CROSS-LINGUISTIC INFLUENCE IN L2 ITALIAN BILINGUAL FAMILIES: A COMPARISON BETWEEN CONVERSATIONS IN A CROATIAN AND TWO ROMANIAN FAMILIES

Marina Castagneto[1]*, Irina Suzana Stan*[2]

1. INTRODUCTION

This work aims to study the cross-linguistic influence phenomena at morphological and pragmatic-discursive level in two linguistic environments (Italian/Croatian and Italian/Romanian respectively) which have never been touched upon by language contact studies.

According to Auer's (1984) framework, bilingualism will be considered «not as a mental disposition of speakers, but a set of complex linguistic activities [...], a feature of interactions, or interactional behaviour, and not of persons» (Auer, 1984: 55). This paper deals with a wide range of cross-linguistic influence phenomena that arise in the sequential development of conversation, as source and context of all the analysed phenomena.

Therefore, the regular Italian loanwords that can conventionally be used as part of the Croatian and Romanian linguistic systems, even in monolinguals' speech, will not be considered. We are interested, instead, in occasional loanwords, and will try to investigate their morphological adaptation in the context of the conversations in the corpus. After analysing the loans, we will move towards contact phenomena such as code-mixing, considering them as a part of the conversation they belong to.

In particular, by looking at the code-mixing phenomena, it can be noticed that in all analysed occurrences the matrix languages (Myers-Scotton, 1997; 2002) are always Romanian and Croatian, the L1 of informants, as the «outside system morphemes» (Myers-Scotton, 2002), especially the verbal inflectional morphemes, come regularly from these languages.

The observed linguistic behaviour of the informants would suggest that the constant selection of a matrix language could point towards a slight dominance of bilinguals in the matrix language of their code-mixing rather than a total balanced bilingualism.

2. THE INFORMANTS

The Croatian family consists of two parents and their two children: a 16 years old son and a daughter who was a university student and had been exposed to Italian since the age of 13. The father had been bilingual for over 20 years, while the other members of the family had been bilinguals for 10 years, since the family moved to Italy.

[1] Università del Piemonte Orientale "A. Avogadro".
[2] Università degli Studi di Milano. In the present paper Irina Suzana Stan wrote sections 2; 3; 5; 6.1; 6.1.2; 6.2.2; 6.3.2 while sections 1; 4; 6; 6.1.1; 6.2; 6.2.1; 6.3; 6.3.1; 7 belong to Marina Castagneto.

The first Romanian family includes a couple of young grandparents (mostly passive Italian bilinguals) and their granddaughter, Alexandra, who was almost 11 years old and for the previous 10 years she had been living in Italy with the grandparents. The second Romanian family consists of a couple of bilingual parents and their two daughters: Lili and Vali of five and respectively around two years old (the latter, due to her young age, did not take part in conversations).

3. THE CORPUS

The corpus consists of two sub-corpora. In the case of Croatian, 125 turns (out of a total of more than a thousand turns from 11 casual conversations) were extracted that presented language transfer phenomena between Italian and Croatian[3]. As for Romanian, 276 turns were selected that were subject to cross-linguistic influence phenomena from three long family conversations (for a total of 6 hours and 48 minutes of recording)[4]. Only 128 turns were finally analysed for code-mixing phenomena, excluding the total of 148 turns of interviews with the two girls. The Croatian sub-corpus contains free conversations among the family members, while the Romanian sub-corpus consists of conversations in which a young Romanian close friend of the families chatted about various themes with Alexandra and Lili. Although there was also interaction with Alexandra's grandmother and with Lili's mother, the two girls were the privileged interlocutors.

4. RATIONALE FOR THE COMPARISON

It seemed appropriate to compare the two sub-corpora on grounds of the numerous sociolinguistic and contextual conditions that the three couples of parents / grandparents shared, such as:

- they share the same native language, but the dominant language is Italian, and they speak their own language to the children: according to Romaine (1995: 183-187), these factors cooperate in creating a peculiar type of early childhood bilingualism (defined as *Non dominant Home Language without Community Support*)[5].
- their L1, besides not being dominant, is not of prestige;
- they are bilingual, but speak in their L1 to each other;
- their attitude towards both languages and bilingualism in general is positive, which leads to an additive and assimilative type of bilingualism;
- they do not exhibit any distribution principles for the selection nor the use of the two languages nor do they seek to direct the children's' choice in the selection of one language rather than another;
- consequently, they appreciate the linguistic creativity of the children, even the mixed one;
- their language production also presents cross-linguistic influence elements; all family members follow the affirmation of Auer (1995: 115-116) that «bilingualism provides

[3] We are grateful to Marina Pejakovic for the Croatian data, taken from Pejakovic (2014).
[4] See Stan, 2011.
[5] In Romaine's classification the "Non dominant Home Language without Community Support" early child bilingualism is classified as the third category of early child bilingualism out of six.

specific resources not available to monolingual speakers for the constitution of socially meaningful verbal activities» [6].

