

Commutazione di codice: approccio "usage-based"

Lukšić, Sonja

Master's thesis / Diplomski rad

2022

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Pula / Sveučilište Jurja Dobrile u Puli**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:137:454408>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-01-13**



Repository / Repozitorij:

[Digital Repository Juraj Dobrila University of Pula](#)



Sveučilište Jurja Dobrile u Puli
Università degli Studi Juraj Dobrila di Pola
Filozofski fakultet
Facoltà di Lettere e Filosofia

SONJA LUKŠIĆ
COMMUTAZIONE DI CODICE: APPROCCIO “*USAGE-BASED*”

DIPLOMSKI RAD / TESI DI LAUREA MAGISTRALE

Pula, 2022. / Pola, 2022

Sveučilište Jurja Dobrile u Puli
Università degli Studi Juraj Dobrila di Pola
Filozofski fakultet
Facoltà di Lettere e Filosofia

SONJA LUKŠIĆ
COMMUTAZIONE DI CODICE: APPROCCIO “*USAGE-BASED*”

DIPLOMSKI RAD / TESI DI LAUREA MAGISTRALE

JMBAG / N. MATRICOLA: 0303071825

Redoviti student / Studente regolare: Sonja Lukšić

Studijski smjer / Corso di laurea: Talijanski jezik i književnost / Lingua e letteratura italiana

Predmet / Materia: Strumenti computazionali per l'analisi linguistica

Znanstveno polje/ Area scientifico-disciplinare: Humanističke znanosti/ Studi umanistici

Mentorica / Relatrice: doc. dr. sc. Nada Poropat Jeletić

Pula, 24. lipnja 2022. / Pola, 24 giugno 2022

IZJAVA O AKADEMSKOJ ČESTITOSTI

Ja, dolje potpisana Sonja Lukšić, kandidatkinja za magistricu Talijanskog jezika i književnosti, ovime izjavljujem da je ovaj diplomski rad rezultat isključivo mogega vlastitog rada, da se temelji na mojim istraživanjima te da se oslanja na objavljenu literaturu kao što to pokazuju korištene bilješke i bibliografija. Izjavljujem da niti jedan dio Diplomskog rada nije napisan na nedozvoljeni način, odnosno da je prepisan iz kojega necitiranog rada, te da i koji dio rada krši bilo čija autorska prava. Izjavljujem, također, da nijedan dio rada nije iskorišten za koji drugi rad pri bilo kojoj drugoj visokoškolskoj, znanstvenoj ili radnoj ustanovi.

Studentica

Sonja Lukšić

Pula, 24.06.2022.

1. Indice	
2. Introduzione	1
3. Lingue in contatto	3
4. Situazione sociolinguistica istriana.....	7
4.1. Quadro storico e geografico.....	7
4.2. Repertori plurilingui: bilinguismo e diglossia	16
4.3. Repertorio italofono istriano.....	17
5. Modalità bilingue e commutazione di codice	24
5.1. Commutazione di codice a livello strutturale.....	26
5.2. Modelli teorici della commutazione di codice.....	31
5.2.1. Il modello della marcatura (<i>markedness model</i>)	32
5.2.2. Il modello della matrice (<i>matrix model</i>)	33
5.2.3. Il modello basato sull'utilizzo (<i>usage-based model</i>)	34
6. Postulati del modello <i>usage-based</i>	36
6.1. Frequenza	37
6.2. Frequenza e ripetizione nelle sequenze plurilessicali.....	38
6.3. Riduzione lessicale	39
6.4. Somiglianza	40
6.5. <i>Chunking</i>	41
6.6. Innovazioni attraverso la prospettiva cognitivista.....	42
6.7. Concettualizzazione	44
7. Strumenti computazionali per l'analisi linguistica e linguistica dei corpora	46
7.1. <i>TalkBank</i>	48
7.2. CHAT e CLAN	48
8. Metodologia della ricerca	52
8.1. Caratteristiche sociodemografiche e sociolinguistiche del campione	54
9. Applicazione del modello <i>usage-based</i> : livello strutturale.....	64
9.1. Dominio dei sintagmi sostantivali	65
9.2. Dominio dei sintagmi aggettivali	66
9.3. Dominio dei sintagmi preposizionali	70
9.4. Dominio dei sintagmi verbali.....	71
10. Applicazione del modello <i>usage-based</i> e realizzazione delle funzioni discorsive.....	73
10.1. Citazioni	73
10.2. Riferimento ad un concetto senza equivalente nella cultura dell'altra lingua.....	74
10.3. Allentamento di tensione e/o suscitazione di effetti umoristici, ironici o sarcastici	75

10.4.	Chiarimenti e riformulazioni	76
10.5.	Aggiunte, connessioni e prese di turno	78
10.6.	Interiezioni	79
10.7.	Sequenze fisse o formulaiche	79
10.8.	Restrizione del messaggio, riduzione fonetica o lessicale	80
10.9.	Rafforzamento di una richiesta	81
11.	Conclusione.....	82
12.	Bibliografia	86
13.	Sitografia.....	97
14.	Allegato 1	99
15.	Allegato 2 - Questionario sociolinguistico somministrato	101
16.	Allegato 3 – Consenso	106
17.	Riassunto.....	108
18.	Summary.....	109
19.	Sažetak.....	110

2. Introduzione

La diversità etnica, culturale e linguistica degli ambienti multiculturali e multilingui implica complessi rapporti inerenti alle questioni di contatto culturale e linguistico, rappresentando uno dei fattori chiave che forgia gli atteggiamenti individuali e collettivi nei confronti delle singole varietà linguistiche e degli usi dello spettro comunicativo funzionale. Il territorio istriano è stato il punto d'incontro e di fusione di numerosi popoli e delle loro culture e lingue lungo tutto l'arco della sua ricchissima storia. Il repertorio plurilingue istriano è contrassegnato da complessi rapporti doppiamente diglossici/poliglossici, nonché dalla prassi abituale dell'uso della commutazione di codice. Quest'ultima rappresenta il cardine del presente lavoro di ricerca, articolato in tre sezioni. Nella prima sezione si metteranno in risalto i momenti più importanti di tipo diacronico che hanno contribuito alla formazione della situazione sociolinguistica odierna insieme ai tratti sincronici che caratterizzano la complessità del suo spettro fenomenologico. In seguito, verranno spiegati alcuni dei più rilevanti approcci allo studio del fenomeno della commutazione di codice, ovvero i tre approcci più usati e più noti, quali il modello della matrice (*matrix model*), il modello della marcatura (*markedness model*), e il modello basato sull'utilizzo (*usage-based model*). In questa sede si è deciso di applicare il modello *usage-based*. La scelta non è stata fatta in base alla qualità o all'efficacia del modello scelto, ma semplicemente perché si tratta del modello in assoluto più recente (che è considerato uno dei più fruttuosi approcci, la cui applicazione vede recentemente una rapida espansione), e di conseguenza finora meno esplorato, ovvero del tutto inesplorato per quanto concerne la realtà linguistica istriana. La parte analitica consisterà pertanto nella presentazione dei risultati di un'indagine sperimentale pionieristica focalizzata sullo studio del fenomeno della commutazione di codice, come parte integrante del corpus del parlato spontaneo di una novantina di parlanti italofoeni istriani. Infatti, il corpus analizzato in questa sede per gli obiettivi della ricerca è stato compilato seguendo i tre criteri chiave delineati da Tognini-Bonelli (2001: 54), ovvero l'autenticità, la rappresentatività e i criteri per il campionamento. L'autenticità del linguaggio del suddetto corpus è garantita dalla spontaneità e naturalezza con cui le persone interagiscono nelle conversazioni, le quali sono state registrate e in seguito trascritte e codificate, mentre la rappresentatività è data dal fatto che i partecipanti coinvolti appartengono alla comunità italofoena istriana, coinvolgendo tutte le aree bilingui della penisola. Infine, i criteri utilizzati per la campionatura della lingua sono conformi ai requisiti necessari per la raccolta, trascrizione e codificazione di dati.

Utilizzando gli strumenti computazionali a nostra disposizione, sono stati in seguito esaminati alcuni dei meccanismi funzionali delineati da Gumperz (1982) e Baker (2010), analizzabili mediante i postulati teorici del modello *usage-based*. Tale modello prevede che le unità siano legate insieme sulla base di co-occorrenza, somiglianza e sovrapposizione o giustapposizione attraverso i processi cognitivi generali di categorizzazione, generalizzazione e schematizzazione. Pertanto, in seguito alla presentazione dell'apparato metodologico adottato (seconda sezione del lavoro), si passerà all'applicazione del suddetto modello a livello strutturale e funzionale, evidenziando e riflettendo sui tratti che appaiono con maggior frequenza nel corpus analizzato (terza sezione).

3. Lingue in contatto

La lingua, oltre ad essere lo strumento di interazione *par excellence*, è il simbolo dell'identità sociale e allo stesso tempo rappresenta un serbatoio di conoscenze culturali (Gumperz, 1974: 786). La comunità sociale, formata da intrecci continui di relazioni, ha il compito di trasmettere di generazione in generazione, di casa in casa, la lingua o le lingue della comunità facendo sì che i comportamenti linguistici vengano accomunati ad attributi personali e identitari. Secondo Thomason e Kaufman (Auer, 2020: 147) il contatto linguistico¹ che si genera in tal modo è condizionato dai fattori sociali, culturali, politici e/o economici che abilitano, incoraggiano e/o forzano i parlanti di lingue diverse a comunicare tra loro (Auer, 2020: 147). I parametri grammaticali (come le risorse linguistiche a disposizione dei parlanti, la sovrapposizione strutturale tra grammatiche e vocabolari, la quantità di variazioni all'interno delle lingue, ecc.) possono avere un impatto sulla quantità e qualità del contatto linguistico. La diversità linguistica è una risorsa culturale di indiscutibile valore che bisogna preservare e il contatto tra lingue è sicuramente un campo di studio di particolare rilevanza in un contesto sociolinguistico composito come quello istriano. Inoltre, esso è caratterizzato dalla convivenza e simbiosi di lingue e culture, ossia da una situazione di bilinguismo/multilinguismo, nonché da relazioni composite di diglossia/poliglossia (vedi capitolo successivo). Secondo il dizionario Treccani, infatti, il bilinguismo rappresenta:²

“la capacità che ha un individuo, o un gruppo etnico, di usare alternativamente e senza difficoltà due diverse lingue (o anche, per estensione, due diverse varietà di una lingua, o la lingua letteraria e il dialetto).”

Da un punto di vista sociolinguistico, il bilinguismo è associato a contesti diversi da quelli strettamente attinenti alla lingua e alla società (alle quali l'individuo è indissolubilmente legato), ovvero ai processi di formazione, non solo di identità collettive ma anche individuali. Berruto (1999: 3-36) spiega che esso rappresenta la situazione in cui due lingue coesistono nella stessa comunità sociale ed entrambe vengono usate senza subordinazione funzionale, sia nel parlato che nello scritto. Tuttavia, gli individui, affinché siano classificati come bilingui, è

¹ Weinreich, nel suo libro *"Le lingue in contatto"* (1963; 1), afferma che due o più lingue sono in contatto se vengono usate alternativamente dalle stesse persone. Thomason (2001: 1) definisce il contatto linguistico come l'uso di più di una lingua nello stesso luogo contemporaneamente. Esiste un consenso generale sul fatto che il contatto linguistico presupponga non solo l'uso della lingua, ma anche un certo tipo di influenza (che può essere unidirezionale o reciproca) delle rappresentazioni cognitive dei parlanti e del loro comportamento. Myers-Scotton (2002:5), infatti, spiega che le lingue sono in contatto nel senso che si trovano l'una accanto all'altra nel lessico mentale dei parlanti e di conseguenza possono esercitare una mutua interferenza.

² <https://www.treccani.it/vocabolario/bilinguismo/> - ultima consultazione il 24/4/2022.

sufficiente che abbiano una competenza minima in due lingue, ma possono avere altresì una competenza molto avanzata. Infatti, Bogliun Debeljuh (2019: 8) individua i tre principali tipi di bilinguismo individuale connessi a livello di conoscenza delle due lingue:

- a) coordinato e complesso – nel bilinguismo coordinato i due sistemi semantici sono indipendenti, mentre in quello complesso non sono indistinguibili;
- b) bilanciato e dominante – la persona bilingue bilanciata ha lo stesso livello di conoscenze delle due lingue, mentre quella dominante padroneggia meglio una delle lingue;
- c) additivo e sottrattivo – nel bilinguismo additivo le nuove acquisizioni linguistiche si aggiungono a quelle vecchie, mentre in quello sottrattivo le nuove acquisizioni linguistiche in una delle lingue vanno progressivamente a scapito dell'altra lingua.

In riferimento al grado di padronanza delle lingue, Milani-Kruljac (1990) differenzia il bilinguismo ricettivo, riproduttivo e quello attivo (si riferisce alla padronanza sia scritta che orale della L2) e passivo (si riferisce alla padronanza parziale della L2). Le competenze in questo senso riguardano: l'ascolto, la lettura, la scrittura oppure la padronanza integrale. Inoltre, se le due lingue sono state acquisite simultaneamente, si parla di bilinguismo simultaneo. Quest'ultima tipologia si contrappone al cosiddetto bilinguismo sequenziale (quando la seconda lingua è stata acquisita dopo aver acquisito la prima).

Dal Negro e Guerini (2007: 18) definiscono il bilinguismo sociale come il bilinguismo che riguarda la comunità dove coesistono due codici, entrambi usati in tutti i settori amministrativi, a livello sia formale che informale. Questo tipo di bilinguismo può ulteriormente essere suddiviso in base al numero delle comunità:

- a) bicomunitario: due comunità condividono un territorio;
- b) monocomunitario: prevede la presenza di una sola comunità bilingue.

Per quanto concerne l'apparato legale, si distingue il bilinguismo *de iure* (ambedue i codici sono riconosciuti legalmente) e quello *de facto* (un codice realmente diffuso e parlato nella comunità). Il bilinguismo croato-italiano si realizza *de facto* a livello sociale e istituzionale ed è riconosciuto *de iure* dal 1994. Infatti, lo Statuto della Regione Istriana, approvato dall'Assemblea della Regione Istriana il 30 marzo 1994, prescrive speciali forme di tutela per la lingua ufficiale della Comunità Nazionale Italiana, ossia la lingua letteraria/standard italiana (Blagoni et al., 2016: 5).

Esistono tre tipologie di bilinguismo nel contesto del contatto linguistico (Dal Negro e Guerini, 2007):

1. situazioni di bilinguismo sostanzialmente stabile;
2. varietà linguistiche autonome;
3. isole linguistiche (in cui la varietà linguistica è usata per secoli senza avere nessun contatto con la corrispondente lingua standard o dialetti simili, rimanendo perciò isolata e circondata da parlate allogene).³

Le lingue in contatto hanno come conseguenza inevitabile l'insorgere di vari mutamenti reciproci che possono assumere forme e dimensioni diverse. Thomason e Kaufman (1988: 74-76) sostengono che nelle lingue possono verificarsi diversi tipi di cambiamenti a seconda dell'intensità e dei tipi di contatto linguistico. Gli autori propongono la seguente "scala di prestito" (*borrowing scale*), che illustra la correlazione tra l'intensità del contatto linguistico e la forma di cambio di lingua che si verifica in determinate situazioni.

Tabella 1. Scala di prestito

<p>1. Nelle situazioni di contatto casuale il bilinguismo non è prominente. I prestiti riguardano il vocabolario non di base, i sostantivi, alcuni verbi, aggettivi e avverbi.</p>
<p>2. Nelle situazioni di contatto leggermente più intenso la presenza del bilinguismo è minima. I prestiti riguardano alcune strutture, ma nulla che possa alterare i tipi di struttura già presenti nella lingua del prestito.</p>
<p>3. Nelle situazioni di contatto più intenso il bilinguismo è presente in maggior misura e l'uso dei prestiti è favorito dai fattori sociali. I prestiti riguardano anche il vocabolario di base e il cambiamento strutturale moderato (es. l'ordine delle parole, alcuni affissi, ecc.).</p>
<p>4. Nelle situazioni di contatto intenso il bilinguismo è molto esteso e i fattori sociali favoriscono fortemente il cambiamento linguistico. I prestiti riguardano il vocabolario di ogni tipo: il cambiamento strutturale è forte in ambedue le lingue.</p>

³ I criteri di definizione dell'isola linguistica sono l'isolamento territoriale, la diversità linguistica e la subordinazione sociopolitica, un certo grado di continuità territoriale e coesione interna. Nei casi in cui una minoranza è distribuita in relazione alla lingua maggioritaria "a macchia di leopardo", si tratta di isole linguistiche discontinue (Blagoni, 2011: 19).

È possibile concludere perciò che maggiore è l'intensità del contatto linguistico, maggiore è la probabilità del cambiamento linguistico, che incide sul numero dei tipi di cambiamento. È stato notato che nelle situazioni in cui il bilinguismo è inesistente o minimo e la maggior parte dei parlanti è monolingue, può verificarsi il cambiamento lessicale (ma non quello strutturale); mentre in situazioni in cui il bilinguismo rappresenta la norma, i cambiamenti sia a livello lessicale sia a livello strutturale sono più frequenti. Il terzo e il quarto livello della tabella rappresentano situazioni di bilinguismo a seconda del grado di moderazione/estensione tra i parlanti. Nel caso istriano, anche se la lingua italiana e quella croata appartengono a due famiglie linguistiche diverse e sono molto differenti, si verifica un contatto linguistico molto intenso soprattutto nell'ambito della comunità italoфона in Istria (visto che i codici italoфoni appartengono all'universo minoritario), perciò tale situazione può essere associata al quarto tipo della scala di Thomason e Kaufman (1988) dove ci si può aspettare un qualsiasi tipo di interferenza linguistica.

4. Situazione sociolinguistica istriana

4.1. Quadro storico e geografico

Grazie alla sua posizione geografica favorevole, la penisola istriana rappresentava da sempre l'intersezione delle rotte commerciali e marittime nel Mediterraneo e formava un collegamento importante con la parte continentale dell'Europa (cfr. Matijašić, 2005). La realtà culturale istriana è storicamente fondata sulle dinamiche millenarie che sono la ragione della convivenza e della permanenza di numerose comunità culturali, etniche e linguistiche, presenti tutt'oggi proprio come lo erano nei secoli passati. L'Istria, anche se può essere percepita geograficamente e storicamente come unitaria, la sua integrità culturale e linguistica invece non presenta le caratteristiche uniformi, ma fonda la propria legittimità nella diversità (Blagoni et al., 2016: 2). La concezione culturalmente e linguisticamente pluralistica di una società democratica implica il riconoscimento di relazioni egualitari di tutte le comunità presenti su un territorio di convivenza a livello di etnia, cultura, nazionalità, lingua, religione, ecc. (Blagoni et al. 2016: 4). Pertanto, l'Istria, per la sua eterogenea composizione etnolinguistica, si può considerare pluriethnica, multilingue e multiculturale. Quando si parla di quest'ultimi, ci si riferisce *in primis* alla convivenza della popolazione croata (maggioritaria) e quella italiana (minoritaria), cioè ai due gruppi etnici che sono in Istria da così tanto tempo da poter essere considerati autoctoni (Banovac, 1998; Orbanić, 2005: 341-373).

La Regione Istriana comprende il territorio del triangolo situato nella parte nord-orientale del Mare Adriatico. La penisola più grande dell'Adriatico è circondata dal mare, mentre il confine settentrionale è disteso tra la Baia di Muggia e il Golfo di Pre luka. La superficie istriana comprende 3.476 chilometri quadrati, dei quali il 90% appartiene alla Repubblica di Croazia e la posizione favorevole è data dal fatto che la regione confina con l'Italia, la Slovenia e con il restante territorio della Croazia. Per capire la situazione sociolinguistica odierna dell'Istria, bisogna ripercorrere le tappe diacroniche più rilevanti e ricordare le ragioni che hanno portato alla situazione attuale.

La penisola deve il suo nome al popolo degli Istri, che la abitavano nell'età del ferro e rappresentava da sempre oggetto di grande interesse grazie alla notevole importanza strategica come punto d'incontro tra vari popoli e culture. Anche se le primissime tracce di vita risalgono alla preistoria, in questa sede, a scopo della presente analisi, si prenderà come punto di partenza l'apporto del dominio romano, che ebbe un ruolo importante nel delineamento della genesi identitaria istriana. La romanizzazione in Istria è intesa e interpretata in modi diversi, e viene

spesso interpretata come un processo di acculturazione inteso come processo in cui una cultura non cancella quella precedente, ma la accoglie creando così nuove forme di convivenza (Matijašić, 1993: 576).

Il primo periodo della romanizzazione risale alla sconfitta degli Istri nel 177 a.C., quando i Romani si stabiliscono nella regione e creano le "*Élite locali*", formate da popolazioni autoctone che avevano adottato le usanze romane e l'uso del latino (Ivetić et al., 2009; Matijašić, 1993). La seconda fase della romanizzazione si considera il processo della costituzione di colonie nell'estremo nord-est d'Italia, necessarie per assicurare un rafforzamento della posizione di Roma sotto continua minaccia. La Colonia Iulia Pola viene fondata intorno al 44 a.C., mentre Parentium e Agida acquisiscono lo status di municipalità. Il processo di romanizzazione era graduale e nel momento in cui l'Istria entrò a far parte della Decima Regione Italiana⁴ quasi tutta la penisola era già in gran parte romanizzata. La fondazione di Aquileia (181 a.C.) e delle successive colonie di Tergeste (Trieste) e di Pola, come nuovi centri della civiltà romana, posero le basi per la romanizzazione più intensa dell'intera Italia nord-orientale. Questo consolidamento di potere sul territorio porta verso il periodo incentrato sull'urbanizzazione, colonizzazione agraria, alfabetizzazione ed introduzione della spiritualità culturale antica, caratterizzato dalla tranquillità e benessere durante il quale gli abitanti dell'Istria vengono eguagliati nei diritti agli abitanti di Roma. Segue la fase della prosperità economica in cui fioriscono l'arte e la cultura romana, che vede la comparsa di ville urbane e suburbane, dei templi e fori, degli edifici pubblici e delle "*villae rusticae*". Il V secolo è segnato da profonde crisi che interessano l'intero Impero Romano d'Occidente a causa delle incursioni barbariche e dei problemi interni, che portano nel 476 alla disgregazione dell'Impero. L'Istria è testimone del susseguirsi di diversi domini in poco tempo e nel corso del VII secolo inizia il graduale processo della pacifica slavizzazione del territorio.

Segue il periodo del feudalesimo in cui gli abitanti delle città che non avevano terre gradualmente impoveriscono, mentre il commercio non è ancora abbastanza sviluppato da poter sostenere l'economia della regione. Con la riunione denominata Placito del Risano⁵ cambiano i rapporti sociali provocando scontri tra la popolazione slava e quella latina, e di

⁴ La Regio Decima, Veneto et Histria comprendeva una vasta area dell'Italia nord-orientale corrispondente al territorio dell'attuale Veneto, Trentino-Alto Adige, Friuli-Venezia Giulia, al quale si aggiungevano a ovest una parte dell'odierna Lombardia e a est la penisola dell'Istria (https://www.treccani.it/enciclopedia/l-italia-romana-delle-regiones-regio-x-venetia-et-histria_%28II-Mondo-dell%27Archeologia%29/ - ultima consultazione il 16/5/2022).

⁵ Tenutasi sul fiume Risano nei pressi di Capodistria.

conseguenza cambia la struttura etnica del territorio. Dal X secolo l'Istria diventa parte del Sacro Romano Impero che, avendo il potere centrale debole, spinge il rafforzarsi del ceto nobile, la creazione dei principati privati e delle signorie. Questa situazione porta alcune città istriane a liberarsi dei domini feudali e ad instaurare l'autogoverno. La fascia costiera cade sotto il dominio della Repubblica di Venezia e si instaura il rapporto basato sui contratti che garantivano alle città istriane la libertà di commercio e la protezione della Serenissima in cambio dell'aiuto militare. L'entroterra invece appartiene all'Impero Asburgico e nel corso del XVI e del XVII secolo le due fasce entrano spesso in conflitto per ragioni di confini, bestiame, terre coltivabili. Alle problematiche esistenti apportano pure le epidemie, le siccità e le calamità naturali.

Figura 1. I confini in Istria nel 1500:⁶



⁶ Istria, Fiume e Dalmazia, cenni storici (https://www.scuolaconfineorientale.it/2021/archivio/seminario/Istria_Fiume_e_Dalmazia.pdf - ultima consultazione il 18/5/2022).

Il trattato di Campoformio del 1797 costituisce la scomparsa della Repubblica di Venezia e i territori veneti insieme all'Istria e alla Dalmazia passano all'Austria.⁷ Tra il 1805 e il 1813 l'Istria si trova sotto il dominio francese e con la sconfitta di Napoleone torna a far nuovamente parte dell'Impero Austriaco. Il XVIII secolo vede un incremento importante della popolazione e con esso il rafforzamento della nazionalità croata. Con la disgregazione della Monarchia austro-ungarica nel 1918, inizia il periodo del dominio italiano in Istria. Gli Accordi di Rapallo del 1920 fanno sì che il territorio istriano venga annesso al Regno d'Italia e solo due anni più tardi prende potere il regime fascista promosso da Benito Mussolini. Nel 1943 la capitolazione dell'Italia nella Seconda guerra mondiale, detto anche l'armistizio di Cassibile, fu il primo passo verso l'unificazione dell'Istria con la Croazia come parte della Jugoslavia, e una delle conseguenze fu la grande ondata migratoria dal 1947 in poi. Contemporaneamente, la popolazione da molte aree dell'ex Jugoslavia immigra nella penisola e la regione deve affrontare nuovamente il cambio di potere. Gli italiani rimasti diventano la minoranza all'interno della Jugoslavia e il loro status di *élite* con potere politico, economico, culturale e sociale viene ridotto drasticamente. L'esodo della popolazione italiana dell'Istria, Fiume e Zara è stato un lungo processo durato dal 1947 al 1956. A prescindere dalle cifre che variano, a seconda dell'interpretazione, l'esodo ha un enorme significato storico perché è un momento di capovolgimento nella storia dell'Istria che ha visto scomparire quasi del tutto un'intera nazione in pochi anni.

Negli anni '60, la penisola conosce il *boom* turistico e diventa una delle regioni più sviluppate della Jugoslavia. Parallelamente, Pola diventa una città militare importante, e perciò viene sottoposta a nuovi movimenti migratori. A Pola arrivano gli ufficiali dell'esercito popolare jugoslavo provenienti da tutte le parti della Jugoslavia con le loro famiglie, e quindi ancora una volta la migrazione influenza fortemente la creazione dell'identità multiculturale della città (Jernej-Pulić, 2016: 12). Pola diventa così una città croata, anche se i croati rappresentavano poco più della metà della popolazione. Gli altri abitanti erano membri della minoranza italiana e una parte eterogenea di varie nazionalità (serbi, bosniaci, ungheresi, albanesi, macedoni, sloveni, rom). Il discorso sul multiculturalismo istriano, come uno dei tratti distintivi diventa particolarmente attuale negli anni '90, che videro un'altra volta la guerra bussare alla porta. Meritevolmente, Pola è stata da molti e giustamente definita "*la Jugoslavia in poche parole*", perché raccoglie tutta la ricchezza etnica e linguistica Jugoslava (Orbanić, 2005: 341-373).

⁷ Colours of Istria (<https://www.coloursofistria.com/it/destinazioni/istria/storia> - ultima consultazione il 16/5/2022).

Milani-Kruljac (1990) invece, nel descrivere la penisola istriana usa il termine “*melting pot*“, il termine che mette perfettamente in evidenza l'importanza della valorizzazione delle diverse identità.

Negli anni '90 dello scorso secolo la presenza degli italiani sul territorio istriano è soprattutto visibile in alcune zone come quella occidentale e meridionale della penisola. Amministrativamente gli italiani sono compresi in tre comuni del territorio istriano che fanno parte della repubblica di Slovenia (Capodistria, Pirano ed Isola) e in dodici città della Repubblica di Croazia (Umago, Buie, Cittanova, Pinguente, Parenzo, Pisino, Rovigno, Pola, Albona, Abbazia, Fiume, Cherso, Lussino) (Milani-Kruljac, 1990). I dati del censimento del 1991 mostrano il raddoppiamento della popolazione italiana in Croazia favorita dalle riforme democratiche e dalla dissoluzione della Repubblica di Jugoslavia. Aumenta il numero delle iscrizioni nelle Comunità degli Italiani e questa svolta contribuisce all'equilibrio demografico della nazionalità.

Tabella 2. Italiani, istriani, jugoslavi e non dichiarati nazionalmente in Istria e a Fiume (1981- 1991) (Giuricin, 2011: 167)

Area	Anno	Totale popolazione	Italiani	Istriani	Jugoslavi	Non dichiarati
Regione istriana	1981	188.292	7.726 (4,1%)	3.619 (1,9%)	21.800 (11,6%)	915 (0,5%)
	1991	204.346	15.306 (7,5%)	37.027 (18,1%)	7.301 (3,6%)	6.014 (2,9%)
Capodistriano	1981	69.591	1.901 (2,7%)	283 (0,4%)	2.942 (4,2%)	373 (0,5%)
	1991	75.929	2.751 (3,8%)	1.854 (2,5%)	1.026 (1,4%)	913 (1,2%)
Regione di Fiume	1981	234.756	2.217 (0,9%)	671 (0,2%)	35.643 (15,1%)	1.154 (0,5%)
	1991	250.846	3.938 (1,6%)	2.358 (0,9%)	9.403 (3,7%)	9.328 (3,7%)
Istria e Fiume	1981	492.639	11.844 (2,4%)	4.573 (0,9%)	60.385 (12,2%)	2.442 (0,5%)
	1991	531.121	21.995 (4,14%)	41.239 (7,7%)	17.730 (3,3%)	16.255 (3%)

Tabella 3. Variazioni del numero di italiani nei censimenti del 1991, 1981, 1971 e 1961 (per aree regionali) (Giuricin, 2011: 169)

Aree	1991	1991-1981	%	1981	1981-1971	%	1971	1971-1961	%	1961
Regione istriana	15.306	+7.580	+98	7.726	-3.776	-32,8	11.502	-2.852	-19,8	14.354
Capodistriano	2.751	+850	+44,7	1.901	-667	-25,9	2.568	+21	+0,8	2.547
Regione di Fiume	3.938	+1.721	+77,6	2.217	-1.229	-35,4	3.446	-355	-9,8	3.801

Il censimento del 2001 mostra un calo importante di tutte le minoranze nazionali dovuto alla pulizia etnica e per la prima volta la Croazia diventa nazionalmente un paese piuttosto omogeneo.⁸

Tabella 4. Popolazione complessiva, nazionalità croata, italiana, serba e appartenenza regionale in Croazia - censimento del 2001 (Giuricin, 2011: 170)

Area	Popolazione complessiva	Croati	Italiani	Serbi	App. regionale
Croazia	4.437.460	3.977.171	19.636	201.631	9.302
Regione Istriana	206.344	148.328	14.284	6.013	8.865
Regione Fiumana*	305.505	258.438	3.539	15.005	150

I seguenti dati mostrano il decremento numerico degli italiani in Croazia dal 1991 al 2001.

