Observations on Croatian as a Heritage Language Across Four Continents

Jim Hlavač
Monash University

Diana Stolac
University of Rijeka

Abstract

Of the approximately 6.67 million Croatian speakers worldwide, about 1.6 million are located in the Croatian diaspora, from Latin America to Western Europe, North America, and Australia. A multi-site project on Croatian as it is spoken in the diaspora was initiated in 2015 which encompassed ten corpora of linguistic data collected in nine different countries across four continents. Backgrounded by an overview of previous research into the speech of Croatian emigrants, this paper defines and explains the notion of 'Croatian as a heritage language'. Our focus then turns to the speakers themselves, and we draw on a combined sample of corpora that consist of recorded speech samples gained from 300 Croatian-origin emigrants. These included first-, second- and third-generation speakers. We provide an overview of features in Croatian as a heritage language regarding the following four areas: pragmatics; lexicon; calques and loan translation; and code-switching. In our presentation of examples from all ten samples we compare data between countries and vintages of migration, and between speakers of older and younger generations. In our examination of examples, we also compare these with forms used in varieties of Croatian that are spoken in the homeland, both non-standard and standard. Our observations provide a contemporary and cross-national description of Croatian as a heritage language. At the same time, we position heritage Croatian within the field of contact linguistics research.

Key words: heritage language, Croatian as a heritage language, Croatian diaspora

Introduction

Over the last 150 years, the emigration of people from Croatian-populated areas in south-east Europe has been substantial and extensive. Today, the number of Croatian-origin people residing in Croatia, Bosnia-Herzegovina, the province of Vojvodina in Serbia, and in the Bay of Cattaro area of Montenegro is approximately 5 million. In comparison,

the number of Croatian-origin people residing in other countries as emigrants or as the descendants of emigrants is approximately 2,900,000 (Central State Office for Croats Abroad, 2019). According to the same source, the countries with the largest numbers of Croatian-origin emigrants are: USA 1,200,000; Germany 350,000, Argentina, Australia and Canada with 250,000 each; Chile 200,000; Austria 90,000; Switzerland 80,000; Italy 60,000; and France and New Zealand with 40,000 each. As stated, these statistics relate not only to emigrants themselves, but to their children, grand-children and subsequent generations.

Not all emigrants or descendants of Croatian emigrants have proficiency in Croatian. Ethnologue (2018) lists the total number of speakers as 6,670,820. Five million of these speakers are located in the 'homeland' countries outlined above. There are clusters of Croatian speakers who are indigenous minorities living in neighbouring countries such Austria, western Hungary and southern Slovakia (Burgenland) (50,000), in southern Hungary (15,000) and in Italy (Molise) (2,000). But the remaining 1.6 million speakers are emigrants, located in the diaspora. As a proportion of the total number of Croatian speakers, those in the diaspora account for about 25% of speakers world-wide.

This paper examines aspects of the Croatian speech of emigrants and their descendants from a sample based on ten corpora from nine different countries across four continents. Of the nine countries represented, eight of them belong to the top eleven countries listed above: USA, Germany, Argentina, Australia, Canada, Austria, Italy and New Zealand. Norway is the ninth country with a comparatively small Croatian-speaking population of 2,000. The aim of this paper is to provide a brief overview of some of the main contact linguistic phenomena that are recorded in the speech of Croatian speakers in the diaspora. This paper is structured in the following way: section 2 backgrounds the situation of speakers in the diaspora by defining the term 'heritage language' and outlines levels of language maintenance in the diaspora; section 3 provides an overview of the main types of language contact phenomena with examples coming from any of the ten corpora; and section 4 generalises some of these phenomena and points to trends and how these relate to our current knowledge of contact linguistics.

Croatian as a heritage language and Croatian language maintenance in the diaspora

The term 'heritage language' became popular in North America in the 1970s following the Ethnic Revival that occurred on that continent from the 1960s onwards. The term refers to a language that is learnt by a child at home in the family, and that is the language of the ethnolinguistic group that the child's parents belong to (or the group of least one of the parents). Thus, the heritage language is usually the child's first language in chronological terms. But with the passage of time and with entry into the formal school system of the country of residence, the child typically then acquires the societally dominant language, that is usually also the language of instruction at school and lingua franca used in interactions with others outside the family and ethnolinguistic group (Rothman, 2009). The societally dominant language often becomes the more heavily used language of the child, and it often becomes the dominant language of the child itself. The child may continue to use and increase their proficiency in the first-acquired language – the heritage language – and this occurs for many children, but not amongst all. Thus, there are a variety of descriptions that can relate to the language use and competence of children, termed here 'second-generation speakers', who are the children of migrants who themselves are 'first-generation speakers'. The proficiency levels amongst second-generation speakers can be very broad and can range from those whose level is indistinguishable from that of homeland speakers (perhaps with the only obvious difference being a narrower repertoire of registers of which they have active command), to those heritage language speakers whose active or passive command is severely limited (Polinsky, 2006; 2007).

The term 'heritage language' is one that is used most commonly in relation to the first generation to be born outside the homeland, i.e., 'second-generation speakers'. We employ the term in this paper to encompass a wider group of speakers, that is, to *any* speaker of Croatian living in the diaspora. Our definition therefore includes first-generation speakers (those born in the homeland), second-generation speakers, as well as third-generation speakers, and speakers belonging to subsequent generations who may not speak Croatian as their chronologically first acquired language, but who naturalistically acquired some proficiency in it, typically in family settings. Our definition therefore relates to those with active proficiency (i.e., ability to speak). It excludes those with passive or listening comprehension skills only. The definition of heritage language speaker used throughout this paper is a functional-linguistic

one: a person residing outside a predominantly Croatian-speaking environment who has oral proficiency in Croatian and whose speech bears forms recognisably attributable to *any* variety of vernacular Croatian (Valdés, 2000).

We have thus defined what a *heritage language* is and how it relates to Croatians in the diaspora. It is important to define the variety against which we will be comparing Croatian as a heritage language. That point of comparison is not *Standard* or *Literary Croatian*. Instead, our point of comparison is the sum total of all and any variety of Croatian, whether standard or non-standard, that is spoken or used in Croatia, or in the areas of origin of Croatian speakers from Bosnia-Herzegovina, Serbia or Montenegro. We label all of these varieties under the hypernym *Homeland Croatian* (hereafter HMLD.Cro). The language most commonly spoken by many if not most Croatian emigrants is their regional variety or dialect, and this is or has been the main model that their children and further generations have acquired. In linguistic terms, HMLD.Cro is a point of contrast or 'baseline' against which data from heritage Croatian speakers is compared. As is discussed below, a monolingual variety of HMLD.Cro is not always the model or the form of input that younger generation heritage speakers receive.

In some countries of the Croatian diaspora, the immigration of Croatian speakers is long-standing, going back 100 years of more. Thus, in some countries such as the United States, Argentina and Chile, and according to the vintage of migration of their forefathers, there are Croatian-origin residents that are members of the fourth or fifth generation. The level of input needed to acquire the heritage language to a functional level typically decreases from one generation to another, usually due to younger speakers' dominance in the language of the wider society which means that in many interactions, even in those involving other Croatian-origin persons, the societally dominant language may be used more so than Croatian. Acquisition of a language is directly related to linguistic input: if a child hears a language infrequently, they will acquire only limited proficiency in it.