5. TYPE OF BILINGUALISM

All children, but the Croatian daughter, exhibit an early bilingualism while the Croatian son is on the border between early and late bilingualism since he has been exposed to Italian from the age of six. All informants actively use two independent linguistic systems which they learned in separate contexts, therefore their bilingualism is coordinated (see Ervin, Osgood, 1954)[7]. Moreover, since the L2 was learned after the L1, and given that the families maintain a relationship with the L1, the type of bilingualism they exhibit is sequential and additive. The acquisition of the second language occurred in L2 environment, so learning was accompanied by acculturation (see Bettoni, 2006: 45); nonetheless, the type of bilingualism under analysis is popular, rather than elitist, integrative and assimilative. None of the three families is part of L1-speaking communities[8] therefore their bilingualism is (mostly) isolated (see Francescato, 1981). Finally, the informants are (almost) balanced bilinguals, at least as far as the youngest family members are concerned and judging from what the corpus allowed us to verify (i.e. in family interactions, diamesically oral and diaphasically linked to a colloquial and informal register, in domains related to the family, food, education)[9]. In the corpus, the youngsters freely alternate the two languages selected in the same amount of speech, while the cross-linguistic influence is bi-directional for all types of phenomena; an exception is represented by loanwords, which in both sub-corpora are always from Italian to Romanian or Croatian, but never vice versa, probably because Italian is the most prestigious. The amount of cross-linguistic influence phenomena in production and perception, the type of influence and the ease of processing in the two languages show a strong bilingual activation.

A difference between the two families is related to the fact that in the Croatian family there is also adequate literacy in the weak language while in the two Romanian families the L1 is used only orally. The two different methods of data collection (an informal interview with the Romanian girls in contrast to free conversations in Croatian) did not create significant barriers as far as comparability of data is concerned. On the contrary, the two sub-corpora were very similar in terms of type of influence and amount of phenomena.

[6] See Moretti, Antonini (1999: 103-104) for an extensive list of the conditions that influence the type and degree of bilingualism.

[7] We are aware that the two polarized notions of *compound* and *coordinated* bilingualism are not very well defined, and even incompatible definitions have been formulated for both. Moreover, some scholars sustain that there is no good experimental evidence to support the notion of these two kinds of bilingualism (see Diller, 2015).

[8] Lili's family entertains some additional contact with Romanian neighbours; Alexandra's grandparents entertain only telephone contact in Romanian.

[9] We are aware that completely balanced bilingualism is considered very rare, because a perfectly balanced bilingual should have the same competence in all language skills (see Romaine, 1995: 12-19), i.e. production and comprehension in written and spoken languages and in different domains. Nevertheless, the corpus did not allow us to widely probe the competence in the two languages of the interviewed speakers. Moreover, some linguistic cues point out that the informants are slightly dominant bilinguals in their L1 (see section 7).

6. CROSS-LINGUISTIC INFLUENCE PHENOMENA

The corpus contains a significant number of occasional loanwords (integrated and non-integrated at both phonological and morphological levels) and code-mixing phenomena.

The analysis will be confined to the more structural types of cross-linguistic influence, starting at a morphological level and then moving towards larger linguistic segments. Occurrences of morphological integration of loans and code-mixing phenomena will therefore be comparatively analysed.

In both languages, morphological integration concerns both the nominal and the verbal system.

6.1. *Morphological integration in the nominal system of loans*

Regarding the nominal system, it should first be remembered that both Romanian and Croatian nouns are inflected by gender, number and case, while Italian has lost the case inflection[10]. As already mentioned, in the sub-corpora there are only borrowings from Italian as a *model language* and Croatian and Romanian as *replica languages*[11], so only how Italian nouns are adapted to these two languages can be verified.

In Romanian, the morphological integration can sometimes not take place by number nor by case; that is, the form of the loanword can be in nominative singular, regardless its syntactic role in the sentence[12]:

(1)[13] Dial. I:

A20: […]

De	**verdur-a**	doar **carot-e**,	varz-ă	şi	salat-ă
of	vegetable-F.SG	only carrot-F.PL	cabbage-F.ACC.SG	and	salad-F.ACC.SG

'(I) only (eat) carrots, cabbage and salad as vegetables'.