Tabella 5. Cittadini di nazionalità italiana in Croazia (1991-2001)

Anno	Nazionalità italiana	Variazioni percentuali
1991	21.303	
2001	19.636	- 7,8%

Al Censimento del 2011⁹, su un totale di 4.290.621 cittadini croati, 17.807 dichiarano l'appartenenza alla nazionalità italiana,¹⁰ di cui 12.543 risiedono nella Regione Istriana. Alla luce dei dati riportati è possibile osservare che la crescita demografica della popolazione appartenente alla nazionalità italiana raggiunge il picco nel 1991, dopodiché segue un decremento costante fino all'anno 2011.

Sebbene dichiarativamente la Regione Istriana sia una regione bilingue, la realtà istriana è di fatto multilingue e pluri-etnica. La ricca diversità linguistica ed etnica si può osservare nei risultati del censimento della popolazione avvenuto nel 2011¹¹, in relazione all'appartenenza nazionale dichiarata che mostra i seguenti risultati: croati (68,33 %), sloveni (0,86%), serbi (3,46%), italiani (6,03 %), bosniaci (2,95%), albanesi (1,15%), rom (0,41%), montenegrini (0,36%), macedoni (0,24%), ungheresi (0,23%), polacchi (0,03%) e rumeni (0,02%). Secondo il censimento le lingue materne dichiarate dei residenti istriani sono le seguenti: croato

⁸ Nel Censimento del 2001 l'89,6% della popolazione dichiara di appartenere alla nazionalità Croata.

⁹ I dati riguardanti il censimento della popolazione sono stati tratti dal portale del "Državni zavod za statistiku Republike Hrvatske" (<http://www.dzs.hr/> - ultima consultazione il 18/5/2022).

¹⁰ Nazionalità e cittadinanza non coincidono nei paesi balcanici: es. essere cittadino croato è un aspetto giuridico-amministrativo (stato civile), mentre la nazionalità dipende da un aspetto "affettivo" e si adotta per libera scelta.

¹¹ I dati riguardanti il censimento della popolazione sono stati tratti dal portale del "Državni zavod za statistiku Republike Hrvatske" (<http://www.dzs.hr/> - ultima consultazione il 18/5/2022).

(86,78%) e italiano (6,83%), serbo (0,93%) e serbo-croato (0,06%), sloveno (0,79%), bosniaco (1,38%), albanese (1,02%), macedone (0,21%), ungherese (0,18%), polacco (0,14%), montenegrino (0,05%) e la lingua dei rom (0,03%).

Lo Statuto della Regione Istriana¹² (2003) prescrive l'ufficialità della lingua croata e di quella italiana nelle istituzioni della regione e a livello degli Statuti delle unità locali di autogoverno con denominazione bilingue. Sono in tutto sette le città e tredici i comuni in cui viene applicato il bilinguismo su tutto o su una parte del territorio.¹³ Le seguenti tabelle indicano il numero degli abitanti appartenenti alla nazionalità italiana nei comuni e nelle città ufficialmente bilingui:

Tabella 6. Popolazione di nazionalità italiana nelle città ufficialmente bilingui (2001)

CITTÀ	NUMERO DI ABITANTI	PERCENTUALE (%)
Buje-Buie	1 261	24,33 %
Novigrad-Cittanova	443	10,2 %
Poreč-Parenzo	540	3,23 %
Pula-Pola	2 545	4,43 %
Rovinj-Rovigno	1 608	11,25 %
Umag-Umago	1 962	14,57 %
Vodnjan-Dignano	1 017	16,62 %

¹² La seduta si è tenuta il 9 aprile 2001; il testo emendato è disponibile in rete al seguente indirizzo web: http://unione-italiana.eu/Backup/documents/Statuto_Regione_Istriana.pdf - ultima consultazione il 13/5/2022.

¹³ Le unità d'autogoverno locale con denominazione bilingue nella Regione Istriana sono: Buje-Buie, Novigrad-Cittanova, Poreč-Parenzo, Pula-Pola, Rovinj-Rovigno, Umag-Umago, Vodnjan-Dignano, Bale-Valle, Brtonigla-Verteneglio, Fažana-Fasana, Funatna-Fontane, Grožnjan-Grisignana, Kaštelir Labinci-Castelliere S. Domenica, Ližnjan-Lisignano, Motovun-Montona, Oprtalj-Portole, Tar Vabriga-Torre Abrega, Višnjan-Visignano, Vižinada-Visinada, Vrsar-Orsera (Consiglio della minoranza nazionale italiana autoctona, 2017).

Tabella 7. Popolazione di nazionalità italiana nei comuni (2001)

COMUNI	NUMERO DI ABITANTI	PERCENTUALE (%)
Bale-Valle	260	23,07 %
Brtonigla-Verteneglio	490	30,14 %
Fažana-Fasana	173	4,76 %
Funtana-Fontane	16	1,76 %
Grožnjan-Grisignana	290	39,4 %
Kaštelir Labinci-Castelliere-S. Domenica	70	4,78 %
Ližnjan-Lisignano	168	4,24 %
Tar Vabriga-Torre Abrega	195	9,80%
Motovun-Montona	98	9,76%
Oprtalj-Portole	122	14,35%
Višnjan-Visignano	155	6,82%
Vižinada-Visinada	88	7,60%
Vrsar-Orsera	59	2,73%

Anche se, come ben risaputo, la Comunità Nazionale Italiana è una comunità etnoculturale “*sui generis*” difficilmente definibile in termini giuridici, politici e socio-antropologici (Blagoni et al., 2016: 5), essa comunque gode del riconoscimento da parte del diritto croato ed è prioritariamente determinata all’autoconservazione delle proprie peculiarità distintive e contraria, pertanto, a una politica di assimilazione perseguitata dalla maggioranza. Di seguito si riporta la raccolta delle disposizioni che possono essere considerate rappresentative nella tutela e nella promozione del bilinguismo nella regione presenti nello Statuto della Regione Istriana:¹⁴

Articolo 6: stabilisce l'uguaglianza della lingua croata e italiana nell'uso ufficiale delle istituzioni regionali nelle unità locali di autogoverno secondo le modalità stabilite dallo Statuto.

Articolo 21: prevede l'uso ufficiale delle due lingue nei procedimenti appartenenti agli ambiti autogovernativi e il riconoscimento della validità dei documenti giuridici privati anche quando sono redatti in lingua italiana.

¹⁴ Statuto della Regione Istriana (<https://www.istra-istria.hr/it/regione-istriana/documenti/statuto-della-regione-istriana/> - ultima consultazione il 16/5/2022).

Articolo 22: prevede lo svolgimento del lavoro dell'Assemblea regionale e della Giunta in lingua croata e in lingua italiana. Assicura il bilinguismo negli atti degli organismi rappresentativi, esecutivi ed amministrativi della Regione e delle persone giuridiche autorizzate ed assicura il recapito del materiale bilingue per le sedute dell'Assemblea e della Giunta regionale, le stesure dei verbali e le pubblicazioni delle conclusioni, gli avvisi ufficiali e gli inviti degli organismi rappresentativi, esecutivi ed amministrativi della Regione, come il materiale per le sedute dell'organismo rappresentativo ed esecutivo.

Articolo 23: si riferisce all'equiparazione ufficiale delle due lingue in parte o su tutto il territorio dei comuni e delle città della Regione Istriana nei quali risiedono gli appartenenti alla Comunità Nazionale Italiana, conformemente ai loro statuti.

Articolo 24: garantisce agli appartenenti alla comunità italiana il diritto d'uso pubblico della propria lingua e scrittura e il diritto alla salvaguardia dell'identità nazionale e culturale. Garantisce inoltre il diritto di organizzare liberamente le proprie attività informative ed editoriali e il diritto di mettere in rilievo le caratteristiche nazionali. Consente l'esposizione della bandiera della Comunità Nazionale Italiana in occasioni solenni e nelle sedute degli organi di rappresentanza.

Articolo 25: prevede l'assunzione di un adeguato numero di dipendenti che possiedono una buona conoscenza della lingua croata e della lingua italiana. Le persone giuridiche che hanno autorizzazioni pubbliche hanno la facoltà di usare la lingua italiana nelle procedure per conto degli appartenenti della comunità italiana.

Articolo 26: assicura il rilascio di documenti e moduli bilingui usati ufficialmente.

Articolo 27: stabilisce che nei territori in cui lo Statuto ha prescritto il bilinguismo verrà assicurato e particolarmente incoraggiato l'insegnamento della lingua italiana quale lingua dell'ambiente sociale.

Articolo 29: prevede il riconoscimento del ruolo e del valore dell'Unione Italiana - quale rappresentante degli appartenenti alla Comunità Nazionale Italiana per la quale assicura i mezzi finanziari nel suo Bilancio, nell'ambito delle proprie possibilità.

4.2. Repertori plurilingui: bilinguismo e diglossia

Per spiegare il concetto del repertorio linguistico bisogna prima spiegare quello della comunità linguistica, i cui membri condividono un certo repertorio. Per Labov (1973: 341) la comunità linguistica sono le persone che condividono un insieme di norme linguistiche e aspettative riguardanti l'uso della lingua condividendo un territorio. Dal Negro (2007: 17) vede il repertorio linguistico come l'insieme delle varietà linguistiche utilizzate in una comunità e questo insieme comprende vari registri, dialetti, stili, accenti, ecc. che esistono in questa comunità.

I repertori linguistici si trovano in rapporti strutturati gerarchicamente sia a livello di standardizzazione sia a livello delle funzioni assegnate. La consapevolezza che gli individui o le comunità potrebbero possedere più di una varietà linguistica ha fatto emergere i termini come "bilinguismo", "trilinguismo", "plurilinguismo" e naturalmente anche "multilinguismo" nel riferimento alla pluralità.

Di seguito saranno descritti i repertori più importanti che sono presenti sul territorio della penisola istriana, ossia il bilinguismo e la diglossia.

Secondo Ferguson (1959) la definizione di diglossia è la seguente:

Diglossia is a relatively stable language situation in which, in addition to the primary dialects of the language (which may include a standard or regional standards), there is a very divergent, highly codified (often grammatically more complex) superposed variety, the vehicle of a large and respected body of written literature, either of an earlier period or in another speech community, which is learned largely by formal education and is used for most written and formal spoken purposes but is not used by any sector of the community for ordinary conversation (Wei, 2000: 75).

Possiamo riassumere il passaggio citato del sociolinguista Charles Ferguson che ha definito la diglossia come una situazione linguistica relativamente stabile in cui, oltre al dialetto che viene usato nelle situazioni informali, c'è una varietà sovrapposta, spesso grammaticalmente più complessa, usata nella letteratura, che viene appresa in gran parte attraverso l'istruzione formale. In altre parole, la diglossia riflette il rapporto verticale tra due varietà linguistiche di cui una è superiore all'altra, come ad esempio il dialetto istroveneto e la lingua italiana, o tra il dialetto ciacavo e la lingua croata. Questo rapporto si differenzia da quello orizzontale che

caratterizza il bilinguismo e che si può trovare nell'esempio della lingua italiana e la lingua croata.¹⁵

Fishman poi nel 1967 aggiunge a questa teoria quattro scenari possibili che mettono in relazione i concetti della diglossia e del bilinguismo (Dal Negro, 2007: 24):

- a) diglossia con bilinguismo (come nel caso dell'italiano e i dialetti italiani)
- b) bilinguismo senza diglossia (il bilinguismo vero cioè tra lingue "alla pari")
- c) diglossia senza bilinguismo (in cui le due varietà sono utilizzate da gruppi della società fortemente separati e non comunicanti)
- d) assenza di entrambi nelle società monolingui isolate

Fishman, inoltre individua tre principali differenze tra bilinguismo e diglossia (Milani-Kruljac, 1990: 46):

1. nel bilinguismo le lingue si trovano sullo stesso livello mentre nella diglossia il rapporto è gerarchico;
2. nella diglossia viene rispecchiata una situazione sociale, mentre nel bilinguismo viene espressa la situazione individuale;
3. nella diglossia l'accento è posto su status diversi, mentre nel bilinguismo i ruoli e gli status non hanno importanza.

4.3. Repertorio italofono istriano

Per quanto riguarda il repertorio linguistico istriano, Blagoni et al. (2016) individuano cinque esternalizzazioni della realtà insulare linguistica in Istria:

1. il primo è l'atollo della dialettologia croata costituito dal dialetto ciacavo;
2. il secondo è l'atollo della dialettologia italiana, costituito dal dialetto istroveneto;
3. il terzo atollo è l'atollo della dialettologia istriota costituito delle varianti locali istriote;
4. la quarta è l'isola della variante locale di Peroj, costituita dal dialetto montenegrino, con circa trecento abitanti;

¹⁵ La dilalìa differisce dalla diglossia nel senso che, anche se la distinzione funzionale è chiara, tra le lingue A e B ci sono domini dove è normale usare, sia l'una che l'altra varietà, alternativamente o congiuntamente (Berruto 1999: 6).

5. la quinta realtà linguistico-insulare è rappresentata dalle lingue appartenenti alle otto minoranze etniche (la minoranza serba, la minoranza bosniaca, la minoranza albanese, la minoranza slovena, la minoranza montenegrina, la minoranza rom, la minoranza ungherese e la minoranza macedone).

Milani-Kruljac (1990) nella sua opera *“La comunità italiana in Istria e Fiume”* contempla il repertorio della comunità italoфона istriana nell’ambito di due ordini:

1. la comunità italoфона minoritaria di primo ordine: il cui repertorio linguistico è costituito dalla parlata minoritaria, cioè l’istroveneto, e lingua italiana standard da una parte e la lingua croata dall’altra;

2. la comunità minoritaria istroromanza di secondo ordine: limitata a poche persone presenti nei territori di Rovigno, Sissano, Gallesano, Valle e Dignano (dove viene tutelata la parlata autoctona istriota).

La situazione linguistica odierna è unica nel suo essere, conseguita da dinamiche diacroniche ed è doveroso in tal senso fare un appunto sui repertori linguistici delle tre generazioni di italoфoni istro-quarnerini individuate da Milani Kruljac (1990: 68):

1. la prima generazione del dopoguerra aveva la conoscenza molto limitata e indispensabile della lingua croata (serbocroata all’epoca) basata sui concetti estranei alla cultura italiana;

2. la seconda generazione è rappresentata dalle persone che usano l’istroveneto negli ambiti familiari o con gli amici; il croato per l’esposizione o la produzione; mentre l’uso dell’italiano standard è limitato a pochissime occasioni e all’esposizione ai programmi televisivi che, secondo Maticchio e Tamaro (2020: 7) potrebbe essere dovuto, ad una maggiore offerta di programmi di spettacolo e di documentari trasmessi dagli emittenti italiani;

3. la terza generazione è composta dai figli nati nei matrimoni misti dove è difficile individuare la madrelingua (perché sono esposti a due lingue dalla nascita). Questa generazione viene sottoposta all’uso dell’italiano standard negli anni di scolarizzazione.¹⁶

Milani-Kruljac (1990: 40) descrive la presenza del bilinguismo croato-italiano e della diglossia nel territorio istriano. In Istria non si può parlare di bilinguismo croato-italiano regionale che

¹⁶ L’uso dell’italiano standard nelle istituzioni è dovuto alla presenza di istituzioni scolastiche e professionali in lingua italiana, in quanto il bilinguismo in Istria è un bilinguismo sia *de facto* che *de iure* (Maticchio e Tamaro, 2020: 8).

implica l'uso sociale di due lingue in modo paritetico. Infatti, la varietà croatofona prevale a livello sociale in gran parte degli ambiti comunicativi pubblici. Il repertorio dei membri della comunità italiana in Istria in linea di massima equivale alla situazione della duplice diglossia. Infatti, come primo punto bisogna menzionare che la lingua italiana standard gode di un prestigio etnolinguistico minore rispetto all'istoveneto, sebbene sia tutelata a livello legislativo. Siccome la diglossia implica il rapporto non paritario delle lingue negli ambiti sociali, la doppia diglossia implica una duplice subordinazione linguistica. Nella situazione istriana è evidente che la lingua italiana si trova in una posizione secondaria rispetto al croato in tutti gli ambiti sociali e amministrativi.

In particolare, Milani-Kruljac (1990: 40) considera la diglossia nell'ambito del repertorio linguistico dei bilingui istriani intesa in due sensi, descrivendo i doppi rapporti diglossici:

- 1) fra l'italiano standard e l'istoveneto;
- 2) fra le L1 e L2 (o due L1), cioè l'italiano e il croato.

L'italiano standard non ha mai raggiunto la capillarità sociolinguistica, il prestigio comunicativo né la vitalità etnolinguistica che equivale a quella dell'istoveneto e rimane tutt'oggi il "privilegio di pochi" (Blagoni et al., 2016: 8). La presenza dialettale in Istria che interessa la comunità italoфона rappresenta pertanto una straordinaria ricchezza e vitalità etnolinguistica. Qui subentra il concetto della commutazione di codice, fenomeno normalmente presente nelle comunità bilingui in generale (Filipi, 1987: 171-179) a livello di modalità bilingue e rappresenta non un fenomeno transitorio, ma un fenomeno piuttosto duraturo.

Il termine "diglossia" può essere sostituito (nel caso istriano secondo la Milani-Kruljac) con il termine "bilinguismo dialettale", visto che la lingua madre di tutti gli italoфoni è difatti il dialetto e l'italiano standard si apprende in un secondo momento. Considerando tutte le definizioni e differenziazioni possibili, Milani-Kruljac (1990: 42) giunge alla conclusione che, siccome la maggior parte degli italoфoni possiede la competenza del dialetto istoveneto, dell'italiano standard e del croato, in Istria ci troviamo davanti alla situazione di una duplice diglossia con un potenziale bilinguismo, che tende alla semplice diglossia una volta finito il periodo di scolarizzazione.

Siccome la classificazione di Fishman risulta non soddisfacente in relazione alla situazione istro-quarnerina, Milani-Kruljac (1990: 41) affianca altre possibilità alle combinazioni indicate sopra:

- a) la triglossia e il bilinguismo che sono presenti a Valle, Rovigno, Gallesano, Dignano, ovvero luoghi dove è presente anche l'uso dell'istoromano;
- b) il bilinguismo e la doppia diglossia che caratterizzano le persone che usano ad esempio due lingue e due dialetti (italiano, croato, ciacavo e istroveneto);
- c) il trilinguismo nei casi in cui il parlante dispone di tre lingue standard (es. croato, italiano, sloveno, ecc.).

La situazione linguistica istro-quarnerina, caratterizzata dal bilinguismo, ossia dalla duplice diglossia, si complica ulteriormente con l'introduzione di ulteriori codici linguistici nell'ambito del repertorio dei parlanti: Milani-Kruljac, infatti, introduce il termine "poliglossia imperfetta" per descrivere le suddette combinazioni. Infatti, il croato rappresenta la "prima lingua Alta", la lingua maggioritaria, usata nei contesti formali e non formali; la "seconda lingua Alta" è l'italiano standard, usato in misura minore nella scrittura e poco nel parlato¹⁷, e la "terza lingua Bassa" è l'istroveneto, che copre tutta la comunicazione orale. Si possono annoverare pure i casi di repertori ancora più compositi, come ad esempio la tetraglossia o la pentaglossia per coloro che usano il dialetto locale, la *koinè* veneta regionale, l'italiano standard ecc.

Uno di questi dialetti locali, l'istrioto, che fa parte dei dialetti microregionali e locali (Maticchio e Tamaro, 2020: 2) si è sviluppato direttamente dal latino volgare parlato in Istria prima della conquista veneziana, più precisamente intorno l'anno Mille per raggiungere il picco verso il Quattrocento. Il termine "istrioto" fu coniato da Graziadio Ascoli che lo usò per denominare i ladini che nell'epoca romana popolavano la regione. Se una volta veniva usato in tutta l'Istria, la comunità che ne faceva uso faceva parte dei ceti bassi come gli agricoltori o i contadini e oggi l'uso delle parlate istriote, tutte con delle caratteristiche diverse, è limitato a cinque località: Rovigno, Sissano, Valle, Dignano, Fasana e Gallesano¹⁸. Le tracce di questo periodo emergono nei risultati delle ricerche effettuate su iscrizioni di lapidi, pietre e altri reperti archeologici istriani (Cergna, 2014: 319). Anche se i primi mutamenti della presenza del latino volgare unitario iniziarono nel IV secolo, l'invasione slava nel secolo VII segnò la rottura definitiva che vide la formazione di tre direzioni linguistiche diverse: il friulano a nord, il dalmatico in Dalmazia e l'istrioto in Istria (che poi venne soppiantato dall'istroveneto).

¹⁷ Blagoni et al. (2016: 10) distinguono 1. l'italiano letterario, scientifico e mediatico, perlopiù scritto e orale, ma mai colloquiale, utilizzato principalmente in contesti formali. 2. l'italiano tipico per comunicazioni occasionali ed episodiche, caratterizzato da numerose tracce di contatti linguistici e 3. l'italiano utilizzato da persone che hanno frequentato le scuole italiane.

¹⁸ Cergna (2014) spiega che l'istrioto è completamente estinto a Fasana, mentre a Dignano non è più parlato attivamente ma è soltanto capito da poche decine di persone.

Figura 2. La cartina geografica dell'Istria che evidenzia l'area dell'uso della lingua istriota oggi:¹⁹



La stima approssimativa dei parlanti dell'istriota non supera duemila persone (Cergna, 2014: 323) e per la mancanza di trasmissione della parlata alle generazioni più giovani è classificato a rischio d'estinzione dall'UNESCO.

Tra gli idiomi estinti bisogna menzionare il tergestino e muglisano, gli idiomi estinti a Trieste e a Muggia tra il XIX e il XX secolo e il vegliotto²⁰, parlato in una zona molto ristretta nell'isola di Veglia, scomparso nel XIX secolo (Filipi, 1989).

L'istoveneto si differenzia dagli altri dialetti veneti in base agli elementi dei dialetti istriani che sono stati introdotti e adattati a livello fonetico, morfologico e sintattico. La sua presenza in Istria risale al secolo XIV grazie all'espansione territoriale della Repubblica di Venezia. Il processo di venetizzazione "linguistica" in Istria si articola in tre fasi:

- 1) dal X al XV secolo, la Repubblica di Venezia sottomette le città istriane al proprio governo. In questo modo la lingua ufficiale diventa il veneto anche se l'uso è limitato all'ambito amministrativo;

¹⁹ Immagine tratta da Muscardin R. (http://www.odos.cloud/images/lingua_istriota.pdf - ultima consultazione il 2/5/2022).

²⁰ Il vegliotto si era sviluppato dal latino delle coste dalmate. Gli scienziati lo denominarono "il dalmatico" (Filipi, 1989).

- 2) dal XV al XIX secolo il veneto si consolida;
- 3) dal XIX secolo in poi il contatto dell'istroveneto e triestino è sempre più intenso e oggi l'istroveneto viene usato dalla stragrande maggioranza degli italofoeni istriani.

Fino al XX secolo il dialetto istroveneto era limitato solamente all'uso orale mentre l'attività letteraria comincia a svilupparsi solo negli anni '60. Grazie all'influenza del dialetto triestino nel XIX secolo sui dialetti dell'Istria, Quarnero e Dalmazia, molti tratti originali si sono persi con il tempo. Dall'altra parte è notevole il numero dei prestiti istroveneti presenti nei dialetti ciacavi.

L'istroveneto nella Regione Istriana riflette la cultura e la vita quotidiana della gente che si contraddistingue grazie a questo fattore importantissimo culturale e identificativo della comunità italoфона istriana. Questo dialetto con le sue caratteristiche intrinseche come la semplicità e l'efficacia, nonché il ventaglio delle funzioni espressive che variano da quelle sociologiche, emotive, storiche e psicologiche, sono la prova della sua forza e stabilità. La bellezza inoltre è data anche dall'originalità e dalla vitalità della quale la comunità italoфона in Istria ne va fiera e che tramanda con grande orgoglio alle generazioni successive. È il codice che oltre a rappresentare il simbolo dell'appartenenza, rappresenta l'anima della collettività che fa parte della grande storia (Milani-Kruljac 1990: 66).

Nel 2021 il Ministero della Cultura, ai sensi dell'articolo 12, comma 1, ha deliberato che il dialetto istroveneto deve far parte del patrimonio culturale immateriale e che bisogna prendere le seguenti misure per proteggerlo:²¹

1. garantire la disponibilità dei beni al pubblico;
2. incoraggiare la trasmissione e il nutrimento dei beni culturali nell'ambiente originario;
3. incoraggiare la partecipazione di individui e della società che ereditano il bene nell'identificare, definire, eseguire e trasmettere il bene culturale;
4. promuovere il bene culturale organizzando incontri professionali, attraverso mezzi elettronici, registrazioni audio e video e in altri modi;
5. assicurare la sostenibilità del bene attraverso l'educazione, identificazione, documentazione, ricerca scientifica, conservazione, tutela, promozione, valorizzazione e la

²¹ Il documento originale si trova allegato alla fine di questa tesi (Allegato 1).

possibilità di trasmettere la tradizione agli eredi attraverso l'educazione formale e non formale e la rivitalizzazione di segmenti abbandonati del bene culturale;

6. educare il personale professionale alla trasmissione di conoscenze e competenze attraverso seminari, laboratori, istruzione formale e non formale;

7. continuare la ricerca del bene culturale, attraverso documentazioni adeguata in tutte le sue forme e attraverso valutazioni professionali e scientifiche;

8. promuovere la funzione e l'importanza del bene culturale nella società e includere la sua protezione nei programmi di sviluppo pianificati;

9. aumentare la consapevolezza che il dialetto costituisca il patrimonio da salvaguardare che garantisce la diversità culturale a livello locale e nazionale e attuare misure di protezione per evitare il rischio di estinzione, distruzione o commercializzazione eccessiva del bene culturale;

10. incoraggiare lo sviluppo del vocabolario e della grammatica del dialetto istroveneto;

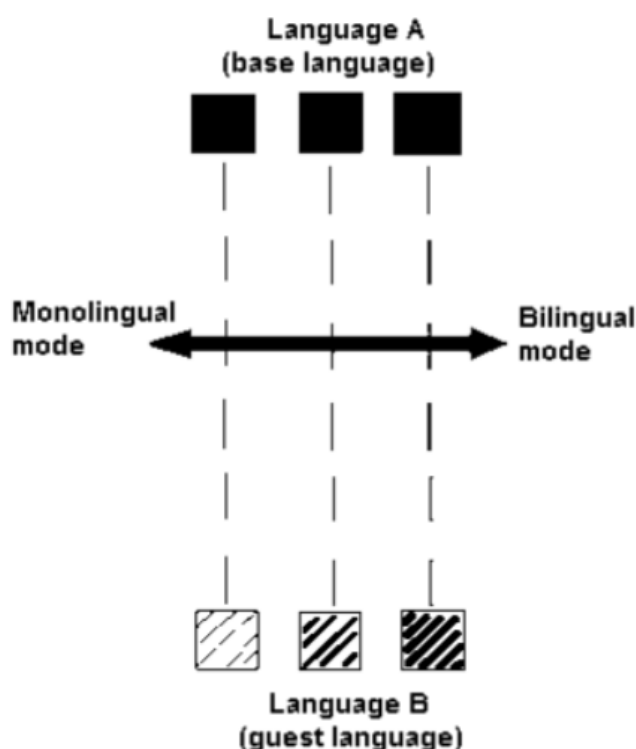
11. includere l'istruzione dell'istroveneto nei programmi di istruzione primaria e secondaria;

12. il detentore del bene è obbligato ad attuare misure di tutela ai fini della conservazione del dialetto ai sensi della legge sulla tutela e la conservazione dei beni culturali e di tutte le norme relative ai beni culturali, aderendo alla sua matrice storico-tradizionale e alla sua occorrenza.

5. Modalità bilingue e commutazione di codice

Negli ultimi decenni si registra un crescente interesse verso la fenomenologia che consegue dal contatto di lingue e culture, in particolar modo verso la commutazione di codice. La comunicazione bilingue, infatti, è caratterizzata da un continuum linguistico composto da situazioni che possono essere definite come monolingui e bilingui. Grosjean (2008: 40) parla della cosiddetta modalità monolingue/bilingue dei parlanti. Le lingue di un parlante bilingue, infatti, possono essere attivate e/o disattivate quando necessario (anche nel corso di uno stesso discorso), per vari motivi, come ad esempio per facilità di espressione oppure per mancanza di un termine specifico. Quando si parla di commutazione di codice bisogna sempre tenere presente che essa implica una lingua base (lingua A) in cui si svolge la maggior parte della conversazione e la lingua B, quella da cui si attinge materiale linguistico a seconda della funzione comunicativa che il parlante desidera attivare. Grosjean afferma che la lingua base può cambiare (all'interno di una modalità bilingue stabile) o che un individuo può passare dalla modalità bilingue a quella monolingue e viceversa, mantenendo la stessa lingua di base. Di seguito si riporta un'illustrazione sintetica del concetto di modalità bilingue e monolingue (Grosjean, 2008: 40).

Figura 3. Il continuum degli stati di attivazione delle lingue



La lingua B è molto utilizzata nella modalità bilingue (maggiore è lo spessore delle linee diagonali all'interno dei quadrati, maggiore è l'uso del linguaggio B) ma, anche se temporaneamente disattivata, può essere richiamata quando necessario. Una delle caratteristiche tipiche della modalità bilingue è il fenomeno della commutazione di codice (*code switching*), termine originariamente coniato da Vogt nel 1954, nella sua recensione del libro di Weinreich "*Languages in Contact*" (1953):

Code-switching in itself is perhaps not a linguistic phenomenon, but rather a psychological one, and its causes are obviously extralinguistic. But bilingualism is of great interest to the linguist because it is the condition of what has been called interference between languages. (Vogt 1954 :368)

Vogt spiega che la commutazione di codice può essere considerata come un fenomeno psicologico, dato che le sue ragioni sono extralinguistiche. Lo studio della commutazione di codice è stato oggetto di studio partendo da tre prospettive distinte:

- a) linguistica,
- b) psicolinguistica,
- c) sociolinguistica.

La prospettiva linguistica si occupa della ricerca di vincoli grammaticali. La prospettiva psicolinguistica si concentra sulla modalità in cui i sistemi linguistici dei bilingui sono archiviati e sulla loro accessibilità nel sistema cognitivo. Infine, la prospettiva sociolinguistica pone l'attenzione sulle motivazioni sociali che si trovano dietro la commutazione di codice. L'inserimento di elementi lessicali da una lingua all'altra alle volte include solo una parola (monolessicale), alle volte più parole (plurilessicale)²².

Shana Poplack (1980: in Carli, 1996: 128) definisce la commutazione di codice come un fenomeno di alternanza di due lingue all'interno di un discorso. La linguista attribuisce molta importanza all'aspetto grammaticale e strutturale dei sistemi linguistici in contatto. Di conseguenza stabilisce che il fenomeno possa avvenire secondo due principi all'interno di una frase:

²² Questi passaggi possono essere scambiati per prestiti, ma secondo Bentahia e Davies (1983: 301-330) il prestito è utilizzato sia da persone monolingui che quelle bilingui e la parola presa in prestito diventa parte del lessico della lingua "ospite", mentre la commutazione di codice è caratteristica solo per le persone bilingui e multilingui. Haugen (1956), inoltre, cerca di distinguere la commutazione dall'interferenza, definendo la prima un inserimento non assimilato, mentre la seconda come un processo comportante la sovrapposizione di due lingue, che ha come conseguenza la deviazione dalle norme di entrambi i sistemi linguistici.