Comparing the countries that are represented in this paper, and the vintage of Croatian migration that each of these countries has witnessed, we see that in some countries, maintenance of Croatian is higher (e.g., Austria and Germany), while in other countries it is not so high (e.g., United States and New Zealand) (Božić, 2000; Filipović, 2001; Hlavac, 2009; Stoffel, 1982). There are factors which are conducive to

maintenance such as proximity to homeland, possibility of return, geographical concentration of speakers, and there are those less conducive such as exogamy, geographical dispersal and pre-migration acquisition of the language of the host society (Čapo *et al.*, 2014; Fishman, 2004; Kresić Vukosav & Thüne, 2019; Stolac, 2017). There are also differences between the rates of maintenance of similar generations from one country to another. We draw on available data on language maintenance from a number of countries: Australia (Hlavac, 2003; Stolac, 2017), Austria (Ščukanec, 2017), Canada (Ćosić, 1992/93; Petrović, 2018), Germany (Kresić, 2011), Italy (Piasevoli, 2007; Županović Filipin & Bevanda Tolić, 2015), the United States (Filipović, 1991; Jutronić, 1976; Jutronić-Tihomirović, 1985) and New Zealand (Stoffel, 1982; 1994), as well as Sweden (Pavlinić-Wolf *et al.*, 1987), Denmark (Pavlinić-Wolf *et al.*, 1988) and Chile (Lasić, 2010). Lastly, we draw on 11 of the 16 chapters that provide empirical data and description of Croatian as a heritage language in an edited volume that is a multi-site and comparative examination of Croatian spoken in the diaspora (Hlavac & Stolac, 2021). On the basis of these studies that include inter-generational and domain-based approaches to language use, we suggest that the likelihood of active proficiency in Croatian across generational groups is likely to be that contained in Table 1.

Table 1 contains a breakdown of generations that includes two sub-categories of first-generation speakers: 'Gen.1a' referring to those who emigrated as late adolescents or adults; 'Gen.1b' referring to those who emigrated as children before the onset of adolescence; and a 'mixed' generation category – 'Gen.2/3' referring to a person whose parents belong to two different generations. By analogy, the likelihood of a 'Gen.3/4' person speaking Croatian is likely to be somewhere between that of a 'Gen.3' and a 'Gen.4' speaker. Table 1's percentages are based on the premise that both parents of a speaker (and all grand-parents, and great grand-parents) are of Croatian origin, that is, that the speaker grows up in an endogamous (extended) family. 'Out-marrying' or exogamy where one parent is not a Croatian-speaker leads to family constellations that can substantially reduce Croatian input and the likelihood of acquisition.

Table 1. Generational membership and active proficiency in Croatian

Generation	Description	Likelihood that they speak Croatian
Gen. 1A	Born in homeland, emigrated as an adult	Very high (100%)
Gen. 1B	Born in homeland, emigrated as a child	Very high (95%)
Gen. 2	Born in the diaspora, both parents born in homeland	High (60%–80%)
Gen. 2/3	Born in the diaspora, one parent born in homeland, other parent born in diaspora	Mid-level (50%)
Gen. 3	Born in the diaspora, both parents born in diaspora, grandparents born in homeland	Mid-level to low (30–50%)
Gen. 4	Born in the diaspora, parents and grandparents born in the diaspora, great-grandparents born in the homeland	Low (10–30%)
Gen. 5	Born in the diaspora, all living preceding generations born in the diaspora	Non-existent to very low (0–10%)

Methodology

In order to examine the speech of Croatian speakers, linguistic data was gathered from informants across three generations in n nine countries. The linguistic data, in most cases, consists of recordings of semi-structured interviews or unstructured conversations between an interviewer interacting with Croatian-speaking informants, either individually or in groups. Some corpora include recordings of informants only speaking amongst themselves in intra-family or intra-group settings without the participation of an interviewer or data-collector. Most of the corpora are recent, i.e., collected in the last five years, while some relate to data collected in previous decades. The data collectors were either

co-residents of the host society and themselves situated in the diaspora setting – Jim Hlavac (Australia), Vesna Piasevoli (Italy), Hanne Skaaden (Norway) Hans-Peter Stoffel (New Zealand) – or they were Croatia-based but with close and long-standing contacts with the diaspora country and the Croatian community in it – Dunja Jutronić (USA), Marijana Kresić Vukosav and Lucija Šimičić (Germany), Ivana Petrović (Canada), Anita Skelin Horvat, Maša Musulin and Ana Gabrijela Blažević (Argentina), Diana Stolac (Australia), Aleksandra Ščukanec (Austria) and Nada Županović Filipin (Italy). Table 2 below contains details of the countries, data collection periods and numbers of recorded informants and the generation that they belong to.

Table 2. Collection periods of the corpora and number of informants from the nine countries according to generation

	USA	Canada	Argentina	Australia	NZ	Austria	Germany	Italy	Norway	**Total**
Collection Period	1980	2007	2016	1996–1997, 2010	1970–1990	2016	2016	2015, 2016	2016	
Gen. 1A		11	5	5	19	11	3	23	6	**83**
Gen. 1B				5	0	3		2	1	**11**
Gen. 2	11	11	4	100	25	7	9	19	3	**189**
Gen. 3			2	3	10			2		**17**
Total	**11**	**22**	**11**	**113**	**54**	**21**	**12**	**46**	**10**	**300**

The recordings were transcribed and linguistic phenomena identified, with a point of comparison or 'baseline' being HMLD.Cro, defined above in Section 2. The context, setting, and repertoires of other Croatian speakers with whom they communicate have direct consequences on the varieties used by the informants in the recordings. For many speakers, particularly those from the second-generation onwards, a monolingual variety of HMLD.Cro is not always the model or the form of input that they received and then acquired. We provided these only for the purpose of linguistic comparison, and it is not suggested that they are forms or constructions that are available to heritage language speakers.

Linguistic data

This section contains a selection of phenomena that are of interest from a contact linguistics perspective. The examples presented here are a selection only and are indicative of but not representative of all of the ten corpora. Examples are selected as they contain forms that diverge from those found in HMLD.Cro. We remind the reader that we focus on these phenomena because they are 'conspicuous' or 'noteworthy', but we also caution that they need not be recurrent within the speech of the same speaker, or in the vernaculars of other members of the same diaspora speech community, let alone in the speech of other diaspora Croatian speakers in other countries. Occurrences of 'one-off' phenomena need to be acknowledged as such, and where possible, recurrent contact phenomena should be foregrounded so that we can see if an innovation is widespread and indicative of a trend or not (Backus, 2015).