In the conversational turn A20, in answer to the question 'What do you like to eat?', the loanword *verdura* is in the feminine and singular form, following the Italian model; moreover, it was not integrated by case, notwithstanding its syntactic role of accusative.

[10] Romanian has five cases (nominative, genitive, dative, accusative and vocative), two numbers (singular and plural) and two genders (masculine and feminine). The existence of the neuter, historically attested but residual, is controversial today, because the Romanian neuter does not have its own forms, except for a specific ending of the plural neuter, *-uri*, e.g. in *becuri* "light bulbs" and in *hoteluri* "hotels" (this ending is quite comparable to the pseudo-neuter suffix *-ora* in several Southern Italo-romance dialects; see Loporcaro, 2009: 130). Moreover, the neuter case in Romanian follows the declination of masculine forms in singular and feminine forms in plural.
Croatian has seven cases, namely the nominative, genitive, dative, accusative and vocative (like Romanian), but it adds two other purely Indo-European cases: instrumental and locative. It has two numbers (singular and plural) and three genders: masculine, feminine, and neuter.

[11] For the concepts of *model language* vs. *replica language* see Weinreich, 1963. For a more recent use of the *model language* / *replica language* dichotomy see Matras, 2009.

[12] For a complete morphological integration (thus also by case) see ex. (8).

[13] Following Myers-Scotton (2002) the object-language elements pertaining to the embedded language (in this study: Italian) are in bold case, while the linguistic elements ascribable to the matrix language (Croatian / Romanian) are in roman case.
Turn marking in the Romanian sub-corpus (Dial. I and II): A=Alexandra, M=grandmother, I=interviewer. In Dial. III: L=Lili, M=mother, I =interviewer.

In Romanian, the feminine singular nouns that are preceded by a preposition have the ending in *-ă* as hallmark of the accusative case (as in *varz̧-ă* and *salat-ă*).

In Croatian, on the other hand, the borrowed nouns take the expected case and number while the gender is sometimes integrated:

(2)[14] Dial. VIII

S98: Is-li smo u **piscin-**u skupa? […]
go-PTCP.PL AUX.1PL in pool-M.LOC.SG together
'Did we go to the pool[15] together?'.

6.1.1. *Morphological integration of the article for Italian loanwords in Croatian*

Since Croatian has no article, it is not surprising that our informants' production in Italian can also lack articles due to cross-linguistic influence.

(3) Dial I:

M15: Sta ti je reka-o **dentist-a?**
What DAT.2SG AUX.3SG say-PTCP.SG.M dentist-SG
'What did the dentist tell you?'.

Articles may be missing before single word switching in Italian when the discourse is in Croatian, as in ex. 4[16]:

(4) Dial. X:

F109: **Sorell-in-a**, rec-i cu ti jedn-u stvar
sister-DIM-F.SG tell-INF FUT.1SG DAT.2SG one-F.ACC.SG thing[F.ACC.SG]
'Little sis, I'll tell you something'
ako ces zna-ti drza-ti **segret-o.**
if FUT.2SG know-INF keep-INF secret-M.SG
'if you know how to keep the secret'.

In the same sub-corpus, the article is present only twice[17]:

(5) Dial. I:

M09: **Alle otto**. I onda ce doc **lo zi-o** […][18]
At.ART.DEF.F.PL eight and then FUT.3SG come.INF ART.DET.M.SG uncle-M.SG
'At 8 o'clock. And then the uncle will come […]'.

In this case, perhaps it is more convenient to consider *lo z̧io,* 'the uncle', as a case of code-mixing, given the overall meaning of 'uncle', which in Italian refers to both maternal

[14] Turn marking in the Croatian sub-corpus: M=mother, P=father, S=sister, F=brother, V=8-year-old cousin.
[15] See Cr. *bazen*, 'pool' (M, SG).
[16] In ex. (4) it is better to consider *segreto* as single word switching (instead of a loan) because it is not conventionally used in Croatian, and it is difficult to imagine that it could occur as such in monolinguals' speech (see Haspelmath, 2009: 40).
[17] The other instance is in ex. (15).
[18] See ex. (13).

and paternal uncle, while in Croatian the semantic field of kinship, which is more complex, indicates a different name for the maternal uncle (*ujak*) and the paternal one (*stric*).

6.1.2. *Morphological integration of the article for Italian loanwords in Romanian*

Romanian, by virtue of its linguistic position within the Balkan zone, presents a postponed definite article (enclitic) which is marked on the noun; therefore, the integration of the Italian loans also passes through the transfer of this morphological mark.