- a) il principio del morfema libero che permette la commutazione dopo ogni costituente purché questo non sia un morfema legato (es. Non permesso= PARLati);
- b) il principio dell'equivalenza che consente la commutazione purché non vengano violate le regole sintattiche delle due lingue (es. QUANDO idemo).

I casi in cui questi vincoli vengono violati sono classificati come prestiti occasionali (*nonce loans*), che sono di carattere individuale e che si differenziano da quelli stabiliti (*established loans*), usati frequentemente. Gardner-Chloros (Carli, 1996: 130) scrive che ci sono più tipi di commutazione di codice che variano da quelli più integrati e quelli meno integrati nelle lingue ospitanti e che le motivazioni possono essere psicologiche o sociologiche. In altre parole, non ci sono norme grammaticali imposte da seguire cambiando il codice e l'alternanza deve avvenire in modo spontaneo e naturale.

5.1. Commutazione di codice a livello strutturale

A livello strutturale, la commutazione può manifestarsi all'interno di una frase (intrafrasale) o tra le frasi (interfrasale) (Myers-Scotton in Carli, 1996: 129). Secondo Poplack la competenza linguistica del parlante influisce sulla scelta del tipo di commutazione di codice. La linguista è dell'opinione che un soggetto con maggior padronanza della lingua interferente preferisce usare la commutazione intrafrasale, mentre un soggetto con minor padronanza della lingua interferente sceglie la commutazione interfrasale nel timore di violare le regole grammaticali di una delle due lingue (Dal Negro e Guerini, 2007: 44). La commutazione interfrasale è quella in cui si verifica l'alternanza tra le frasi, alla fine della frase o tra frasi separate da una virgola, un punto o una congiunzione. Siccome le due lingue sono correlate solo semanticamente, in questo caso non ci sono problemi di compatibilità grammaticale o strutturale. Vediamo un esempio di commutazione interfrasale estratto dal corpus:

Doc. 700.

Doc. *INV: dijemo del lavor che i [/] che i fa.

*INV: [- hrv] **od svakodnevnog hm posla.**

*INV: e invece (.) altre persone (.) le legi perché no(n) le ga +.

*INV: ++ compagnia.

Trad. Diciamo del lavoro che fanno.

Del lavoro quotidiano.

E invece altre persone leggono perché non hanno compagnia.

La commutazione intrafrasale è quella in cui si verifica l'alternanza all'interno di una frase, di solito senza pause, esitazioni o interruzioni. In questo caso spesso si verificano incompatibilità a livello strutturale (ad esempio per quanto concerne l'ordine delle parole, ecc.) o morfologiche (ad esempio le differenze flessive, ecc.). Vediamo un esempio estratto dal corpus:

Doc. 699.

*REB:pol eser **zdravo@s** quanto che ti vol.

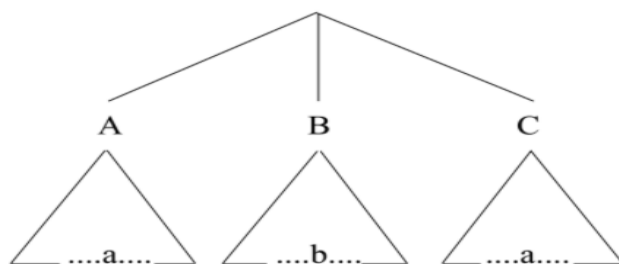
Trad. Può essere sano quanto vuoi.

I linguisti adottano prospettive diverse per definire e teorizzare la commutazione di codice e spesso volte non sono concordi nella definizione della commutazione di codice in relazione ad altri fenomeni del contatto linguistico, come ad esempio i prestiti. Lewandowski (1976: 415) intende per prestito una parola integrata in un'altra lingua. Poplack (1980) sostiene che i prestiti occasionali si differenziano da quelli stabiliti sulla base della frequenza d'uso e del grado di integrazione (Dal Negro e Guerini, 2007: 44). I prestiti stabiliti sono integrati dal punto di vista morfologico, fonologico e sintattico e sono ampiamente usati in una determinata comunità, mentre quelli occasionali oltre a non essere integrati, vengono usati a livello creativo individuale. La studiosa stabilisce che la commutazione di codice è più frequente se si tratta di due codici tipologicamente simili, dove il passaggio può avvenire senza la violazione di modelli strutturali, mentre si preferiscono i prestiti se i due codici presentano importanti differenze tipologiche. Myers-Scotton (Carli: 132) distingue tra prestiti culturali (*cultural borrowings*), che servono a esprimere concetti per mancanze lessicali in una lingua e prestiti essenziali (*core borrowings*), che si usano per esprimere una particolare sfumatura espressiva o per aumentare il prestigio di ciò che designano. Il pensiero di Gardner-Chloros (Carli: 132) è in linea con quello di Myers-Scotton nel considerare i due fenomeni, quello della commutazione e i prestiti, molto simili. Secondo la studiosa un'innovazione può essere classificata come commutazione che diventa prestito se usata frequentemente (Dal Negro e Guerini, 2007: 44).

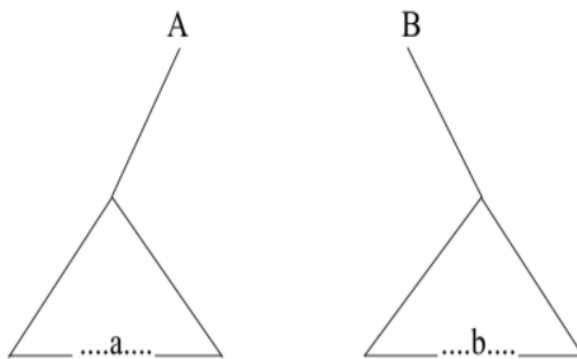
Un'altra distinzione che deve esser fatta è quella tra commutazione di codice ed enunciazione mistilingue (*code mixing*). La differenza principale è stata riassunta da Seckin (2014), che percepisce la commutazione di codice come un fenomeno che avviene intenzionalmente da una

persona (implicando dunque le sfumature funzionali che le risorse linguistiche offrono), mentre l'enunciazione mistilingue implica l'uso dell'altro codice involontariamente, spesso volte perché non si conosce la parola o la frase corretta o nel corso dell'apprendimento linguistico (si pensi all'interlingua, ad esempio). Auer (1999: 309-332) definisce l'enunciazione mistilingue come giustapposizione di due lingue, dove l'uso di queste due lingue ha un'importanza per gli interlocutori più globale che locale e in questo caso vengono usati gli schemi ricorrenti. Nel ventaglio dei fenomeni delle alternazioni linguistiche, Auer pone in centro l'enunciazione tra i due poli opposti che sarebbero la commutazione di codice da una parte e i “*fused lects*” dall'altra. Secondo il linguista la commutazione è legata al contesto del momento e le intenzioni sono del carattere pragmatico. I *fused lects* invece non prevedono l'uso libero dell'alternazione ma delle regole che uniscono le forme alle funzioni. L'enunciazione mistilingue, quindi, rappresenta la fusione dei due poli e si può spiegare come il fenomeno dove le regole vengono usate liberamente e con creatività. Auer (1999: 309-332) e Alvarez-Cáccamo (1998: 29-50) considerano l'enunciazione mistilingue priva di funzionalità interazionali nel senso che il parlante “mischia” due o più codici all'interno della frase, senza che il passaggio da un codice ad un altro svolga una particolare funzione nella conversazione. Muysken (2000: 8) differenzia i tre principali tipi di enunciazione mistilingue che sono presenti nelle comunità linguistiche bilingui: l'inserimento, l'alternanza e la lessicalizzazione congruente, stabilendo inoltre che solitamente uno di questi tipi domina nella conversazione, mentre gli altri due sono altrettanto presenti.

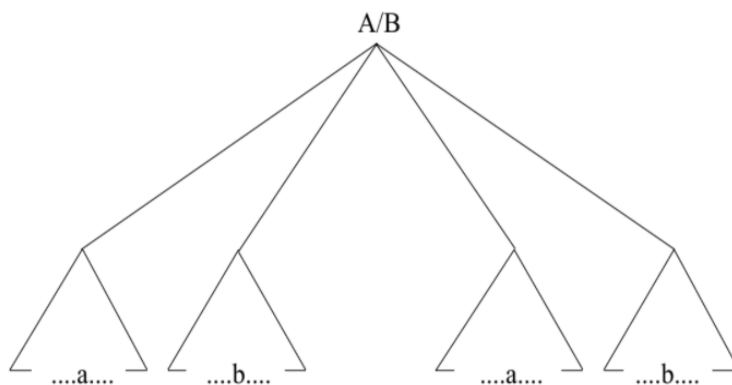
- 1) L'inserimento prevede innanzitutto la determinazione di una lingua nella struttura complessiva in cui si trovano i costituenti inseriti dell'altra lingua (es. un prestito spontaneo dove un'unità viene intromessa in una frase in un altro codice).



- 2) Nell'alternanza entrambe le lingue sono utilizzate alternativamente (a seconda dei vincoli strutturali di ciascuna lingua). Si differenzia dall'inserimento per il numero delle unità mescolate.



- 3) Nella lessicalizzazione congruente la struttura grammaticale è condivisa da entrambe le lingue e le parole sono inserite più o meno casualmente (Muysken, 2000: 8).



Secondo i linguisti le motivazioni per le quali avviene la commutazione di codice sono diverse e ci sono altrettanto diversi approcci metodologici allo studio del fenomeno. Siccome si pone il problema dell'interpretazione funzionale e la commutazione di codice risulta strettamente legata al contesto, di conseguenza ogni generalizzazione diventa difficile. È difficile stabilire una regolarità nelle ricorrenze del fenomeno dato che l'uso dipende dall'esperienza comunicativa concreta. Blom e Gumperz (1972: 407-434) distinguono tra la commutazione di codice situazionale e quella metaforica. La commutazione situazionale avviene nelle situazioni di un diverso contesto istituzionale associato a un codice distinto e viene spiegata nel seguente modo:

Distinct varieties are employed in certain settings (such as home, school, work) that are associated with separate, bounded kinds of activities (public speaking, formal negotiation, special ceremonials, verbal games, etc.) or spoken with different categories of speakers (friends, family members, strangers, social inferiors, government officials, etc.). (Gumperz, 1982: 60)

Il passaggio sopracitato spiega che le varietà distinte vengono impiegate in determinati contesti che sono associati a diversi tipi di attività e che si utilizzano con categorie di parlanti specifiche. Questo tipo di commutazione di codice è tipico nelle società diglossiche dove i codici sono strettamente legati ai rapporti interpersonali, a ruoli sociali, a particolari attività o all'uso istituzionale. La commutazione di codice metaforica non dipende dalle situazioni o contesti diversi ma si verifica nelle situazioni dove c'è un'intenzione precisa come, ad esempio, per dare un tocco personale alla conversazione. Essa viene spiegata nel seguente modo:

The context in which one of a set of alternates [codes] is regularly used becomes part of its meaning, so that when this form is then employed in a context where it is not normal, it brings in some of the flavor of this original setting. (Bloom e Gumperz, 1972: 425)

Il passaggio spiega che il contesto in cui un codice viene usato regolarmente diventa parte del suo significato, così che quando questa forma viene poi impiegata in un contesto in cui non ci si aspetta che avvenga, porta con sé l'impronta di questa impostazione originale e creativa. Successivamente Gumperz usa il termine "commutazione conversazionale" al posto del termine "commutazione metaforica" e osserva che le strategie comunicative dipendono dalle conoscenze linguistiche e aspettative condivise degli interlocutori, il che li distingue dai parlanti che non possiedono tali conoscenze. Il linguista, infatti, traccia una distinzione tra le diverse funzioni conversazionali della commutazione di codice (Gumperz, 1982: 75-81). Si riporta di seguito l'elenco delle stesse:

1. Citazioni dirette o discorsi riportati (indiretti) (vedi Županović Filipin e Bevanda Tolić: 2015)
2. Specificazione del destinatario (vedi Cerruti, 2004)
3. Interiezioni (vedi Poropat Jeletić et al., 2021)

4. Ripetizioni, enfatizzazioni, chiarimenti, rafforzativi (vedi Županović Filipin e Bevanda Tolić, 2015)
5. Restrizioni strutturali del messaggio(vedi Županović Filipin e Bevanda Tolić, 2015)
6. Personalizzazione vs. oggettivazione (vedi Cerruti, 2004).

Nella bibliografia consultata, le stesse funzioni possono essere descritte in modi diversi da autori diversi. Di seguito si riporta l'elenco proposto da Baker (2010: 56-57), che indica 12 funzioni comuni che sono state identificate per attivare la commutazione di codice (per ulteriori approfondimenti vedi Županović Filipin e Bevanda Tolić, 2015 e Cerruti, 2004):

1. Enfaticizzazione di una parola o di una frase particolare
2. Un parlante non conosce una parola o una frase in una lingua
3. Riferimento ad un concetto senza equivalente nella cultura dell'altra lingua
4. Rafforzamento di una richiesta
5. Chiarimento
6. Indicazione di deferenza, amicizia o legame familiare
7. Intervenzione conversazionale
8. Allentamento della tensione e/o suscitazione dell'umorismo
9. Cambiamento di atteggiamento o di relazione
10. Creazione di distanza dall' ascoltatore, o relazione più formale
11. Esclusione delle persone da una conversazione
12. Introduzione di determinati argomenti

In conclusione, bisogna notare che l'elenco dei valori pragmatici della commutazione di codice non è limitato ai tipi sopra menzionati, è anzi aperto e molto ricco. Una volta individuate le funzioni, nella seconda parte analitica della tesi ci si occuperà del modo in cui queste alternanze linguistiche ricorrono, prendendo in considerazione l'ottica cognitiva del modello *usage-based*.

5.2. Modelli teorici della commutazione di codice

I tre modelli teorici recenti che cercano di spiegare il funzionamento della commutazione di codice maggiormente citati nella bibliografia consultata sono il modello della marcatura

(*markedness model*), il modello della matrice (*matrix model*) e il modello basato sull'utilizzo (*usage-based model*). L'applicazione di diverse teorie nel corso degli anni ha ispirato molti lavori nell'ambito della sintassi, morfologia e fonologia, portando spesso a dei risultati diversi (Nilep, 2006: 1-22). Gli ultimi due decenni hanno visto un progresso notevole negli approcci empiristici al linguaggio e l'impegno verso lo studio del linguaggio attraverso il realismo psicologico. Ai modelli *matrix* e *markedness* è stato affiancato il modello *usage-based* che è considerato uno dei più fruttuosi approcci la cui applicazione vede recentemente una rapida espansione.

5.2.1. Il modello della marcatura (*markedness model*)

Myers-Scotton (1993; in Shin, 2018: 131-133) nel suo *markedness model* sostiene che i parlanti bilingui sono consapevoli delle conseguenze sociali, della scelta di una lingua specifica o della combinazione di lingue in un contesto particolare. La linguista afferma che in una data circostanza, una particolare varietà linguistica usata rappresenta la scelta non marcata, mentre l'altra varietà, ossia la lingua inaspettata, rappresenta la scelta marcata. Myers-Scotton scrive che i parlanti bilingui sono consapevoli del fatto che la marcatezza è associata alle pratiche previste in una data comunità e che le scelte inaspettate o marcate vengono prese per ragioni specifiche. Traendo l'esempio dalle società multilingui postcoloniali, spiega che il passaggio alla lingua etnica locale per parlare della propria famiglia sarebbe la scelta non marcata, nel senso che non c'è nulla di straordinario in questa scelta linguistica, mentre l'uso della lingua etnica locale in un discorso pubblico costituirebbe la scelta marcata, ovvero solleverebbe dubbi sulle intenzioni del parlante. Nel *markedness model*, gli enunciati sono prodotti per esprimere le intenzioni del parlante nei confronti di un pubblico specifico con il quale si condivide la lingua e ci si capisce. Il rapporto tra il parlante e l'ascoltatore basato sulla definizione dell'intenzionalità dei codici scelti struttura la premessa del *markedness model*. In riferimento a quanto spiegato, il modello della marcatura è limitato a due domini: in primo luogo, è fortemente basato sulla conoscenza extralinguistica del linguaggio colloquiale e in secondo luogo dipende in gran parte dalle aspettative dell'ascoltatore sulle intenzioni per soddisfare le esigenze dello scambio linguistico (Nilep, 2006: 1-22). Burt (2002: 993-1001) spiega che la scelta di alcune forme previste in alcune comunità, ad es. *tu/lei* in italiano viene usato con la famiglia e gli amici, mentre *lei* viene usato più formalmente per rivolgersi a individui di rango o potere superiori oppure per esprimere cortesia verso una persona. In altre parole, *lei* non è

marcato nelle interazioni formali, e *tu* non è contrassegnato in quelle meno formali. Secondo Myers-Scotton la L1 della persona bilingue è più probabile che sia quella non marcata, anche se la L2 può essere utilizzata come non marcata nei contesti formali. Alla luce delle spiegazioni riportate possiamo concludere che la lingua non marcata coincide con la lingua matrice, la quale, come sostiene Myers-Scotton, corrisponde alla madrelingua della persona bilingue, mentre la lingua non marcata viene usata nei contesti specifici (Shin, 2018: 131-133).

5.2.2. Il modello della matrice (*matrix model*)

Il modello *matrix* di Myers-Scotton (2002: 8) implica che nella commutazione di codice esiste un sistema asimmetrico tra le due lingue coinvolte: una è la lingua matrice (*ML-matrix language*) e l'altra è la lingua incorporata (*EL-embedded language*). La lingua cosiddetta matrice fornisce il quadro morfo-sintattico della frase, mentre quella incorporata è composta dalle parole inserite nella lingua matrice. Il modello consente l'inserimento anche delle isole linguistiche, ossia dei gruppi sintattici, se questi soddisfano le condizioni di corretta costruzione sia della lingua matrice sia della lingua incorporata. Il modello prevede una classificazione basata sul dominio linguistico che distingue la commutazione di codice classica da quella composita (Myers-Scotton, 2002, in Selvaggi, 2012: 80). Il modello MLF (*Matrix Language frame*) non è applicabile alla commutazione composita perché la cornice è costituita da più di una lingua, mentre la commutazione classica implica l'incorporazione costituita da morfemi lessicali, misti o di gruppi sintattici. Il modello MLF si basa su due principi di alternanza di codici (Selvaggi, 2012:81):

- a) *the morpheme order principle* o il principio dell'ordine del morfema: stabilisce che una lingua determina l'ordine dei morfemi all'interno dell'enunciato bilingue. Normalmente si tratta della lingua matrice, ma all'interno della lingua incorporata l'ordine dei morfemi è determinato dalla lingua incorporata;
- b) *the system morpheme principle* o il principio del morfema di sistema: stabilisce che una tipologia di morfema di sistema proviene sempre dalla lingua matrice.

Nel modello *matrix* si parla di quattro tipi di morfemi riassunti nella seguente tabella:

Tabella 8. Il modello di classificazione dei morfemi, il 4-M Model (Selvaggi, 2012:81).

Morfemi contenuto o lessicali	Morfemi di sistema
1. la testa del morfema composto	1. precoci (<i>early system morphemes</i>)
	2. ponte (<i>bridges</i>)
	3. estranei (<i>outsiders</i>)

Nei morfemi di sistema rientrano gli affissi, gli articoli e i pronomi. I morfemi ponte collegano le parti tra di loro, mentre i morfemi estranei sono quelli di accordo soggetto-oggetto ecc.

5.2.3. Il modello basato sull'utilizzo (*usage-based model*)

La linguistica tradizionale di contatto descrive i fenomeni di contatto come fonologici, lessicali, strutturali e pragmatici. In contrasto con questa tendenza, l'approccio *usage-based* promuove un metodo leggermente diverso per la descrizione dei dati perché le descrizioni tradizionali non sono sempre riconosciute come teoricamente rilevanti.

“Grammar can be viewed as the cognitive organization of ones experience with language.” (Bybee, 2006; in Hoffman, Trousdale, 2013: 49)

Come spiega la citazione sopra riportata, la grammatica può essere vista come l'organizzazione cognitiva della propria esperienza con il linguaggio e quindi si ipotizza che la competenza linguistica di una persona rappresenti l'inventario che un individuo possiede, composto da abbinamenti di forme e significati indicati come unità. Quindi potrebbe essere giusto dedurre che nessuna unità possa avere la connotazione esclusivamente fonologica, morfologica, sintattica o semantica. Il concetto di unità risale alla nozione saussuriana del segno linguistico che unisce la forma, la cui lunghezza è variabile ovvero può includere morfemi, collocazioni ed espressioni fisse e un significato (Langacker 1987; in Hakimov e Backus, 2021). La visione di base del modello *usage-based* del linguaggio si basa sul fatto che le strutture linguistiche non ci sono state date alla nascita ma che tutte le costruzioni, le strutture, le regole e le categorie vengono create attraverso l'apprendimento e tramite i processi cognitivi generali (Bybee, 2010). Le unità apprese vengono memorizzate insieme a tutte le informazioni che possono essere percepite in un dato contesto, permettendo in questo modo la creazione di un ricco deposito di esperienze con tutti i dettagli inerenti a queste esperienze nella nostra memoria.

Non appena si verificano delle somiglianze tra due parole o costruzioni, queste vengono collegate tra loro formando delle reti che combinano sequenze basate su diverse proprietà formali e semantiche (Hakimov e Backus 2021). In questo modo le nuove costruzioni o schemi linguistici, con le caratteristiche simili a quelle già esistenti, vengono memorizzate e depositate vicino a quelle conosciute. Ad un certo punto questi schemi, grazie alla ripetizione continua, cominciano ad essere produttivi e ad essere usati indipendentemente, creando delle sequenze nuove.

6. Postulati del modello *usage-based*

L'approccio *usage-based* posa sulla teoria della rappresentazione del linguaggio attraverso il processo cognitivo. Indipendentemente dal fatto che l'attenzione sia rivolta all'acquisizione o all'uso alternato delle lingue, secondo il modello la conoscenza di una lingua si basa sull'uso linguistico effettivo e sulle generalizzazioni dedotte in base agli eventi di utilizzo (Langacker, 1987; Goldberg, 2006; Bybee 2010). L'idea di base è che diventano automatizzati i processi, ovvero le prassi dell'uso degli elementi che si verificano ripetutamente in congiunzione e che vengono utilizzati coerentemente per la stessa funzione. Secondo quest'idea si può dedurre altresì che la complessità del linguaggio emerge altrettanto attraverso l'interazione di cognizione e di uso. Nonostante nella linguistica moderna sia sorta solo recentemente l'idea che gli schemi linguistici ripetitivi basati sull'utilizzo possano modellare la rappresentazione cognitiva della lingua, le basi di quest'idea sono state gettate ancora nel 1880 da Hermann Paul nei suoi *“Principi di storia del linguaggio”*. Il seguente passaggio di Paul descrive la sua versione di questo modello:

[...] Every act of speaking, listening, or thinking adds something new. Even by an exact repetition of a previous act, some forces of the already existing mental grammar are strengthened. And even if somebody can look back on a rich linguistic life there is always the opportunity for something new: Aside from the introduction of things that were previously unusual in the language, at least a new variant to old elements may be added. [...Both the attenuation and the strengthening of the old elements as well as the addition of new ones constantly shift the associations within this system of representations. (in Hakimov 2021: 49)

Possiamo concludere dal passaggio sopracitato che l'esperienza aggiorna costantemente le costruzioni o schemi esistenti e che la ripetizione di un atto consolida la grammatica mentale già esistente. L'attenuazione e il rafforzamento dei vecchi elementi, così come l'aggiunta di nuovi elementi, spostano costantemente le associazioni all'interno del sistema di rappresentazioni nella memoria linguistica, che tra l'altro può anche decadere e quindi può far scattare la riorganizzazione delle reti (cfr. Bybee e Beckner 2010). In questo modo vengono spiegate le costruzioni che presentano delle peculiarità di significato, come ad esempio “il piano B” e quelle che sono composte in espressioni convenzionali, come “buon anno”. Nel modello *usage-based* ci sono diversi processi cognitivi che influenzano l'uso e lo sviluppo della struttura linguistica. Essi verranno analizzati brevemente in seguito.

6.1. Frequenza

Nei modelli basati sull'utilizzo, la frequenza di utilizzo gioca un ruolo importante nella produzione e nella comprensione del linguaggio. La frequenza con cui vengono utilizzati un elemento linguistico o una costruzione linguistica è correlata alla sua forza rappresentativa e alla sua facilità di attivazione. Ovvero, maggiore è la frequenza d'uso di una struttura linguistica, un elemento o una costruzione, più forte è la sua rappresentazione, e più probabile sarà la sua selezione nella conversazione. Ci sono due tipi di frequenza:

- 1) la frequenza *token*: è la frequenza con cui un elemento o una frase specifica si verifica nel testo;
- 2) la frequenza *type*: è la frequenza di elementi distinti che possono essere utilizzati in un *pattern*.

Per quanto riguarda l'impatto della frequenza di utilizzo delle parole sull'elaborazione di parole polimorfemiche e sequenze plurilessicali, la frequenza lessicale ha un impatto diretto sul riconoscimento di parole polimorfemiche o complesse. In tal senso ci sono due modi di riconoscimento delle parole polimorfemiche (Hakimov, 2021):

1. in cui la parola si può scomporre nei suoi costituenti;
2. in cui la scomposizione della parola è impossibile e viene riconosciuta direttamente come intera.

Maggiore è la frequenza dell'uso della parola polimorfemica, maggiore è la probabilità che si acceda direttamente a questa parola nel lessico mentale e la si usi in modo autonomo. Possiamo citare l'esempio della sequenza commutata riguardante la parola polimorfemica **zološki@svrt@s** (inserita in una sequenza in istroveneto estratta dal Doc. 30 del corpus su cui verranno basate tutte le nostre analisi). La sequenza commutata avviene grazie alla giustapposizione e può essere analizzata in due modi. Il primo riguarda il livello plurilessicale dove le due parole non sono intercambiabili, ovvero la parola "**zološki**" implica l'affiancamento della parola "**vrt**" in croato. L'altro modo si riferisce alla scomponibilità della parola "**zološki**", prendendo in considerazione la derivazione della parola "**zoologia**" o "**zoologija**". Questa costruzione è caratterizzata dalla cosiddetta "*partial productivity*", ossia essa non può essere estesa a qualsiasi elemento lessicale ma è soggetta a vincoli di carattere semantico come spiegato da Goldberg e Suttle (2011: 1238). Comunque, l'accettabilità di costruzioni usate è influenzata sia

dalla frequenza di ricorrenza delle combinazioni nella lingua, sia dalla similarità delle unità meno attestate con le unità più frequenti:

The high-frequency expressions and the low-frequency expressions that were semantically related to the high-frequency ones were judged more acceptable than the isolated (low frequency unrelated) combinations. This result indirectly confirms the centrality of the higher-frequency expressions in forming similarity clusters, as both the high-frequency and the lower-frequency related expressions were judged to be more acceptable than the lower-frequency isolated expressions. (Bybee e Eddington, 2006:352)

In altre parole, le espressioni ad alta frequenza e quelle a bassa frequenza semanticamente correlate a quelle ad alta frequenza sono state giudicate più accettabili delle combinazioni definite semplicemente bassa frequenza. Questo risultato conferma indirettamente la centralità delle espressioni ad alta frequenza nella formazione di *cluster* di somiglianza.

6.2. Frequenza e ripetizione nelle sequenze plurilessicali

Un altro aspetto interessante che riguarda le sequenze plurilessicali sono gli studi circa la distribuzione complessiva delle varie esitazioni e disfluenze nella conversazione spontanea. Ad esempio, Kapatsinski (2005; in Hakimov, 2021: 66) nota che quanto viene ripetuto in una pausa di esitazione o insicurezza è influenzato dalle informazioni distributive al di là delle singole parole. La ripetizione di solito coinvolge l'ultima parola, come in (a), ma a volte interessa due o anche più parole (b).

a. I really appreciated [the, +the] whole, uh, English class.

b. The crime level is not as high as it is in other areas [of the, +of the] city. (Kapatsinski 2005: 481)

La ripetizione secondo il modello *usage-based* viene studiata nella maniera in cui è determinata dalla circoscrizione sintattica e nella maniera in cui è basata sulla frequenza d'uso. Gli studi hanno dimostrato che il parlante comincia a ripetere l'unità lessicale o plurilessicale dal confine sintattico più vicino, ma attraversa il confine scegliendo la parola più o meno scontata o aspettata (Hakimov, 2021). Intesa in questo modo, la ripetizione può essere considerata come coesione che gioca un ruolo importante nella produzione del parlato, specialmente come è stato

dimostrato nelle analisi delle disfluenze dove si può osservare come il processo del *chunking*, che verrà spiegato più avanti, avviene gradualmente.

6.3. Riduzione lessicale

Gli studi che indagano la produzione di parole polimorfemiche nei corpora linguistici si basano sull'osservazione della frequenza lessicale, la quale determina la produzione di parole nel discorso spontaneo. Di norma, secondo Bybee, le parole ad alta frequenza sono soggette a riduzione fonetica (Hakimov, 2021: 64) e tendono ad essere pronunciate in maniera più breve rispetto alle parole a bassa frequenza. Gli autori concludono che queste parole subiscono un processo di erosione, che è segnato da una perdita di struttura morfologica e da uno sviluppo verso forme monosillabiche. La riduzione avviene anche nelle sequenze plurilessicali e gli autori sostengono che questo è possibile perché l'alta frequenza di queste sequenze conferisce loro lo status di unità nel deposito mentale delle persone (Hakimov, 2021: 65).

Di seguito riporto l'elenco di ipotesi sull'interazione della frequenza di utilizzo con la riduzione lessicale (Bybee, File Muriel e De Souza, 2016).

1. La riduzione segue gli andamenti articolatori già presenti nel linguaggio. Tali tendenze sono accelerate in alcune frasi ad alta frequenza grazie all'automatizzazione delle attività neuromotorie.
2. La riduzione è foneticamente graduale, caratterizzata da un continuum di variazione fonetica.
3. Con un'elevata frequenza di utilizzo, anche le sillabe accentuate possono essere ridotte, le sequenze con più sillabe accentuate sono raggruppate come singole unità, portando a uno stress prominente.
4. La riduzione lessicale non è un segno di linguaggio sciatto o pigro, ma piuttosto è un comportamento finemente sintonizzato che è accuratamente modellato all'interno di un discorso.
5. Una riduzione può essere un precursore dei cambiamenti sonori futuri.
6. Una riduzione è un'ulteriore prova dell'ipotesi che il cambiamento fonetico influisce su parole e frasi a velocità diverse, a seconda di come spesso la parola o la frase ricorre nei contesti che favoriscono il cambiamento, includendo non solo il contesto fonetico, ma anche quello funzionale, lessicale e anche il contesto grammaticale.