Our examples are categorised and sequenced according to conventional linguistic categories. These are, in order: pragmatics, lexicon with sub-categories according to part of speech, allocation of gender (for nouns) and morphological and/or phonological integration; calques and loan translations. Instances of code-switching, i.e., the contribution of lexical or other forms from two languages within the same utterance or clause are presented at the end. English glosses are provided with all examples, and where the morpho-syntactic features are of interest, a narrow gloss of these is provided as well. Where available, demographic information about the speaker of the example is given that typically includes generation, gender and age. The following acronyms are employed that identify which corpus and which language combination examples are taken from: ARG.Cro, AUS.Cro, AUT.Cro, CAN.Cro, GER.Cro, ITAL.Cro, NOR.Cro, NZ.Cro, USA.Cro. Where transfers are phonologically integrated into Croatian, we present them according to Croatian orthography, e.g. *ju no*; when they are not integrated, we present them according to the orthography of the source language, e.g. *you know*.

Pragmatics

Pragmatics deals with communicative structure and the use and meaning of words or utterances in relation to their situation. In short, pragmatics looks at *what* a speaker wishes to communicate, and *which means* s/he employs to achieve this. The most conspicuous examples of speakers' use of pragmatics are discourse markers, i.e., sequentially dependent elements that demarcate units of speech such as *oh well*, *I mean* or *you*

know in English. Structurally, discourse markers traverse linguistic boundaries relatively easily because they are morpho-syntactically untied to other elements. Many studies statistically show thatthey can be the most frequent type of transfer from one language, e.g., (Matras, 2000; Moder, 2004). In the Croatian speech of second-generation Croatian-Australians, Hlavac (2006: 1888) shows how *you know* and *so* are more frequent than their Croatian equivalents *znaš/znate* and *te/pa/tako da*, but that *kao* 'like' as an approximator or hedge is more frequent than *like*. From the predominantly Anglophone settings, we record a number of discourse markers that are themselves very common in English speech:

AUS.Cro: *yeah, like, anyway, orajt* 'all right', *dazan meta* 'doesn't matter'.

CAN.Cro: *so, yeah, anyway, doesn't matter*

NZ.Cro: *ju no* 'you know', *ja minin* 'I mean', *really, vel* 'well', *eh* (Maori English)

We can see that some are phonologically integrated amongst some speakers in some countries, e.g., *dazan meta*, while the same form is unintegrated amongst others, e.g., *doesn't matter*. The last discourse marker *eh* in NZ.Cro is unique to New Zealand, and comes originally from Maori, via Maori English. It is a turn-final marker that invites confirmation or a response from the other speaker, analogous to the Croatian tag-question markers *ne*? or *jel'*? From other settings we record the transfer of the following markers from Italian, German and Spanish respectively:

ARG.Cro: *bueno* 'okay/right/well', *claro* 'clear/of course', *sí* 'yes'

GER.Cro: *also* 'so', *nun* 'now', *nǝ* 'eh' (similar to Maori English *eh*)

ITA.Cro: *ma no* 'but no', *anzi* 'on the contrary', *comunque* 'however', *capisci* 'you understand'

Lexicon

This section presents lexical items that have entered Croatian from other languages. As stated, these are the most widely studied category of all contact linguistic phenomena in diachronic as well as synchronic studies of Croatian. Our presentation of lexical items commences with nouns with examples given that also show gender allocation and conventions of phonological and/or morphological integration. This is followed by

adjectives and adverbs and verbs. Presentation of parts of speech in this sequence reflects their reported levels of frequency in contact linguistics studies in general (Winford, 2003: 51).

Nouns
Looking at predominantly Anglophone societies first, the following are some of the nouns located in these corpora:

CAN.Cro: *đinđerela* 'ginger ale', *imigrejšn* 'immigration', *bankrupsi* 'bankruptcy', *demiđ* 'damage', *nejber* 'neighbour', *šapa* – shop, *kena* 'can'

(G1,F,72)   *... ja bi volila popit čašu vina sa malo **đinđerelom**...*

'…I would like to have a glass of wine with some **ginger ale**…'

NZ.Cro: *resa* 'race', *goma* 'gum', *marketa* 'market', *gambelja* 'gamble', *tanga* 'water tank', *šanda* 'shanty'

*u nas je bilo... **timber** za učinit mu **šandu.***

'we had… **timber** for us to build him a **shanty**'.

USA.Cro: *donsi* 'doughnuts', *kort hauz* 'courthouse', *runavej* 'runway', *damidž* 'damage', *kozin* 'cousin', *blankete* 'blankets', *šifta* 'shift', *tičerka* 'female teacher'

*bija je na **runaveju**. lipe **kops** imate*

'he was on the **runway**'. 'you have nice **cups**'.

Many of the lexical items are realia characteristic of the diaspora setting such as *goma* 'gum' and *šanda* 'shanty', while others such as *donsi* 'doughnuts' may have apparent equivalents in Croatian such as *uštipci*, but the appearance (and taste!) of them may be specific to the diaspora context that accounts for their occurrence. The following are reported from Spanish-, Italian-, German- and Norwegian-language settings:

ARG.Cro: *barcito* 'little coffee bar', *asado* 'roast meat', *cuadra* 'neighbourhood'

*Ovisi u kojoj si **empresi**.* (G1,F,74)   *Sam imala ovako jednu **bolsitu*** (G1,F,88)

'It depends which **firm** you work in.'   'I, cos' I had this like **little bottle**.'

ITA.Cro: *šufite* 'attics', *tonno* 'tuna', *mercurio* 'mercury', *concorsi* 'competitions'

Prete *dođe i reče.* (G1,F,72) *Radila sam kao* **ragioniere.** (G1,F,74)

'The **priest** came and said.' 'I worked as an **accountant.**'

AUT.Cro: *šulrat* 'school council', *Krankengeschichte* 'medical history',

Schlüsselkraft 'key position', *Hindernis* 'obstacle', *Einstellung* 'attitude'

ja sam išla na svoj **mündliche Matura** *otišla sa* **Fieber.** (G2,F,22)

'I went for my **final year oral exam**, leaving home with a **fever**'.

NOR.Cro: *kruške vadi, i stavlja u.. u øm ..., kako se zove* **kurv?** (G2,F,25)

'She takes out the pears and puts them in… in, um.. how do you say **basket**?'

We can see from a number of these examples, particularly those from English and German, languages with a high percentage of lexemes have word-final consonants, that transfers are allocated masculine gender when integrated into Croatian, e.g. *nejber* 'neighbour', *runavej* 'runway' and *damidž* 'damage'. But a number of items with word-final consonants attract an *-a* suffix, which marks them as feminine gender nouns, e.g., *resa* 'race', *marketa* 'market' and *šifta* 'shift'. A larger number of these are found in the speech of older vintage emigrants in NZ.Cro, compared to CAN.Cro and USA.Cro that have speakers also from more recent migration waves. It is not clear why some speakers affix a feminine suffix; it may be that the legacy of Italian as a source language for many transfers in the speech of NZ.Cro-speakers from Dalmatia is an influence that is applied also to those from English. Overall, and across all corpora, a tendency is apparent for speakers to allocate masculine gender to transferred nouns. In the AUS.Cro corpus, where the gender of the nouns is ascertainable, 79% of transferred nouns are masculine, with the remaining 21% feminine, and there are no neuter nouns identified. These trends are in line with Surdučki (1978: 288), Jutronić-Tihomirović (1985: 33) and Filipović (1986: 130) who all reported a tendency to

integrate transfers as masculine nouns amongst speakers in North America.

