out of the
30 seven cases. F is the most transparent gender, as the syncretism it has with the other genders is

1 within case only in ins in the nominal paradigm and never within case in the adjectival
2 paradigm.

3

3.2.1 The acquisition of the Croatian gender system

5 The acquisition of Croatian is overall understudied, and thus reference to numerous studies is
6 not possible. Croatian is a language with rich derivational morphology which enhances the
7 transparency of lemmas, and this morphological richness could act as a booster for the
8 acquisition of the inflectional paradigms (Kovačević et al., 2009:153). Kovačević et al. (2009:
9 have analyzed corpus data from one of the children present in the Kovačević (2004) corpus
10 from ages 1;3-2;8. The study focuses mostly on the child's incremental acquisition of the case
11 system and only marginally looks at gender, focusing mainly on the use of nouns in relation to
12 the three genders rather than the gender markings on agreeing elements. They noticed that the
13 distribution of the three genders reflects the distribution of the nouns in child directed speech
14 of the same corpus: 35% M, 59% F, and 6% N. Consequently, the child produced mostly F
15 nouns (58%) followed by M (36%) and N (6%). Other than that, there are no notes on the gender
16 agreement of these nouns, nor are there any indications of how accurate agreement is. However,
17 in relation to the acquisition of the case paradigm, the authors state that the first opposition of
18 inflected forms are found in regular F nouns between the Nominative and Accusative case
19 (Kovačević et al., 2009:172). They find that the child grasps the complex case system easily
20 and is using all seven cases, with different frequencies, at age 1;10. This implies that all the
21 children that participated in our task should be able to use the full declensional paradigm. The
22 corpus results from Kovačević et al. (2009:165) report the nominative and accusative reaching
23 together 78% of the child's production at age 2;5. This prevalence of the two aforementioned
24 cases is also evident in the input. Recalling table 4, NOM and ACC are not syncretic at any
25 level, and considering the very low frequency of N nouns (6%), Croatian children might still
26 be at the stage in which the syncretism between M and N is not evident. Since the corpus
27 contains data only until 2;8, evidence of a more distributed use of the case paradigm or a more
28 frequent use of N nouns is not provided. We might assume that the neutralization of M and N
29 becomes more evident with increased exposure and usage of the full case paradigm, and we can
30 also speculate that this might be reflected in the acquisition of gender, namely that it will take
31 children longer to master M and N because it will take them longer to become aware that these
32 nouns belong to different genders.

33

3.3 Main differences between Italian and Croatian

35 Here we will concisely summarize the crucial differences between Italian and Croatian which
36 are relevant for the predictions that we make (section 4). The facilitative factors are shaded in
37 grey.

	Number of gender values	Even distribution of gender values in the language	Obligatory determiner	Opaque nouns (assignment)	Syncretism in the paradigm
Italian	2	yes	yes	yes	limited
Croatian	3	no	no	no (yes)5	extensive

38 Table 6: Crucial differences between Italian and Croatian that can influence the degree of

---

[5] The -*i* class F nouns are opaque, but not part of the current study

1 gender transparency

2 It is obvious that Italian is the more transparent gender system: it has fewer gender values and
3 these are more evenly distributed, it has the obligatory gender marked determiner (the article),
4 and the syncretism of the article is limited to the definite article l' used when the noun begins
5 in a vowel. Both languages have a set of nouns that are considered opaque: for Croatian this is
6 the -i declension class, but it is excluded from the task in this study giving Croatian a less
7 opaque gender assignment system for the purposes of this particular study, while Italian has
8 what we have referred to as C-class of nouns (Gudmundson, 2010:14). The gender these nouns
9 is however clearly disambiguated with the obligatory gender-marked article. The Italian
10 nominal paradigm is more limited than the Croatian one because it does not include case. Thus,
11 when comparing Croatian noun to Italian ones only in the nominative, the two systems seem
12 transparent to the same degree. However, when all cases in Croatian are taken into account, it
13 becomes clear that M and N are part of the same declension class and thus the paradigmatic
14 differences are often neutralized, leaving F the least ambiguous gender value. Additionally, the
15 Italian M and F values are quite evenly distributed in speech (Costa et al., 2003:186), as they
16 are in Croatian, however the N here is significantly less frequent than the remaining two gender
17 values (Kovačević et al., 2009:161) which could affect the time that it takes for children to
18 master it.

19 Going back to figure 1 presented in section 2, we may now place Croatian on the transparency
20 scale. From what has been said in the current section, Croatian is less transparent than Italian
21 and should thus be placed to the right of it. On the scale presented by Kupisch et al. (2018:,
22 Russian is placed next to Italian. We deem Croatian more transparent than Russian because
23 Croatian does not have ambiguities in the assignment system of the gender as the ones described
24 by Rodina and Westergaard (2017:. The updated scale is presented in figure 2.

25 Figure 2: updated transparency scale

26 It is interesting that Italian and Croatian are next to each other on this scale, and thus we will
27 be observing differences in the acquisition of two gender systems which are transparent to a
28 very similar degree.