Thus, the Romanian sub-corpus includes forms like *cappellul*, in which the article *-ul* was agglutinated to the Italian word *cappello* 'hat':

(6) Dial II:

A27: **Cappell**=ul ăla albastr-u şi cu rochi-a […]
hat=ART.DET.M.SG DEM.DIST.M.SG blue-M.SG and with dress.F.NOM.SG
'That blue hat and the dress […]'.

The following is a somewhat more complex instance:

(7) Dial. II:

A52: […]
Are chiar **cord-e**=le **vocal-i da Gianna Nannini**.
have.PRS.3SG even chord-F.PL=ART.DEF.F.PL vocal-F.PL like Gianna Nannini
'She has the vocal chords like Gianna Nannini'.

Here the Italian loanword *corde vocali* (Rom. *coarde/corzi*[19] *vocale*) is declined by plural number, and with its marks of gender and number influences the selection of the article in Romanian. Despite the fact that it is a multi-word expression, it has been disassembled and the Romanian postponed definite article has been placed only on the head of the NP.

The Romanian undefined article is proclitic, as in Italian. In the utterance:

(8) Dial. I

A4: […]
a veni-t **cugin**=a mea şi vroi-am
AUX.3SG come-PTCP.PST.3SG cousin=ART.DEF.F.SG POSS.F.1SG and want- IPFV.1SG
'My cousin came and I wanted'
să=i **racont** o **stori**-e cu o
SBJV=DAT.3SG tell ART.INDF.F.SG story-F.ACC.SG with ART.INDF.F.SG
'to tell her a story with a'
pecorell-ă şi cu o fat-ă
sheep-ACC.F.SG and with ART.INDF.F.SG girl-F.ACC.SG
'sheep and a girl'

it can be seen how the two Italian loans, two singular feminine nouns, are preceded by the indefinite feminine article *o* 'a' (the indefinite masculine article is *un*, just like in Italian). As for *storie,* the Italian mark for gender and number *-a* has been replaced with *-e*, which is the ending of the corresponding feminine noun in Romanian, *povest-e*. The final *-ă* of

[19] Both forms are currently accepted.

pecorellă (pronounced [ə]) is again an integration in the Romanian nominal paradigm, since *-ă* is one of the accusative marks of the feminine noun declination.

6.2. *Morphological integration in the verbal system of loans*

Regarding the verbal system, there is a constant presence of morphological integration, so the borrowed Italian verbs are inflected according to the paradigm of the two replica languages. Here the most significant differences depend on the diversity in the verbal system of Croatian and Romanian.

6.2.1. *Integration of the verb in Croatian*

As expected, in Croatian the inflected forms of the Italian verb are replaced by the corresponding inflected forms. The occurrences of integration of the verbs in the corpus must be considered as cases of code-mixing (see section 6.3).

According to Myers-Scotton (1997; 2002), in code-mixing[20] the morpho-syntactical framework of the sentence is governed and determined by the *matrix language*[21]. More specifically, as for the verbal morphology, in Myers-Scotton's works all inflectional affixes are considered *system morphemes* supplied by the matrix language. In the first version of the *Matrix Language-Frame Model* (1993) tense and aspect are considered as syntactic categories involving the feature [+Quantification], a property characterizing only some system morphemes[22], because they involve quantification across events (tense morphemes, for instance, select one time-frame rather than another). Subject-verb agreement is instead [-Quantification], but it is still expressed by a system morpheme, because it cannot potentially assign θ–role or receive θ–role. The subject-verb agreement too, therefore, is selected in the morpho-syntactic frame of the matrix language.

All types of inflectional morphemes of the verbs are still considered system morphemes in the subsequent *4M-model* (Myers-Scotton, 2002), where they are categorized as *outside late system morphemes*: they are defined as «outside» because they refer to grammatical information outside the Maximal Projection of Head, and «late» because they «are only accessed when the lemmas underlying content morphemes send directions to the Formulator about how larger constituents are to be assembled» (*ib.*: 301). Late system morphemes, then, are structurally assigned, and not directly linked to the speaker's intentions, neither are they salient at the level of mental lexicon (*ib.*: 73-76). Also in the 4M- model the System Morpheme Principle applies and, in mixed constituents, all the outsider late system morphemes must come from the matrix language[23].

[20] It is important to underline that Myers-Scotton draws no distinction between code-switching and code-mixing, so intersentential and intrasentential code-switching are both labelled as code-switching.

[21] For the sake of precision, Myers-Scotton defines the matrix language by the role it plays in the Matrix Language - Embedded language hierarchy, realized in the Morpheme Order Principle and in the System Morpheme Principle (Myers-Scotton, 2002: 59). This last principle is stated as it follows: «In Matrix Language + Embedded Language constituents, all system morphemes which have grammatical relations external to their head constituent (i.e. which participate in the sentence's thematic role grid) will come from the matrix language» (Myers-Scotton, 1997: 83).