Many of the transferred nouns presented above are phonologically and/or morphologically integrated, while others are not. It is no coincidence that integration occurs more frequently in the speech of Gen.1 speakers, and this is evident where information is available of the generational membership of the speaker. Conversely, unintegrated forms occur more frequently in the speech of Gen.2 and Gen.3 speakers. Below are some more examples of integration, all coming from Gen.1 informants:

ITAL.Cro   *bila sam izvan **šagome*** (G1,F,75)

            'I defied the rule' (Ital.: *sagoma* 'model')

NZ.Cro: ***švanap*** 'swamp', *u **švampiman*** 'in the swamps' (G1)

            šal(ad) 'shallow area', *kopat na **šal(ad)i*** 'to dig on shallow land'

CAN.Cro: *... rekao je da sam radila u tako nekom **ofisu***... (G1,F,43)

            'and he said that I worked in some office…'

When integration occurs, it is usually phonological. Above we see some conventions of integration that involve vowel epenthesis (runway > *runavej*) or the replacement of phonemes not contained in Croatian such as the labio-velar approximant /w/ with a labio-dental fricative /v/ with assimilation of /s/ to /ʃ/ (swamp > *švanap*). Morphological integration is apparent only where the structural role of the noun requires overt markers. Many, such as singular nouns in subject position, do not typically attract morphological markers. Studying the speech of Gen.1 gumdiggers in New Zealand, Stoffel (1991) records well over 200 different words and expressions specific to this field of work and almost are completely adapted to the phonological system of NZ.Cro.

Below are some examples of unintegrated nouns, both from Gen.2 speakers:

AUS.Cro: *... ima baš onaj **sauce** na njega što je najbolje...* (G.2,F,25)

            ... it has that **sauce** on it which is the best…

AUT.Cro: *To je bio njoj **Hindernis** što je bila bolesna.* (G.2,F,22)

            'She was not feeling well and this was her **obstacle / excuse**.'

Overall, there are a number of trends in relation to phonological integration. 'Conspicuous interventions' in the phonological integration of transfers (e.g. *runavej, švanap*) tend to occur amongst older vintages of Gen.1 migrants, amongst whom proficiency in the host society language often remained low. Their lack of proficiency in the phonology of the host language led to changes to align transfers to the phonotactic structures of their own Croatian variety. Phonological integration amongst younger or recent-vintage Gen.1 speakers tends to be restricted to changes in vowels, often diphthongs being rendered as monophthongs with or without *-j*, e.g., Eng: *overs* [oʊvəz] > AUS.Cro: *overi*, Eng: *state* [steɪt] > USA.Cro: *štejt*, or those consonantal sounds not in Croatian being replaced by their closest Croatian equivalents, Eng: *rent* [ɹent] > AUS.Cro: *rent*.

Amongst Gen.2 and Gen.3 speakers the situation is different. They have native-like pronunciation of forms from the host language which for most is their L1. They need not integrate transfers to the phonological system of Croatian as their bilingual repertoires encompass the phonological inventories of *both* languages, and integration according to Croatian phonological (and morphological) marking is optional, not involuntary. This is why, in general terms, amongst Gen.1 speakers (and Gen.1A speakers in particular), incidence of phonological (and morphological) integration is typically very high, around 90%. Amongst Gen.2 speakers and those of subsequent generations, it is 30%–40%. In relation to the differences in frequency of integration, Albijanić (1982: 18) reports that amongst his Gen.1 speakers, 99% of English-origin transfers are 'assimilated', while amongst his Gen.2 informants, it was between 11% and 50%. Morphological integration generally mirrors the trends for phonological integration. Morphological without phonological integration is less common.

Adjectives and adverbs
Adjectives and adverbs are parts of speech that are the next most likely to be transferred after nouns in contact situations (Field, 2002: 35, 38).

ARG.Cro: *peligroso* 'dangerous', *ridculos* 'ridiculous', *tremendo* 'tremendous'

CAN.Cro: *bizi* 'busy', *fultajm* 'full-time', *laki* 'lucky', *najs* 'nice', *pjur* 'pure'

USA.Cro: *bjutiful* 'beautiful', *nu* 'new', *šur* 'sure', *krezi* 'crazy', *najs* 'nice'

Transferred adjectives can appear as attributives to a noun:

USA.Cro: ***Front*** *soba je bila...* 'The **front** room was...'

AUS.Cro: ***Rong*** *je čaša!* 'That's the **wrong** cup!'

But it appears they more commonly occur as predicates:

ITAL.Cro: *… hrana nije tako **sana**… zdrava k'o u Italiji.* (G3,M,9)

'… the food is not as ***health***… healthy as in Italy.'

*Sestra je bila **brava**.* (G1,F,64).

'My sister was good at school…'

ARG.Cro: *…gospođe su tolko pažljivi i... **paciente**, kako se kaže.*

'… the ladies are so considerate and… **patient,** as they say.'

In some instances, comparative forms of adjectives occur, e.g. NZ.Cro: *streta* 'straighter'. In another instance, the comparative suffix *-ije* is affixed to a predicate adjective:

NZ.Cro: *To je malo **hardije**.* (G1)

'This one [pointing to a piece of gum] is a bit **harder**.'

In another instance, the word final *-i* ending of a geographical location results in it being inflected as an adjective:

NZ.Cro: *U **Sidnom*** (ADJ-LOC.M.SG) *smo stali.* 'We stopped off in **Sydney**.' (G1)

Adjectives can also combine with adverbs that further quantify them:

USA.Cro: *On je bil **so skerd***. 'He was **so scared**'. (G2)

Transferred adverbs can occur on their own, here quantifying a Croatian constituent:

ARG.Cro: *Sam ga našo i da, ja.... i ništa. Okaj, **directamente** ništa.*

'I found him and yeah, I… and it was nothing. Okay, **directly** nothing.

Verbs

After nouns, adjectives and adverbs, the part of speech that is the next

most likely to be transferred is verbs. Verbs play a central role in the morpho-syntactic grid of clauses. It is this central role that accounts for why verbs are much more likely to attract morphological (and phonological) markers. This applies also to Gen.2 speakers: in an Australian-Croatian corpus of 2,000 English lexical transfers, only about 15% are integrated, but for verbs, this percentage rises to over 70%. We give examples here in their infinitive form, which was either the form they were used in, or the infinitive suffix was recognisable from the finite form of the verb used:

CAN.Cro: *agrijati se* 'to agree', *tičati* 'to teach', *fonati* 'to phone'

USA.Cro: *bildati* 'to build', *forgetati* 'to forget', *stapat* 'to stop', *pentati* 'to paint'

AUS.Cro: *stopati* 'to stop', *fiksati* 'to fix', *tajpovati* 'to tape', *overhaulati* 'to overhaul'

NZ.Cro: *špirat* 'to spear', *oškrepat* 'to scrape'.