29
30 **4. The current study**
31
32 The importance of this study lies in the fact that we are comparing two languages that have
33 transparent gender systems, which has been repeatedly found to be the key to an early
34 acquisition of gender (Levy, 1983). Based on this premise, both Italian and Croatian children
35 are expected to have an early grasp of their respective gender systems. As Caselli et al.
36 (1993:379) put it '*it is widely accepted that inflectionally rich languages promote an early
37 acquisition of morphology*', including Italian in this list of languages. Kovačević et al. (2009)
38 also supposed that the morphological richness of Croatian may promote the early development
39 of inflection.
40   However, the two target languages differ in other relevant features which may affect the
41 timing of acquisition of the system and the individual gender values, even if gender is acquired
42 early in both groups. We thus expect a high rate of correct agreement, but the intra-language

1 differences are expected to affect the timing of and perhaps the order in which a particular
2 gender is mastered. Thus, the relevance of this study lies in the fact that we are analyzing how
3 different factors in transparent languages can affect gender acquisition. Additionally, the
4 acquisition of the gender system in Croatian is relatively new territory, with only one study
5 taking it into consideration: Kovačević et al. (2009) but not from the perspective of agreement.
6 It will thus offer a starting point for further investigation of the acquisition of gender in
7 Croatian, according to which more specific research questions could be tackled through more
8 precisely designed tasks.
9 　　　We have decided to elicit adjectives because the adjectival paradigms because of their
10 optional status in both languages and the fact that in both languages the declension paradigm is
11 regular and thus offers good grounds for comparison between the two languages. The study
12 aims to answer the following research questions:

13 　　1. Are Croatian children slower than Italian children to acquire the gender system?
14 　　2. Is the most regular gender (feminine) acquired first in both languages?
15 　　3. How do the acquisition paths differ in the two languages?

16
17 Based on the properties of the two gender systems and what has been found so far in previous
18 studies, we predict that Croatian will be acquired more slowly when compared to Italian. This
19 prediction is primarily based to the syncretic distribution of infections of M and N in some parts
20 of the case system (table 4) accompanied by the low frequency of N nouns in CDS. Whereas in
21 Italian the presence of the obligatory gender-marked article is a relevant cue and children have
22 been found to acquire gender unproblematically.
23 　　　If regularity and transparency are relevant factors, F should be the gender that is acquired
24 first or with more accuracy in both languages. Recall from sections 3.1 and 3.2 that F is the
25 more regular gender in both languages as in Italian the F article is more regular than M due to
26 its simpler realization, and in Croatian F nouns do not have within-case syncretism with other
27 gender values, while M and N show extended similarities.
28 　　　We expect the three Croatian genders to be acquired in different stages, as we expect the
29 low frequency of N to result in a later acquisition of this gender value.
30 　　　Overall, this study will compare the acquisition of two easily acquirable gender system,
31 and it will thus reveal the nuances of how the gradience of gender transparency affects accuracy.
32
33 **5. Methodology**
34
35 The task consisted of adjective elicitation using images of animals and objects denoting
36 referents of different genders. We have chosen to elicit adjectives since they are optional and
37 agree in gender with the noun in both languages.
38
39 **5.1 Participants**
40
41 The research was conducted on two samples of children, native speakers of Italian (n=30, 11
42 male) and native speakers of Croatian (n=30, 16 male). Each language group included two age
43 groups of 15 children each. The children that were chosen for this research were required to be
44 native and monolingual speakers with both parents of either Italian or Croatian nationality. The
45 parents/caregivers were informed about the testing and had to sign a consent form in order for
46 their child to participate.
47 　　　Ensuring age balance across the language groups was not entirely possible due to the
48 availability of children and parental consent. The language groups are still close in age and we
49 are able to make relevant statistical comparisons. The mean age of the younger Croatian group
50 (n=15, 6 male) was 2;10, whereas the older group (n=15, 10 male) had a mean age of 4;2. The

1 mean age of the younger Italian group (n=15, 3 male) was 3;0, and the older group (n=15, 8
2 male) had a mean of 3;10. This means that the age range of the Croatian children is wider than
3 the one of the Italian participants, which is also convenient as we expect this group to acquire
4 gender at a slower pace than the Italian children and with a wider age range, we will be able to
5 observe this potential difference more clearly.
6   Two participants were excluded from the younger groups due to their general lack of
7 adjective production; both children, one from the Italian and one from the Croatian group, were
8 2;3. This does not automatically suggest that gender is not or cannot be acquired before age 2;6,
9 but it indicates how children are not producing enough adjectives at this age.
10
11 **5.2 Materials**
12
13 Thirty images depicting animals and inanimate objects were used. The images were
14 downloaded from the Internet from open source websites. These images were selected on the
15 basis of the grammatical gender of the depicted noun in Italian and Croatian. The same set of
16 images was used for both the Croatian and Italian groups. This means that in Italian fifteen
17 images represent an M noun and the other fifteen represent an F noun. In Croatian, these were
18 distributed as follows: ten are M, ten are F, and ten are N. The test items were chosen for both
19 languages tracking in parallel the genders they were assigned to in each language. So for
20 example, when the item *tree* has been chose as stimuli, it was counted as an M noun for Italian
21 and N for Croatian, conversely the item *snail* took a F slot for Italian and M for Croatian, the
22 item *cat* took a M slot in Italian and a F slot in Croatian. As we went along with the choice of
23 the stimuli, these had to fill specific gender slots in both languages depending on how many
24 were left for each gender. Unfortunately, due to matching the referents to the genders cross-
25 linguistically, the Italian C-class nouns (n=6) are contained only within M. However, as the
26 previous literature has stated, the article is a very reliable marker of gender (Chini, 1995) and
27 Italian children have been found to master these types of nouns easily (Belletti and Guasti,
28 2015).
29   All the items used are displayed in table 6 and are divided by language and gender. Due
30 to a considerable amount of N referents denoting abstract concepts, nouns such as *sky* had to be
31 included. We did not explicitly check for frequency of these nouns in the two languages, but
32 during the piloting of the task all the participants were able to name all the images included.
33 The only somewhat problematic item was *wing*, but it could not be replaced due to a low
34 availability of nouns that were N in Croatian and F in Italian.
35
36