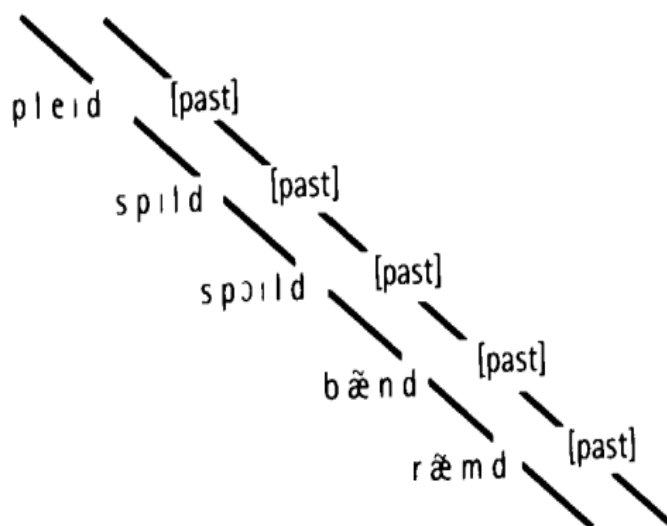
6.4. Somiglianza

Vi è un accordo generale nel contesto dell'approccio *usage-based* sul fatto che le parole siano associate ad altre parole da legami di somiglianza fonetica e semantica:

Learners are fairly cautious in producing utterances based on generalizing beyond the input. They can only be expected confidently to use a new verb in a familiar pattern when that new verb is relevantly close in meaning to verbs they have already heard used in the pattern. (Goldberg, 2006:99)

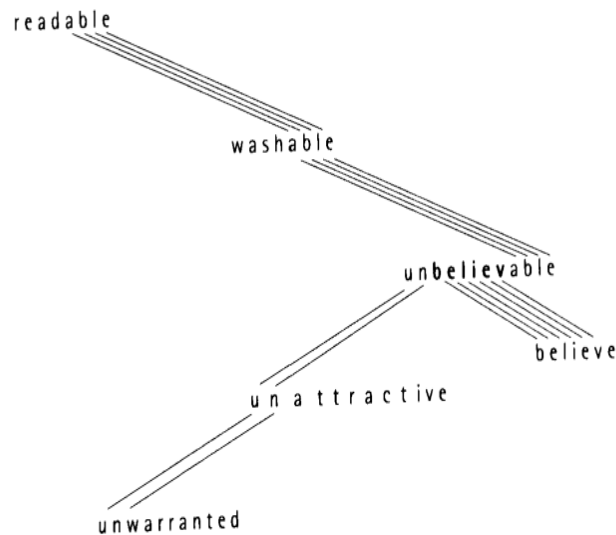
Il passaggio sopracitato spiega che c'è una certa cautela nel produrre espressioni basate sulla generalizzazione, ovvero che ci si può aspettare l'uso di un nuovo verbo in uno schema familiare solo quando quel nuovo verbo è vicino nel significato ai verbi già usati nello schema. Le seguenti figure rappresentano le connessioni delle parole basate sulla somiglianza:²³

Figura 4. Connessione fonologica e semantica dei verbi in Past Simple inglese



²³ Bybee J.L., Beckner C. (2010: 835).

Figura 5. Esempio di connessione della struttura interna della parola *unbelievable* con altre parole



Come sostenuto da Bybee (Hakimov, 2021: 50), le relazioni di somiglianza tra le parole danno origine all'emergere di relazioni morfologiche. Ad esempio, la struttura interna della parola *rivitalizzare* è il risultato delle sue relazioni con altre parole aventi caratteristiche fonetiche e semantiche simili, tra cui, *vitalità* o *vitale*. In termini di un approccio *usage-based*, la memorizzazione di esempi di parole e sequenze di parole che sono simili in una rete, consente l'emergere di gruppi di esemplari. Le unità possono variare da quelle molto specifiche a quelle più generali e da unità molto piccole a quelle a livello di discorso. Maggiore è il grado di somiglianza in cui le unità sono state usate precedentemente, più i parlanti saranno sicuri di usare queste unità, come spiega il seguente passaggio:

The greater the degree to which previously attested instances fill a semantic space that include the potential target instance, the more confident speakers will be in using the target instance. (Goldberg, 2006: 100)

6.5. *Chunking*

“*To chunk a text*” significa letteralmente spezzare un testo segmentandolo in una sequenza non strutturata di unità di testo organizzate sintatticamente chiamate blocchi (Abney, 1991: 257-278). Perciò il *chunk* sta per blocco o pezzo che, unendosi ad altri pezzi o blocchi, formano delle unità più grandi. *Chunking* implica la capacità di costruire tali strutture in modo ricorsivo, questo portando ad un'organizzazione gerarchica della memoria. Secondo Ellis (Hakimov,

2021: 58) il *chunking* è il meccanismo cognitivo per cui un insieme di connessioni associative permanenti vengono sviluppate nella memoria a lungo termine. Grazie a questo meccanismo la persona può produrre le sequenze velocemente e con facilità. Secondo Bybee (2010; in Hoffman e Trousdale, 2013: 49) il processo di *chunking* usato frequentemente ci aiuta a capire perché le persone migliorano a livello neuromotorio e cognitivo. Recentemente è emersa l'idea che il comportamento consueto si consolida con l'esperienza. Alcuni studiosi sono dell'opinione che la frequenza con cui le unità linguistiche sono percepite, o prodotte, rafforza le associazioni tra loro nella memoria a lungo termine (Hakimov, 2021: 160). Un altro meccanismo è suggerito da Diessel (Hakimov, 2021: 160) che sostiene che il processo alla base dell'emergere di unità sequenziali è l'automatizzazione, la quale è correlata ma non identica al *chunking*.²⁴ Descrive l'automatizzazione come il meccanismo cognitivo attraverso il quale i processi controllati si trasformano in processi automatici. Quasi tutte le attività sequenziali iniziano come processi controllati, ma vengono spesso trasformati in processi automatici attraverso la ripetizione. Questo è un fenomeno cognitivo molto comune coinvolto in molte attività quotidiane. Secondo lui l'automatizzazione consente alle persone di eseguire attività sequenziali complesse con poco sforzo. Diessel (2016: 20) afferma che questi processi si completano a vicenda; in altre parole, l'automatizzazione è il processo responsabile dell'elaborazione degli elementi sequenziali e della loro trasformazione in unità, mentre il *chunking* è il processo che sta alla base della loro archiviazione e organizzazione nella memoria.

6.6. Innovazioni attraverso la prospettiva cognitivista

La produttività linguistica si basa sul fatto che la lingua evolve nel corso del tempo modificando tutti gli aspetti strutturali che i parlanti, cercando di rispondere a sempre nuove esigenze, usano dei modi sempre più creativi. Essere in grado di comprendere e produrre un grande numero di espressioni linguistiche nuove con le appartenenti regole, significa conoscere una lingua. La produzione linguistica e il processo di selezione sintattica sono considerati come meccanismi complessi d'interazione fra innovazione e ripetizione e il riconoscimento di questo aspetto è fondamentale per comprendere la facilità con cui i parlanti, adulti e bambini, acquisiscono una lingua.

²⁴ Bybee usa i due termini come sinonimi (Hakimov 2021: 160).

Speakers are at once impressively creative and impressively repetitive. The creative aspect of language allows speakers to express new ideas in new contexts. [...] The formulaic aspect presumably simplifies the speakers' task, in that many utterances or parts of utterances can be pulled directly off the mental shelf without having to build every utterance anew, from scratch. (Goldberg, in Hoffman e Trousdale, 2013: 26)

Il passaggio citato spiega che nel linguaggio la ripetitività e la creatività giocano un ruolo importante. L'aspetto creativo del linguaggio consente ai parlanti di esprimere nuove idee in contesti nuovi, mentre la ripetitività semplifica i messaggi che possono essere estratti direttamente dal deposito mentale senza dover essere ricostruiti da capo.

Dabrowska e Lieven (2005: 437-474) in uno studio sull'acquisizione della L1 dei bambini hanno osservato che le nuove espressioni vengono prodotte dagli stessi processi cognitivi. La produzione di nuove parole coinvolge la combinazione delle unità simboliche: giustapposizione e sovrapposizione. Giustapposizione rappresenta la composizione lineare di parole, una dopo l'altra e nella sovrapposizione le unità delle due cornici si sovrappongono.

Es. Le due costruzioni possono generare una terza costruzione.

Shall I PROCESS? + Open that = Shall I open that

Dunque, al centro di questo modello c'è il concetto di costruzioni che emergono con l'esposizione crescente in una lingua. Queste costruzioni vengono allargate attraverso l'aggiunta delle parole ad un'espressione formando così un pezzo congelato (*frozen chunk*) fino a quando la persona non arrivi ad usare uno dei suoi componenti in combinazione con una parola nuova.

Es. *Give me that.*

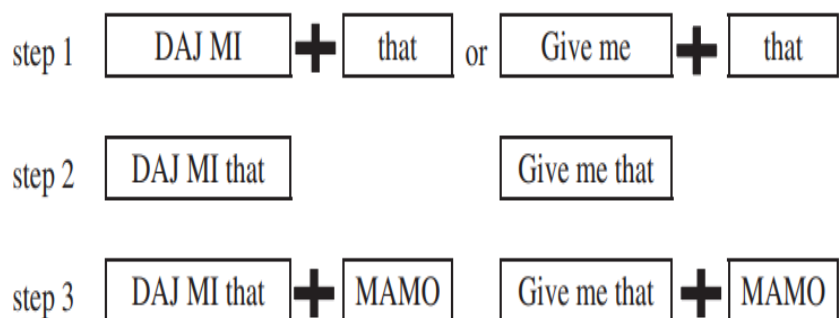
Give me X.

“*Give me*” viene usato come la cornice dentro la quale possono essere usate in un certo posto X varie parole.

Secondo Gaskins (2019) se due lingue sono separate nell'input e la persona continua ad alternare le lingue, la commutazione di codice deve essere evolutiva. Il primo passo di tale uso è la prima istanza di parole provenienti da una o entrambe lingue combinate spontaneamente.

Il secondo passo è la ripetizione che porta col tempo all'automatismo. Il terzo passo combina il pezzo (*chunk*) alle parole riempitive di entrambe le lingue.

Figura 6. Modello usage-based nella commutazione di codice dei bambini studiati da Levin e Dabrowska (in Gaskins et.al. 2019).



6.7. Concettualizzazione

I linguisti concordano sul fatto che le costruzioni astratte sono rappresentate nel lessico e nella grammatica mentale. Il significato si basa sulla concettualizzazione (Langacker, 1991: 117) ma ciò non implica automaticamente che il significato si riferisca ad un concetto. Le persone concettualizzano l'esperienza in molteplici modi e la stessa situazione può essere vista diversamente. Ogni costruzione semantica, lessicale o grammaticale, trasmette una diversa concettualizzazione. Ad esempio:

Tieni il volante.

Afferra il volante.

Queste due frasi evocano diverse concettualizzazioni dell'intensità della situazione. Come sottolineato da Diessel (2017: 10), concettualizzazioni ricorrenti delle esperienze uguali o simili danno origine a convenzioni semantiche.

Nell'ottica del modello *usage-based* la variazione linguistica è concepita come una caratteristica inseparabile dal linguaggio (The "Five Graces Group" et al., 2009). La variazione linguistica è considerata il risultato del processo di ripetizione delle strutture linguistiche nella comunicazione umana ma nei modi diversi. Le lingue spesso hanno due o più modi simili di esprimere lo stesso concetto, magari più moderni. Quando in una lingua esistono costruzioni vecchie e quelle più recenti o moderne, si preferisce l'uso di quelle nuove, mentre le costruzioni o parole più antiche vengono preservate principalmente in contesti ad alta frequenza (Bybee e

Beckner, 2010). In conclusione, le teorie basate sul modello *usage-based* sostengono che l'uso del linguaggio è un processo in cui i parlanti prendono delle decisioni di strategie comunicative in continuazione (Hakimov, 2021: 72). Alla base delle decisioni stanno i processi di selezione e adattamento del linguaggio.

7. Strumenti computazionali per l'analisi linguistica e linguistica dei corpora

In questo capitolo viene tracciata una panoramica concettuale inerente alla linguistica dei corpora e alla linguistica computazionale, con particolare attenzione alla descrizione degli strumenti computazionali usati nell'analisi del corpus istroveneto-croato, utilizzato in questa sede. La linguistica computazionale si occupa dell'elaborazione automatica del linguaggio naturale e ha come compito principale la costruzione di programmi per computer per elaborare parole e testi in linguaggio naturale. La linguistica dei corpora, invece, si dedica allo studio del linguaggio attraverso l'analisi empirica di grandi database di linguaggi naturali, ovvero dei corpora testuali. Tognini-Bonelli riporta alcune definizioni del corpus (2001):

- Il corpus è una raccolta di testi che si presume siano rappresentativi di una data lingua, dialetto o altro sottoinsieme di una lingua da utilizzare per l'analisi linguistica. (Francis, 1992: 17 citando Francis 1982: 7).
- Il corpus è una raccolta di testo in lingua naturale, scelto per caratterizzare uno stato o una varietà di una lingua. (Sinclair, 1991: 171).
- Un corpus è un sottoinsieme di un ETL (*Electronic Text Library*) costruito secondo i criteri di progettazione espliciti per uno scopo specifico (Atkins, Clear e Osler, 1992:1)
- Per corpus si intende una raccolta di campioni di testo scorrevole. I testi possono essere in forma orale, scritta o intermedia e gli esempi possono essere di qualsiasi lunghezza. (Jan Aarts, 1991: 45).
- Un corpus è una raccolta di frammenti di linguaggio che vengono selezionati e ordinati secondo criteri linguistici espliciti per essere utilizzati come campioni di una lingua. (Aquile 1994: 2.1).

Nell'ambito di una teoria contestuale del significato, per avere una maggiore comprensione del materiale studiato, bisogna soffermarsi sulla spiegazione delle differenze principali tra il testo e i corpora. La principale differenza consiste nel fatto che un testo esiste in un contesto comunicativo e unificato, mentre il corpus riunisce molti testi diversi e quindi non può essere identificato con un unico e coerente evento comunicativo. Di seguito riporto la tabella con delle differenze principali tra un testo e il corpus identificate da Tognini- Bonelli (2001: 3):

Tabella 9. Differenze tra il testo e il corpus:

Il testo	Il corpus
Si legge intero	Si legge frammentato
Si legge orizzontalmente	Si legge verticalmente
Si legge per contenuto	Si legge per schemi formali
Viene letto come evento unico	Viene letto come eventi ripetuti
Viene letto come un atto di volontà individuale	Viene letto come un campione di pratica sociale
L'istanza di parole	Fornisce approfondimenti sulla lingua
Evento comunicativo coerente	Evento comunicativo non coerente

La tabella sopra delineata ha lo scopo di differenziare due fonti testuali che possono sembrare simili ma che comportano fasi analitiche molto diverse. Mettendo in relazione i due concetti possiamo dire che il corpus, dato il diverso quadro metodologico di analisi, può fornire degli approfondimenti su un testo specifico.

L'analisi dei corpora comporta l'uso del computer per la ricerca rapida del *database*, che comprende una qualsiasi raccolta di lingua scritta o parlata trascritta. Grazie al computer è possibile con molta facilità assemblare i corpora selezionati secondo le varietà linguistiche, i generi, gli argomenti o specifiche funzioni dell'analisi per una data applicazione. Se prima questo tipo di indagine analitica era riservata solo a pochi studiosi, oggi è disponibile su ampia scala ed è particolarmente rilevante per quelli che operano nel settore linguistico (insegnanti, traduttori, lessicografi ecc.).

Il termine “*spoken language corpus*” (corpus del parlato) si riferisce a una raccolta di campioni linguistici di comunicazione che servono come strumento per comprendere la natura del linguaggio umano e della comunicazione (Kuvač Kraljević e Hržica 2016: 88). Un corpus di questo genere si basa su discorsi spontanei che normalmente contengono esitazioni, ripetizioni o errori non intenzionali. La lingua parlata, infatti, è un fenomeno multidimensionale definito da una varietà di stili, registri e dialetti. Siccome non è possibile catturare tutte le possibili variazioni in una singola conversazione, i corpora sono progettati a seconda degli obiettivi del ricercatore. Tognini-Bonelli (2001: 54) definisce i tre criteri base per la progettazione di un corpus:

1. l'autenticità del linguaggio nel corpus,
2. la rappresentatività di lingua nel corpus,
3. i criteri utilizzati per campionare la lingua.

7.1. TalkBank

La *TalkBank* è il risultato di un progetto di ricerca interdisciplinare in fieri finanziato dalla *National Science Foundation*. L'obiettivo del progetto è supportare la condivisione dei dati e l'accesso diretto a livello di comunità alle registrazioni e alle trascrizioni della comunicazione umana, ovvero del parlato. Il concetto è emerso da due iniziative che si erano rivelate importanti per le rispettive comunità di utenti. La prima è il *Linguistic Data Consortium* che ha pubblicato circa 288 grandi corpora negli ultimi dieci anni e la seconda è il sistema CHILDES che ha costruito un database di trascrizioni delle interazioni genitore-figlio in 20 lingue. *TalkBank* contiene i database in 12 aree diverse di ricerca e offre i seguenti servizi (MacWhinney et al., 2004)²⁵:

1. le linee guida per la condivisione etica dei dati,
2. i metadati e l'infrastruttura per identificare i dati disponibili,
3. i formati comuni e ben specificati per testo, audio e video,
4. gli strumenti per la trascrizione e l'annotazione allineate nel tempo,
5. un formato di scambio comune per le annotazioni,
6. l'infrastruttura basata su rete per supportare una collaborazione efficiente (in tempo reale),
7. l'educazione dei ricercatori all'esistenza di dati condivisi, strumenti, standard e buone pratiche.

7.2. CHAT e CLAN

In passato le analisi comparative dell'acquisizione di una lingua si basavano in gran parte su confronti di dati ottenuti in modi diversi e per ciascuna lingua individualmente. Al fine di ottenere maggiori somiglianze possibili nelle analisi ed eliminare le differenze che emergono confrontando i dati ottenuti con metodi diversi, gli esperti americani della Carnegie-Mellon University di Pittsburgh nell'ambito del progetto CHILDES (*Child Language Data Exchange*

²⁵ I servizi riportati sono stati tradotti dall'inglese.

System)²⁶ hanno sviluppato negli anni '80 il sistema di trascrizione CHAT (*Codes for the Human Analysis of Transcripts*)²⁷, che consente l'armonizzazione dei dati trascritti per qualsiasi lingua. Il materiale linguistico così preparato e trascritto può quindi essere analizzato con l'ausilio di un programma del pacchetto CLAN (*Computerized Language Analysis*)²⁸. Esso comprende, tra gli altri, i programmi di analisi fonologica, morfologica e sintattica. Di seguito verrà presentato brevemente il sistema CHAT e il metodo di analisi dei dati con l'ausilio dei programmi CLAN.

Siccome le trascrizioni delle registrazioni audio (di discorsi orali) vengono eseguite manualmente, di conseguenza i materiali disponibili per le analisi sono limitati, dato che la trascrizione è un processo molto lungo e impegnativo. Al fine di consentire il confronto interlinguistico e la verifica dei propri dati, è stata rilevata la necessità di uno scambio di materiale tra scienziati, ma una delle condizioni necessarie per l'elaborazione dei dati è l'uniformità delle registrazioni. Nel 1984 E. Bates, C. Snow e B. MacWhinney hanno per l'appunto cercato di creare un database con campioni di discorsi spontanei dei bambini (CHILDES) che fossero uniformi. Contemporaneamente sono stati sviluppati i codici di trascrizione audio univoci (CHAT) e i programmi di analisi linguistica computerizzata (CLAN). I sistemi di trascrizione e analisi dei dati sono stati inizialmente utilizzati solo per le ricerche inerenti all'acquisizione della L1, ma con il tempo, grazie alla loro semplicità, accessibilità e applicazione universale (applicabile indipendentemente dal sistema linguistico, scrittura, ecc.) sono stati introdotti in altre aree come, ad esempio, per la ricerca di disturbi del linguaggio, nell'apprendimento delle lingue straniere, per l'analisi conversazionale ecc. Oggi il database CHILDES, il sistema CHAT e il programma CLAN sono indispensabili per chiunque sia coinvolto nello studio e nell'analisi linguistica, tanto che circa 1.400 scienziati provenienti da una quarantina di paesi hanno contribuito alla creazione del database che oggi contiene campioni di una trentina di lingue.

Ogni trascrizione trascritta e codificata in sintonia con CHAT deve iniziare con un'intestazione che contiene informazioni sulle circostanze della raccolta dei dati, ossia:

@Languages: (es. *vec*, *hrv*) – vengono indicate le lingue che vengono usate all'interno della registrazione. Nel nostro caso *vec* indica l'istoveneto e *hrv* il croato.

²⁶ Trad. "Scambio di dati sul linguaggio infantile".

²⁷ Trad. "Codici per l'analisi umana delle trascrizioni".

²⁸ Trad. "Analisi computerizzata del linguaggio".

@Participants: CAR Carla Target_Adult, MAR Target_Adult, ELI Elisa Target_Adult – qui vengono elencati tutti i partecipanti che hanno partecipato alla registrazione trascritta, ovvero alla creazione del file audio. I nomi sono fittizi (per garantire l’anonimato) e si segnano solo le prime tre lettere del nome inventato. Ad ogni nome viene aggiunta la categoria alla quale appartiene, ovvero se si tratta di un bambino, di un giovane o di un adulto.

@ID: vec, hrv|vec-hrv Corpus|CAR||||Target_Adult|| - l’ID è una specie di carta d’identità, in cui troviamo tutte le informazioni di un preciso interlocutore

@Birth of CAR: 05-APR-1989 – indica la data di nascita del partecipante

@Birth of MAR: 05-JUN-1989– indica la data di nascita del partecipante

@Birth of ELI: 26-JUL-1989– indica la data di nascita del partecipante

@Date: 15-APR-2012 – indica la data della registrazione

@Time Duration: 00:16:23 - indica la durata della registrazione

@Situation: Spontaneous speech - indica la tipologia specifica del discorso (nel nostro caso si tratta di conversazioni spontanee)

@Transcriber: Paolo Rossi - indica la persona che ha stilato la trascrizione.

Dopo l’intestazione segue la trascrizione vera e propria, ovvero le enunciazioni ordinate secondo i turni dai singoli interlocutori, segmentate secondo specifici criteri sintattici (MacWhinney et al., 2004). I codici principali di CHAT utilizzati in tal senso sono i seguenti (Vuk, 2016):

- 1) I due punti (:) segnano un suono prolungato
- 2) Il segno [/] viene messo tra le parole che si ripetono mentre il segno [//] viene usato nelle ripetizioni con delle correzioni
- 3) Il segno [///] viene usato nella riformulazione dell’enunciato
- 4) Il segno del punto (.) indica la pausa
- 5) Il segno delle tre x consecutive (xxx) viene usato quando una parola o una sequenza non possono essere decifrate

- 6) Il segno +// indica l'interruzione dalla parte dell'interlocutore
- 7) Il segno +/ indica l'interruzione dalla parte dell'interlocutore e il passaggio del turno all'altro interlocutore
- 8) Il segno +... indica la frase incompleta
- 8) Il segno +“/ indica il discorso diretto
- 9) Il segno +... indica l'interruzione di un enunciato dopo il quale segue un enunciato differente che non è legato a quello che lo precede
- 10) Il segno ++ indica la continuazione dalla parte di un interlocutore dell'enunciato dell'altro interlocutore l'interlocutore
- 11) Il segno +< indica le sovrapposizioni degli interlocutori
- 12) Il segno [*] indica gli errori
- 13) Il segno @s viene usato per indicare le parole commutate

8. Metodologia della ricerca

Per raggiungere l'obiettivo della tesi prefissato, ovvero quello dell'applicazione dell'analisi della commutazione di codice nell'ambito del repertorio bilingue istriano, abbiamo dovuto scegliere un corpus che fosse di libero accesso e rispondesse a tutti i criteri base precedentemente annoverati e descritti da Tognini-Bonelli (2001: 54). In tal senso, abbiamo scelto un corpus la cui progettazione è ancora in fieri, ovvero il corpus del parlato bilingue istroveneto-croato che si sta sviluppando nell'ambito del progetto scientifico “*Approccio multilivello²⁹ all'analisi del discorso nello sviluppo del linguaggio*” / “*Multilevel approach to spoken discourse in language development*” (2018-2022) finanziato dalla Fondazione Croata per la Scienza e coordinato dalla professoressa Gordana Hržica dell'Università degli studi di Zagabria.

Prima di passare alla parte analitica, risulta doveroso spiegare la metodologia che sta alla base della creazione di tale corpus, passare in rassegna le caratteristiche sociodemografiche e sociolinguistiche del campione (parlanti bilingui istriani) e i presupposti per l'applicazione del modello *usage-based* sul corpus.

La progettazione e costruzione del corpus corrisponde ai presupposti scientifici inerenti all'autenticità del linguaggio (le conversazioni registrate avvengono in un contesto non monitorato e spontaneo, in presenza di familiari o amici), alla rappresentatività della lingua (nel nostro caso del parlato spontaneo tra interlocutori bilingui istriani di tre generazioni, provenienti da tutte le aree bilingue della penisola istriana), ai criteri utilizzati per campionare la lingua (relativi alle norme standard delle piattaforme *TalkBank*, ovvero CHAT e CLAN). Le registrazioni sono state raccolte nelle varie parti dell'Istria per assicurare la rappresentatività del corpus in relazione a diverse variabili sociodemografiche e sociolinguistiche (età, genere, luogo di provenienza, madre lingua, ecc.). Per ottenere un corpus il più possibile autentico gli esaminatori sono stati scelti tra gli appartenenti della comunità bilingue locale e sono stati istruiti metodologicamente. La durata delle conversazioni registrate varia dai 15-20 minuti ciascuna e il corpus complessivo è composto da 32 registrazioni, che corrispondono a una decina di ore di materiale registrato. Siccome il corpus verrà reso pubblico mediante la pubblicazione sulla piattaforma *TalkBank* alla fine dell'anno 2022, una volta ultimato il progetto, tutte le registrazioni sono state trascritte e codificate seguendo le norme standard della

²⁹ L'approccio multilivello si riferisce agli approcci teorici applicati al discorso ossia, l'approccio formalistico-strutturalista, l'approccio funzionale e quello cognitivo, che implicano il livello macrostrutturale e quello microstrutturale.

banca dati *TalkBank*³⁰ (CHAT e CLAN), ovvero una delle banche dati più rilevanti contenente corpora dal parlato. Nell'ambito della banca dati *TalkBank* si trovano tre sottobanche dedicate alla raccolta di corpora e allo studio del multilinguismo, ovvero la *Second Language Tutors*, la *SLABank* e la *BilingBank*. Il corpus a cui noi facciamo riferimento farà parte della *BilingBank*, che al momento include i seguenti corpora bilingui (con i rispettivi coordinatori):

Tabella 10.³¹

CORPUS	LINGUE	COORDINATORI
Anadolu	Turco-danese (Anadolu Corpus)	Fatma Özcan, Ilknur Keçik, Jens Normann Jörgensen
BlumKulka/Snow	Ebreo-inglese (in Israele e New York)	Shoshana Blum-Kulka and Catherine Snow
Díaz Collazos	Giapponese-spagnolo (in Cali, Columbia)	Ana Maria Collazos
Eppler	Tedesco-inglese (a Londra)	Eva Eppler
Gardner-Chloros	Greco-inglese (in Israele and New York)	Penelope Gardner-Chloros
Hatzidaki	Greco-francese (in Francia)	Aspa Hatzidaki
Køge	turco-danese (in Danimarca)	Jens Normann Jörgensen
MLE-MPF	Francese-inglese (a Parigi)	Penelope Gardner-Chloros
Silver	Inglese-cinese (a Singapore)	Rita Silver
Miami	Spagnolo-inglese (a Miami)	Margaret Deuchar
Patagonia	Gallese-spagnolo (in Argentina)	Margaret Deuchar
Pilot	Gallese-inglese (in Wales)	Margaret Deuchar
Siarad	Gallese-inglese (in Wales)	Margaret Deuchar

Il corpus del parlato bilingue istroveneto-croato si adatta ai fini della nostra analisi. Infatti, ci siamo proposti di individuare gli schemi ricorrenti sia strutturali che di interazione tra i codici di commutazione (istroveneto e croato) e di valutare l'applicazione del modello *usage-based* nell'ambito di esempi concreti tratti dal corpus. Nella parte analitica, infatti, verranno compresi gli esempi più rilevanti e frequenti del corpus, che riflettono le funzioni individuate da Gumperz (1982) e Baker (2010), in virtù dell'ottica del modello *usage-based*. Inoltre, verranno

³⁰ I corpora contenuti nella *TalkBank* rappresentano delle risorse molto rilevanti per lo studio del discorso spontaneo non monitorato, del discorso bilingue e dell'apprendimento di una seconda lingua/lingua straniera, del linguaggio infantile, del linguaggio delle persone con vari disturbi (balbuzie, disturbi appartenenti allo spettro autistico, demenza, danni cerebrali e trauma cranici, ecc.).

³¹ <https://biling.talkbank.org/access/> - ultima consultazione il 16/6/2022.

spiegate le innovazioni linguistiche secondo i processi di giustapposizione (*juxtaposition*) e sovrapposizione (*superimposition*). L'approccio *usage-based* è l'approccio più recentemente sviluppato per teorizzare la commutazione di codice. La scelta dello stesso non è stata fatta in base alla qualità o efficacia ma semplicemente perché è il meno conosciuto e di conseguenza anche il meno applicato negli studi, soprattutto per quanto concerne la combinazione linguistica istroveneto-croato. L'applicazione del suddetto modello implica la soddisfazione di alcuni criteri:

1. la possibilità dell'analisi quantitativa dei dati: dato che il modello *usage-based* si basa sulla frequenza d'utilizzo, il nome del modello stesso implica l'analisi quantitativa dei dati;
2. la possibilità di utilizzare il metodo nei discorsi spontanei: il corpus analizzato consiste in trascrizioni di conversazioni spontanee e il modello *usage-based* risulta adeguato per questo tipo di corpus;
3. la comparazione che include lingue diverse (nel nostro caso la combinazione istroveneto-croato).

8.1. Caratteristiche sociodemografiche e sociolinguistiche del campione

Il numero dei partecipanti coinvolti nello studio è 91 e tutti hanno consentito la trattazione dei propri dati ai fini della ricerca firmando un consenso predefinito che gli esaminatori hanno consegnato ad ogni parlante (vedi Allegato numero 3). I partecipanti hanno inoltre compilato il “Questionario di (Auto)Valutazione del Parlante Bilingue”³² (vedi Allegato numero 2), contenente domande di natura sociodemografica, socioeconomica e sociolinguistica, ovvero:

1. Dati generali:
 - età
 - sesso
 - luogo di nascita
 - luogo di residenza
 - livello di istruzione
 - professione

³² Il questionario originale si trova allegato alla fine di questa tesi (Allegato 2).