In the last example, we see how an English-origin verb is not only phonologically integrated, but it also attracts a prefix '*o-*' that renders it a perfective verb. (The limits of this paper prevent us from looking at aspect as a semantic-syntactic category in more detail.) Below are some examples of verbs in their finite form:

ARG.Cro: *ako se manjine ne **pueblaju** sa Argenticima, uvijek će..*' (G1,74)

'.. if minorities don't populate (=reproduce) with Argentinians, it'll always..'

ITA.Cro: *On je bil **ferito** della Prima Guerra… bio je star*. (G1,F,74)

'He was **wounded** in World War I… he was old.'

The last example with a past participle is less common but shows that transferred verbs can sometimes occur in compound tense constructions. Further, the valence of a transferred verb can vary from one community to another. For example, in AUS.Cro, *ringat* 'to ring' has a succeeding (animate) object in ACC case, e.g. *ringat ću te* 'I'll ring you-ACC'. In NZ.Cro the same verb requires DAT marking for an animate object, e.g. *ringat ću ti* 'I'll ring you-DAT'.

Calques and loan translations

This section presents examples of calques that refer to the transference of properties (structural or semantic) of a donor language word onto its Croatian equivalent. A calque can be defined as "*a more or less faithful reproduction of a foreign-language item* [or items] *via an item* [or items] *in the lexicon of the recipient language*" (Turk, 2013: 45). Typically, calques are shorter, one- to three-word NP constructions while the term 'loan translation' suggests a construction which is longer, such as a VP. We draw on Backus & Dorleijn's (2009: 77) definition of loan translations being the "*usage of morphemes in Language A that is the result of literal translation of one or more elements in a semantically equivalent expression in Language B*". An example of the transfer of the semantic properties of a donor language verb onto a Croatian one is given below:

GER.Cro: *i tamo je **radio** zanat* (Ger. 'und da **machte** er seine Ausbildung') (G2,M,41)

'and that's where he **did** his vocational training.'

Here, the polyfunctionality of German *machen* 'to do/make' influences its Croatian equivalent, *raditi*. German *machen* can also mean 'complete', where Croatian *raditi* does not have this meaning. (The verb *odraditi* 'to have completed the task [of doing something]' does have this sense, but is more specialised than German *machen*.) Thus, the semantic field of *machen* influences a German-Croatian bilingual's use of *raditi*.

Compound constructions such as complex NPs are another type of calque where a sequence of constituents that assumes a particular meaning in one language is replicated in another by the same sequencing of equivalent Croatian forms. In the example below, a second-generation speaker is enquiring about scholarships:

ITA.Cro: *Imate li vi kakvu **burzu od studija***? (G2,F,26)

'Do you have any **scholarships** [lit. 'stock exchanges of study']'

(Ita.: **borsa di studio**; HMLD.Cro: stipendija)

The above listed examples of calques are congruent to those recorded in previous studies of Croatian as a heritage language, e.g., *meko piće* 'soft drink', *stranorođen* 'foreign'. They represent loan translations where phraseological expressions from one language are replicated in another.

In the following examples, the meaning of the loan translation is still probably apparent to those without proficiency in the donor language:

CAN.Cro: *...a ovde naši ljudi **rade duge sate** i mizeran poso..* (G1,F,55)

'…our people here **work long hours** at tedious jobs…'

*…je li voliš ovi majica, je li voliš.. o ne volim kako **to stoji na nju***… (G2,F,22)

'…do you like this shirt, do you like it.. oh I don't like how **it looks on her**...'

AUS.Cro: *… onda mama nije htjela **doći mene dobit**.* (G.2,M,22).

'… then mum didn't want **to come and get me**'.

GER.Cro: *Ali opet kažem sve **leži do** roditelja.* (G2,F,42)

'And I'm saying again that everything **is up to** the parents.'

[Ger. … alles **liegt an** den Eltern; HMLD.Cro: sve **je do** roditelja.]

But in the following examples, it is likely that the fellow interlocutor requires a knowledge of the other language, and the bilingual norms of speakers of the speech community to understand the meaning of the phrase:

NZ.Cro: *ić **u slike*** 'go **to the pictures**' (= 'go to the cinema') (G2)

*Ja **gledan za njega** u garden.* 'I am **looking for him** in the garden' (G1)

ITA.Cro: ***Titula** te knjige mi **bježi**.* (G2,M,18)

'**The title** of the book **escapes** me' = 'I can't remember it'.

(Ita: **il titolo** del libro mi **sfugge**. HMLD.Cro: **ne sjećam se naslova** knjige)

ARG.Cro: *I jezik je… **nema ništa vidjeti sa** španjolskim* (G2,M,69)

'And the language is… has nothing to do with Spanish.'

(Spa: **nada que ver con**… HMLD.Cro: **nema nikakvu vezu sa**...)

Surdučki (1966: 132) also records analogous loan translations of verbal phrases such as *praviti novac* 'make money', *skakati ka zaključku* 'jump to a conclusion' or *dati nekome kredit* 'give someone credit' (=acknowledging someone). Other examples of loan translations recorded in English-Croatian diaspora settings are *to je u redu sa mnom* 'that's alright with me' (Škvorc, 2006: 20) and *kuća u dvanaejs nogu dugu* 'a house twelve feet long' (Gasiński, 1986: 37).

Code-switching

The last category of language contact phenomena that we look at is code-switching. Code-switching refers to "*the use of material from two (or more) languages by a single speaker in the same conversation*" (Thomason, 2001: 132). This paper adopts a division of code-switching according to the point at which the switch in language takes place in regard to clause boundaries. Thus, we distinguish 'extra-clausal' code-switching which refers usually to the insertion of discourse-specific elements, such as discourse markers that are grammatically independent of the structural features of the clause.

ARG.Cro: *Čokljava, ćorava, šepava, **no importa**, ja idem.* (G1,F,74)

'Lame, blind, limping, **doesn't matter**, I go.'

AUS.Cro: *Osveta, **that's right**, on je njega ubio i tako..*(G2,M,20)

'Revenge, **that's right**, he killed him and so…'

In the above two examples, the extra-clausal code-switches appear as an evaluative aside and as an acknowledgment marker respectively. Both of them are structurally untied to the remaining constituents in the utterance. Aside from extra-clausal code-switching, we further distinguish between 'intra-clausal' code-switching that takes place within the same clause, often referred to 'insertions' in Muysken's (2008) terms, and 'inter-clausal' code-switching that takes place at clause boundaries, which is equivalent to Muysken's term of 'alternation'. Below are selected examples of intra-clausal code-switching:

AUT.Cro: *jesam se pripremala ali **natürlich** je to za mene bilo **eine Woche vor**.* (G2,F,22)

'I did do some preparation, but **of course** for me that was **one week earlier'**

ITA.Cro: *šta ima **gente** radit u Bosni **durante l'anno** kad nema nikog?* (Gen.2,F,21)

'what can **people** do in Bosnia **during the year** when there's nobody there?'

ARG.Cro: *Ovi su vam **mejores** koji imaju **dulce de leche**. To je **dulce**, to je slatko od mlijeka, znate.* (G2,M,43)

'These are better those that have **dulce de leche** ['sweetened caramel milk'].   That is **dulce** ['sweet'], that is sweet from the milk, you know.