English translation	Italian noun	Gender in Italian	Croatian Noun	Gender in Croatian
Lion	Leone	M	Lav	M
Frog	Rana	F	Žaba	F
Sun	Sole	M	Sunce	N
Grapes	Uva	F	Grožđe	N
Penguin	Pinguino	M	Pingvin	M
Snail	Lumaca	F	Puž	M
Tree	Albero	M	Stablo	N
Cat	Gatto	M	Mačka	F
Giraffe	Giraffa	F	Žirafa	F
Sheep	Pecora	F	Ovca	F
Sea	Mare	M	More	N
Dog	Cane	M	Pas	M
Butterfly	Farfalla	F	Leptir	M
Apple	Mela	F	Jabuka	F
Egg	Uovo	M	Jaje	N
House	Casa	F	Kuća	F
Mouse	Topo	M	Miš	M
Star	Stella	F	Zvijezda	F
Heart	Cuore	M	Srce	N
Bull	Toro	M	Bik	M
Ladybug	Cocinella	F	Bubamara	F
Moon	Luna	F	Mjesec	M
Eye	Occhio	M	Oko	N
Monkey	Scimmia	F	Majmun	M
Flower	Fiore	M	Cvijet	M
Pear	Pera	F	Kruška	F
Ear	Orecchio	M	Uho	N
Candle	Candela	F	Sviječa	F
Sky	Cielo	M	Nebo	N
Wing	Ala	F	Krilo	N

Table 6: Distribution of the test materials across the genders in the two languages

1 All the images were printed in black and white on thick paper. Some of the example images are
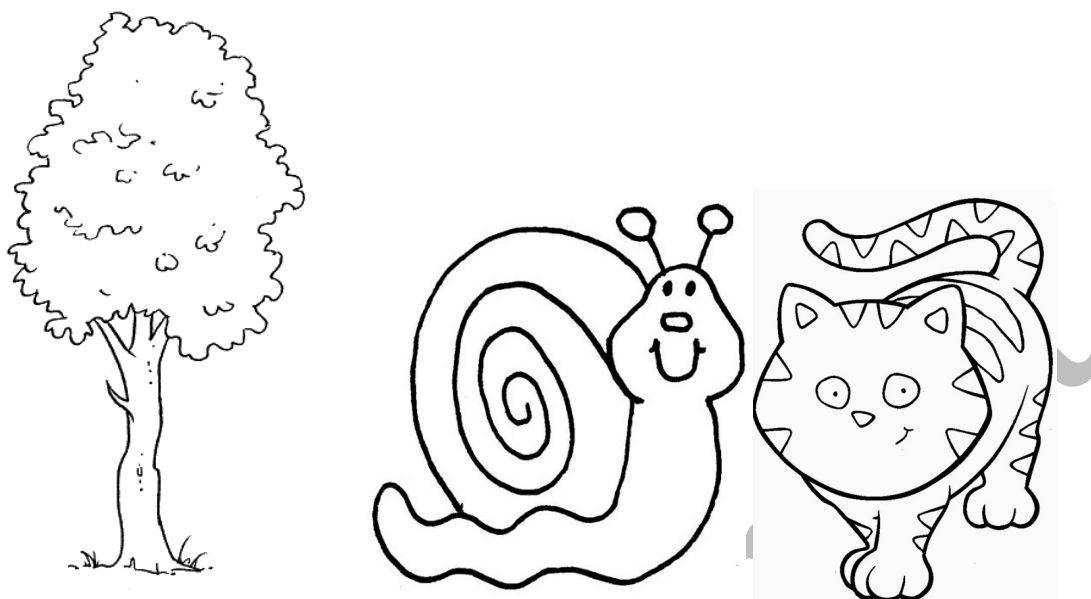2 provided in figure 3.
3

4
5 *Figure 3: Some of the images used in the task: Tree- M in It, N in Cro; Snail, F in It, M in Cro; and Cat- M*
6 *in It, F in Cro*

7
8 **5.3 Procedure**
9
10 All the participants were interviewed individually. Interviews were held in a slightly isolated
11 area of the kindergarten in order to avoid acoustic interferences. Interviews lasted from five to
12 fifteen minutes.
13   Each child was asked to describe what was represented in the image. The children were
14 explicitly asked before the testing to say whether the referents of the images were *beautiful,*
15 *good, small*[6] etc. until it was clear to the participant that the description had to include an
16 expression of a quality of the shown object. After making sure that the children understood the
17 task, the testing proceeded. The images were shown in a random order, which was
18 accomplished by taking them from a bag, and the participant was asked to describe them. The
19 interviews were recorded with a digital recorder Panasonic RR-US430 either held in the
20 interviewer's hand or laid on the desk. The uttered adjectives were manually transcribed on a
21 paper during the interview, and this was then cross-referenced with the recording.
22   When the child produced incorrect agreement, the question was repeated in order to
23 establish if that was just a simple distraction or if it was a non-target-like gender production. If
24 the child made a correct agreement the second time, the response was considered correct,
25 otherwise it was scored as incorrect.
26
27 **6. Results**
28
29 First, we will look at the distribution of correct responses in the data divided per group. An
30 answer was counted as correct if the adjective gender matched the gender of the noun, it was
31 incorrect if it did not. The non-applicable (NA) category includes the following cases: no
32 response (n=155), no adjective produced (n=54), and use a non-gendered adjective (only for
33 Italian children) (n=35).

---

[6] All of these adjectives agree in gender in Italian

1

Target gender	Correct	Incorrect	NA	Total
M	165 (73%)	/	60 (27%)	225
F	166 (74%)	4 (2%)	56 (25%)	225

2 Table 7: Distribution of correct/incorrect answers in the Italian younger group

3

Target gender	Correct	Incorrect	NA	Total
M	205 (91%)	/	20 (9%)	225
F	201 (89%)	/	24 (11%)	225

4 Table 8: Distribution of correct/incorrect answers in the Italian older group

5

Target gender	Correct	Incorrect	NA	Total
M	123 (82%)	8 (5%)	19 (12%)	150
F	117 (78%)	8 (5%)	25 (16%)	150
N	110 (73%)	17 (11%)	23 (15%)	150

6 Table 9: Distribution of correct/incorrect answers in the Croatian younger group

7

Target gender	Correct	Incorrect	NA	Total
M	137 (91%)	6 (4%)	7 (5%)	150
F	145 (97%)	1 (0,06%)	4 (3%)	150
N	135 (90%	9 (6%)	6 (4%)	150