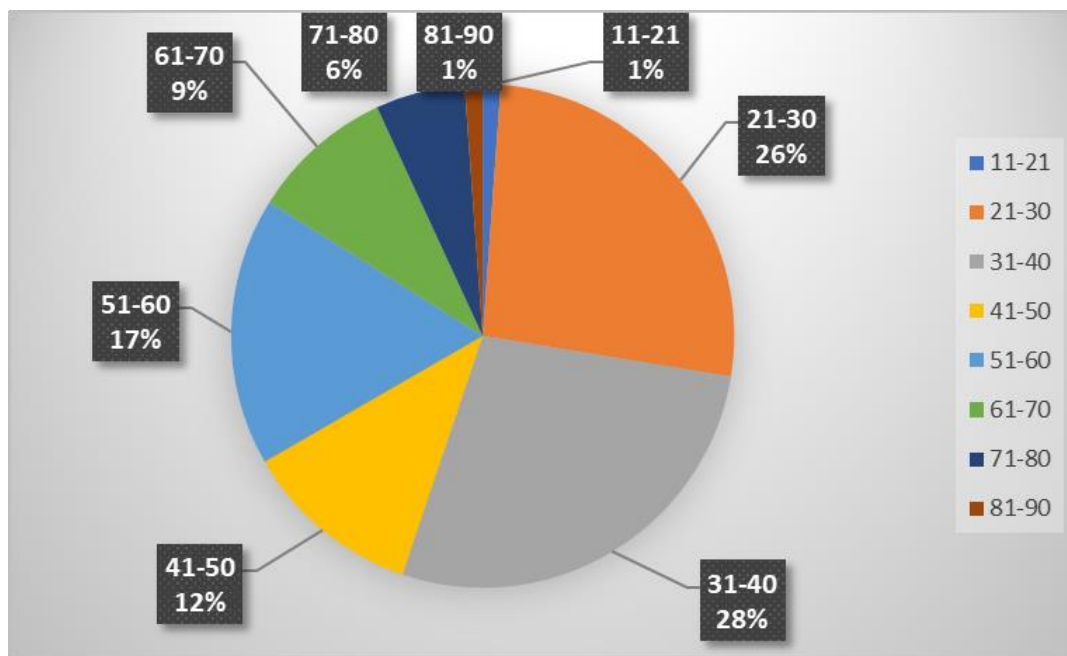
2. Competenze linguistiche:

- a) lingua italiana
- b) lingua croata
- c) dialetto italiano
- d) dialetto croato
- e) lingua o lingue straniere/altre lingue;

- contesto di acquisizione e d'uso delle lingue e età in cui l'acquisizione/apprendimento è stato avviato.

I dati dell'indagine sociolinguistica condotta attraverso il questionario sono stati inseriti nelle tabelle del programma Excel dalle quali sono stati creati i grafici che indicano l'età dei partecipanti, il sesso, il luogo di nascita, il luogo di residenza e l'autovalutazione delle competenze linguistiche per poter confrontare i risultati ottenuti.

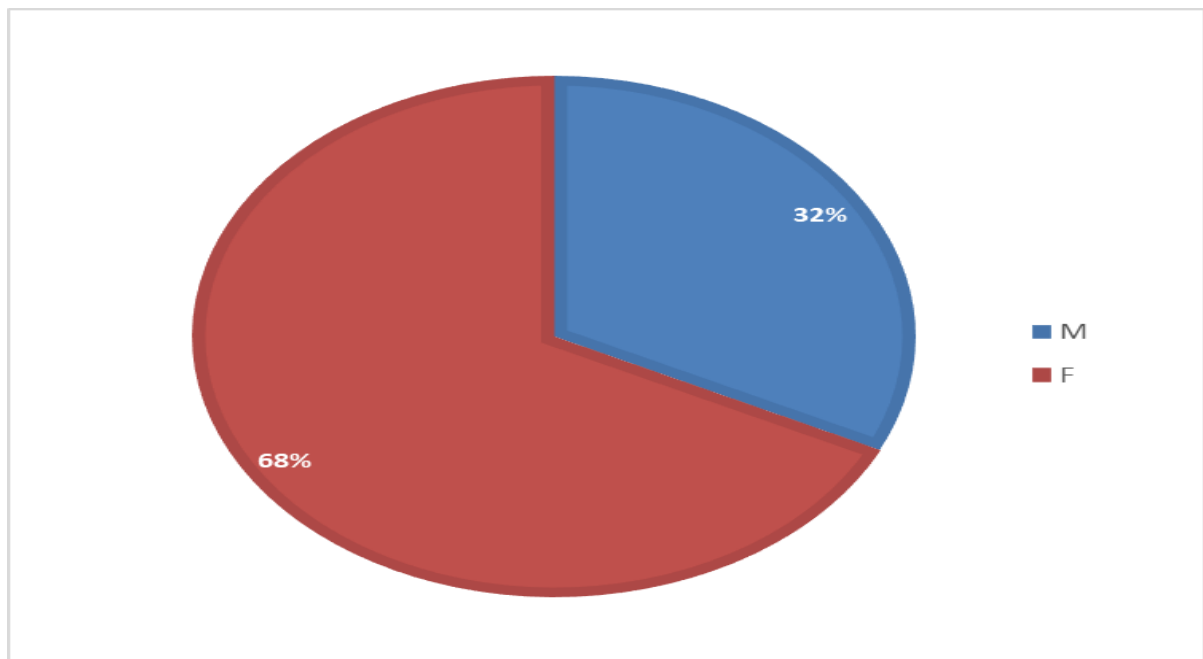
1. Grafico inerente all'età dei partecipanti



Il primo grafico indica le fasce d'età di 91 partecipanti (l'intero campione). Si può notare che la maggior parte dei partecipanti ha tra i 31 e i 40 anni (il 28%); al secondo posto c'è la fascia d'età che va dai 21 ai 30 anni (26%); al terzo posto si colloca il gruppo di persone che hanno tra i 51 e i 60 anni (17%) e poi c'è il gruppo tra i 41 e i 50 anni. Il 6% appartiene alla fascia d'età che va dai 61 ai 70 anni e il 9% invece appartiene ai partecipanti che hanno tra i 61 e i 70 anni. Infine, l'1% degli intervistati ha tra gli 11 e i 20 anni e l'1% tra i 81 e i 90 anni. Possiamo

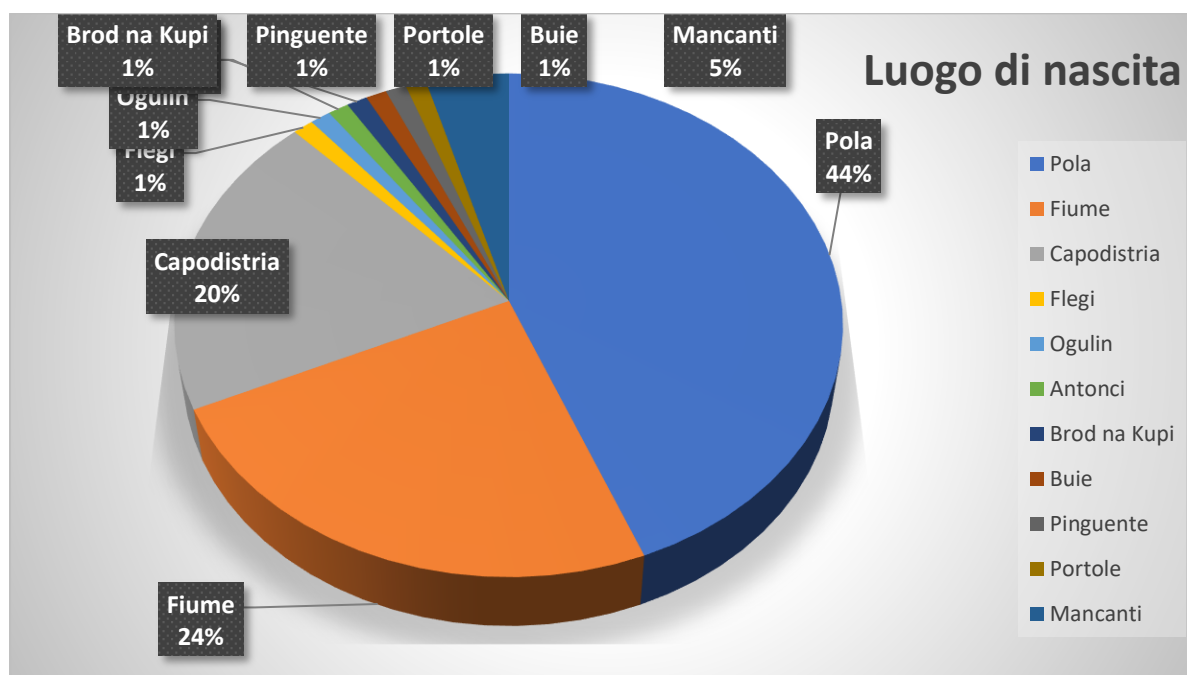
concludere che la maggior parte del campione appartiene alla fascia d'età più produttiva, che va dai 21 ai 40 anni, mentre la percentuale più bassa appartiene alle fasce d'età più vulnerabili, ossia ai bambini e agli anziani. Una possibile ragione può essere che la maggior parte degli esaminatori appartiene altrettanto alla fascia di età della maggioranza, ragione per cui hanno scelto di includere i partecipanti con i quali hanno rapporti di amicizia.

2. Grafico inerente al sesso dei partecipanti



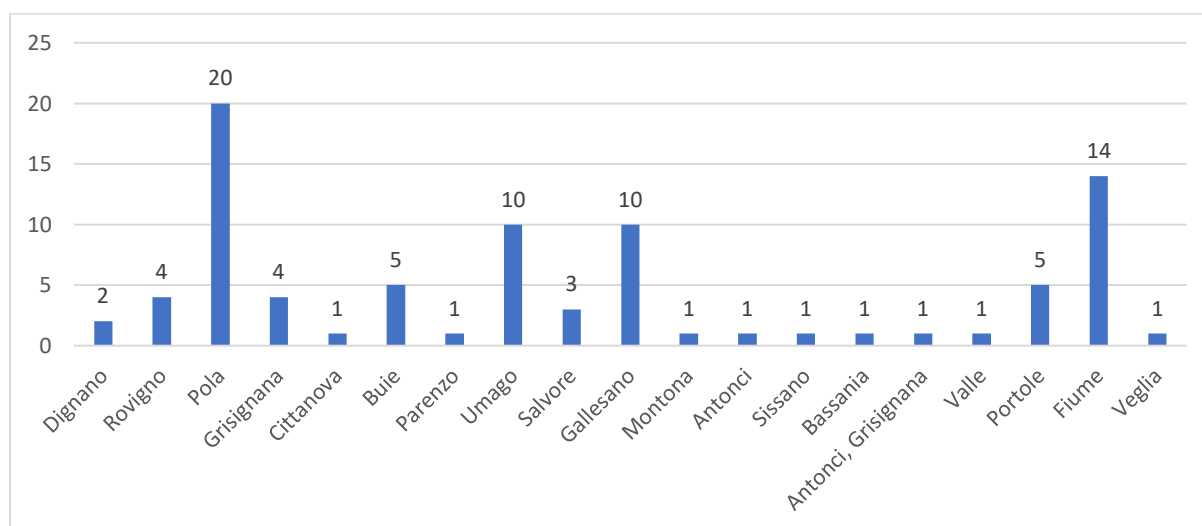
Il secondo grafico rappresenta la suddivisione dei partecipanti per sesso. Possiamo notare che il 68% dei partecipanti appartiene al sesso femminile, mentre il 32% appartiene al sesso maschile, perciò c'è una netta prevalenza del sesso femminile tra i partecipanti.

2. Grafico inerente al luogo di nascita



Nel terzo grafico è possibile osservare la percentuale delle persone che sono nate nelle città sopra indicate. Su 91 partecipanti, il 44 % (40 persone) è nato a Pola, il 24 % (22 persone) a Fiume, il 20% a Capodistria (18 persone) e l'1% (1 persona) rispettivamente nelle restanti località. I dati non sono disponibili per il 5% dei partecipanti (siccome non hanno risposto alla domanda). Possiamo osservare che la maggior parte delle persone intervistate è nata nelle città più grandi come Pola e Fiume.

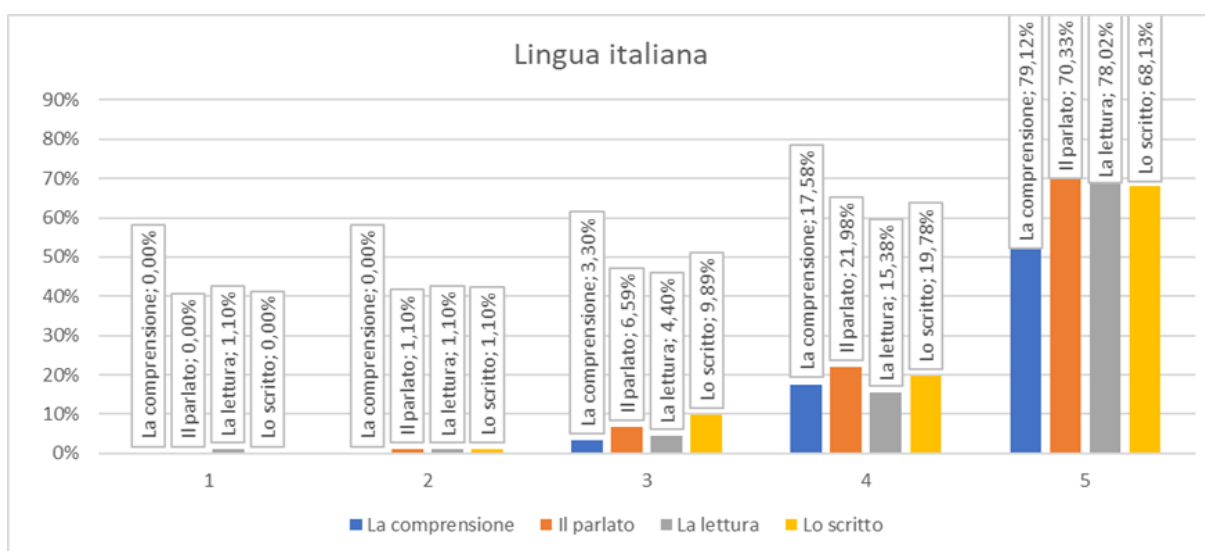
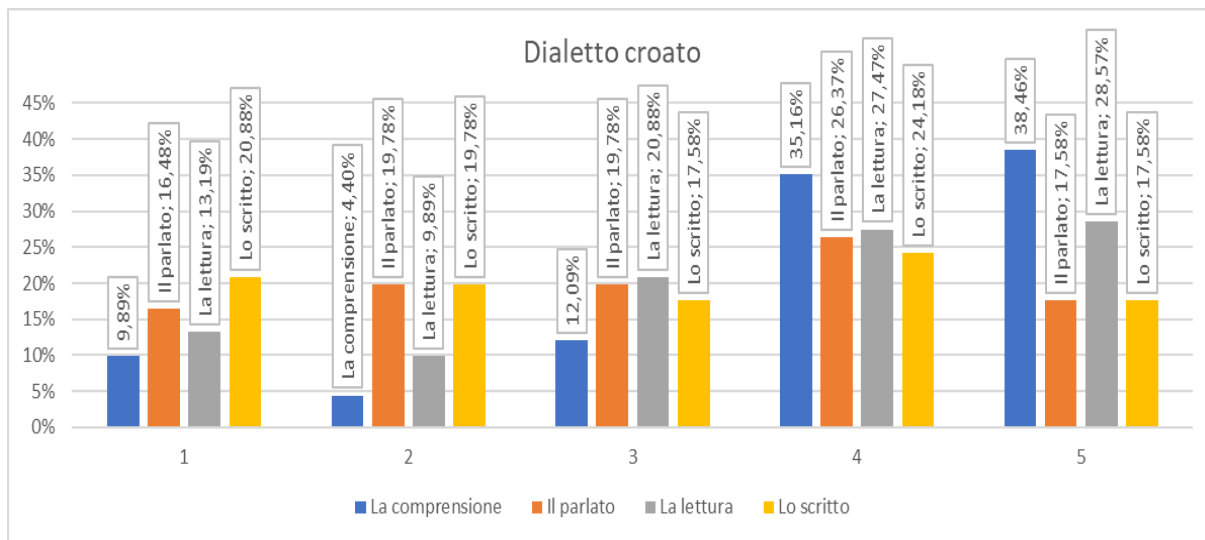
3. Grafico inerente alla residenza

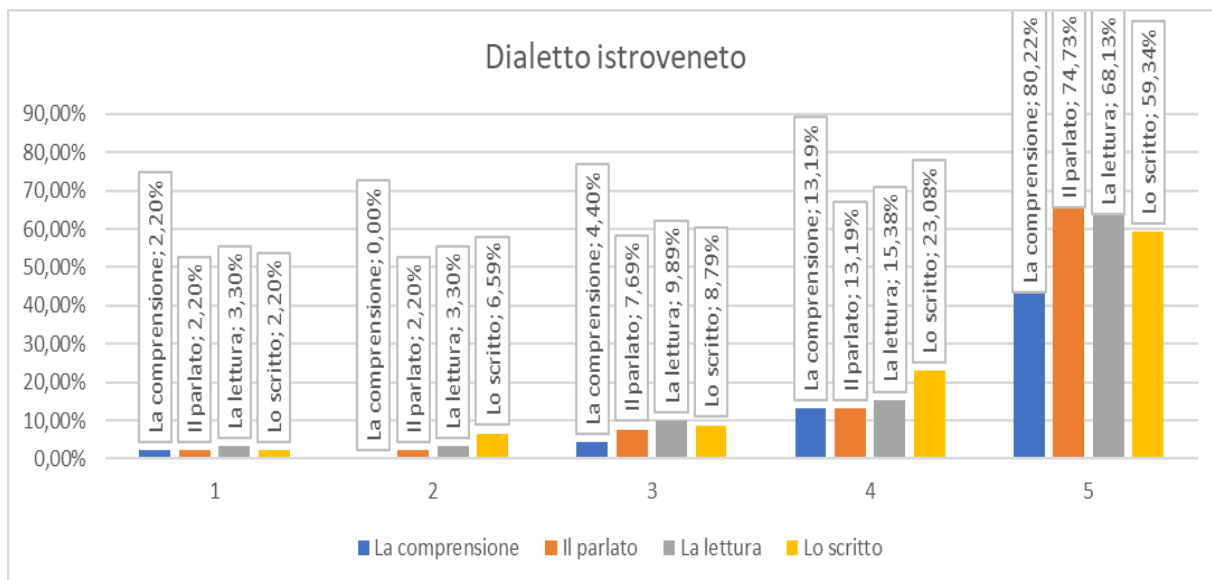


Nel quarto grafico possiamo osservare la residenza dei partecipanti. La maggior parte dei partecipanti abita a Pola e a Fiume. Seguono Galleseano con 10 partecipanti, Umago con 10,

poi Buie e Portole con 5 parlanti. Rispettivamente Rovigno e Grisignana ne contano 4 ciascuna e 2 persone abitano a Dignano. Cittanova, Parenzo, Montona, Antonci, Sissano, Bassania, Grisignana, Valle e Veglia ne contano 1 ciascuno. Quindi possiamo concludere che la maggior parte delle persone intervistate abita nelle città più popolate come Pola o Fiume e questo fatto non sorprende dato che in queste città è presente un numero maggiore degli appartenenti alla minoranza nazionale italiana.

4. Grafici inerenti alle abilità linguistiche dei partecipanti

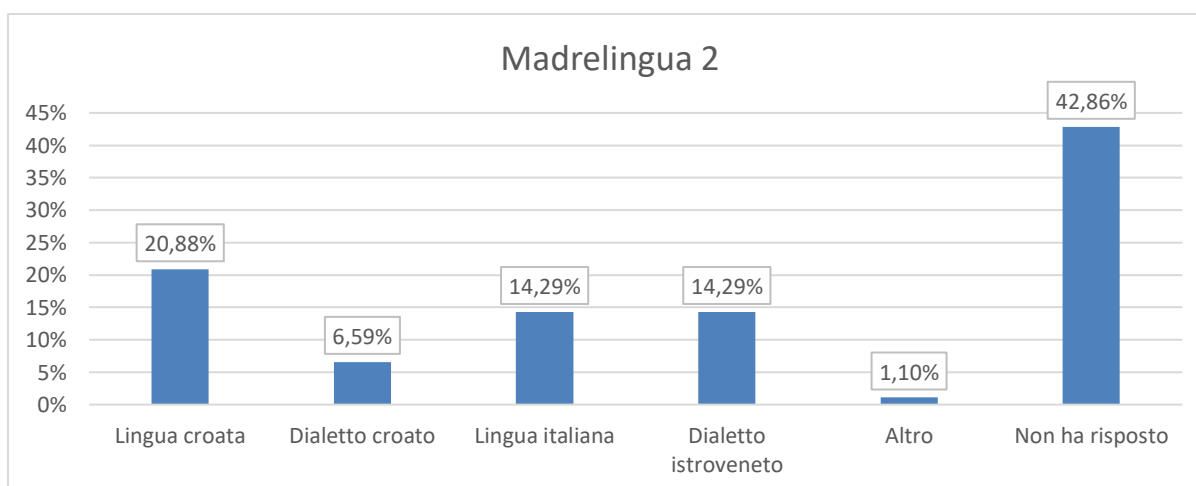
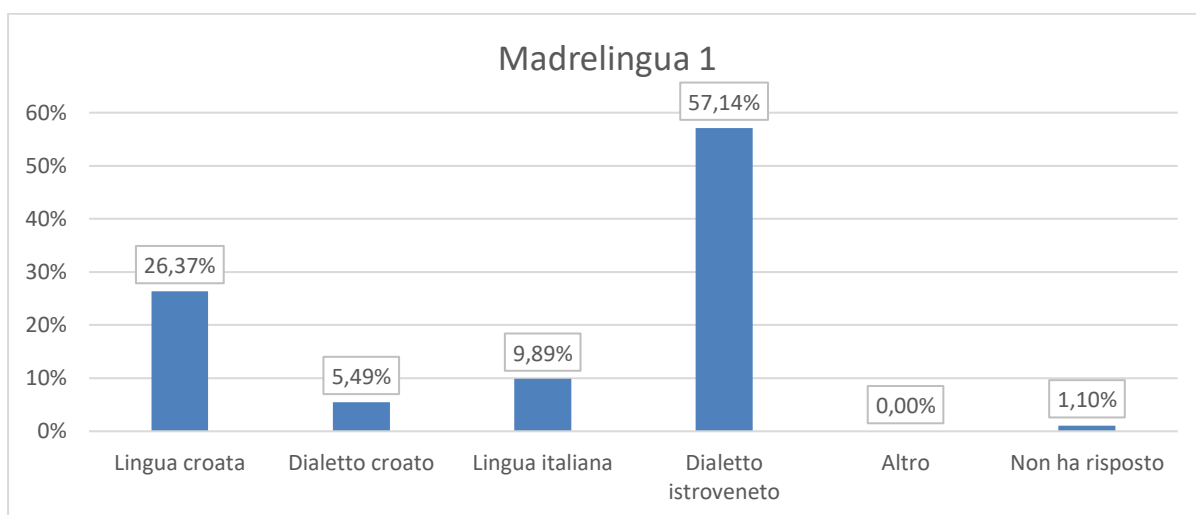


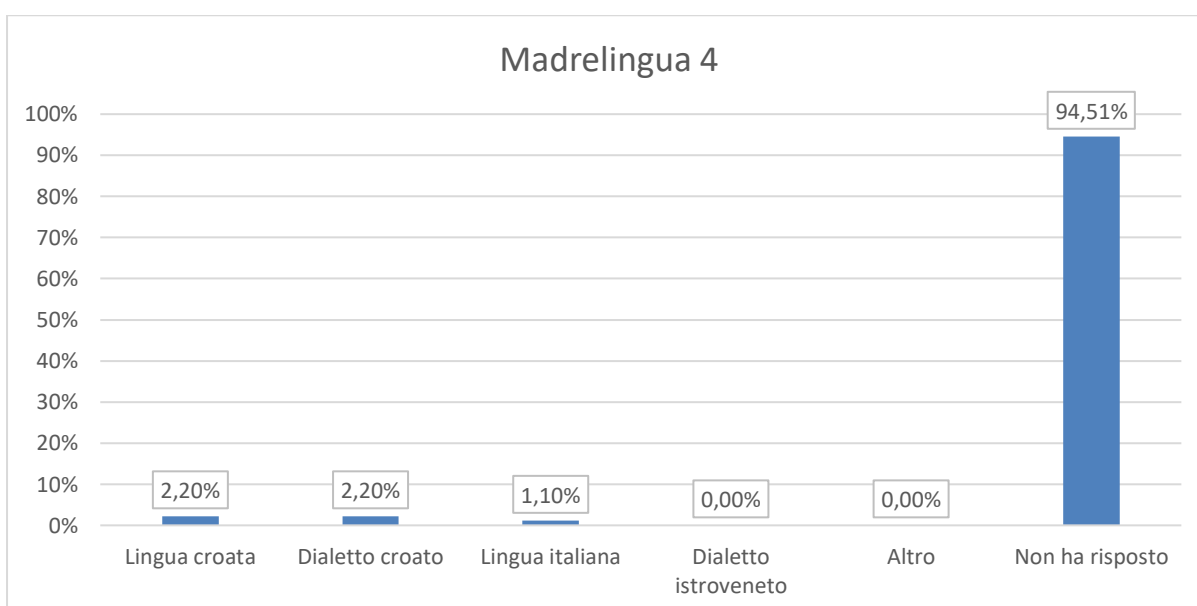
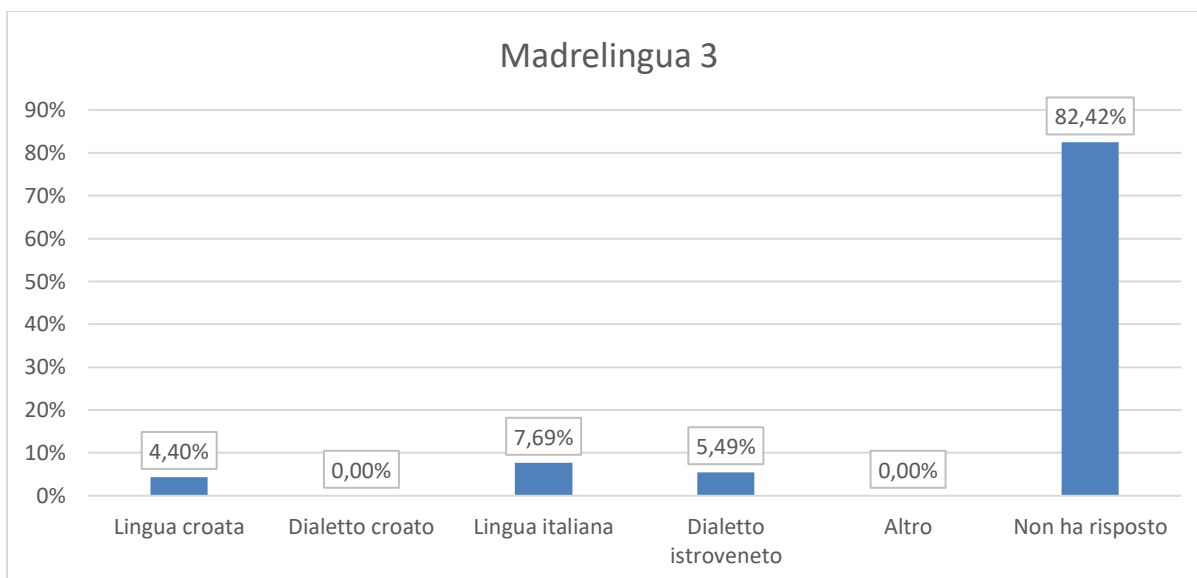


I grafici sopra delineati rappresentano l'autovalutazione dei partecipanti riguardo alle proprie abilità linguistiche. Sono rappresentati quattro grafici che indicano le quattro varietà linguistiche: si è optato per una ripartizione del questionario in quattro competenze linguistiche misurate su scala *Likert*, ricodificate in 1 come abilità minima, fino a 5 come abilità massima. Bisogna prendere in considerazione la relatività dei risultati che potrebbero non rispecchiare la realtà dato che si basano su risposte soggettive che poggiano sull'auto-percezione individuale. Come si evince dal primo grafico, i partecipanti presentano i valori più o meno uniformi per quanto riguarda la conoscenza della lingua croata. La maggior parte dei partecipanti, cioè l'80%, dichiara di comprendere la lingua croata benissimo, il 61% dichiara di parlare il croato benissimo, il 70% è molto competente nella scrittura, mentre il 58% affida il voto 5 all'abilità della lingua scritta. Si può osservare che una grande parte dei partecipanti definisce la propria conoscenza della lingua croata come ottima e che nessuno pensa di essere insufficiente nelle cinque abilità (solo il 2% valuta insufficiente la propria abilità di espressione scritta). Il secondo grafico rappresenta le abilità dei partecipanti ad esprimersi in dialetto croato e si può osservare che i risultati sono più diversificati rispetto al grafico precedente. Qui si nota che la percentuale dei partecipanti che non conosce il dialetto croato sale rispettivamente per tutte le competenze. Il 38% ha riferito di avere un ottimo livello di comprensione del dialetto croato, il 17% del parlato, il 28% della lettura e il 17% della scrittura. Il livello minimo invece è stato riferito per la comprensione dal 10%, per il parlato dal 16%, per la lettura dal 28% e per lo scritto dal 17%. Il terzo grafico rappresenta le quattro abilità nella lingua italiana e i risultati sono simili a quelli ottenuti nel primo grafico inerente alla lingua croata. Il 79% ritiene di possedere un ottimo livello di comprensione, il 70% del parlato, il 78% della lettura e il 68% della scrittura.

La percentuale che percepisce la propria conoscenza della lingua italiana come minima o insufficiente è pari a zero (solo l'1 % riguardo la lettura). Il quarto grafico rappresenta la conoscenza dei partecipanti del dialetto istroveneto. Anche nell'ultimo grafico si può osservare una netta prevalenza del livello ottimo nelle quattro abilità ma è presente, anche se minimamente, una percentuale che ritiene di avere un basso livello nelle competenze. L'80 % definisce il proprio livello di comprensione come ottimo e rispettivamente anche il 74% il parlato, il 68 % la lettura e il 59 % la scrittura.