[22] «Any lexical item or affix which is a member of a syntactic category specified as plus for [Quantification] is a system morpheme» (Myers-Scotton, 1997: 100). Content morphemes are always [-Quantification], while system morphemes can be [± Quantification].

[23] See Myers-Scotton's *4-M model* (2002: 302).

The corpus seems to respect the System Morpheme Principle, especially if the verbal inflectional morphology is looked at, e.g.:

(9) Dial. I

M29: […]
Ako **sistem-**am **tutto non lo so**.
if set_up-PRS.1SG everything NEG M.3SG know.PRS.1SG
'I don't know whether I will set everything up or not'.

S30: Ako ne uspije-s **sistem-are tutto**, vrati-t ces se […]
if NEG succeed-PRS.2SG set_up-INF everything return-INF FUT.2SG REFL
'If you do not manage to set everything up, you will go [again to Zagreb]'.

Here the Italian verb *sistemare* has taken the morpheme *-am* of ISG.PRS.IND. in place of the analogous Italian morpheme *-o*.

Unlike in Romanian (see 6.2.2), the infinitive forms can be integrated as such, taking the infinitive morpheme.

It is to be noted, however, that regardless of the conjugation of the Italian verb, the borrowed verb is always integrated in the first conjugation of Croatian, taking the morpheme *-ati*. In Croatian as well as in Italian, the integration of loans is always in the first conjugation (compare, for example, loans in Italian from English: *taggare*, *chattare* etc) for frequency reasons.

In Croatian a morphological integration of complex forms is also found, e.g. reflexive forms in compound tenses:

(10) Dial. XI

S124: A misli-m da im-a
but think-PRS.1SG that have-PRS.3SG
'But I think there is'
ni-sam se **accorg**-il-a da ne-ma
NEG-AUX.1SG REFL realize-PTCP.PST-F that NEG-have.PRS.3SG
'I did not realize that there was not[24]'.

6.2.2. *Integration of the verb in Romanian*

The Romanian language regularly replaces the infinitive forms with the subjunctive in the dependent clauses, as in the following example:

(11) Dial. I

A89: […]
Ne-a zis să **descrive**-m **un' amic-a**
DAT.1PL-AUX.3SG tell.PTCP.PST.3SG SBJV describe-PRS.1PL ART.IND.F.SG friend- F.SG
'We were told to describe a friend […]'.

Here it can be seen that the Italian verb *descrivere* (Rom. *a descrie*) has been integrated into the subjunctive mood, which in Romanian is formed through the rule of ante position of the morpheme *să* in the present tense (*noi descriem*) due to the loss of infinitive, a

[24] Here *nisam se accorgila* can be literally translated: "I think there is, not was I aware" (the verb is in part. pass. fem. form).

linguistic phenomenon that characterizes the Balkan zone. Following Berruto (2005: 86-89), *descrivem* could be considered as a hybrid lexical form constructed from the lexical morpheme *descriv-*, taken from the Italian language system, and the Romanian inflectional morpheme of the subjunctive mood. Ex. (11), in Muysken's (2000) theoretical frame, could be considered as an instance of congruent lexicalization. It. *decriv(-ere)* and Rom. *a desri(-e)* can be assumed as *homophonous diamorphs*; in contexts of strong bilingualism in which two languages are similar in grammar and lexicon and they are very close to the activation threshold, it is even possible that the distinction between the two codes may be neutralized exactly where a pair of homophonous diamorphs appears (see Clyne, 1967). This is because the Italian *descriv(em)* has triggered the code-mixing and therefore it is not surprising that the NP *un'amica* is its object. Moreover, in the Romanian sub-corpus, cases of morphological integration of verbal loans are found in the present tense, imperfect tense and reflexive forms.

6.3. *Code-mixing*

Code-mixing has been defined by Poplack (1980) as «intrasentential code-switching». In code-mixing the shift occurs in the middle of a sentence, with no interruption, hesitation or pauses indicating a shift.

Unlike code-switching, for code-mixing no pragmatic and discursive functions have been recognized yet[25], so this form of linguistic mixing seems to be more linked to the inter-penetrability of the grammars of the two involved languages rather than to context, social network, socio-symbolic values and the selection of pragmatic strategies or structuring of speech by the participants.