8 *Table 10: Distribution of correct/incorrect answers in the Croatian older group*

9 We can see that both language groups had a high accuracy rate and we can safely assume that
10 they are aware of the nominal category of gender and use it accordingly.
11 At first glance, it seems that the Croatian groups produce more errors. Subsequent statistical
12 analyses will show whether this difference is significant. The Italian children have more NA
13 data than the Croatian groups, however this is not due to their low response rate, but to the fact
14 that some adjectives in Italian do not agree in gender and these errors amount to a total of 35
15 items in the dataset.
16   In the following sections we will address the research questions laid out in section 4 more
17 specifically. However, for the rest of the analyses we will focus on the binary distinction
18 consisting of correct/incorrect gender agreement, and we will thus not consider the NA
19 responses.
20
21 **6.1 Timing of gender mastery**
22
23 In figure 3 the data points used in the subsequent statistical analyses are displayed. The
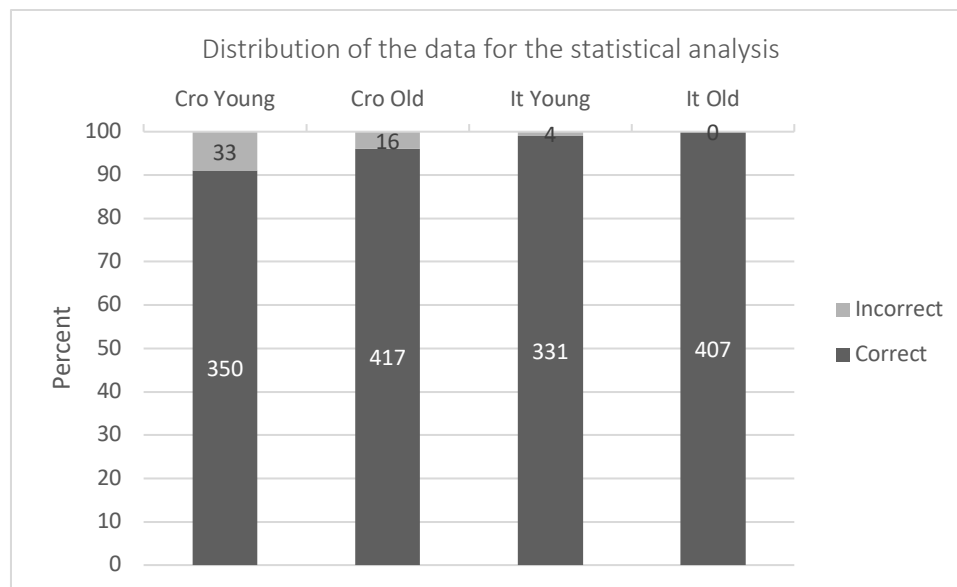24 percentiles are displayed on the y-axes, but the stacked columns contain the raw data.

Figure 3: Distribution of the data for the statistical analysis

Looking purely at the percentiles, it seems that the Italian children have acquired gender already in the younger group whereas the Croatian children have not. Nonetheless, figure 3 does not tell us the significance of these different error rations and does not show the error ratio per gender value. Thus, a statistical analysis is in order.

We have conducted a linear regression using Jamovi (Love et al., 2018). Here, group (young vs. old), language (Croatian vs. Italian), and target gender (M, F, and N[7]) were set as factors. The outcome of the analysis is presented in table 11.

Predictor	Estimate	SE	t	p
Intercept	0.95704	0.00504	190.006	< .001
old – young	0.03154	0.00904	3.489	< .001
croatian – italian	-0.03623	0.01004	-3.610	< .001
F – M	0.00138	0.00992	0.139	0.890
N – M	-0.05309	0.01411	-3.763	< .001

Table 11: Linear regression of all groups

The intercept was set to the values of *young*, *Italian*, and *M*. The fact that the Intercept shows a highly significant value means that the distribution of correct and incorrect responses is significant, and from the raw data in tables 7-10, it is obvious that correct responses were predominant in the task. Furthermore, the significance at the group and at the language level shows that the four groups we tested show significant differences in their responses, however the nature of these differences is not yet evident from the current analysis. The fact that there is no difference in the responses to the F and M gender means that the children master it to the same extent in both languages; the difference between the N and M refers only to the Croatian group and we can see that the children are significantly less accurate with one of the genders.

---

[7] Note that N was a value only available in the Croatian groups.

1 In light of the raw data, it is evident that N agreement is less accurate, but more precise statistical
2 analysis will reveal whether this difference is significant.

3   The next step in our analysis is to look at the language groups by conducting a linear
4 regression on the Italian and Croatian groups separately. This separation will shed light on the
5 source of the significant difference in table 11. We will first look at the Italian group (table 12)
6 and then proceed to the Croatian group (table 13).

Predictor	Estimate	SE	t	p
Intercept	0.9940	0.00269	369.49	< .001
old – young	0.0118	0.00538	2.19	0.29
F – M	-0.106	0.00536	1.98	0.48

7 Table 12: Linear regression of the Italian groups
8
9 Apart from the intercept, the statistical analysis does not find any significant differences in the
10 distribution. This means that there are no significant differences between the two age groups of
11 Italian children and that both genders are acquired equally well. This means that the Italian
12 children master gender by the age of 2;6, which is the age of the youngest participant taken into
13 consideration. The task does not offer insight into gender mastery prior to 2;6.
14   The same analysis was conducted for Croatian children and is presented in table 13.