GER.Cro: *… svi koji su tamo na sajmu radili, isto su bili kao Nijemci meni. Isto su svoj posao **haben den Ernst genommen**.* (G2,F,31)

'… all who worked there at the trade fair, they were also like Germans to me. Also, their work, they **took it seriously.**'

AUS.Cro: *... kao... kako on... um... **comes to terms with**... ta žena koja je umrla..* (G2,F,17)

' ... like... how he... um... **comes to terms with.**.. that woman who died...'

We see in the five examples above that in some cases the code-switches contain individual items only, in others they are ADV phrases, and in some others they are larger parts of VPs. In general, the code-switch itself does not appear to be significant, i.e., the change of language does not signify a change in the footing, pitch or conversational implicature of the speaker's utterance. Instead, for the speakers who produced these utterances, code-switching is an unremarkable and unmarked occurrence when interacting with bilinguals with similar linguistic repertoires to themselves. The profiles of other interlocutors determine the sociolinguistic situation such that these speakers then switch to 'bilingual mode' (Grosjean, 2013) and the variety of speech that can often be used in this kind of situation is one that draws on two linguistic codes, i.e., 'classic code-switching' (Myers-Scotton, 2002). In only one case does the code-switch itself appear to be meaningful – *dulce de leche* – which is a speciality peculiar to Argentina, and without an apparent equivalent in Croatian. Here, 'lexical need' is the motivator; but in the other instances this is not the case. A topical area of study in relation to intra-clausal code-switching is whether constraints or facilitating factors

exist, i.e., whether speakers can freely change languages at certain points in an utterance, or whether there are structural properties that inhibit this happening between certain combinations of words or morphemes. Space prevents us from looking into this here, and we refer the interested reader to descriptions of code-switching, e.g., Gardner-Chloros (2009) and models that explain its structure, e.g., Myers-Scotton (2002).

We now move to instances of inter-clausal code-switching.

CAN.Cro: *...ako, ako se neko radovao, **they had some mental problems**, to je svak mrzio. **I'm telling you***...

'...and if, if somebody was looking forward to it, **they had some mental problems**, everybody hated that. **I'm telling you**…' (G2,M,22)

In the above example, each switch in the language choice occurs when there is a shift in footing about what the speaker is talking about. After the first clause in Croatian, the alternation to English *they had some mental problems* functions as a descriptive aside to the situation that has just been recounted. After switching back to Croatian, the last code-switch into English is a listener-focused speech act *I'm telling you*. The code-switches themselves augment the conversational shifts in this passage – it is not just a change of topic focus that the speaker employs, it is a change of linguistic code that is 'pressed into service' to achieve this. In the example below, the speaker is retelling what she dreamt last night. She is addressing her daughter and her daughter's boyfriend who are both also Croatian-Italian bilinguals and she commences her turn in Italian:

ITA.Cro: ***Poi mi ha fatto arrabiare papà perché non mi ha fatto la foto.***

Ma digni se, ajde! Uglavnom, stvarno baš gluposti koji puta sanjam... (G1,F,48)

'**Then dad made me angry because he didn't take the picture**. Come on, get up! Anyway, sometimes I really have stupid dreams…'

Similar to the above example from Canada, the code-switch co-occurs with a shift in conversation. From the description of the dream in Italian, the mother then shifts to Croatian to issue an instruction to her daughter, using the imperative, *Ma, digni se*! 'Come on, get up!'. She then continues in Croatian in making a self-evaluation of her dream, which is

directed both at herself as well as the other two interlocutors. Sometimes inter-clausal code-switching can occur due to a speaker quoting verbatim what another person said. The following example is from Germany, and the speaker starts her turn in German:

GER.Cro: ***Die gute Frau Kempel***. *To je bila stara njemačka čistačica, reče samo, **"Hää"**. A ja njoj, **"Was stellen Sie hier dar**?"* (G1,F,70)

'**Good old Mrs. Kempel**, she was an old German cleaning lady, and she just said "**Whaat**?" and I replied "**What do you represent here?**"'

The topic of the turn is introduced in German which topicalises the main protagonist, *Frau Kempel*. The code-switch into Croatian is a different speech act that evaluates the protagonist, and her direct speech is introduced overtly via *reče samo* 'she just said'. The code-switch to German is a verbatim description of what Frau Kempel said. The code-switch back to Croatian is another speech act that foregrounds the response, and the response given in German at the time is rendered again in German. This is an example of how quoting verbatim the speech of others can precipitate inter-clausal code-switching. In the following example, it is two different sets of interlocutors that account for why the speaker code-switches from English into Croatian:

NZ.Cro: ***Oh, you've bought a section*** ['sekʃn]... *Oni su kupili sekšon* ['sekʃon]. (G2,M)

'**Oh, you've bought a section** (plot of land)... They have bought a section'

The first utterance above is addressed to English-speaking acquaintances who are physically present and the speaker is addressing them directly via 2PL, *you*. The speaker's parents are next to him, and relays, using third person, the same information to his parents. He repeats the word *section* used when addressing the English-speakers, and phonologically integrates it in a way characteristic of the speech of his Gen.1 parents.

Conclusion

The aim of this paper is to report on common and prominent contact linguistic phenomena from a multi-site project examining features of heritage Croatian as it is spoken in nine different countries across four continents. We contextualised existing findings of Croatian as a heritage

language in our presentation of linguistic phenomena and also applied a contact linguistics approach to our analysis of examples at the same time. We have examined corpora from each ot the studies according to linguistic categories and presentedselected examples from the following four areas: pragmatics, lexicon, calques/loan translation and code-switching.

The transfer of pragmatic markers is a common occurrence in language contact scenarios. Discourse markers readily cross linguistic boundaries due to their frequency or immediacy in the speech patterns of speakers who otherwise are frequently employing the societally dominant language. We record a number of pragmatic forms originating from other languages in the speech of Croatian speakers. There are forms such as *ju no* 'you know', *ja minin* 'I mean', and *really* that have ready Croatian equivalents such as *znaš/znate*, *hoću reći* and *stvarno*. But there are others such as *anyway* or *orajt* 'alright' that are perhaps more characteristic of Anglophone discourse patterns, and their occurrence represents not only a transference of the form of the discourse markers, but also a transference of Anglophone discourse patterns *together with* the forms themselves. In other instances, the polyfunctionality of English *yeah* and Spanish *bueno* that perform a wide variety of functions (e.g., agreement-marker, back-channelling signal, turn-terminator, hedge etc.) in English and Spanish respectively may account for why they frequently find their way into Croatian speech as well. Overall, we observe that these discourse markers are additive features to speakers' repertoires and do not displace their Croatian equivalents, at least not completely.