For Berruto code-mixing (in his term: «enunciazione mistilingue») is defined as «passaggio all'interno di una frase (o di una struttura riconducibile a una frase) da una lingua o varietà ad un'altra lingua o varietà senza che vi sia concomitanza con mutamenti nel flusso della situazione, e senza quindi che sia attribuibile al segmento frammisto una sua microfunzione in cui sia messo in evidenza o in gioco il significato sociale o il valore simbolico della varietà interessata: non vi è pertanto intenzionalità a scopi socio-comunicativi, e il segmento frammisto non coincide con un atto linguistico, bensì è definibile *solo* in termini di categorie morfosintattiche e non pragmatico-discorsive» (Berruto, 1990: 112). As the mixed segments are definable just in terms of morpho-syntactic categories, it is maybe not surprising if code-mixing is mainly investigated by theoretical linguistics, while code-switching falls into the research fields of pragmatics and conversational analysis.

6.3.1. *Analysis of code-mixing (Croatian)*

All Croatian informants have a strong bilingual competence, so they produce a low number of code-mixing phenomena (16 occurrences) in the 125 turns analysed.

[25] The definition of code-switching and code-mixing is controversial (see Alfonzetti, 1992: 19-21), and many scholars do not distinguish between them. Among the authors who operate this distinction, some adopt the superordinate term *language mixing*, which includes both code-switching and borrowings (Grosjean, 1990: 108), while others consider the code-mixing a subordinate category of code-switching (McClure, 1981: 86). Following the normally accepted formal criterion, code-switching is intersentential, while code-mixing is intrasentential (e.g. Muysken, 1984). In this work this definition will also be accepted.

The mother, who represents the only case of Croatian dominant bilingual, does the most code-mixing (6 out of 16 instances in the corpus in the 32 turns attributable to her, that is, she produces code-mixing in 18.75% of the conversational turns ascribed to her)[26].

Code-mixing can be triggered by priming, i.e. from an immediately preceding integrated loanword (mostly a verb, in 4 out of 6 cases), as priming obviously lowers the threshold of activation of the other language. Therefore, it triggers the code mixing depending on the selection of an element of the mental lexicon belonging to the primed language. In the following two turns, the triggering element is the integrated verb *sistemam*:

(12) Dial. I

M29: […]
Ne ja cu [long pause] ima-m **appuntament-o** u srijed-u.
NEG 1SG FUT.1SG have-PRS.1SG appointment-M.SG in Wednesday-ACC
'No I will [long pause] I have an appointment on Wednesday'.
Ako **sistem-**am **tutto non lo so**.
if set_up-PRS.1SG everything NEG M.3SG know.PRS.1SG
'I don't know whether I will set everything up or not'.

As always, code-mixing can refer to the insertion of *embedded language islands*[27] consisting of different types of syntactic structures. Following the dictates of Myers-Scotton's works (1997; 2002), the embedded language islands in the corpus contain well-formed constituents according to the embedded language grammar, and show internal structure dependency relations (Myers-Scotton, 1997: 78). Moreover, the placement of the embedded language islands within the Projection of Complementizer depends on the matrix language procedures (Myers-Scotton, 2002: 58). Embedded language islands in the corpus consist of:

– simple prepositional phrases:

(13) Dial. I

M09: […]
Alle otto. I onda ce doc **lo zi-o.**
At.ART.DEF.F.PL eight and then FUT.3SG come.INF ART.DET.M.SG uncle-M.SG
'At 8 o'clock. And then your uncle will come'
da me odvez-e na **a Lampugnano**
that ACC.1SG take-PRS.3SG on to Lampugnano
'to take me to Lampugnano'[28].

– noun phrases functioning as direct objects of the verb:

(14) Dial. X

F111: Misli-o sam napravi-t
think-PTCP.M.SG AUX.1SG make-INF
'I was thinking of having'.

[26] The highly competent Italian speaking daughter operated code-mixing in 9.09% of the turns attributed to her.

[27] «Embedded language islands are full constituents consisting only of E.L. morphemes occurring in a bilingual CP [Projection of Complementizer] that is otherwise framed by the matrix language» (Myers-Scotton, 2002: 139).

[28] In ex. (13) the preposition governs a toponym.

un-a fest-a di sorpresa per il compleann-o di Mario
ART.IND-F.SG party-F.SG of surprise for ART.DEF.M.SG birthday-M.SG of Mario
'a surprise party for Mario's birthday'.