Predictor	Estimate	SE	t	p
Intercept	0.9384	0.00826	113.652	<.001
old – young	0.0491	0.01652	2.971	0.003
F – M	0.0171	0.02017	0.846	0.398
N – M	0.0624	0.02022	3.084	0.002

15 Table 13: Linear regression of the Croatian groups
16
17 As in the previous analyses, the significant difference in the intercept means that there are
18 significantly more correct answers, which was already evident from the raw data. The Croatian
19 group also has a significant age difference which means that the correct answers increase
20 significantly with age (p=0.003). Again, there is no difference between M and F, but there is a
21 significant difference between M and N (p=0.002). This suggests that N is acquired later than
22 the other two genders in Croatian. In order to check for this, we have to make further analyses.
23 We have thus tested another dataset from which we excluded N (table 14). If the difference in
24 the age groups is no longer significant, it would mean that the low accuracy of N is the sole
25 reason for the observed difference in table 13.
26

Predictor	Estimate	SE	t	p
Intercept	0.9567	0.00860	111.18	< .001
old – young	0.0380	0.01721	2.21	0.028
F – M	0.0172	0.01718	1.00	0.316

Table 14: Linear regression of the Croatian groups with no neuter gender

1   The analysis showed that the difference between the younger and older children is still
2   significant, but to a lesser degree, when N is not accounted for. This means that N strongly
3   contributes to the age difference of Croatian children, but it is not the only factor and Croatian
4   children acquire gender at a slower pace than Italian children do (statistical difference between
5   the two age groups not observable for the Italian group).

6   **6.2 The acquisition of F**

7   Due to a greater regularity of F throughout the paradigm, both in assignment and agreement in
8   the two languages, we have hypothesized that it will be the first gender to be mastered. Figures
9   4 and 5 show the distribution of the correct/incorrect responses divided by gender in the two
10 languages and display the percentage of correct answers per gender and per group.
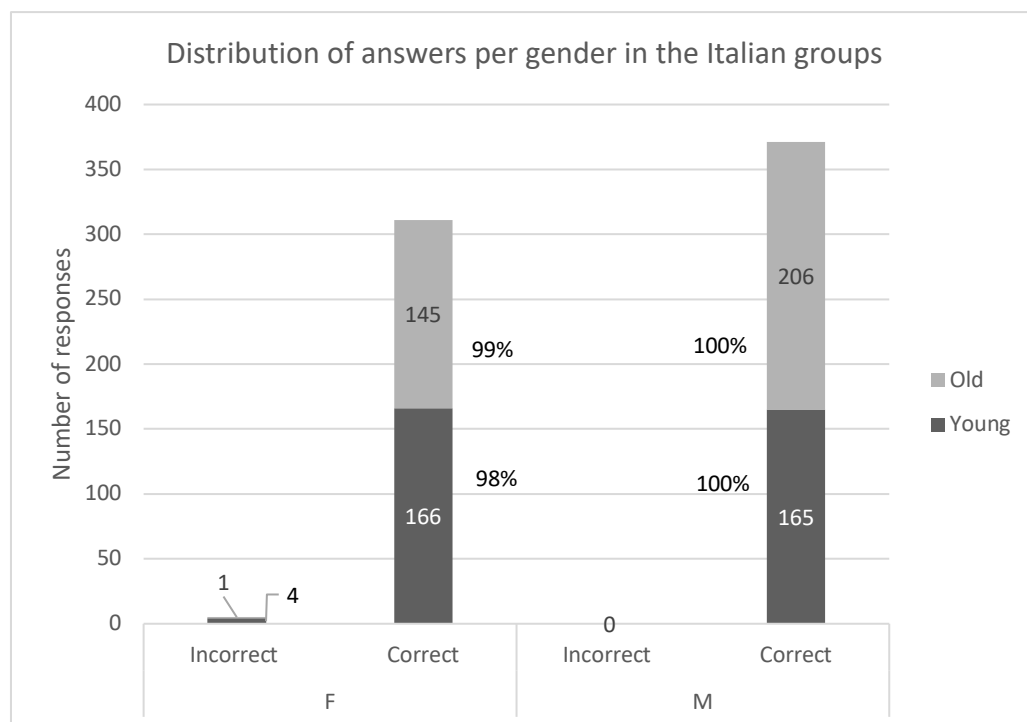

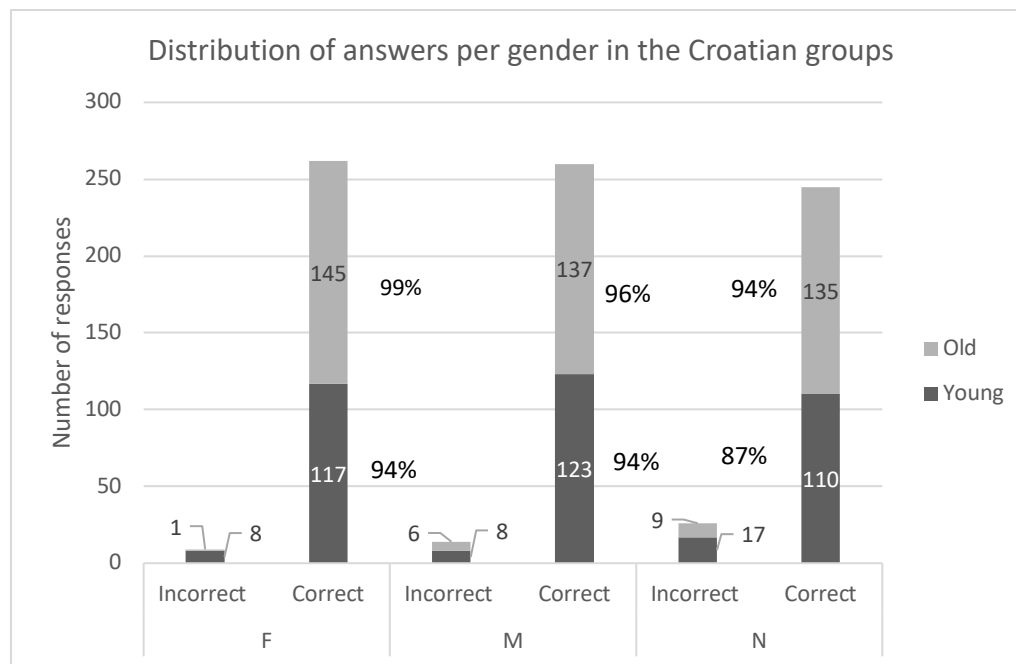12 Figure 4: Distribution of answers per gender in the Italian groups

Figure 5: Distribution of answers per gender in the Croatian groups

We can see that the Italian children are very accurate in both genders even in the young group, but nevertheless the difference between F and M is evident as the latter is at ceiling-level already in the younger group. This does not fit the previous findings for the article system, but it fits the explanation from Pérez-Pereira (1991: that M is the unmarked gender in Spanish.