5. Grafici inerenti alla lingua madre

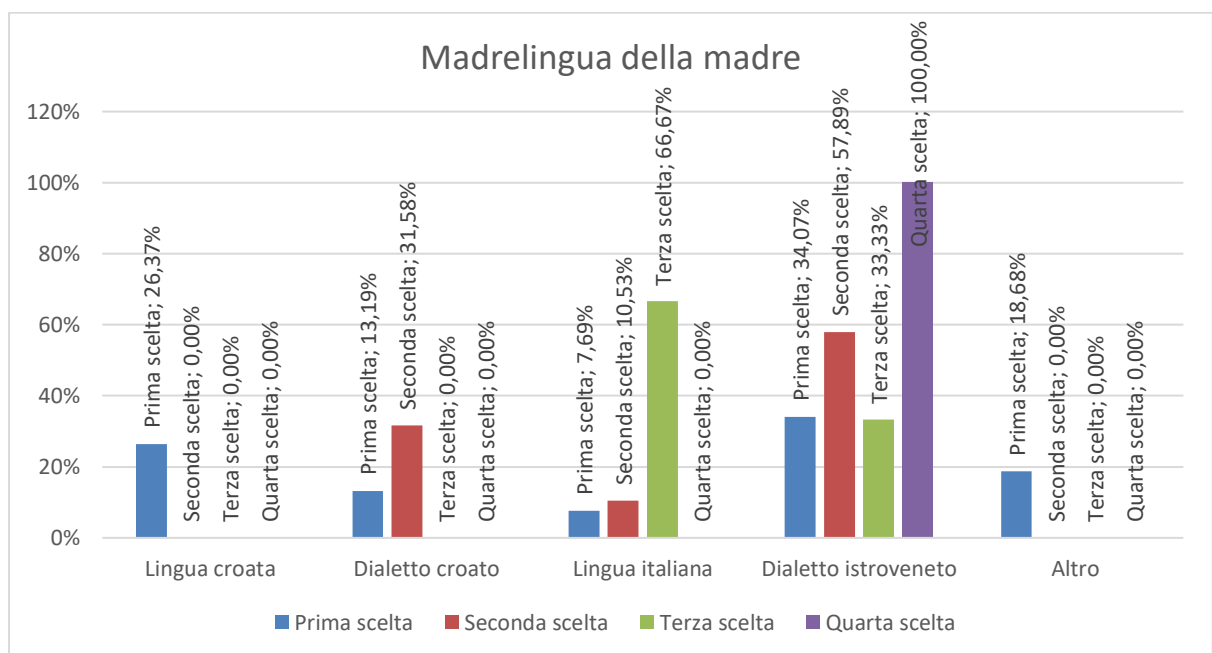


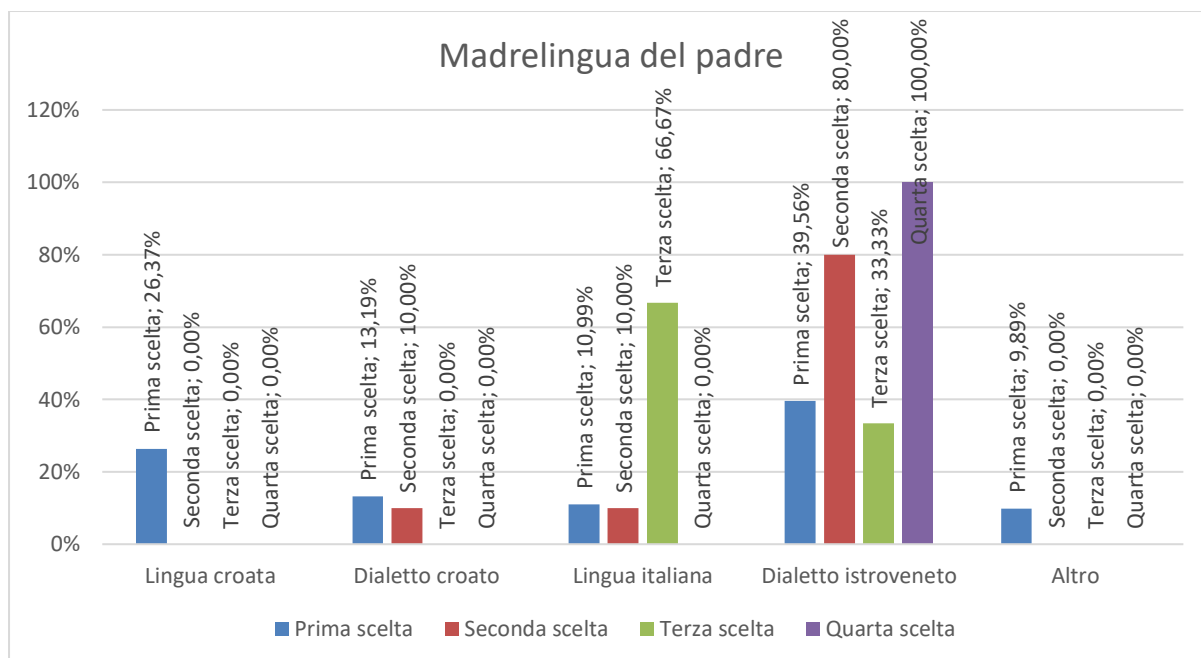


Il primo grafico rappresenta la prima madrelingua acquisita da parte dei 91 partecipanti. Possiamo notare che la stragrande maggioranza dichiara di avere l'istoveneto come prima lingua o madrelingua 1, cioè il 57% (che corrisponde a 52 persone), mentre il 26,37% o 24 persone dichiarano di avere la lingua croata standard come prima lingua. Infine, 9 partecipanti (9,89%) dichiarano la lingua italiana standard come prima lingua e 5 invece il dialetto istriano (5,49%). Il 20,88% (19 partecipanti) dichiara di avere la lingua croata come seconda lingua, la lingua italiana e il dialetto istoveneto occupano la stessa posizione nel grafico con il 14,29% ciascuna (13 partecipanti), mentre il dialetto croato è dichiarato da 6 persone (6,59%). Solo una persona ha riportato la lingua slovena come seconda lingua. In totale 39 partecipanti non hanno risposto alla suddetta domanda (il 42,86% del campione complessivo). Il terzo grafico presenta le risposte sull'eventuale esistenza di una terza lingua o lingua straniera. I risultati ottenuti

indicano una bassa percentuale della conoscenza di una terza lingua. Il dialetto istroveneto è stato annoverato da 5 persone (5,49%), la lingua italiana da 7 persone (7,69%), la lingua croata da 4 persone (4,40%), mentre 75 persone non hanno risposto alla domanda. Alla domanda sull'eventuale quarta lingua 86 persone (94,51%) non hanno risposto, confermando di non avere una quarta lingua nel proprio repertorio, mentre 5 hanno scelto la lingua croata (2), la lingua italiana (2) e il dialetto croato (1). Alla luce dei risultati presentati possiamo concludere che la maggior parte delle persone, cioè 65 con esattezza dichiarano l'istroveneto come lingua prima o lingua seconda e 22 persone dichiarano l'italiano standard come lingua 1 o lingua 2. Dunque, in totale 87 persone su 91 indicano una varietà italoфона da sola o in combinazione con una varietà croatofona.

6. Grafici inerenti alla lingua madre dei genitori





Gli ultimi due grafici presentano i valori inerenti alla madre lingua dei genitori dei partecipanti. Data la diversità linguistica ed etnica in Istria, la madrelingua dei genitori può differire da quella dei figli. Dai risultati ottenuti possiamo dedurre se il canale della trasmissione generazionale della lingua è paterno o materno. Osservando i due grafici possiamo notare che i risultati sono molto simili con valori poco più superiori nel grafico inerente alla madrelingua del padre. I risultati ottenuti ci permettono di constatare che il dialetto istroveneto è stato trasmesso in quasi uguale misura dal canale sia materno che paterno.

9. Applicazione del modello *usage-based*: livello strutturale

Nella sezione analitica (capitolo 9) ci focalizzeremo sull'analisi delle caratteristiche strutturali della commutazione di codice nel contesto del modello *usage-based*. Prima di passare all'analisi del corpus, premettiamo che uno dei limiti principali della ricerca svolta si riferisce al tempo limitato che abbiamo avuto a disposizione per condurre la ricerca. Infatti, la complessità e composità del fenomeno della commutazione di codice meriterebbe una trattazione molto più approfondita e longeva, che tenga conto delle particolarità di ogni singolo parlante bilingue che ha apportato alla creazione del corpus analizzato, ovvero all'analisi dettagliata dell'autobiografia sociolinguistica di ogni parlante specifico, e di un'indagine qualitativa approfondita focalizzata sull'analisi delle spiegazioni soggettive dei parlanti inerenti ad ogni singola scelta linguistica commutata, per poter affermare con certezza quale sia il livello di competenza linguistica/comunicativa dei parlanti in ognuna delle lingue del loro repertorio, ossia la specifica funzione discorsiva implicata nelle sequenze discorsive enunciate in modalità bilingue. Pertanto, in questa sede, ci limiteremo ad un'indagine molto più superficiale, ossia rivolta all'articolazione della commutazione di codice in termini strutturali analizzabili mediante i postulati dell'approccio *usage-based*, con qualche appunto esiguo inerente agli esempi più ricorrenti e palesi della funzionalità discorsiva riscontrata nel corpus. Sarebbe senz'altro opportuno allargare e approfondire tale indagine (quale punto di partenza) in termini di ricerca sociolinguistica qualitativa in un futuro prossimo. In tal senso, non avendo condotto un'indagine qualitativa preliminare approfondita, a livello di generalizzazione dei dati sociolinguistici ottenuti nel questionario compilato dai bilingui del corpus, la loro competenza inerente all'istoveneto e al croato viene considerata come alta o molto alta (siccome si tratta delle loro lingue materne o L1/L2 utilizzate in numerosi domini d'uso quotidianamente). In tal senso, le porzioni discorsive commutate rappresentano sequenze stabilite (Poplack 1980) sulla base della frequenza d'uso e del grado di integrazione (Dal Negro e Guerini, 2007: 44), ossia integrate dal punto di vista morfologico, fonologico e sintattico e ampiamente e volutamente usate quotidianamente. Per tracciare un parallelo con la terminologia utilizzata da Myers-Scotton (Carli: 132), possiamo dunque parlare dei cosiddetti prestiti culturali (*cultural borrowings*) e prestiti essenziali (*core borrowings*). Infatti, come afferma Myers-Scotton un'innovazione linguistica può essere classificata come commutazione che diventa prestito se usata frequentemente (Dal Negro e Guerini, 2007: 44). Pertanto il nostro corpus appare essere ricco dei cosiddetti prestiti culturali (*cultural borrowings*) che servono per esprimere concetti

per mancanze lessicali nell'altra lingua e prestiti essenziali (*core borrowings*), che si usano per esprimere una particolare sfumatura espressiva o per aumentare il prestigio di ciò che designano.

9.1. Dominio dei sintagmi sostantivali

Nella commutazione di codice i sostantivi sono la classe più frequentemente usata rispetto ad altre classi di parole. Molti linguisti infatti collocano i sostantivi in cima alle gerarchie dei prestiti (Hakimov, 2021:132). Il motivo per cui i nomi occupano una posizione così importante in tutti i processi di alternanza linguistica è la loro natura semantica. Secondo van Hout e Muysken (1994: 39-62) i sostantivi sono molto usati perché sono collegati a contenuti culturali. Backus sostiene che la ragione principale sta nel fatto che alcuni sostantivi denotano dei concetti molto specifici per i quali non esiste la parola adatta nella lingua ospitante (Hakimov, 2021: 133). I sostantivi hanno tra l'altro un'elevata libertà sintagmatica, il che significa che sono meno legati ad altre parole nella frase e possono quindi essere inseriti in qualsiasi tipo di frase. I tre seguenti passaggi evidenziano la scelta che riguarda l'uso del sostantivo come categoria meno vincolata nella commutazione di codice e sono analizzabili secondo i principi del modello *usage-based*. La struttura delle frasi in tutti gli esempi è presentata dalla combinazione delle costruzioni sovrapponibili, ovvero quelle che includono gli elementi fissi pronunciati in istroveneto e le parole pronunciate in croato che sono sostituibili.

Doc. 10.

*CAR: te speta la bira davanti **el market@s**.

Trad. Ti aspetta la birra davanti al negozio.

La parola **“market”** potrebbe essere sostituita da qualsiasi altra parola che denoti un referente davanti al quale si possa “aspettare”. Potenzialmente, il parlante avrebbe potuto scegliere un'altra parola, come *trgovina*, *dućan*, *butiga*, *bottega*, *negozio*. Sicuramente la scelta è connessa allo specifico contesto fisico (ad esempio la dimensione del supermercato, le insegne e denominazioni visibili, ecc.). Nell'ambito di cornici sovrapponibili come questa, la base fissa è la sequenza che precede la parte sostituibile. Possiamo notare che la parte fissa può essere composta dall'uso dell'articolo istroveneto davanti al sostantivo croato. Gli esempi mostrano un'elevata dipendenza dall'articolo determinativo italiano “il” (*el* in istroveneto), siccome gli articoli non esistono nella lingua croata. Gli articoli sono presi come facenti parte delle cornici

perché rappresentano una categoria di classe chiusa cioè non vengono mai usati da soli. Uno dei principi sui quali poggia l'approccio *usage-based* riguarda la riduzione fonetica o lessicale delle parti commutate. Infatti, i linguisti affermano che spesso le parole ad alta frequenza sono soggette a tali tipi riduzione. Nel prossimo esempio la parola “*stan*” è più breve dell'equivalente “appartamento” o “quartier”; la parola “*faks*”³³ è più breve di “facoltà”, mentre la parola “*x-ica*” risulta più breve di “*tessera universitaria*”.

Doc.11.

CAR: e devo preparar tuto **el stan@s** per voi che vegnì.

Trad. E devo mettere a posto l'appartamento per voi che venite.

Doc. 10.

*ELI: mi me go cominciado con ti farne družit(i)**@s** quel giorno che ti son
vegnuda sola davanti **el faks@s**.

Trad. Io ho cominciato a frequentarti quel giorno che eri venuta davanti alla facoltà.

Doc. 14.

*DIN: e ma no(n) ti dovevi anche far **la x-ica@s**?

Trad. Ma non dovevi fare anche la tessera universitaria?

9.2. Dominio dei sintagmi aggettivali

Nel contesto del modello *usage-based* possiamo notare che nei seguenti passaggi tutti gli aggettivi (spesse volte aggettivi croati che si riferiscono ad un sostantivo in istroveneto) hanno un significato abbastanza astratto e generico a differenza della costruzione aggettivo + sostantivo croato. La frequenza d'utilizzo degli aggettivi astratti è alta e questo tipo di aggettivi viene applicato regolarmente in croato. Per quanto riguarda la struttura, gli aggettivi astratti sono posizionati più frequentemente nella posizione finale della frase, anche se non si può escludere un posizionamento diverso. La struttura delle frasi emerge dalla sovrapposizione delle cornici in cui l'aggettivo può essere sostituibile a qualsiasi altro aggettivo.

³³ Abbreviazione della parola *fakultet*, usata nel gergo giovanile croato.

Doc. 11.

*CAR: se ti ti son d'acordo perché jednostavno@s ne sarà troppo **naporno@s**.

Trad. Se tu sei d'accordo perché semplicemente ci sarà troppo faticoso.

Doc. 64.

*AUR: ma mi so che ti ti son un pò [//] un pochetin **skeptična@s**.

Trad. Ma io so che tu sei un pò scettica.

Doc. 64.

*ELE: proprio me piaceva perché almeno go: [/] go fato iskoristit(i)@s el giorno.

*ELE: a@s difemo **kvalitetno@s**.

Trad. Proprio mi piaceva, perché almeno ho passato una bella giornata diciamo di qualità.

Doc. 64.

*ELE: però su internet je un pochetin **zeznuto@s** perché dopo no(n) ti son

sicuro per el: [//] per la mifura (.).

Trad. Però su Internet è un po complicato perché dopo non sei sicuro per la misura.

Un caso molto frequente e molto interessante riguarda l'uso di “*ful*” nell'ambito del discorso in istroveneto. Strutturalmente, il modello che si verifica risulta essere il seguente: “*ful*” + “aggettivo in croato”. In termini del modello *usage-based*, il posto sostituibile è rappresentato dell'aggettivo.

Doc.10.

*CAR: che semo tuti **ful@s bliski@s**.

Trad. Che siamo molto vicini. (parlando degli zii)

Doc. 10.

*CAR: savè quando che je le ferie o no(n) savè?

*CAR: perché ve fa **ful@s mal**?

Trad. Sapete quando cominciano le ferie o no?

perché vi fa molto male?

Doc. 10.

*CAR: perché mi in quel tempo iero ancora **ful@s efulberante**.

Trad. perché io all'epoca ero molto esuberante

*ELI: me iera **ful@s super**.

Trad. Mi sono trovata molto bene.

La parola *ful* è stata presa dall'inglese (*full*) e si presenta come un *nonce loan*, cioè prestito occasionale, anche se nel caso dei giovani bilingui istroveneto-croati ne possiamo notare un'alta frequenza d'utilizzo. Viene usata sempre davanti all'aggettivo, siccome serve da base nella costruzione sovrapposta dove l'aggettivo può essere intercambiabile. La gamma di aggettivi che si verifica nella posizione finale è limitata e la categoria semantica a cui appartengono è vincolata.

Numerosi studi hanno dimostrato che, sia nei casi di enunciazione mistilingue che in quelli di commutazione di codice, la combinazione aggettivo + sostantivo appare regolarmente indipendentemente dall'ordine usato nelle due lingue (Hakimov:117). Lo confermano pure gli esempi che seguono (aggettivo+sostantivo nell'ambito del sintagma aggettivale). Notiamo pure la presenza della giustapposizione (quale tratto dell'approccio *usage-base*), in cui la sequenza è lineare: ad esempio la parola "**terenska**" implica la parola "**nastava**", mentre la parola "**dopunska**" implica la parola "**isprava**".

Doc. 14.

*ENR: go anche la **terenska@s nastava@s**.

Trad. Ho anche il lavoro sul campo.

Doc.14.

*ENR: scrivi che me servi el diploma e la **dopunska@s isprava@s** che go finido el preddiplomski.

Trad. Scrive che mi serve il diploma e il supplemento al diploma che ho finito lo studio triennale.

Nel corpus è altrettanto presente un alto numero di costruzioni composte da aggettivi italiani e sostantivi croati, secondo il processo di sovrapposizione, che vede la base composta dall'aggettivo dimostrativo e la parte intercambiabile in croato. Di seguito si riportano alcuni esempi:

Doc 13.

NO*MAR: ti ga **(que)sto kiosk@s**.

*MAR: e un unico numero de cioè al mefe i ciapava.

Trad. C'è questa edicola e un unico numero di Cioè avevano al mese.

Doc. 61.

*ELE: +< dove ti me ga [///] dove ti ga comprado **quela: tegla@s** per beber?

Trad. Dove mi hai, dove hai comprato quella pinta per bere?

Doc.61.

*ANT: ti ga [///] ti te compri **quela spužva@s** +...

Trad. Hai... comprato quel materassino.

Doc. 64.

*ELE: e cosa ti me diševi per **(que)sto trening@s**?

Trad. E cosa mi dicevi di questo allenamento?

9.3. Dominio dei sintagmi preposizionali

Ai sintagmi preposizionali è stato dato molto spazio nell'ambito del modello *Matrix Language Framework* (MLF) di Myers-Scotton (Hakimov, 2021: 178). Come è stato esposto nei capitoli precedenti, il principio principale di questo modello è che uno dei linguaggi in contatto fornisce la struttura di base della frase, mentre il ruolo dell'altra lingua è limitato alla fornitura di elementi insertivi. Nel contesto del modello *usage-based*, l'occorrenza di una preposizione in un discorso può portare a una selezione ripetuta di questa preposizione nel discorso successivo, a condizione che venga mantenuta la compatibilità semantica. In altre parole, una volta prodotta una preposizione, è molto probabile che venga nuovamente selezionata in un altro discorso (Hakimov, 2021: 195). Negli esempi che seguono si nota l'uso dei modelli "preposizione articolata istroveneta" + "sostantivo croato". Siccome per l'istoveneto nello scritto non vengono utilizzate le geminate, graficamente la preposizione risulta essere staccata dall'articolo (*su la* per *sulla*; *de la* per *della*, ecc.).

Doc.10.

*ELI: mi lo gavevo visto la prima volta per la festa de l'uva.

*ELI: che (e)l cantava.

*CAR: **su la bina@s** (e)l iera in marzo.

Trad. Io lo avevo visto la prima volta alla festa dell'uva.

Che cantava sul palcoscenico in marzo.

Doc. 20.

*LUC: +" va vedi le foto che go **sul mobitel@s**.

Trad. Va a vedere le foto che ho sul cellulare.

Doc. 10.

*CAR: iero tuta ful@s xxx **de la srednja@s**.

Trad. Io ero molto xxx della scuola media superiore.

Doc. 10,

*ELI: e là faria la parte **del matičar@s**.

Trad. E lì farei la parte del notaio pubblico/pubblico ufficiale.

Inoltre, per ritornare al dominio dei sostantivi, le scelte citate negli esempi esposti dimostrano una preferenza per la riduzione lessicale, in quanto la parola “*bina*” è più breve rispetto alla parola “palcoscenico”, la parola “*mobitel*”, a parte essere più breve della parola cellulare, è foneticamente vicina all’inglese “*mobile phone*” (scelta lessicale più moderna); la parola monolessicale “*matičar*” è più ridotta rispetto all’equivalente italiano (sequenza plurilessicale “notaio pubblico” o “pubblico ufficiale”) e la stessa cosa vale per la parola “*srednja*” (più ridotta rispetto all’equivalente plurilessicale italiano “scuola media superiore”).

9.4. Dominio dei sintagmi verbali

Nel corpus sono molto frequenti i verbi croati all’infinito che sono preceduti dal verbo *far* istroveneto, che assume diverse forme nell’ambito dei paradigmi verbali e innesca l’attivazione di verbi specifici in croato. I verbi più probabili che vengono innescati, possono implicare a loro volta la sovrapposizione delle cornici. Data l’alta frequenza d’uso di tale costrutto linguistico, si potrebbe definire tale sequenza ripetitiva, sempre secondo i postulati dell’approccio *usage-based* come una potenziale sequenza ripetitiva che con il tempo potrebbe diventare un’unità a sé stante nell’abito della modalità bilingue (ad esempio Langacker, 1987: 59; Bybee e Scheibman, 1999: 581).

Doc.10.

*ELI: mi me go cominciado con ti **farne družit(i)**@s quel giorno

Trad. Io ho cominciato a frequentarti quel giorno.

*ELI: che **ti te ga fato riješit(i)**@s de ela.

Trad. Che ti sei sbarazzata di lei.

Doc. 20.

*VER: (a)deso **me go fa(to) premjestit(i)**@s perché xxx.

no(n) go posto.

Trad. Adesso mi sono trasferita perché xxx.

Nel seguente passaggio possiamo notare due costruzioni che formano un unico schema. Una è composta dal verbo fare + il verbo all'infinito (questa volta inglese) e l'altra dalla preposizione + da un sostantivo. In questo caso si tratta dell'allargamento dello schema grazie al *chunking* che, secondo il modello *usage-based*, implica un insieme di connessioni associative. In altre parole, le connessioni associative tra gli elementi di una sequenza diventano più forti e riflettono direttamente la probabilità basata sulla frequenza della loro ricorrenza (Bybee e Scheibman, 1999: 581).

Doc. 10.

*ELI: te **farò pasteat su face** che ti vedi.

Trad. Te lo incollo su Facebook così lo vedi.

Come abbiamo già ribadito, nel corpus prevale l'istoveneto (rappresenta la lingua base) e numerosi sono i passaggi di codice dall'istoveneto al croato. Ci sono comunque dei rari momenti in cui la commutazione avviene pure nella direzione opposta. Sarebbe opportuno analizzare dettagliatamente pure questo aspetto in un futuro prossimo. Nel primo esempio riportato sotto possiamo notare che la sequenza commutata rispetta le regole grammaticali della lingua ospitante. La parola istoveneta "lezion" (lezione), che corrisponde alla parola croata "*sat/predavanje*", è preceduta dalla preposizione croata *na* (andare a lezione) e non assume nessuna desinenza, mentre la parola "palestra", preceduta dalla preposizione *u* (in palestra) assume la desinenza *-u* dell'accusativo croato. Pure la parola "cena" assume il morfema grammaticale *-u*, tipico nell'accusativo dei sostantivi femminili. Nelle costruzioni citate, seppur si tratti della direzione commutazionale opposta (dal croato all'istoveneto), possiamo notare che le preposizioni vengono usate come parti fisse nella sovrapposizione delle cornici e che le parole istovenete in tal caso risultano essere sostituibili e intercambiabili.

Doc. 11.

MAR: ok (.) super@s ali@s rađe@s ćeš@s ici@s na@s **lezion nego@s u@s

palestru@ita+hrv.

*ELI: zoveš@s svih@s na@s **cenu**@ita+hrv.

*ELI: [- hrv] ljudi si ne razbijaju glavu šta će obući(i).

Trad. Inviti tutti a cena. La gente non deve preoccuparsi del vestiario.

10. Applicazione del modello *usage-based* e realizzazione delle funzioni discorsive

Come abbiamo già menzionato, la commutazione di codice viene usata per realizzare numerose funzioni conversazionali tra i parlanti bilingui, rivestendo un valore pragmatico-comunicativo. Nei sottocapitoli che seguono vengono riportati alcuni degli esempi ricorrenti delle funzioni conversazionali riscontrate da Gumperz (1982) e Baker (2010) e vengono evidenziati i postulati del modello *usage-based* implicati.

10.1. Citazioni

Una delle funzioni conversazionali che molto frequentemente viene realizzata tramite sequenze commutate è quella della citazione di porzioni di discorso enunciate nell'altra lingua e pertanto riportate integralmente nella stessa. Questa scelta è correlata al principio di fedeltà nel riportare il messaggio nella varietà linguistica originale. Negli esempi che seguono si nota che la sequenza pronunciata in croato viene introdotta dai modelli tipici che introducono il discorso: "mi te go dito" (io ti ho detto) "e ghe faso" (e le faccio), "ela la ga dito" (lei ha detto), "el me ga dito" (lui mi ha detto). Secondo il modello *usage-based* questi passaggi sono esempi di sovrapposizione perché la sequenza che precede la citazione serve da cornice, mentre la parte citata è quella che può essere sostituita da qualsiasi combinazione di elementi.

Doc. 10.

*ELI: +" poderiimo anche andar magnar la piza.

*ELI: iera tipo le due de dopo pranzo.

*ELI: ela la ga dito +"/.

*ELI: +" ma sì [/] sì kao@s (.) može@s.

Trad. Potremmo anche andare a mangiare la pizza. Erano tipo le due del pomeriggio.

Lei ha detto: "Ma sì, come per dire va bene. "

Doc. 10.

*ELI: mi te go dito +"/.

*ELI: +" ciao (.) **kako@s to@s da@s si@s sama@s?**

Trad. Io ti avevo detto: "Ciao, come mai sei da sola?"

10.2. Riferimento ad un concetto senza equivalente nella cultura dell'altra lingua

Sono frequenti gli esempi di commutazione per referenti che non trovano un equivalente nell'altra cultura. Alcuni degli esempi tipici sono "**domovnica**" oppure "**e-gradani**", due costrutti presenti nella cultura amministrativo-burocratica croata e assenti da quella italiana ("**domovnica**" si traduce spesso come "*certificato di cittadinanza*", in termini descrittivi).

*ENR: me servi la **domovnica@s** e el certificato de nascita.

Trad. Mi serve il "certificato di cittadinanza" e il certificato di nascita.

*ENR: dove poso andar per ritirar la **domovnica@s** e (que)sto certificato?

Trad. Dove si richiede il certificato di cittadinanza e questo certificato?

*DIN: per cior la **domovnica@s** ti devi andar al **matični@s ured@s**.

Trad. Per il certificato di cittadinanza devi andare all'ufficio anagrafe.

*DIN: con (que)sta carta de identità ti pol aceder al sito **e-gradanin@s**.

Trad. Con questa carta d'identità puoi accedere al servizio dell'identificazione elettronica dei cittadini.

*ENR: ma: per cosa servi (que)sto **e-gradanin@s**?

Trad. Ma a cosa serve questo servizio all'identificazione elettronica dei cittadini?

*ENR: che se pol far la domanda tramite el sito **e-gradanin@s** se ti ga la

carta de identità col chip cioe' la **e-osobna@s**.

Trad. Che si può effettuare la richiesta online se hai la carta d'identità con il chip cioè la carta d'identità elettronica.

10.3. Allentamento di tensione e/o suscitazione di effetti umoristici, ironici o sarcastici

Una delle funzioni di commutazione di codice individuata da Baker si riferisce alla suscitazione di effetti umoristici, ironici o sarcastici. Un esempio è rappresentato dal termine “*baba*” nel seguente esempio, usato in modo ironico per riferirsi ad una persona curiosa o una spiona. Il focus cade sul sostantivo commutato che viene fatto seguire dall’articolo indeterminativo seguendo il modello ironico in croato **baba jedna**, per dare all’effetto ironico maggiore rilevanza.

Doc. 40.

*LUC: *baba@s* una subito la je vignuda spiar.

Trad. Nonna una è uscita subito a spiare.*

Nel seguente passaggio il parlante usa la parola onomatopeica commutata “*štocati*” preceduta dal verbo “fare” all’indicativo presente della prima persona singolare per rimproverare all’interlocutore l’uso eccessivo del cellulare. La costruzione sintattica inoltre risulta essere un calco sintattico dal croato (**uvijek štocaš**). Per quanto riguarda l’analisi inerente al modello *usage-based*, in tutti i suddetti esempi si può notare la struttura che coinvolge la sovrapposizione delle cornici. Nel primo caso la parola “*baba*” potrebbe essere sostituita da qualsiasi altro sostantivo che come categoria è la meno vincolata. La parola “*štocati*” occorre nella costruzione specifica fare (istoveneto) + verbo (croato) dove questo verbo può essere sostituito da qualsiasi altro verbo all’infinito.

Doc. 744.

*GIU: sempre ti fa *štocat(i)@s* sun quel *mobitel@s*.

*LUI: sì (.) <butilo via> [!] (.) in graia.

Trad. Stai sempre su quel telefonino.

Buttalo via nel cespuglio.

Nel prossimo passaggio viene usato l’aggettivo in sloveno e il sostantivo in croato. In questo caso l’aggettivo “*smučarske*” e “*alpsko nešto*” con intento umoristico, in riferimento alla Slovenia come il luogo più vicino all’Istria dove gli istriani trascorrono le vacanze invernali. In questo caso, nel contesto del modello *usage-based* la concettualizzazione gioca un ruolo

importante nell'uso della varietà non istroveneta per denotare un concetto sloveno. La costruzione può essere analizzata secondo il modello nell'ambito della combinazione aggettivo + sostantivo come una delle combinazioni più frequenti nella commutazione di codice.

Doc. 13.

*CAR: xxx si **smučarske jakne**.

*MAR: je **alpsko nešto**.

Trad. Xxx si le giacche da sci.

Qualcosa di tipo alpino.

10.4. Chiarimenti e riformulazioni

Nel prossimo passaggio possiamo notare che la parola “*mislim*” è stata usata per riformulare un pensiero, facilitando così la comprensione alle interlocutrici. Questo è uno dei segnali discorsivi usati molto frequentemente. Secondo l'approccio *usage-based* possiamo notare che strutturalmente la parola/sequenza commutata, che ha la funzione di chiarimento o riformulazione di una porzione di testo enunciata in precedenza, si posiziona sempre nella parte iniziale (di apertura dell'eloquio) e viene seguita dalla riformulazione o spiegazione. Questa costruzione segue dunque la formula della sovrapposizione, in cui la parola “*mislim*” serve da cornice e la parte seguente è quella intercambiabile e soggetta al *chunking*, grazie alle possibilità infinite di estensione strutturale. Possiamo notare perciò che il processo attraverso la ripetizione diventa automatico nel voler chiarire un punto parafrasando o riformulando l'idea non intesa correttamente, in questo modo restringendone il significato.