In the following example:

(15) Dial II

F37: Ma sta su **le** zamjenic-e Mari?
but what be.PRS.3PL ART.DEF.F.PL pronoun-F.NOM.PL Mari
'But what are pronouns, Mari?'

it can be seen how the article *le*, "the" (feminine, plural), is the object of code-mixing and therefore carried into Croatian which lacks articles, regardless of the fact that there is a wrong selection of the article due to the agreement with the vocal ending of the Croatian lexeme. Ex. (15) is particularly interesting because the empty position of the determiner in Croatian has been filled by an Italian element: it is also possible that the activation of a functional element of Italian could have been enhanced exactly because there was an empty position in Croatian. Yet, examples (3) and (4) show the opposite syntactic behaviour and exclude the article before Italian loanwords. Instances as in ex. (15) represent a serious Achilles' heel of the Matrix Language-Frame Model, since there is a head noun phrase in which the article does not belong to the matrix language (while the other constituents of the sentence do); the main problem is that the article is a typical system morpheme assigned by the matrix language, as it is [+Quantifier], [- θ-role assigner] [- θ-role receiver][29]. Berruto (2005: 89 ff.) discusses similar instances in code-mixing between Italian language and its dialects, also signalling further occurrences of this kind in Franceschini (1998: 58) and Alfonzetti (1992: 175).

An even stronger violation is present in:

(16) Dial. III

F45: **Non ho fame**
NEG AUX:PRS.1SG hunger
'I am not hungry'.

S46: Kako ne-mas **fame, vien-i**!
how NEG-have.PRS.2SG hunger come-2SG
'What do you mean you are not hungry, come on!'.

Here, the code-mixing occurs within a "support verb" construction (see Ježek, 2005: 181-183), in other words inside a multi-word lexical unit whose components are not entirely free from a syntactic point of view, while the verb, whose meaning is widely generic, only provides the *Aktionsart*. Moreover, ex. (16) violates a constraint mentioned by Gumperz: «when a phrase is seen as an idiomatic whole, it cannot ordinarily be broken as a switch» (1982: 89).

Within the same dialogue there is the following:

[29] Yet, it should be said that in Myers-Scotton's *4-M model* determinants are considered «early system morphemes», salient in the level of the mental lexicon, and, «because of their link to the content morphemes [they] may move with them in various contact phenomena» (Myers-Scotton, 2002: 300, Hypothesis 4). However, in ex. (15), the violation of the model still persists, as the subsequent content morphemes is in Croatian (*zamjenice).*

(17) Dial. III

S44: Daj ajde dodj-i vecera=t,
come_on.IMP.SG come_on.IMP.PL come-IMP.SG dinner=for
'Come on, come to eat'
vecera **è** **pront-a**
dinner[F.NOM.SG] be.PRS.3SG ready-F.SG
'dinner is ready'

This is a rather interesting case because, if strictly syntactical considerations are followed, it can be noticed that the INFL morphemes are in Croatian in the first part of the turn (on the inflected verb *vecerat*) and in Italian in the second part, being conveyed by copula *è*, a system morpheme that should be assigned by the matrix language, as it is [-θ-role assigner]. Probably the matrix language is Croatian, because *vecera è pronta* 'dinner is ready' seems to be a calque on the Italian model *la cena è pronta* 'dinner is ready'. The sentence *vecera è pronta* is considered an embedded language island, as it is formulaic in its structure (as embedded language island often are). The lemma *vecera* 'dinner' could have been selected (instead of *cena* 'dinner') because of the priming of *vecerat*, with the same word root.

6.3.2. *Code-mixing analysis (Romanian)*

In the Romanian sub-corpus there are 13 code-mixing phenomena (11 by Alexandra, 1 by her grandmother, 1 by Lili).

Half of them (6 occurrences) belong to the two longer conversational turns of Alexandra talking to her grandmother (dial. II). Nevertheless, on the whole, the code-mixing phenomena within the Romanian sub-corpus are not particularly relevant to the current analysis and will be further exploited elsewhere.

7. FURTHER RESEARCH

An interesting observation, which deserves deeper analysis, is that in all instances that contain code-mixing in the Croatian sub-corpus the inflectional categories are always conveyed by verbal morphemes in Croatian, and Croatian is therefore the matrix language. Even in the Romanian sub-corpus, the inflection categories in the code-mixing are always assigned to Romanian, at least as far as the main clauses are concerned: Romanian is the matrix language too.

After having analysed the entire corpus, it can be concluded that the bilingualism of the youngsters would seem objectively balanced (at least in the domains that were observable and in the colloquial linguistic register of the families). However, even if they seem to show an equal proficiency in both their first and second language, the constant presence of system morphemes (especially outside late system morphemes) belonging to their L1 could reveal that, as a matter of fact, they are not totally balanced bilinguals, but dominant bilinguals (Myers-Scotton, 2008) in their L1.

This hypothesis, however, needs to be filtered through a wider number of occurrences (including other corpora, perhaps in other languages) in order to be verified.

REFERENCES

Alfonzetti G. (1992), *Il discorso bilingue. Italiano e dialetto a Catania*, FrancoAngeli, Milano.