Our prediction holds for Croatian since F is the first gender to become error-free in our task. For the other two genders, the errors are still present in the older group, even if diminished. The possible reasons for this are discussed in the next section where we look at the error distribution in the two language groups. Based on this, we could argue that gender agreement in Croatian is not fully acquired in the younger group, i.e. by age 3;4 (age of oldest participant).

We have conducted ANOVAs on the distribution of answers for each gender to see how it changes with age. Tables 15-18 display the results of each ANOVA. It was not possible to conduct an ANOVA for M in Italian as there were only correct answers.

	Sum of Squares	df	Mean Square	F	p
Group	0.0510	1	0.0510	4.82	0.029
Residuals	3.9059	369	0.0106		

Table 15: ANOVA age group comparison of responses in the Italian feminine

	Sum of Squares	df	Mean Square	F	p
Group	0.220	1	0.2200	6.98	0.009

	Sum of Squares	df	Mean Square	F	p
Residuals	8.481	269	0.0315		

1   Table 16: ANOVA age group comparison of responses in the Croatian feminine

2

	Sum of Squares	df	Mean Square	F	p
Group	0.0250	1	0.0250	0.512	0.475
Residuals	13.2597	272	0.0487		

3   Table 17: ANOVA age group comparison of responses in the Croatian masculine
4

	Sum of Squares	df	Mean Square	F	p
Group	0.344	1	0.3436	3.99	0.047
Residuals	23.162	269	0.0861		

Table 18: ANOVA age group comparison of responses in the Croatian neuter

5   The statistical results from above need to be discussed by considering the distributions shown
6   in figures 3 and 4. For Italian it is obvious that F is not the gender that is mastered first as there
7   are no errors with M, whereas for F a significant improvement can be noted. When it comes to
8   Croatian, the results have revealed a significant difference, i.e. improvement of F, likely due to
9   the reduction of the errors from 8 to 1. The difference is not present for M as it reflects that M
10   errors are still present in a similar proportion. When it comes to N, the group difference is
11   significant, which means that the children have improved their agreement with N, but it still
12   remains the gender with the most errors.

13   **6.3 Error patterns**

14   In this section, we will look a bit more closely into the errors that the children make, more
15   precisely, which gender is used instead of the target gender. The answer is straightforward for
16   the Italian children as no mistakes are made with M and since the language has only two genders
17   this means that all the errors made were M agreement on a F target.

18   The error pattern in Croatian might reveal different factors at play in the two age groups. The
19   responses are summarized in table 19, the shaded cells marking a target-like response.

20

		Response gender		
Target gender	group	F	M	N
F	old	144	0	1
	young	117	3	5
M	old	2	137	4

	young	1	123	7
N	old	3	5	134
	young	3	14	109

1 Table 19: Distribution of gender responses in Croatian children

2 F has a similar distribution of errors among M and N. Within the errors of the other two genders
3 we can see that there is tendency to mistake N with M and vice versa, but F realization for both
4 genders is also present. This tendency is likely due to the syncretism that the two genders have
5 across the case paradigm. With respect to the age groups, M and F proceed together at an early
6 stage, with the N lagging in accuracy. Whereas at a subsequent stage, F is acquired while M
7 and N are at the same level. The possible reasons for this will be considered in the discussion.

8 **7. Discussion**

9 With this task we have strived to reveal whether Italian and Croatian children differ in the time
10 course of the acquisition of their gender system and whether individual gender values are
11 acquired at a different pace due to the differences present in the two systems. In this section we
12 will outline the results in relation to the literature presented throughout the paper and our
13 predictions in order to identify the possible implications for the field of gender acquisition
14 related to transparency.

15 We have predicted that the Italian children will have a higher accuracy rate due to a higher
16 degree of transparency which is provided by (i) less ambiguous declension classes and (ii)
17 presence of an obligatory gender-marked article which acts as a gender cue in case the noun
18 itself does not provide it. The results confirmed this prediction as there were no significant
19 differences between the two age groups in Italian, but there were in Croatian. This entails that
20 the Italian children have mastered adjectival agreement already at age 2;6. This is in line
21 with previous research regarding the acquisition of gender in Italian (Pizzuto and Caselli, 1992;
22 Kupisch et al., 2002). Thus, from a theoretical perspective, the fact that the Italian gender
23 system is morphologically transparent and has a salient gender cue on the article, makes the
24 Italian gender system very easy to acquire. For Croatian, we cannot provide a confirmation of
25 the mastery of adjectival gender agreement by age 3;6 due to only F being error-free. The
26 obtained result indicates that the transparency level of e gender system is related to its
27 acquisition and it can thus be used to make predictions related to the timing and ease of
28 acquisition.

29 The results that we have obtained for Croatian are more central to the discussion, as no
30 previous study has investigated gender acquisition in Croatian from an agreement perspective.
31 We found that Croatian children also make few errors, which is expected as the system is
32 transparent. However, the errors that the children made reveal that the Croatian gender system
33 is acquired in at least two stages. These stages will be discussed in relation to our third research
34 question below.