Doc. 62.

*ELE: **mislim**@s el ga (que)sta faccia perché el je +...

Trad. Voglio dire, ha questa faccia perché è...

Un'altra parola usata spesso nel corpus in funzione di riformulazione dell'enunciato è la parola “*znači*”, corrispondente alla parola italiana “allora” o “dunque”, segnalando l'innescamento di una apertura, avente spesso la funzione di divisione dell'enunciato o turno comunicativo a metà o comunque in due parti. Applicando l'analisi *usage-based* si può notare l'alta ricorrenza della parola, e anche se la posizione non sembra fissa, la posizione iniziale sembra quella più

frequente. Un altro fattore riguardante l'analisi *usage-based* è la struttura che vede la parola “znači” seguita sempre dalla sequenza plurilessicale croata. Siccome quest’ultima non viene usata mai da sola, indica spesso un ideale inizio per estendere le costruzioni attraverso il *chunking*.

Doc. 11.

*MAR: fin le tre e un quarto (.) **znači@s nemamo@s xxx.**

Trad. Fino le tre e un quarto. Dunque non abbiamo xxx.

Doc. 11.

*ELI: **znači@s** uni canta questa canzon (.) la bafe +...

Trad. Dunque alcuni cantano questa canzone.

Doc. 11.

*CAR: **znači@s** la ga una camereta tipo dove che +...

*CAR: sa de quele camerete che je sempre fredo +//.

Trad. Dunque ha una cameretta dove... sai una di quelle camerette dove fa sempre freddo

10.5. Aggiunte, connessioni e prese di turno

Per intervenire in una conversazione i parlanti si servono degli elementi linguistici che, svuotandosi in parte del loro significato originario, assumono dei valori aggiuntivi che servono a sottolineare la strutturazione del discorso, a connettere elementi frasali, interfrasali, extrafrasali e ad esplicitare la collocazione dell'enunciato in una dimensione interpersonale, sottolineando la struttura interattiva della conversazione (Bazzanella, 1995:225).

In tal senso la parola “**uglavnom**“ viene usata come dispositivo linguistico per introdurre un'aggiunta o una presa di turno. Il modello *usage-based* vede la frequenza di questi marcatori discorsivi sempre in posizione iniziale che, secondo il processo di sovrapposizione, comunque servono da base che precede l'inizio di un discorso nuovo.

Doc. 62.

*ELE: +< e **uglavnom**@s niente ierimo in (que)sta scuola.

*ELE: iera super.

Trad. Comunque niente, eravamo in questa scuola. Era fantastico.

Doc. 62.

*ELE: +< **ma uglavnom**@s per (que)sta roma dio mio.

Trad. Comunque, parlando di Roma, Dio mio.

Nel corpus sono stati individuati molti marcatori verbali espressi in croato in funzione della presa di turno. Questi segnali discorsivi vengono espressi spesso utilizzando dei verbi all'imperativo alla II persona singolare o plurale come allocutivo di cortesia. Un esempio di apertura che segnala la presa del turno è la parola “**pazi**“. Come nel caso delle parole con funzione di chiarimento, è possibile notare che secondo l'approccio *usage-based* la ripetizione delle parole (in posizione iniziale della frase), porta spesso all'uso automatico della stessa nelle situazioni in cui si vuole in qualche modo intervenire in una conversazione, spiegando il proprio punto di vista oppure per introdurre una nuova informazione (il che implica il processo di *chunking*).

*ELI: e **pazi**@s una volta iera cusì bustine.

Trad. E guarda che una volta c'erano le bustine.

10.6. Interiezioni

Gumperz sostiene che la commutazione realizzata con accento o intensità fonetica può avere diverse funzioni, come ad esempio per sostenere l'interazione. La parola “*kužiš*” (capisci/hai capito) ad esempio, risulta essere usate frequentemente nella lingua croata per sostenere l'interazione. L'intonazione interrogativa indica la domanda retorica, nel senso che la conferma della comprensione è data per scontata. L'approccio *usage-based* in questo caso vede l'apparizione della parola “*kužiš*” come connettivo tra una prima e una seconda spiegazione e la ripetizione è automatica, data l'alta frequenza d'utilizzo.

Doc. 30.

*ELI: *kužiš@s* (.) (*z*)*nači@s* iera lori due e i kumovi@s.

*ELI: e i se ga spojado.

Trad. Hai capito, allora, c'erano loro due e i testimoni. E si sono sposati.

10.7. Sequenze fisse o formulaiche

Risultano essere frequenti le sequenze fisse o formulaiche commutate (Hakimov, 2021: 55). L'approccio *usage-based* vede queste sequenze come delle espressioni usate talmente spesso da sostenere che esse siano radicate nella mente dell'individuo in modo da poter essere utilizzate con molta facilità. In questi casi avviene il processo di giustapposizione: il parlante, infatti, combina gli elementi ponendoli in sequenze e creando così un nuovo messaggio. Si riportano di seguito alcuni esempi.

Doc. 10.

*MAR: e me scrivi xxx.

*MAR: +" cos(a) ti fa *ovo@s ono@s*?

Trad. E mi scrive: Cosa fai e così avanti?

Doc. 13.

*CAR: *to@s je@s to@s* per (a)deso.

Trad. Questo è tutto per adesso.

Doc. 11.

*ELI: no (.) ela **ovako@s i@s onako@s** la comprerà.

Trad. No. Lei la comprerà comunque.

Doc. 11.

*MAR: mi anderò **kako@s tako@s** a lezion.

Trad. Io andrò comunque a lezione.

Doc. 700.

*INV: no(n) ti po: [///] **iz@s vedra@s neba@s**.

Trad. Ma non puoi... un fulmine a ciel sereno.

10.8. Restrizione del messaggio, riduzione fonetica o lessicale

Come già spiegato in precedenza, spesse volte le parole ad alta frequenza sono soggette a riduzione fonetica (Bybee 2000) e tendono ad essere pronunciate in maniera più rapida rispetto alle parole a bassa frequenza, in sintonia con l'efficienza e l'economia linguistica. Spesso questo porta le strutture plurilessicali a diventare monolessicali. Questo succede perché l'alta frequenza di queste sequenze conferisce loro lo status di unità nel deposito mentale (Hakimov, 2021). Nei seguenti casi possiamo osservare come le strutture plurilessicali, come "acqua minerale" o "stazione degli autobus" vengono sostituite da quelle monolessicali "*mineralna*" e "*busna*" (le versioni ridotte). Le due parole croate sono degli aggettivi che, grazie all'alta frequenza di utilizzo, assumono la funzione del sostantivo che implica e possono essere sostituite da qualsiasi sostantivo nel processo della sovrapposizione.

Doc. 10.

*MAR: +" bevemo una **mineralna@s** a metà?

Trad. Beviamo un'acqua minerale a metà?

Doc. 500.

*FIL: bom vedemo per (que)sta **busna@s** allora qua (a)speta.

Trad. Va bene, vediamo per questa stazione dell'autobus.

10.9. Rafforzamento di una richiesta

Nell'elenco di Baker (2010) è inserita la funzione della commutazione di codice volta al rafforzamento di una richiesta. Nei seguenti passaggi c'è l'introduzione della parola ciacava “**nu**“, che in italiano/istoveneto corrisponde alla parola “su“ o “dai“ usata con l'intenzione di istigare l'interlocutore ad agire in un dato modo e spesso accompagnata da qualche diversità di tono. Nel contesto del modello *usage-based* possiamo notare che la parola “su“ viene usata ogni volta tra il verbo al modo imperativo e l'oggetto della richiesta. Tale costruzione lascia spazio al *chunking*, che prevede l'estensione delle richieste e il consolidamento di queste sequenze attraverso l'esperienza. Inoltre, si può osservare la somiglianza sia fonetica che semantica delle parole “nu“ e “su“.

Doc. 501.

*FIL: sempre i parla de robe inutili.

*BRU: bom [/] bom lasemo star.

*BRU: varda **nu@s** io scominsià (que)sto telegiornale che sentimo (a)ncora qualche cosa de novo.

Trad. Parlano sempre di cose inutili.

Bene, lasciamo stare. Guarda su, sta per cominciare il telegiornale, sentiamo qualcosa di nuovo.

*FIL: forse andarà più facile.

*BRU: bom picia dame **nu@s** un toco de carta.

Trad. Forse sarà più facile. Bene piccola, dammi su un pezzo di carta.

11. Conclusione

In situazioni in cui comunità linguistiche diverse si trovano in simbiosi nella stessa area e interagiscono costantemente, il contatto linguistico è un evento comune e le sue conseguenze sono visibili nell'ambito di frequenti alternanze e commutazioni linguistiche. L'intento della presente tesi è stato quello di analizzare la commutazione di codice che si manifesta nel parlato spontaneo non monitorato dei bilingui istriani. Nella parte introduttiva si è cercato di delineare la situazione sociolinguistica in prospettiva diacronica e sincronica e una cornice teorica inerente al fenomeno della commutazione di codice (e ai modelli teorici recenti che lo contemplano), quale prassi comune e frequente nell'ambito della modalità bilingue. Per le analisi ci siamo serviti di un corpus rappresentativo e autentico secondo i criteri di Tognini-Bonelli (2001), avente l'istoveneto come lingua matrice (il codice identificativo della comunità italoфона istriana, avente un grande prestigio etnolinguistico), con costanti sequenze commutate in lingua croata (la prima lingua ufficiale della Regione istriana, ossia il codice che potenzialmente può coprire tutte le funzionalità discorsive). Nel corpus la commutazione di codice si manifesta in più modi: attraverso l'introduzione di singole parole o frasi ed espressioni intere, nel mezzo della frase, in posizione iniziale o finale; a volte a livello interfrasale, a volte intrafrasale. Nel passaggio da una varietà all'altra è possibile notare una netta prevalenza del tipo intrafrasale, in cui non vengono violati i modelli strutturali e che secondo Poplack (1980) rappresenta l'indice dell'alta padronanza linguistica dei due codici. Di conseguenza i motivi per l'uso della commutazione di codice dei parlanti non sono considerati un segno di incompetenza linguistica, ma piuttosto un segno di capacità discorsive superiori (i bilingui sfruttano semplicemente tutte le risorse linguistiche che hanno a disposizione, come fanno pure i monolingui), perché assolutamente volontaria e di natura funzionale discorsiva (mediante la commutazione di esplicano le funzionalità interazionali).

Nella parte teorica vengono spiegati i tre approcci teorici recenti più noti e più frequenti nelle analisi dei corpora bilingui, ovvero il modello della marcatura, il modello della matrice e il modello *usage-based*. Questo studio si è posto l'obiettivo principale di analizzare, attraverso un'indagine di carattere sperimentale, il fenomeno della commutazione di codice nella comunità italoфона istriana adottando la prospettiva *usage-based*. La scelta di trattare il modello *usage-based* non è stata fatta in base alla qualità o all'efficacia del modello scelto, ma semplicemente perché è il modello più recente e di conseguenza il meno esplorato, con l'intento di stabilire se i postulati del modello fossero applicabili al corpus istriano.

Dato che la visione di base del modello *usage-based* si basa sul fatto che le unità apprese vengono memorizzate insieme a tutte le informazioni che possono essere percepite in un dato contesto, creando un deposito di esperienze con tutte le informazioni inerenti a queste esperienze nella nostra memoria, il concetto dell'uso è inevitabilmente radicato nella conoscenza umana del linguaggio. Il suo contributo generale, infatti, è stato quello di offrire un quadro linguistico alternativo il cui obiettivo non è specificare l'appartenenza ad un insieme di frasi grammaticali, ma capire come la cognizione renda possibile la creazione e la modifica delle strutture nell'utilizzo del linguaggio. Grazie all'integrazione di diverse discipline, dalla psicologia all'informatica, si è dimostrato che livelli di analisi profondamente diversi come la fonologia e la sintassi possono essere compresi in termini basati sull'uso. Negli esempi tratti dal nostro corpus abbiamo in tal senso visto come possono essere analizzati i vari processi cognitivi che influenzano l'uso e lo sviluppo della struttura linguistica. Il modello *usage-based* prevede che le unità siano legate insieme sulla base di co-occorrenza, somiglianza e sovrapposizione o giustapposizione, attraverso i processi cognitivi generali di categorizzazione, generalizzazione e schematizzazione. L'analisi del corpus ha confermato l'attendibilità di tali postulati pure nell'ambito del discorso bilingue istroveneto-croato, in generale spiegando le concordanze evidenti dei modelli ripetitivi che rappresentano le preferenze di combinazioni lessicali specifiche che coinvolgono l'istoveneto e il croato.

La frequenza con cui viene utilizzato un elemento linguistico, o una costruzione linguistica, secondo il modello *usage-based* rappresenta la base della conoscenza linguistica. Nel corpus analizzato si è potuto osservare che l'alta frequenza d'uso facilita l'attivazione delle strutture linguistiche e che la forte rappresentazione causata dall'uso porta alla loro più probabile selezione nella conversazione. Per quanto concerne le sequenze plurilessicali come i costrutti formulaici o le forme fisse usate frequentemente nella lingua croata abbiamo potuto osservare che queste tendono ad apparire nel linguaggio bilingue preservando la loro integrità. Questo fenomeno trova la spiegazione offerta dal modello *usage-based*, ossia si riferisce al fatto che le sequenze usate frequentemente hanno accesso diretto nel lessico mentale e acquistano una rappresentanza autonoma assegnata (es. verbo *far* in istroveneto + verbo all'infinito in croato). Nella parte analitica sono stati riportati numerosi esempi di costruzioni marcate da giustapposizione e sovrapposizione, costituite da una parte fissa e una parte soggetta a variazioni, che può essere intesa come il processo di *chunking*, il quale implica la capacità di costruire tali strutture in modo ricorsivo. Sono stati messi in risalto pure i frequenti casi in cui le parole commutate sono state associate ad altre parole tramite legami di somiglianza fonetica

e/o semantica. Inoltre, si è potuto osservare che molte scelte vengono applicate in sintonia con i principi di economia e efficienza linguistica: la riduzione lessicale avviene frequentemente e per realizzarla si fa uso della commutazione di codice, con l'intento costante di trasformazione delle sequenze plurilessicali in quelle monolessicali commutate e le parole plurisillabiche in monosillabiche.

L'elenco delle funzioni conversazionali della commutazione di codice individuate da Gumperz (1982) e Baker (2010) come tipiche della commutazione di codice nelle comunità bilingui è stato ripercorso in ottica *usage-based*, per valutarne il funzionamento nell'ambito dei postulati di quest'ultimo. In tal senso abbiamo analizzato la commutazione di codice nell'ambito delle funzioni discorsive ricorrenti nel corpus. Si è potuto osservare che tali funzioni, realizzate per mezzo della commutazione di codice, possono essere individuate come vere e proprie strategie conversazionali. Inoltre, si è notato che le proprietà distributive delle categorie grammaticali, come la frequenza di occorrenza e il rapporto con altri elementi nelle frasi, influenzano in modo decisivo la struttura del discorso bilingue. Pertanto, in conclusione possiamo affermare che l'approccio basato sull'uso si presenti come un modello in grado di spiegare adeguatamente i principi della commutazione di codice propria del corpus analizzato, seppur sia doveroso notare che risulta essere difficile stabilire una regolarità assoluta nelle ricorrenze del fenomeno, visto che dipendono dalle scelte individuali e creative di ogni singolo parlante, a seconda delle esperienze e abitudini conversazionali, dalla miriade dei singoli contesti sociolinguistici, stili comunicativi e finalità specifiche preposte. Infatti, la creatività linguistica va esaminata nel contesto della conoscenza umana, del giudizio e della capacità di risoluzione dei problemi, piuttosto che all'interno dei confini di una grammatica ristretta e autosufficiente (Langacker, 1987:73). Si tratta dunque di uno spettro multifattoriale che implica un modello altamente flessibile e sempre relativo, che non poggia le fondamenta su norme prestabilite. Possiamo dunque concludere che le strategie conversazionali implicate, da una parte definiscono l'insieme dei comportamenti linguistici e dall'altra forniscono mezzi e risorse comunicative da usare secondo le preferenze e le intenzioni dei parlanti (Cerruti, 2003: 123).

Bisogna specificare inoltre che l'intera analisi è stata eseguita sulla base di interpretazioni soggettive ragionate dell'autrice e il risultato ottenuto rappresenta il frutto di osservazioni personali, basate su pochi studi finora effettuati aventi come oggetto di studio altre lingue, che indagano lo sviluppo della cognizione sociale e del ragionamento pragmatico, ossia che esaminano l'unione pragmatico-sociale con la grammatica. Questo lavoro pertanto rappresenta uno dei primi tentativi, se non il primissimo, di studiare il fenomeno della commutazione

attraverso l'approccio *usage-based* in Istria. Numerosi sono dunque i limiti dell'indagine condotta in questa sede. Si necessita di ulteriori indagini sociolinguistiche qualitative e approfondimenti che consentiranno di trattare il fenomeno nelle sue molteplici sfaccettature, nell'ambito delle strutture concettuali ed eventi di utilizzo ripetuti, in grado di offrire delle risposte inerenti alla modalità in cui si presenta il fenomeno della commutazione di codice tra i parlanti bilingui istriani.

12. Bibliografia

Abney S. (1991) Parsing by Chunks. In: Principle-Based Parsing. Dordrecht: Kluwer Academic Publishers, 257-278.

Alfonzetti G. (1992) Il discorso bilingue. Pavia: Franco Angeli.

Alvarez-Cáccamo C. (1998) From 'switching code' to 'code-switching': Towards a reconceptualisation of communicative codes. In Auer, P. Code-Switching in Conversation. Language, interaction and identity. Londra: Routledge, 29-50.

URL:https://www.academia.edu/971097/From_switching_code_to_code_switching_Towards_a_reconceptualisation_of_communicative_codes

Anđel M., Blaži D. (1994): CHAT i CLAN- Alati za računalnu obradu jezičnih transkripata, Laboratorij za psiholingvistička istraživanja. Edukacijsko-rehabilitacijski fakultet, Zagreb.

Arduini S., Damiani M. (2016) Linguistica applicata. Padova: libreriauniversitaria.it edizioni .

Auer P. (2020) Language contact. Pragmatic factors. Freiburg: Routledge.

URL: https://www.researchgate.net/publication/346912381_Language_contact

Auer P. (1999) From code-switching via language mixing to fused lects. Towards a dynamic typology of bilingual speech. International Journal of Bilingualism 3: 4, 309-332.

Backus, Ad. (1996) Two in one: Bilingual speech of Turkish immigrants in the Netherlands (Studies in Multilingualism 1). Tilburg, Netherlands: Tilburg University Press.

Baker C. (2010) Code-switching. In K. Malmkjær (Ed.), Linguistics Encyclopedia. (3rd ed.), 55-60.

Banovac B. (1998) Društvena pripadnost, identitet i teritorij. Sociološko istraživanje regionalne pripadnosti u Istri. Rijeka: Pravni fakultet Sveučilišta u Rijeci .

Bazzanella, C. (1995) I segnali discorsivi. In: Grande grammatica italiana di consultazione III, Tipi di frase, deissi, formazione delle parole, a cura di L. Renzi, G. Salvi, A. Cardinaletti, Bologna: Mulino, 225-257.

Bentahia A., Davies E.E. (1983) The Syntax of Arabic-French Code-switching. Lingua 59., 301-330.

Berruto G. et al. (1996) Introduzione all'italiano contemporaneo: la variazione e gli usi, a cura

- di Sobrero Albert A., 2a ed., Roma- Bari: Editori Laterza.
- Berruto G. (1999) Le varietà del repertorio. In: Introduzione all'italiano contemporaneo. Vol. I: La variazione e gli usi, a cura di Alberto Sobrero, Roma–Bari: Editori Laterza, 3–36.
- Blagoni R. (1999) O nekim aspektima lingvističkog atlasa u okvirima znanosti o jeziku. In: *Annales. Series historia et sociologia* 16., 29-44.
- Blagoni R. (2004) Istra u zrcalu povijesnih smjerenja Identitet i strategije tumačenja odnosa čovjeka, jezika i svijeta. Pula: Filozofski fakultet u Puli.
- Blagoni R. (2011) Metafora jezičnog otoka. In: *Tabula* 9., 16-27.
- Blagoni R., Poropat Jeletić N., Blecich K. (2016) *The Italophone Reefs in the Croatophone Sea*. Wien: Peter Lang Verlag.
- Blom, J. P., Gumperz J. (1972) Code-switching in Norway. In: Gumperz J. and Hymes D. (eds.), *Directions in Sociolinguistics*. New York: Holt, Rinehart and Winston, 407-434.
- Bogliun Debeljuh L. (1994) *L'identità Etnica – gli italiani dell'area istro-quarnerina*, Etnia. Trieste-Rovigno: Centro di Ricerche Storiche Rovigno.
- Bogliun L., Burra A., Debeljuh A. (2019) *Uso pubblico della lingua italiana e attuazione del bilinguismo integrale nelle città e nei comuni bilingui della Regione Istriana*. Ricerca scientifica.
- Bolshakov I., Gelbukh A. (2004) *Computational linguistics, Models, Resources, Applications*. Mexico: Instituto Politécnico Nacional. Libro disponibile in rete al seguente indirizzo web: <https://www.gelbukh.com/clbook/Computational-Linguistics.pdf>
- Burt, S. (2002) Maxim confluence. *Journal of Pragmatics*. 34., 993-1001.
- Bybee, J.L. (1985) *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins Publishing Company.
- Bybee J.L. Beckner C. (2010) Usage based theory. New Mexico: University of New Mexico, 827-855.
- Bybee J.L., Scheibman J. (1999) The effect of usage on degrees of constituency: The reduction of don't in English. *Linguistics* 37., 575–96.

Bybee J. L. (2000) The phonology of the lexicon: Evidence from lexical diffusion. In: Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*. Stanford, CA: CSLI Publications, Center for the Study of Language and Information, 65–85.

Bybee J.L. (2002) Phonological evidence for the exemplar storage of multiword sequences. *SSLA* 24 (2)., 215-221.

Bybee J. L. (2006) From usage to grammar: The mind's response to repetition. *Language* 82(4)., 711–733.

Bybee, J. L., Eddington D. (2006) A Usage-Based approach to Spanish verbs of becoming. In: *Language* 82(2), 323-355.

Bybee J.L., Beckner C. (2010) Usage-based theory. Capitolo 32. University of New Mexico. URL: <https://www.unm.edu/~jbybee/downloads/BybeeBeckner2010UsageBasedTheory.pdf>

Bybee J., File-Muriel R., De Souza R.N. (2016) Special reduction: a usage-based approach. In: *Language and Cognition* 8., Cambridge University Press, 421– 446.

URL:<https://www.cambridge.org/core/journals/language-and-cognition/article/special-reduction-a-usagebased-approach/84DECBA10841AA5BC599C9473A40E648>

Capasso D., Russi R. (2010) Il sogno italiano, atti del II convegno internazionale di studi dell'Aiba (Associazione degli Italianisti nei Balcani), Banja Luka (18-19 giugno): Aonia edizioni.

Carli A. (1996) Il fenomeno della commutazione di codice, in *Miscellanea* 3. Università di Trieste: EUT Edizioni, 127-146.

URL: https://www.openstarts.units.it/bitstream/10077/8032/1/Carli_miscellanea_03.pdf

Cergna S. (2014) L'istrioto: cenni storici, *Atti, Centro di Ricerche Storiche, Rovigno*, vol. XLIV, 317-332.

URL: <https://hrcak.srce.hr/file/351227>

Cerruti, M. (2003) Il dialetto oggi nello spazio sociolinguistico urbano. *Indagine in un quartiere di Torino*, RID 27.

URL: https://www.researchgate.net/publication/304170698_Aspetti_pragmatico-funzionali_della_commutazione_di_codice_italiano-dialetto_un'indagine_a_Torino

Cerruti M., Regis R. (2005) Code switching e teoria linguistica: la situazione italo-romanza, *Rivista di linguistica*, Torino: Università di Torino.

Consiglio della minoranza nazionale italiana autoctona (2017), a cura di Silvano Zilli. *La Comunità Nazionale Italiana autoctona negli Statuti delle Città e dei Comuni bilingui della Regione Istriana*. Libro disponibile al seguente indirizzo web: <https://drive.google.com/file/d/1dvWLJuNKpk0w2-tX1NNdpXsN6wW2lBun/view>

Dąbrowska E., Lieven E.V.M. (2005) Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics* 16(3), 437–474.

Dal Negro S. (2005a) Minority languages between nationalism and new localism: the case of Italy, *International Journal of the Sociology of Language*, n. 174., 119.

Dal Negro S., Guerini F. (2007) *Contatto. Dinamiche ed esiti del plurilinguismo*, Roma: Aracne.

Davies P. (2010) Identifying word-order convergence in the speech of Welsh-English bilinguals. *Dissertazione di dottorato, The School of Linguistics and English Language, Bangor University*.

Delibera del Ministero della Cultura, Direzione per la Tutela dei Beni Culturali (2017).

Diessel, Holger (2016) Frequency and lexical specificity: A critical review. In Heike Behrens & Stefan Pfänder (eds.), *Experience counts: Frequency effects in language (Linguae & Litterae 54)*, Berlin: de Gruyter, 209–237.

Di Sciullo A., Muysken P., Singh R. (1986). Government and code-mixing, *Journal of linguistics*.

URL:<https://repository.uibn.ru.nl/bitstream/handle/2066/14611/3901.pdf;jsessionid=AE7D29D940324451408683A734FC51B5?sequence=1>

Ellis N. C. (1996) Sequencing in SLA: Phonological memory, chunking and point of order. *Studies in Second Language Acquisition* 18., 91–126.

Filipi G. (1987) Alcuni aspetti diacronici del bilinguismo. *Scuola nostra*, 19, 171-179. Rijeka: Edit.

Filipi G. (1989) Situazione linguistica Istro-quarnerina. In: *Ricerche sociali* 1., 73-82.

Fishman J. (1973) Bilingualism with and without Diglossia, Diglossia with and without Bilingualism. *The Journal of Social Issue*, XXIII, 1967, n. 2., 2 8-38.

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

Gaskins D. et al. (2019) Towards a usage-based model of early codeswitching: Evidence from three language pairs, *Applied Linguistics Review*.

Giuricin E. (2011) La comunità italiana nei censimenti jugoslavi, *Ricerche Sociali* 18., Trieste-Rovigno, 139-186.

URL:<https://crsrv.org/wp/wp-content/uploads/2020/03/Ezio-Giuricin-La-comunita-italiana-nei-censimenti-jugoslavi-croati-e-sloveni-1945-2011.pdf>

Goldberg A. E. (1995) *Constructions: A construction grammar approach to argument structure*, Chicago/London, University of Chicago Press.

Goldberg, A. E. (2006) *Constructions at work: The nature of generalization in language*, New York, Oxford University Press on Demand.

Goldberg A.E., Suttle L. (2011) The partial productivity of constructions as induction. Psychology Department, Princeton University.

URL:<https://adele.princeton.edu/wpcontent/uploads/sites/277/2015/01/SuttleGoldbergLinguistics11.pdf>

Grosjean, F. (1982) *Life with Two Languages*, Cambridge: Harvard University Press.

Grosjean, F. (2008) *Studying bilinguals*, Oxford/New York: Oxford University Press.

Gumperz J.J. (1982) *Code switching in Conversation in Discourse strategies*. Cambridge: Cambridge University Press.

Hakimov N. (2021) *Explaining Russian-German code-mixing. A usage-based approach*, Berlin: Language Science Press.

Hakimov N, Backus A. (2021) Usage-Based Contact Linguistics: Effects of Frequency and Similarity in Language Contact, *Journal of Language Contact*..

URL: https://brill.com/view/journals/jlc/13/3/article-p459_459.xml?language=en#R000053

Halwachs S., D. W. (2007) Sociolingvistika. Uvod u lingvistiku, ur. Glovački-Bernardi Z., Kovačec, A., Mihaljević, M., Halwachs, D.W., Sornig, K., Penzinger, C. i R. Schrodtt, Zagreb: Školska knjiga, 217-245.

Haugen, E. (1956) *The Norwegian Language in America: a study in bilingual behaviour*. Philadelphia: University of Pennsylvania Press.

Hoffman T., Trousdale G. (2013) *The Oxford handbook of construction grammar*. Oxford: University Press.

Ivetić E. (2009) *Istra kroz vrijeme, Pregled povijesti Istre sa osvrtom na grad Rijeku*, Rovigno: Centro di Ricerche Storiche.

Jernej-Pulić M. (2016) *Jezik kao kapital u istarskom multikulturalnom okruženju*. Dissertazione di dottorato, Filozofski fakultet u Zagrebu.

Kapatsinski, V., M. (2005) *Measuring the relationship of structure to use: Determinants of the extent of recycle in repetition repair*. Berkeley Linguistics Society.

Kapatsinski, V, M. (2010) *Frequency of use leads to automaticity of production: Evidence from repair in conversation*. Language and Speech.

Kuvač Kraljević J., Hržica G. (2016) *Croatian Adult Spoken Language Corpus (HrAL)**, FLUMINENSIA, Zagabria, 87-102.

URL: <https://hrcak.srce.hr/file/256835>

Kruszewski M. (1995) *Writings in General Linguistics: On Sound Alternation (1881) and Outline of Linguistic Science (1883)*. Edited by E. F. Konrad Koerner. Amsterdam-Philadelphia, PA: Benjamins.

Labov W. (1973) *Sociolinguistic patterns* Philadelphia: University of Pennsylvania Press.

Langacker, Ronald W. (1987) *Foundations of Cognitive Grammar: Theoretical Prerequisites*. Vol. 1. Stanford, CA: Stanford University Press.

Lewandowski Th (1976) *Linguistische Worterbuch, 2.*, Heidelberg: Quelle and Meyer.

Lubello S. (2016) *Manuale di Linguistica Italiana*. Berlin: De Gruyter Publisher.

MacWhinney B., Snow, C. (1990) *The Child Language Data Exchange System: An update*. Journal of Child Language 17., 457-472.

URL: <https://childes.talkbank.org/access/Eng-NA/0docs/MacWhinney1990.pdf>

MacWhinney et al. (2004) *TalkBank: Building an Open Unified Multimodal Database of Communicative Interaction*, Lisbon: European Language Resources Association.

URL: <http://www.lrec-conf.org/proceedings/lrec2004/pdf/392.pdf>

Maticchio I., Tamaro S. (2020) Italiano, istroveneto e istrioto. Un'indagine sugli usi delle lingue minoritarie romanze in Istria. In: *Lingue minoritarie tra localismi e globalizzazione*, a cura di Marra A. e Dal Negro S, Milano: studi AItLa.

Marcarino A. (1995) *Sociologia dell'azione comunicativa*, Napoli: Guida editori.

Matijašić R. (1993) Starija povijest Istre i neki njezini odrazi na novije razdoblje. *Društvena istraživanja*, n. 4-5., 569-585.

Matijašić, R. (2005) *Od iskona do prevlasti Serenissime, Pula, Tri tisućljeća mita i stvarnosti*, a cura di M. Bertoša et al., Pula: C.A.S.H., 9-45.

Matijašić, R. (2005) *Istra, Istarska enciklopedija*, a cura di M. Bertoša e R. Matijašić, Zagreb: Leksikografski zavod Miroslav Krleža, 337-339.