Auer P. (1984), *Bilingual Conversation*, Benjamins, Amsterdam.

Auer P. (1995), "The Pragmatics of Code-switching: A Sequential Approach", in Milroy L., Muysken P. (eds.), *One Speaker, Two Languages: Cross-Disciplinary Perspectives on Code-switching*, Cambridge University Press, Cambridge, pp. 115-135.

Berruto G. (1990), "Italiano regionale, commutazione di codice e enunciati mistilingui", in Cortelazzo M. A., Mioni A. M. (eds.), *L'italiano regionale*, Bulzoni, Roma, pp. 105-130.

Berruto G. (2005), "Dialect/standard convergence, mixing and models of language contact: the case of Italy", in Auer P., Hinskens F., Kerswill P. (eds.), *Dialect Change. Convergence and Divergence in European Languages*, Cambridge University Press, Cambridge, pp. 81-95.

Bettoni C. (2006), *Usare un'altra lingua*, Laterza, Bari-Roma.

Clyne M. (1967), *Transference and triggering: observations on the language assimilation of postwar German-speaking migrants in Australia*, Martinus Nijhofff, The Hague.

Diller K. C. (1970), "'Compound' and 'Coordinate' Bilingualism: A Conceptual Artifact", in *Word*, 26, 2, pp. 254-261: https://www.researchgate.net/publication/287847207_Compound_and_Coordinate_Bilingualism_A_Conceptual_Artifact.

Ervin S. M., Osgood C. (1954), "Second language learning and bilingualism", in *Journal of Abnormal and Social Psychology*, 58, pp.139-145.

Francescato G. (1981), *Il bilingue isolato*, Minerva Italica, Bergamo.

Franceschini R. (1998), "Code-switching and the notion of code in linguistics: proposals for a dual focus model", in Auer P. (ed.), *Code-Switching in Conversation. Language, interaction and identity*, Routledge, London-New York, pp. 51-72.

Grosjean F. (1990), "The psycholinguistics of language contact and code-switching: concepts, methodology and data", in *Papers for the workshop on concepts, methodology and data*, European Science Foundation, Strasbourg, pp. 105-119.

Gumperz J. J. (1982), *Discourse Strategies*, Cambridge University Press, Cambridge.

Haspelmath M. (2009), "Lexical borrowing: Concepts and issues", in Haspelmath M., Tadmor U. (eds.), *Loanwords in the World's Languages. A Comparative Handbook*, De Gruyter Mouton, Berlin, pp. 35-54.

Ježek E. (2005), *Lessico*, il Mulino, Bologna.

Loporcaro M. (2009), *Profilo linguistico dei dialetti italiani*, Laterza, Bari-Roma.

Matras Y. (2009), *Language Contact*, Cambridge University Press, Cambridge.

McClure E. (1981), "Formal and functional aspects of the codeswitched discourse of bilingual children", in Durán R. P. (ed.), *Latino Language and Communicative Behavior*, Ablex, Norwood, N.J.

Moretti B., Antonini F. (1999), *Famiglie bilingui. Modelli e dinamiche di mantenimento e perdita di lingua in famiglia*, Osservatorio della Svizzera Italiana.

Muysken P. (1984), "Linguistic dimensions of language contact" in *Revue Québécoise de Linguistique*, 14, 1, pp. 49-76.

Muysken P. (2000), *Bilingual speech. A typology of Code-Mixing*, Cambridge University Press, Cambridge.

Myers-Scotton C. (1997), *Duelling Languages. Grammatical Structure in Codeswitching*, Clarendon, Oxford.

Myers-Scotton C. (2002), *Contact linguistics. Bilingual Encounters and Grammatical Outcomes*, Oxford University Press Oxford, Oxford.

Myers-Scotton C. (2008), *Multiple Voices: An Introduction to Bilingualism*, Blackwell, Australia.
Pejakovic M. (2014), *Fenomeni di interferenza linguistica tra italiano e croato nell'interazione familiare*, Università degli Studi del Piemonte Orientale "A. Avogadro", Vercelli, Undergraduate dissertation (unpublished).
Poplack Sh. (1980), "Sometimes I'll start a sentence in Spanish y termino en español": toward a typology of code-switching, in *Linguistics,* 18, 7/8, pp. 581-618.
Romaine S. (1995), *Bilingualism*, Second Edition, Blackwell, Oxford.
Stan I. S. (2011), *Bilinguismo precoce italiano-rumeno in contesto italiano: un case study*, Università degli Studi del Piemonte Orientale "A. Avogadro", Vercelli, Undergraduate dissertation (unpublished).
Weinreich U. (1963), *Languages in Contact. Findings and Problems*, Mouton, The Hague.