35 We have also speculated that F will be the first gender to be mastered in both languages,
36 due to its stronger regularity when compared to the other genders. This has indeed shown to be
37 true for Croatian as the agreement patterns for F reach ceiling level in the older group, whereas
38 this does not happen for the M and N. A possible reason for this finding might be (i) the
39 syncretism of M and N in oblique cases and (ii) the considerably lower frequency of N in the
40 input for which children require more time and exposure to grasp. This prediction does not hold

1 for Italian, even though the literature on the acquisition of articles would suggest otherwise.
2 There are two possible reasons for this: (i) the Italian adjectival agreement does not have
3 irregularities with M as the article system does; and (ii) the adjective is acquired at a later stage
4 than the article, what Bottari et al. (1993) refer to as the *morphological insertion strategy* which
5 includes the insertion of the correct morpheme.

6   Finally, we wanted to see if the differences in the gender systems resulted in differences
7 in the acquisition of individual genders. The Italian children are basically at ceiling for their
8 adjective production at the age we tested. On the other hand, we can see that the Croatian
9 children go through at least two stages which could be observed with the included age groups.
10 The first stage consists of similar error rates with F and M, but significantly higher error rates
11 with N. We can summarize the stage as follows: (F=M)<N. This is likely due to the lower
12 frequency of N: 6% in child directed speech (Kovačević et al., 2009). The second stage consists
13 of F being at ceiling and the error rates with M and N being similar; this also entails that the
14 agreement accuracy for N has significantly improved since it was much more error prone than
15 M in the younger group. This stage can be summarized like this: F<(M=N). The improvement
16 of N agreement is probably due to a longer exposure to N nouns and their patterns. A plausible
17 reason for M not improving as much as F is the syncretism between M and N. As the child's
18 exposure to and usage of the case paradigm increase, the similarity between M and N becomes
19 more evident. If we were to test an older group, it is likely that these difficulties due to
20 syncretism would resolve.

21   What the results suggest is that the degree of transparency of the gender system matters.
22 We cannot look only at a manifestation of gender in isolation, but at the full agreement
23 paradigm to make more accurate predictions of how a gender system might be acquired.
24 However, the full Croatian paradigm was not investigated here as the elicitation proceeded in
25 NOM, yet we see that the effects of syncretism in oblique cases were reflected in the agreement
26 error rates. The Croatian gender system is acquired more gradually when compared to Italian.
27 Nevertheless, the errors made by Croatian children are quite low in both age groups, which
28 means that gender is grasped quite easily.

29   This study, among others, shows the relevance of placing transparency on a continuum.
30 Below the revised figure (figure 2) is repeated. Here we can see how both languages
31 investigated here are in the higher end of the transparency scale and also next to each other,
32 which makes the relative differences in transparency rather low, but these differences still lead
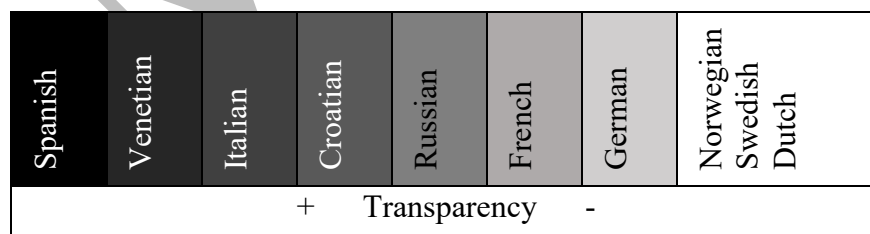33 to dissimilarities when it comes to mastery of the gender system.

Spanish	Venetian	Italian	Croatian	Russian	French	German	Norwegian Swedish Dutch
+			Transparency			-	

34 Figure 2: updated transparency scale

35 We also argue that in order to evaluate the transparency of the gender system of a given
36 language the full paradigms of the agreeing elements have to be taken into consideration.
37 Frequent and clear cues contribute greatly to a fast mastery of the gender system (i.e. the Italian
38 article), while syncretism leading to ambiguous gender information and low frequency in the
39 input (Croatian N) hinder this process.

8. Conclusions

This study has found differences in the time course of acquisition which can crucially be attributed to the different degrees of transparency present in Croatian and Italian. Both languages have transparent gender systems and are acquired easily. However, the gender-marked article in Italian, the syncretism of M and N in Croatian as well as the low frequency of N, make it so that the Italian system is more transparent and thus more easily acquired.

Consequently, Italian children are target-like in their realization of adjectives for both genders already in the younger group . Croatian children have overall more errors in adjectival agreement and we can recognize two stages with distinct error patterns: in the first stage they are equally accurate with F and M, and significantly less accurate with N; in the second stage F is error-free, whereas the accuracy of the M and N is roughly the same. We have attributed this to the low frequency of N, as well as the syncretism of M and N across the case paradigm. This is not a confirmation of the mastery of the full gender system in Italian, but the results nevertheless show how degree of transparency matters and how it reflects on the acquisition of gender values.

This study thus contributes to research on formal cues in gender acquisition by considering transparency as a continuum. It shows how even in two gender systems that are considered transparent, gender mastery does not proceed at an equal pace. The data presented here argues quite clearly in favor of treating transparency as more than a binary feature between transparent and opaque, and it suggests the need for a transparency scale and how to place the languages on a continuum. As we have shown here, in order to detect a language's place on this continuum, full paradigms of the agreeing arguments should be taken into consideration.

This study also a valuable endeavor into the acquisition of Croatian since it is an understudied language when it comes to language acquisition and it reveals how also the full case paradigm can affect how each gender is acquired. The different times of mastery of each gender value could represent a decisive starting point for future research that could include a wider age range of the participants and, more importantly, testing agreement patterns and accuracy on the full case paradigm.

References

Aljović, N. (2002). Long adjectival inflection and specificity in Serbo-Croatian. *Recherches linguistiques de Vincennes* (31)**,** 15.

Audring, J. (2014). Gender as a complex feature. *Language Sciences* 43**,** 5-17.