The transfer of lexical items is a well-studied topic in the field of philology in general, as well as in contact linguistics in particular. Nouns are the most common lexical item to be transferred. Some lexical items relate to realia of the diaspora setting that have no ready equivalents in Croatian. It is therefore 'lexical need' that accounts for their occurrence. This applies to forms like *šanda* 'shanty', *đinđerala* 'ginger ale' or *šulrat* 'school council'. But most are not specific to the new environment, and instead occur as items that are maybe 'more immediate', 'more easily retrieved' or 'more amenable' than their Croatian counterparts, e.g., *fultajm* 'full-time' – HMLD.Cro: *puno radno vrijeme*; *overhaulati* 'to overhaul' – HMLD.Cro: *remontirati/temeljito pregledavati*; *laki* 'lucky' – HMLD.Cro: *imati sreću*; *imigrejšn* 'immigration' – HMLD.Cro: *ured za useljavanje*. We also see that very often there is no apparent reason for the occurrence of lexical transfers and this observation is in line with other bilingual data sets of diaspora populations that report that 'bilingual

speech' is a default variety of speech for many speakers in many intra-group settings, and a monolingual variety of speech (employing either language only) would be more marked and conspicuous than one that encompasses two linguistic codes.

Comparing parts of speech, we report that nouns are the most numerous groups of lexical items, followed by adjectives and adverbs, followed then by verbs. Where morphological or attributive features show the gender of a transferred noun, we observe that most are allocated masculine gender. There are phonotactic reasons for this based on the form of nouns in the source languages where consonant-final words attract masculine marking. This is particularly the case with nouns transferred from English and German that commonly feature consonant-final endings. Amongst some groups of heritage speakers, affixation of the suffix *–a* renders consonant-final transfers as feminine nouns. It is possible that pre-emigration conventions in the assimilation of Italian loanwords in the vernaculars specific to those emigrants – specifically older vintage migrants from Dalmatia – are re-applied to English-origin transfers that result in these *-a* suffixed feminine transfers. Differences emerge in relation to the frequency of phonological and morphological integration of nouns, less so adjectives and adverbs. Gen.1 speakers almost always integrate other-language items into Croatian, both phonologically and morphologically. Gen.2 and Gen.3 speakers generally do not do this; overall, their rates of phonological and morphological integration of transfers into Croatian are between 30 and 40%.

Calques and loan translations are a less transparent contact linguistic category as they consist of Croatian constituents only. A small number of calques and instances of semantic transference are recorded; loan translations are more common. Some are based on other-language models that are not so different from the Croatian equivalents, so that they still remain comprehensible to a Croatian speaker without knowledge of the other language, e.g., *..a ovde naši ljudi **rade duge sate*** '…our people here **work long hours'** – HMLD.Cro: *ovdje naši ljudi **dugo ostaju na poslu**; … mama nije htjela **doći mene dobit*** '…mum didn't want **to come and get me'** **–** HMLD.Cro: *mama nije **me htjela pokupiti**.* Others are not so transparent and their use is evidence of recurrence and widespread bilingualism in intra-group settings that allows loan translations to become popularised, e.g. *Ja **gledan za njega** u garden* 'I am **looking for him** in the garden' – HMLD.Cro: ***Tražim ga** u vrtu*; *jezik **nema ništa vidjeti sa***… 'the language **has nothing to do**

with…' HMLD.Cro: *jezik **nema nikakvu vezu sa**….* Such loan translations (and calques) may appear unusual or incongruous. In fact, calques and loan translations, whether produced unconsciously or consciously, can often be a source of mirth and for some they are a folk-linguistic exercise that show how diaspora speakers can 'play with their two languages'. But we also suggest that they may serve another purpose. Although some calques and loan translations may look 'odd', they are a strategy for speakers 'to keep speaking Croatian' so that that they produce constructions or passages in Croatian when otherwise they would produce these constructions or passages in the other language via code-switching or use of lexical transfers. It may be that calques and loan translations are a 'maintenance strategy' that enables speakers to 'keep speaking Croatian'.

The last category of contact linguistic phenomena that we have examined are code-switches. In structural terms, code-switching can occur within clauses and at clause boundaries. Extra-clausal code-switches such as pragmatic or discourse markers commonly punctuate speakers' repertoires, and these record a high frequency. Intra-clausal code-switches or multi-item insertions are also frequent. Similar to lexical transfers, their incidence can be motivated by context, theme or situation. At the same time, there need not be a reason or motivation for their occurrence: speakers regularly draw on their other code as an unmarked feature when interacting with other in-group members. Inter-clausal code-switches also commonly occur. In many cases, the effect of a change in footing or change of speech is augmented by the code-switch where the shift in conversational locus is amplified by the concurrent change in linguistic code.

Our overview of some language contact phenomena from a multi-site project of various Croatian-speaking communities in the diaspora reveals many expected outcomes, and also some unexpected ones. We see how a contact linguistics approach to various corpora enables not only a comparison of data between Croatian communities living in different countries but also with other heritage language groups in similar diaspora settings. We look forward to further studies of the speech of Croatian emigrants as a contribution to our knowledge of *all* varieties of spoken Croatian. We also look forward to these studies positioning Croatian as a more frequently studied language in contact linguistics research.

Bibliography

Albijanić, A. (1982). San Pedro revisited: Language maintenance in the San Pedro Yugoslav Community. In R. Sussex (Ed.), *The Slavic Languages in Emigre Communities*, *Linguistic Research Inc.,* 11–22.

Backus, A. (2015). A usage-based approach to code-switching: The need for reconciling structure and function. *Code-switching Between Structural and Sociolinguistic Perspectives* 43, 19–38.

Backus, A. & Dorleijn, M. (2009). Loan translations versus code-switching. In *The Cambridge Handbook of Linguistic Code-switching*, Cambridge University Press, 75–94.

Božić, S. (2000). *Kroaten in Wien. Immigranten und Integration im Zusammenhang mehrschichtiger ethnischer Beziehungen.* Naklada Jesenski i Turk.

Central State Office for Croats Abroad (2019). Hrvatski iseljenici u prekomorskim i europskim državama i njihovi potomci. https://hrvatiizvanrh.gov.hr/hrvati-izvan-rh/hrvatsko-iseljenistvo/hrvatski-iseljenici-u-prekomorskim-i-europskim-drzavama-i-njihovi-potomci/749

Čapo, J., Hornstein Tomić, C., & Jurčević, K. (Eds.) (2014). *Didov san. Transgranična iskustva hrvatskih iseljenika.* Institut za etnologiju i folkloristiku & Institut društvenih znanosti Ivo Pilar.

Ćosić, V. (1992/93). Sociolingvistički status hrvatskih iseljenika u Quebecu. *Radovi razdjela filoloških znanosti* 22(23), 27–45.

Ethnologue. (2018). Languages of the World: Croatian. https://www.ethnologue. com/language/hrv

Field, F. (2002). *Linguistic borrowing in bilingual contexts.* John Benjamins.

Filipović, R. (1986). *Teorija jezika u kontaktu.* Školska knjiga.

Filipović, R. (1991). Neposredni jezični dodiri u hrvatskim dijalektima u SAD. *Senjski zbornik* 18, 31–40.

Filipović, R. (2001). Croatian dialects in the United States: sociolinguistic conditions for the maintenance of a dialect. *International Journal of the Sociology of Language* 147, 51–63.

Fishman, J. (2006). Language Maintenance, Language Shift, and Reversing Language Shift. In T. Bhatia & W. Ritchie (Eds.), *The Handbook of Bilingualism,* Blackwell, 406–436.

Gardner-Chloros, P. (2009). *Code-switching.* Cambridge University Press.

Gasiński, Th. (1986). English elements in the speech of the Croatian immigrant community of Santa Clara Valley, California. *Zbornik Matice Srpske za filologiju i lingvistiku* 29(2), 31–45.

Grosjean, F. (2013). Bilingual and monolingual language modes. In C. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics.* Blackwell, 1–9.

Hlavac, J. (2003). *Second-generation speech. Lexicon, code-switching and morpho-syntax of Croatian-English bilinguals.* Peter Lang.

Hlavac, J. (2006). Bilingual discourse markers: Evidence from Croatian-English code-switching. *Journal of Pragmatics* 38, 1870–1900.

Hlavac, J. (2009). Hrvatski jezik među Australcima hrvatskog podrijetla. In J. Granić (Ed.), *Jezična politika i jezična stvarnost*. Croatian Applied Linguistics Society, 84–94.

Hlavac, J. & Stolac, D. (2021). *Diaspora Language Contact. The Speech of Croatian Speakers Abroad.* Mouton De Gruyter.

Jutronić, D. (1976). Language maintenance and language shift of the Serbo-Croatian language in Steelton, Pennsylvania. *General Linguistics* 16, 166–186.

Jutronić-Tihomirović, D. (1985). *Hrvatski jezik u SAD.* Logos

Kresić, M. (2011). Deutsch-kroatische Sprachidentitäten. In Z. Glovacki-Bernardi (Ed.), *Deutsch in Südost- und Mitteleuropa: Kommunikationsparadigmen im Wandel.* FF Press, 93–104.

Kresić V., M. & E.-M. Thüne. (2019). Remigration as a linguistic experience: language-related aspects of 'homecoming'. In I. Fabijanić, L. Štrmelj, V. Ukic Košta, & M. Bregovic (Eds.), *Migrations: Literary and Linguistic Aspects.* Peter Lang, 297–324.

Lasić, J (2010). Jezični identitet hrvatske iseljeničke zajednice u Čileu. *Croatian Studies Review* 6, 163–172.

Matras, Y. (2000). Fusion and the cognitive basis for bilingual discourse markers. *International Journal of Bilingualism* 4(4), 505–528.

Moder, C. L. (2004). Discourse across cultures, across disciplines: an overview'. In C. Moder & A. Martinovic-Zic (Eds.), *Discourse Across Languages and Cultures.* John Benjamins, 1–11.

Muysken, P. (2008). *Functional categories.* Cambridge University Press.

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

Pavlinić-Wolf, A., Anić, J., & Ivezić, J. (1987). Jezik jugoslavenske djece koja žive u Švedskoj (Preliminarni rezultati sociolingvističkog istraživanja). *Slavica Lundensia* 11, 159–173.

Pavlinić-Wolf, A., Brčić, K., & Jeftić, N. (1988). Supplementary mother-tongue education and the linguistic development of Yugoslav children in Denmark. *Journal of Multlingual and Multicultural Development* 9(1–2), 151–167.

Petrović, I. (2018). Croatian as a heritage language in Canada. *Zbornik radova Filozofskog fakulteta Sveučulišta u Splitu* 11, 59–72.

Piasevoli, V. (2007). La comunità Croata di Trieste ed altre associazioni. In D. Murković (Ed.), *I croati a Trieste*. Comunità croata di Trieste, Tergeste grafica e stampa 495–512.

Polinsky, M. (2006). Incomplete acquisition: American Russian. *Journal of Slavic Linguistics* 14(2), 191–262.

Polinsky, M. (2007). Reaching the end point and stopping midway: different scenarios in the acquisition of Russian. *Russian Linguistics* 31, 157–199.

Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism* 13, 155–163.

Stoffel, H.-P. (1982). Language Maintenance and Language Shift of the Serbo-Croatian Language in a New Zealand Dalmatian Community. In R. Sussex (Ed.), *The Slavic languages in emigre communities.* Linguistic Research Inc. 121–139.

Stoffel, H.-P. (1991). Common Trends in the Morphological Adaptation of Loanwords in Serbo-Croatian Migrant Dialects. In V. Ivir & D. Kalogjera (Eds.), *Languages in Contact and Contrast.* Mouton, 417–429.

Stoffel, H.-P. (1994). Dialect and Standard Language in a Migrant Situation: The Case of New Zealand Croatian. *New Zealand Slavonic Journal* 21, 153–170.

Stolac, D. (2017). Odnos nasljednoga jezika i standardnoga hrvatskog jezika u govoru povratnika. In M. Sopta, V. Lemić, M. Korade, I. Rogić & M. Perić Kaselj (Eds.), *Hrvatska izvan domovine II.* Centar za istraživanje hrvatskog iseljeništva, 675–683.

Surdučki, M. (1978). *Srpskohrvatski i engleski u kontaktu.* Matica Srpska.

Ščukanec, A. (2017). Nov(ije) generacije Hrvata u Austriji: razmišljanja i iskustva. In M. Sopta, V. Lemić, M. Korade, I. Rogić, M. Perić Kaselj (Eds.), *Hrvatska izvan domovine II.* Centar za istraživanje hrvatskog iseljeništva, 433–439.

Sažetak

Od otprilike 6,67 milijuna hrvatskih govornika širom svijeta oko 1,6 milijuna pripada hrvatskoj dijaspori, od Latinske Amerike do zapadne Europe, i od Sjeverne Amerike do Australije. Međunarodni projekt kojim se istražuju obilježja hrvatskoga jezika govorenoga u dijaspori započeo je 2015. godine obuhvativši deset korpusa lingvističkih podataka prikupljenih u devet zemalja na četiri kontinenta. Temeljen na spoznajama iz prethodnih istraživanja govora hrvatskih emigranata ovaj rad definira i objašnjava pojam *hrvatski jezik kao nasljedni jezik*. Nakon teorijskoga uvoda naš se fokus usmjerava na govornike i donose se podatci iz terenskih istraživanja i zabilježenih snimljenih govora sakupljenih između 300 iseljenika hrvatskoga porijekla. Ispitanici pripadaju prvoj, drugoj i trećoj generaciji iseljenika. Donosi se pregled značajki hrvatskoga kao nasljednoga jezika u odnosu na četiri područja: pragmatika, leksik, jezično kalkiranje te prebacivanje kodova. U našem predstavljanju primjera iz svih deset prikupljenih korpusa uspoređujemo podatke između pojedinih država i iseljeničkih valova, te između govornika starijih i mlađih generacija. U raščlambi primjera također ih uspoređujemo s oblicima u uporabi u hrvatskome jeziku u domovini, kako sa standardnom, tako i s nestandardnim idiomima. Naše zamjedbe pružaju suvremen i međunacionalni opis hrvatskoga kao nasljednog jezika. Istovremeno se hrvatski kao nasljedni jezik smješta u područje istraživanja kontaktne lingvistike.