Barić, E., Lončarić, M., Malić, D., Pavešić, S., Peti, M., Zečević, V., et al. (2005). "Hrvatska gramatika". 4th ed. (Zagreb: Školska knjiga).

Belletti, A., and Guasti, M.T. (2015). *The Acquisition of Italian.* John Benjamins Publishing.

Bottari, P., Cipriani, P., and Chilosi, A.M. (1993). Protosyntactic devices in the acquisition of Italian free morphology. *Language Acquisition* 3(4)**,** 327-369.

Caselli, M.C., Leonard, L.B., Volterra, V., and Campagnoli, M.G. (1993). Toward mastery of Italian morphology: a cross-sectional study. *Journal of Child Language* 20(2)**,** 377-393. doi: 10.1017/S0305000900008333.

Chini, M. (1995). *Genere grammaticale e acquisizione: aspetti della morfologia nominale in italiano L2.* Franco Angeli.

1   Chini, M. (1998). Genuserwerb des Italienischen durch deutsche Lerner. *Wegener, Heide: Eine*
2   *zweite Sprache lernen: empirische Untersuchungen zum Zweitspracherwerb. Tübingen,*
3   *Narr*, 39-60.
4   Corbett, G.G. (1991). *Gender.* Cambridge: Cambridge University Press.
5   Costa, A., Kovačić, D., Franck, J., and Caramazza, A. (2003). On the autonomy of the
6   grammatical gender systems of the two languages of a bilingual. *Bilingualism:*
7   *Language and Cognition* 6(3), 181-200. doi: 10.1017/S1366728903001123.
8   De Marco, A. (2005). *Acquisire secondo natura: lo sviluppo della morfologia in italiano.*
9   FrancoAngeli.
10   Gudmundson, A. (2010). *L'acquisizione del genere grammaticale in italiano L2: Quali fattori*
11   *possono influenzare il grado di accuratezza.* Doctoral degree, University of Stockholm.
12   Karmiloff-Smith, A. (1981). *A functional approach to child language: A study of determiners*
13   *and reference.* Cambridge University Press.
14   [Dataset] Kovačević, M. (2004). *Croatian Kovacevic Corpus.* The CHILDES Project: Tools
15   for Analyzing Talk. 3rd
16   Edition. . Available: https://childes.talkbank.org/access/Slavic/Croatian/Kovacevic.html.
17   Kovačević, M., Palmović, M., and Hržica, G. (2009). The acquisition of case, number and
18   gender in Croatian. *Development of nominal inflection in first language acquisition: A*
19   *cross-linguistic perspective*, 153-177.
20   Kupisch, T., Akpınar, D., and Stöhr, A. (2013). Gender assignment and gender agreement in
21   adult bilinguals and second language learners of French. *Linguistic Approaches to*
22   *Bilingualism* 3(2), 150-179.
23   Kupisch, T., Geiß, M., Mitrofanova, N., and Westergaard, M. (2018). "Gender cues in L1
24   Russian children acquiring German as an early L2. ", in: *EuroSLA.*).
25   Kupisch, T., Müller, N., and Cantone, K.F. (2002). Gender in monolingual and bilingual first
26   language acquisition: Comparing Italian and French. *Lingue e linguaggio* 1(1), 107-150.
27   Levy, Y. (1983). It's frogs all the way down. *Cognition* 15(1-3), 75-93.
28   Love, J., Dropmann, D., and Selker, R. (2018). "Jamovi project". jamovi version 0.9 ed.).
29   Matthews, P. (2007). *The concise Oxfortd dictionary of linguistics.* Oxford: Oxford Univeristy
30   Press.
31   Mihajlović, M. (2014). A computational morphology approach to Croatian noun inflection—
32   the case of gender assignment. *Wiener Linguistische Gazzette* (Special Issue 78A), 120-
33   127.
34   Noccetti, S. (2002). *Pre-and Protomorphology in Language Acquisition: an Italian case study.*
35   Edizioni Plus srl.
36   Pérez-Pereira, M. (1991). The acquisition of gender: what Spanish children tell us. *Journal of*
37   *Child Language* 18(3), 571-590. doi: 10.1017/S0305000900011259.
38   Pizzuto, E., and Caselli, M.C. (1992). The acquisition of Italian morphology: Implications for
39   models of language development. *Journal of child language* 19(3), 491-557.
40   Rodina, Y. (2008). Semantics and morphology: The acquisition of grammatical gender in
41   Russian.
42   Rodina, Y., and Westergaard, M. (2015). Grammatical Gender in Norwegian: Language
43   Acquisition and Language Change. *Journal of Germanic Linguistics* 27(2), 145-187.
44   doi: 10.1017/S1470542714000245.
45   Rodina, Y., and Westergaard, M. (2017). Grammatical gender in bilingual Norwegian–Russian
46   acquisition: The role of input and transparency. *Bilingualism: Language and Cognition*
47   20(1), 197-214. doi: 10.1017/S1366728915000668.
48   Smoczyńska, M. (1985). *The acquisition of Polish.* Lawrence Erlbaum Associates, Inc.
49   Trenkić, D. (2004). Definiteness in Serbian/Croatian/Bosnian and some implications for the
50   general structure of the nominal phrase. *Lingua* 114(11), 1401-1427.

1   Velnić, M. (2015). Of good thieves and old friends: An analysis of Croatian adjectival forms.
2   *Poljarnyj vestnik* 18**,** 18-54.
3   Voghera, M., Iacobini, C., Savy, R., Cutugno, F., De rosa, A., and Alfano, I. (Year). "VoLIP: a
4   searchable Italian spoken corpus, in Complex Visibles Out There.", in: *Olomouc*
5   *Linguistics Colloquium: Language Use and Linguistic Structure.*, eds. L. Veselovská &
6   M. Janebová), 628-640.
7   Vuletić, D. (1991). *Istraživanje govora.* Fakultet za defektologiju Sveučilišta u Zagrebu.
8