Milani-Kruljac N. (1990) *La Comunità Italiana in Istria e a Fiume fra diglossia e bilinguismo*, Etnia I, Unione degli Italiani dell'Istria e do Fiume, Trieste-Rovigno: Università popolare di Trieste.

Marcarino A, (1995) *Sociologia dell'azione comunicativa*, Napoli: Guida editori.

Matras, Y. (2009) *Language contact (Cambridge textbooks in linguistics)*. Cambridge: Cambridge University Press.

Moravcsik, E. A. (1978) *Language contact*. In: Joseph H. Greenberg, Charles A. Ferguson e Edith A. Moravcsik (eds.), *Universals of human language: Method and theory*. Stanford: Stanford University Press, 93–122.

Muscardin R. *La lingua istriota nella letteratura istro-quarnerina*.

URL: http://www.odos.cloud/images/lingua_istriota.pdf

Muysken, P. (2000) *Bilingual Speech: A Typology of Code-mixing*, Cambridge: Cambridge University Press.

Myers-Scotton C. (1990) Codeswitching and borrowing: Interpersonal and Macrolevel meaning, in Jacobson R., Codeswitching as a Worldwide Phenomenon. New York: P Lang, 85- 105.

Myers-Scotton C. (1992) Comparing Codeswitching and Borrowing, Journal of Multilingual and Multicultural Development, University of South Carolina.

URL:https://www.researchgate.net/publication/239468245_Comparing_Codeswitching_and_Borrowing

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

Myers-Scotton C. (2006) Multiple voices: An Introduction to Bilingualism, Oxford: Blackwell

Nilep C. (2006) Code switching in Sociocultural Linguistics. Colorado Research in Linguistics. 19:1-22.

Orbanić S. (2005) Pula na prijelazu tisućljeća. In: Pula. Tri tisućljeća mita i stvarnosti, ur. M. Bertoša et al., Pula: C.A.S.H., 341-373.

Orlić O. (2008) Mnogoznačje istarskog multikulturalizma. Etnološka tribina: godišnjak Hrvatskog etnološkog društva 38., 39-59.

Pivčević M., Commutazione del codice nella Cronaca del nostro Piccolo paese di Miljenko Smoje: Università di Zara.

URL:[file:///C:/Users/sonja/Downloads/969753.55 Chapter_Manuscript-856-1-1020200131%20\(1\).pdf](file:///C:/Users/sonja/Downloads/969753.55%20Chapter_Manuscript-856-1-1020200131%20(1).pdf)

Poplack, S. (1980) Sometimes I start a sentence in English Y TERMINO EN ESPAGNOL; toward a typology of codeswitching, Linguistics 18., 582-618.

Poropat Jeletić N., Moscarda Mirković E., Bortoletto A. (2021) Anales: Annali di Studi istriani e mediterranei. Incidenza e implicazioni di alcuni tratti formali pertinenti tipici del discorso bilingue istriano: i casi di commutazione di codice, Ser. hist. sociol. 31., Capodistria, 329-340.

Rumici G. (2020) Istria, Fiume e Dalmazia. Cenni storici, testi ed immagini, Associazione Nazionale Venezia Giulia e Dalmazia, Roma: Nca Web.

URL:<https://www.scuolaeconfineorientale.it/2021/archivio/>

seminario/Istria_Fiume e_Dalmazia.pdf

Schneider U. (2014) Frequency, chunks and hesitations: a usage-based analysis of chunking in English. Dissertazione inaugurale, Albert-Ludwigs-Universität Freiburg.

URL:https://www.academia.edu/43838547/Frequency_Chunks_and_Hesitations_A_Usage_based_Analysis_of_Chunking_in_English

Seckin E. (2019) Code Switching: Definition, Types, and Examples.

URL:<https://owlcation.com/humanities/Code-Switching-Definition-Types-and-Examples-of-Code-Switching>

Selvaggi D. (2012) Il bilinguismo. Da anomalia a costante antropologica. Roma: Aracne editrice S.r.l.

Shin J. (2018) Bilingualism in Schools and Society: Language, Identity, and Policy. 2a ed. New York: Routledge.

Shohamy, E. (2006) Language Policy. Hidden Agendas and New Approaches. London: Routledge.

Siemund P. (2008) Language Contact and Contact Languages, Hamburg: John Benjamins Publishing Company.

URL:https://www.academia.edu/17869846/Language_Contact_and_Contact_Languages

Sirbu A. (2015) Significance of language as a tool of communication, Mircea cel Batran, Naval Academy Scientific Bulletin, Volume 8.

The “Five Graces Group”, Beckner C., Blythe R., Bybee J.L., Christiansen M.H., Croft W., Ellis N.C., Holland J., Ke J., Larsen-Freeman D. and Schoenemann T. (2009) Language is a complex adaptive system: Position paper. *Language Learning* 59., 1–26.

Thomason, S.G. (2001) *Language contact: an introduction*, Edinburgh: Edinburgh University Press.

Thomason, S.G., Kaufman T. (1998) *Language contact, creolization, and genetic linguistics*. Berkeley, CA: University of California Press.

Tikka M. (2009) Manifestarsi plurilingui a tavola. La commutazione di codice di una famiglia italo-svedese. Dissertazione di dottorato, Department of French, Italian and Classical Languages, Stockholm University.

Tognini-Bonelli E. (2001) *Corpus Linguistics at Work*, Studies in Corpus Linguistics. Amsterdam/Philadelphia: John Benjamins Publishing Company.

van Hout R., Muysken P. (1994) Modeling lexical borrowability. *Language Variation and Change* 6(1). 39–62.

Vrbanec N. (2021) *Variazione dialettale: istroveneto e fiumano a confronto*. Tesi di laurea magistrale, Sveučilište Jurja Dobrile u Puli.

Vuk A. (2019) *Progettazione e costruzione di un mini-corpus de parlato bilingue: struttura e raccolta di dati*. Tesi di laurea triennale, Sveučilište Jurja Dobrile u Puli.

Weber, J. (2012) Horner K., *Introducing Multilingualism. A social approach*. London-New York: Routledge.

Wei Li. (2013) *Conceptual and Methodological Issues in Bilingualism and Multilingualism Research*. *The Handbook of Bilingualism and Multilingualism*, a cura di T. K. Bhatia e W. C. Ritchie, USA-UK: Wiley Blackwell, 26-51.

Weinreich U., Martinet A. (1963) *Languages in Contact. Findings and problems*. Paris: The Hague, Mouton and C.

Whitney, William D. (1881) *On mixture in language*. *Transactions of the American Philological Association.*, 12. 5–26.

Wiechmann, D. (2008) *On the computation of collocation strength: Testing measures of association as expressions of lexical bias*. *Corpus Linguistics and Linguistic Theory* 4(2). 253–290.

Županović Filipin N., Bevanda Tolić, K. (2015) *Tra due sponde linguistiche: Commutazione di codice in due generazioni di una famiglia bilingue*, Università di Zagabria: Facoltà di Lettere e Filosofia, 55-83.

URL:file:///C:/Users/Korisnik/Desktop/SRAZ_60_Zupanovic_Filipin_Bevanda_Tolic.pdf

Županović Filipin N, Bevanda Tolić, K. (2015) La commutazione di codice nel bilinguismo croato-italiano: analisi di un tipo testuale, Università di Zagabria: Facoltà di Lettere e Filosofia.
URL: <http://doi.fil.bg.ac.rs/pdf/journals/italbg/2015-1/italbg-2015-1-2.pdf>

13. Sitografia

<https://www.collinsdictionary.com> (ultima consultazione il 30/3/2022)

https://dizionari.corriere.it/dizionario_italiano/C/chip_1.shtml (ultima consultazione il 28/4/2022)

https://dizionari.corriere.it/dizionario_inglese/Inglese/F/full_1.shtml (ultima consultazione il 7/4/2022)

<https://www.dizionario-italiano.it/dizionario-italiano.php?lemma=PIDGINIZZAZIONE100>
(ultima consultazione il 4/6/2022)

<http://childes.psy.cmu.edu> (ultima consultazione il 15/4/2022)

<https://www.coloursofistria.com/it/destinazioni/istria/storia> (ultima consultazione il 16/5/2022)

<http://www.dzs.hr/> (ultima consultazione 18/5/2022)

<https://www.istrapedia.hr/it/natuknice/1852/istroveneto> (ultima consultazione il 10/3/2022)

<https://www.istra-istria.hr/it/regione-istriana/documenti/statuto-della-regione-istriana/> (ultima consultazione il 16/5/2022)

<http://ldc.upenn.edu> (ultima consultazione il 10/3/2022)

<http://www.multidis.erf.hr/o-projektu.html> (ultima consultazione il 9/6/2022)

<http://rjecnik.neologizam.ffzg.unizg.hr/2016/02/01/x-ica/> (ultima consultazione il 20/3/2022)

https://www.sapere.it/sapere/dizionari/traduzioni/latino_italiano/M/mercatus,%C5%ABs.html
(ultima consultazione il 28/3/2022)

https://www.scuolaeconfineorientale.it/2021/archivio/seminario/Istria_Fiume_e_Dalmazia.pdf
f (ultima consultazione il 18/5/2022)

<http://talkbank.org> (ultima consultazione il 16/6/2022)

https://www.treccani.it/enciclopedia/1-italia-romana-delle-regiones-regio-x-venetia-et-histria_%28Il-Mondo-dell%27Archeologia%29/ (ultima consultazione il 16/5/2022)

<https://www.treccani.it/vocabolario/bilinguismo/> (ultima consultazione l'8/3/2022)

https://www.treccani.it/enciclopedia/1-italia-romana-delle-regiones-regio-x-venetia-et-histria_%28Il-Mondo-dell%27Archeologia%29/ (ultima consultazione il 16/5/2022/)

http://unione-italiana.eu/Backup/documents/Statuto_Regione_Istriana.pdf (ultima consultazione il 13/5/2022)

<https://web.stanford.edu/class/linguist289/inMemoriamMauriceGross.pdf> (ultima consultazione il 28/3/2022)

14. Allegato 1



REPUBLIKA HRVATSKA
MINISTARSTVO KULTURE I MEDIJA

UPRAVA ZA ZAŠTITU KULTURNE BAŠTINE

KLASA: UP/I-612-08/21-06/0113

URBROJ: 532-06-02-02/1-21-1

Zagreb, 17. rujna 2021.

HRVATSKA UNIJA RIJEKA
20.10.2021.
377

Ministarstvo kulture i medija na temelju članka 12. stavka 1. Zakona o zaštiti i očuvanju kulturnih dobara („Narodne novine“, broj 69/99, 151/03, 157/03, 87/09, 88/10, 61/11, 25/12, 136/12, 157/13, 152/14, 44/17, 90/18, 32/20, 62/20) i članka 20. stavaka 1. i 2. Pravilnika o obliku, sadržaju i načinu vođenja Registra kulturnih dobara Republike Hrvatske („Narodne novine“, broj 89/11 i 130/13), donosi

RJEŠENJE

1.

Utvrđuje se da **Istromletačko narječje** ima svojstvo nematerijalnoga kulturnog dobra u smislu članka 9. stavka 1. alineje 1. Zakona o zaštiti i očuvanju kulturnih dobara.

2.

Za kulturno dobro iz točke 1. ovoga Rješenja utvrđuje se sljedeći sustav mjera zaštite:

- osigurati dostupnost dobra javnosti;
- poticati prenošenje i njegovanje kulturnoga dobra u izvornoj sredini;
- poticati sudjelovanje pojedinaca, grupa i zajednice koje baštine dobro u identificiranju, definiranju, izvodenju i prenošenju dobra;
- popularizirati i promovirati kulturno dobro održavanjem stručnih skupova, putem elektroničkih medija, audio i video zapisa i na drugi način;
- osigurati održivost dobra kroz edukaciju, identificiranje, dokumentiranje, znanstveno istraživanje, očuvanje, zaštitu, promicanje, povećanje vrijednosti, mogućnost prenošenja tradicije nasljednicima putem formalnog i neformalnog obrazovanja, te revitalizaciju napuštenih segmenata dobra;
- educirati stručni kadar za prenošenje znanja i vještina putem seminara, radionica, formalnog i neformalnog obrazovanja;
- nastaviti sa istraživanjem dobra, primjerenim dokumentiranjem u svim vidovima i načinima suvremenog bilježenja te stručnim i znanstvenim vrednovanjem;
- promicati funkciju i značaj dobra u društvu, te uključiti zaštitu dobra u planirane razvojne programe;
- popularizirati i promovirati nematerijalno kulturno dobro te time nastojati ojačati svijest da je riječ o jedinstvenom blagu koje na lokalnoj i nacionalnoj razini osigurava kulturnu raznolikost te provoditi mjere zaštite kako bi se izbjegla opasnost od nestajanja, uništenja ili pretjerane komercijalizacije kulturnoga dobra;
- potaknuti izradu zavičajnog rječnika i gramatike;
- uključiti njegovanje materinskog govora u nastavne planove i programe osnovnoškolskog i srednjoškolskog obrazovanja;
- nositelj dobra dužan je provoditi mjere zaštite radi njegova očuvanja, sukladno Zakonu o zaštiti i očuvanju kulturnih dobara i svim propisima koji se odnose na kulturna dobra, pridržavajući se njegove povijesno-tradicijske matrice i pojavnosti.

3.

Nositelji su sve pravne i fizičke osobe evidentirane na Popisu nositelja koji je sastavni dio ovoga Rješenja. O svim promjenama nositelja nadležno tijelo je dužno obavijestiti Stručno povjerenstvo za utvrđivanje svojstva kulturnoga dobra, koje će sukladno izmjenama revidirati Popis nositelja.

4.

Na navedeno nematerijalno kulturno dobro iz točke 1. izreke ovoga Rješenja primjenjuje se Zakon o zaštiti i očuvanju kulturnih dobara i svi propisi koji se odnose na kulturna dobra.

5

Predmetno kulturno dobro upisat će se u Registar kulturnih dobara Republike Hrvatske - Listu zaštićenih kulturnih dobara pod brojem Z - 7470.

6.

Žalba ne odgađa izvršenje ovoga Rješenja.

O b r a z l o ž e n j e

Istromletačko narječje obuhvaća italoromanske govore mletačkoga tipa u hrvatskoj i slovenskoj Istri te u manjem dijelu Istre koji se danas nalazi u granicama Republike Italije. Ono predstavlja materinski ili prvi jezik najvećega dijela pripadnika talijanske nacionalne manjine u Republici Hrvatskoj. Ono se može definirati kao skup mletačkih govora kojima se govori u Istri. Istromletačko narječje pripada talijanskome kulturnom krugu, ali i specifičnom višejezičnom i višekulturnom jadranskom kulturnom krugu. U smislu klasifikacije kulturnih dobara, unutar nematerijalne baštine, pripada u kategoriju usmena predaja, izričaji i govori.

Istromletačko je narječje u granicama Republike Hrvatske jezik svakodnevnoga sporazumijevanja u Istarskoj županiji, i to na području gradova Buje/Buie, Novigrad/Cittanova, Poreč/Parenzo, Pula/Pola i Umag/Umago te općina Brtonigla/Verteneglio, Grožnjan/Grisignana, Oprtalj/Portole, Tar-Vabriga/Torre-Abrega i Višnjan/Visignano. U obuhvat područja na kojem se koristi istromletački mogu se uključiti područja gradova Rovinj/Rovigno i Vodnjan/Dignano te općina Bale/Valle, Fažana/Fasana i Ližnjan/Lisignano (na čijem se području nalazi Šišan). Nadalje, nezanemariv broj govornika istromletačkog nalazi se na području općina Funtana/Fontana, Kanfanar, Kaštelir – Labinci/Castelliere – S. Domenica, Medulin, Motovun/Montona, Vižinada/Visinada i Vrsar/Orsera.

Mletački govori u Istri, za razliku od istroromanskih (istriotskih) govora, nisu plod neprekinuta razvoja lokalne latinštine. Ondje su se učvrstili zahvaljujući teritorijalnoj ekspanziji Mletačke Republike, koja je u drugoj polovici XIV. st. definitivno ovladala istarskim obalnim gradovima, te kolonizaciji kugom i malarijom opustošena prostora zapadne Istre. Mletački se tako širi područjima koja su izvorno pripadala govornicima istroromanskog i furlanskog tipa. U razdoblju od XV. do XIX. st. mletački je postao dominantan jezik zapadnoistarske obale i priobalja, iako je u istome razdoblju dio toga prostora temeljito kroatiziran doseljavanjem iz Dalmacije. Podaci austrijskih popisa o govornicima talijanskoga jezika kao Verkehrssprache, koji u pojedinim gradovima i općinama na zapadu Istre prelaze 90% stanovništva, kriju zapravo govornike mletačkog. U razdoblju između dvaju svjetskih ratova, obilježeno ukidanjem hrvatskih institucija te suzbijanjem javne uporabe slavenskih varijeteta, vrijednost mletačkog kao sredstva svakodnevne usmene komunikacije svakako je porasla. Do dramatična pada broja govornika mletačkog dolazi s egzodusom istarskih Talijana, ponajviše u razdoblju od 1945. do 1954. godine. U posljednjem se stoljeću broj govornika znatno smanjio.

Sve do početka XX. st. mletački se u Istri prenosio isključivo usmenim putem, a književna aktivnost intenzivnije se razvija istom od 60-ih godina XX. st. Od XIX. st. primjetan je utjecaj tršćanskoga mletačkog govora na mletačke govore u Istri i Kvarneru, ali i u Dalmaciji, što je za posljedicu imalo gubitak niza izvornih dijalektalnih obilježja. Udio mletačkih posuđenica u tradicionalnom leksiku čakavskih govora znatan je.

Uz književnu aktivnost, poznavanju mletačkoga u Istri i na Kvarneru doprinose leksikografska i dijalektološka djela, no još smo uvijek uskraćeni za iscrpan opis tih mletačkih govora. Javna vidljivost istromletačkog postiže se, osim spomenutim publikacijama, manifestacijom Festival dell'istoveneto koju organizira Talijanska unija u suradnji s Istarskom županijom.

Povjerenstvo za nematerijalnu kulturnu baštinu Ministarstva kulture na sjednici održanoj 10. rujna 2020. godine ocijenilo je da **Istromletačko narječje** ima svojstvo nematerijalnoga kulturnog dobra kako to definira Zakon o zaštiti i očuvanju kulturne baštine u članku 9. stavku 1. alineji 1.

Na osnovi predočene dokumentacije i iznesenih činjenica, a prema preporuci Povjerenstva za nematerijalnu kulturnu baštinu, Stručno povjerenstvo za utvrđivanje svojstva kulturnog dobra na sjednici održanoj 17. rujna 2021. godine utvrdilo je da **Istromletačko narječje** ima svojstvo kulturnoga dobra.

Donošenjem ovoga Rješenja, sukladno članku 12. stavku 1. Zakona o zaštiti i očuvanju kulturnih dobara na predmetno dobro primjenjuju se citirani Zakon, kao i svi drugi propisi koji se odnose na kulturna dobra.

Sukladno članku 12. stavku 4. istoga Zakona, točkom 5. izreke ovoga Rješenja, određena je obveza upisa predmetnoga dobra u Registar kulturnih dobara Republike Hrvatske – Listu zaštićenih kulturnih dobara.

Sukladno članku 12. stavku 5. Zakona o zaštiti i očuvanju kulturnih dobara, žalba protiv ovoga Rješenja ne odgađa njegovo izvršenje.

15. Allegato 2 - Questionario sociolinguistico somministrato

UPITNIK (SAMO)PROCJENE DVOJEZIČNIH GOVORNIKA

Poštovana/i, ljubazno Vas molimo da pažljivo pročitate i popunite upitnik koji je anoniman.

Prikupljeni podatci služe isključivo za znanstveno istraživanje. Unaprijed zahvaljujemo.

1. Opći podatci

Spol: M / Ž Godina rođenja _____ Mjesto rođenja _____ Mjesto stanovanja _____

Koji je Vaš najviši završeni stupanj obrazovanja?

osnovna škola srednja škola viša škola/prvostupnik

fakultetska diploma ili magistar struke poslijediplomski studij

Koji se jezik koristio u vrtiću kojeg ste pohađali? hrvatski talijanski _____

Koji je bio nastavni jezik osnovne škole koju ste pohađali? hrvatski talijanski

Koji je bio nastavni jezik srednje škole koju ste pohađali? hrvatski talijanski

Koji je bio nastavni jezik fakulteta koji ste pohađali (ako ste ga pohađali)?

hrvatski talijanski _____

U kojem ste gradu pohađali fakultet? _____

Obrazovanje majke: osnovna škola srednja škola viša škola/prvostupnica

fakultet/magistrica struke poslijediplomski studij Je li Vaša majka zaposlena? Ako je odgovor potvrđan, navedite profesiju: NE DA (profesija: _____)

Obrazovanje oca: osnovna škola srednja škola viša škola/prvostupnik fakultet/magistar

struke poslijediplomski studij Je li Vaš otac zaposlen? Ako je odgovor potvrđan, navedite profesiju:

NE DA (profesija: _____)

Jeste li zaposleni (ako je odgovor potvrđan, navedite svoju profesiju): NE DA (profesija:

_____) Student/učenik Koliko osoba živi u Vašoj užoj obitelji?

_____ Koji je prosječan godišnji prihod (netto) Vaše uže obitelji, s obzirom na podatke tržišne agencije GfK o prosječnom godišnjem prihodu kućanstva (2012.) od otprilike 94 000 HRK*. Uzmite u obzir cjelokupnu zaradu svih članova obitelji.

znatno manji manji prosječan viši znatno viši

(< od 70.000 HRK) (71 – 90.000 HRK) (91 – 100.000 HRK) (101 – 120.000 HRK) (>120.000) **

Prosječan godišnji prihod je uz obrazovanje članova uže obitelji mjera socioekonomskog statusa obitelji i zato je često uključena u istraživanjima o jezičnom razvoju. Ona će nam omogućiti usporedbu među obiteljima, ali se ovaj podatak neće nigdje eksplicitno navoditi.

1. Jezično znanje

U narednim pitanjima zamolit ćemo Vas da procijenite koliko dobro znate neke jezike. Procijenite ocjenama od 1 do 5. Dakle, ako mislite da nešto nedovoljno znate, upišite 1, ako mislite da nešto znate odlično, upišite 5. Možete upotrebljavati i čitav raspon ocjena između navedenih (2 za dovoljan, 3 za dobar i 4 za vrlo dobar).

Primjerice, ako netko izvrsno razumije talijanski jezik, ali smatra da griješi pri govoru, može svoje razumijevanje procijeniti s 5, a govor nižom ocjenom.

	Hrvatski jezik (1 do 5)	Hrvatski dijalekt (1 do 5)	Talijanski jezik (1 do 5)	Talijanski dijalekt (1 do 5)	Neki drugi jezik (1 do 5)	Neki drugi jezik (1 do 5)
Razumijem						
Govorim						
Čitam						
Pišem						

3. Jezici u okolini

U prvom stupcu navedite ime svog materinskog jezika/(svih) svojih materinskih jezika (hrvatski književni/standardni jezik, hrvatski dijalekt, talijanski književni/standardni jezik, talijanski dijalekt, nešto drugo – upišite što). Ako imate dva ili više materinskih jezika (na primjer, talijanski dijalekt, hrvatski književni/standardni jezik), u drugom stupcu pored svakog jezika/dijalekta napišite s koliko mjeseci života je započelo njihovo usvajanje.

Ime materinskog jezika/dijalekta

Mjesec početka usvajanja

(ako je od rođenja, upišite 0)

1. _____
2. _____
3. _____
4. _____

Materinski jezik/jezici majke:

Materinski jezik/jezici oca:

hrvatski književni/standardni jezik

hrvatski književni/standardni jezik

hrvatski dijalekt

hrvatski dijalekt

talijanski književni/standardni jezik

talijanski književni/standardni jezik

talijanski dijalekt

talijanski dijalekt

U narednim pitanjima zamolit ćemo Vas da odredite koliko se Vi ili netko drugi služite određenim jezikom.

Zamislite da je ukupna uporaba svih jezika 100 %. Pokušajte odrediti koliko postotaka čini svaki od jezika.

Primjerice, za osobu koja govori samo hrvatskim dijalektom, upisalo bi se 100 % kraj toga tog dijalekta.

Za osobu koja govori otprilike podjednako hrvatskim književnim/standardnim jezikom i talijanskim dijalektom upisalo bi se 50 % kod hrvatskog standardnog/književnog jezika, a 50 % kod talijanskog dijalekta.

Ostala bi polja ostala prazna

Važno je da ukupan broj postotaka u svakom pitanju bude 100 %

	hrvatski književni /standardni (%)	hrvatski dijalekt (%)	talijski književni /standardni (%)	Talijski dijalekt (%)	Neki drugi jezik _____ (%)
Kojim se jezikom najčešće služite u kući (uža obitelj)?					
Kojim se jezikom najčešće služite s ostalom rodbinom?					
Kojim se jezikom najčešće služite s prijateljima?					
Kojim se jezikom najčešće služite u školi/na fakultetu/na poslu?					
Kojim se jezikom najčešće služite u javnim ustanovama?					
Kojim se jezikom najčešće služite u provođenju slobodnog vremena (izlasci, razgovori sa susjedima, itd.)?					
Na kojem jeziku brojite?					
Na kojem jeziku mislite?					
Na kojem jeziku sanjate?					
Na kojem jeziku gledate TV?					
Na kojem jeziku čitate knjige/časopise?					
Na kojem jeziku surfate internetom?					
Koji su se jezici govorili u kućanstvu u kojem ste odrasli?					
Koje su jezike govorili drugi članovi obitelji, susjedstva ili kruga prijatelja s kojima ste češće dolazili u kontakt pri odrastanju?					

Navedite koji su članovi kućanstva u kojem ste odrasli govorili kojim jezikom.

hrvatski književni/standardni jezik

hrvatski dijalekt _____

talijski književni/standardni jezik _____

talijanski dijalekt _____

_____ - _____

Navedite koji su drugi članovi obitelji, susjedstva ili kruga prijatelja s kojima ste češće dolazili u kontakt pri odrastanju govorili kojim jezikom.

hrvatski književni/standardni jezik _____

hrvatski dijalekt _____

talijanski književni/standardni jezik _____

talijanski dijalekt _____

_____ - _____

Navedite koji su članovi kućanstva u kojem sada živite govore kojim jezikom.

hrvatski književni/standardni jezik

hrvatski dijalekt _____

talijanski književni/standardni jezik

talijanski dijalekt _____

_____ - _____

Navedite koji su drugi članovi obitelji, susjedstva ili kruga prijatelja s kojima često dolazite u kontakt govore kojim jezikom.

hrvatski književni/standardni jezik _____

hrvatski dijalekt

talijanski književni/standardni jezik _____

talijanski dijalekt _____

_____ - _____

16. Allegato 3 – Consenso

OBAVJEŠTENI PRISTANAK

za sudjelovanje u istraživanju

PROJEKT: Višerazinski pristup govornom diskursu u jezičnom razvoju (UIP-2017-05-6603)

CILJ U PROJEKTU: Izgradnja Hrvatskog pripovjednog korpusa

IZVORI FINANCIRANJA: Hrvatska zaklada za znanost (UIP-05-2017)

NOSITELJ PROJEKTA: Edukacijsko-rehabilitacijski fakultet Sveučilišta u Zagrebu

PROVODITELJI PROJEKTA: Fakultet za interdisciplinarne, talijanske i kulturološke studije Sveučilišta

Jurja Dobrile u Puli, Edukacijsko-rehabilitacijski fakultet Sveučilišta u Zagrebu

VODITELJ PROJEKTA: doc dr. sc. Gordana Hržica

Projekt *Višerazinski pristup govornom diskursu u jezičnom razvoju* (UIP-2017-05-6603) provodi se u svrhu rasvjetljivanja prirode jezične obrade na razini diskursa. U znanstvenim istraživanjima koja proizlaze iz projekta ispitanici su djeca i odrasli s urednim jezičnim statusom, jednojezični ili dvojezični govornici hrvatskoga te osobe s razvojnim jezičnim teškoćama u odrasloj dobi. Jedna od glavnih aktivnosti projekta jest prikupljanje uzoraka govornog jezika osoba kako bi se izgradio Hrvatski dvojezični korpus, prvi ovakav korpus za hrvatski jezik. Korpus hrvatskog jezika i istromletačkog dijalekta u okviru hrvatsko-talijanske dvojezičnosti u Istri omogućit će nam da saznamo više o osobinama jezika dvojezičnih govornika talijanskog i hrvatskog, ali i da proučimo i očuvamo specifične osobine istromletačkog dijalekta.

Ovo ispitivanje služi u svrhu prikupljanja jezičnih uzoraka govora dvojezičnih odraslih govornika. Ispitivač će snimiti zvučni zapis razgovora dvojezičnih govornika u trajanju od 15 minuta. Prijepisi snimki bit će dio Hrvatskog dvojezičnog korpusa koji će se rabiti za istraživanja u ovom projektu, ali će biti otvoren i za druge stručnjake. No, prijepisi će sadržavati anonimizirane podatke o ispitanicima (prezimana se neće navoditi, imena će biti promijenjena, ali bit će dostupni podatci o gradu/mjestu snimanja i godinama sudionika).

Razina stresa i/ili neugode u ovom istraživanju nije veća od one koju se doživljava u vrlo uobičajenim svakodnevnim situacijama. U ovom istraživanju nećete biti izloženi nikakvom specifičnom riziku.

Naglašavamo da je sudjelovanje u ovom istraživanju dobrovoljno i da imate pravo bez ikakvih posljedica odustati od sudjelovanja ili se iz njega povući. Ako imate neke pritužbe na provedeni postupak ili ste zabrinuti zbog nečega što ste tijekom istraživačkog postupka doživjeli, molimo vas obratite se na ghrzica@erf.hr.

Povjerljivost informacija o Vašem identitetu je u istraživanju zajamčena. Pristup podacima imat će samo istraživači. Ako nalazi ovog istraživanja budu javno objavljeni, objavit će se samo za čitavu grupu sudionika, te se neće objavljivati individualni odgovori.

Potvrđujem da sam pročitao/pročitala ovu obavijest za gore navedeno znanstveno istraživanje te sam imao/imala priliku postavljati pitanja. Pristajem na sudjelovanje u navedenom znanstvenom

istraživanju.

Ime i prezime ispitanika: _____

Datum rođenja ispitanika: _____

Potpis: _____

Mjesto i datum: _____

17. Riassunto

La prima sezione di questo lavoro mette in risalto i momenti più importanti di stampo diacronico che hanno contribuito alla formazione della situazione sociolinguistica odierna nel territorio istriano. In seguito vengono spiegati alcuni dei più rilevanti approcci allo studio della commutazione di codice, quali il modello della matrice (*matrix model*) il modello della marcatura (*markedness model*) e il modello basato sull'utilizzo (*usage-based model*). La parte analitica consiste nella presentazione dei risultati di un'indagine sperimentale focalizzata sullo studio del fenomeno della commutazione di codice nel contesto del modello *usage-based* nell'ambito del corpus istroveneto-croato, comprendente trascrizioni di conversazioni spontanee la cui lingua matrice è l'istroveneto, caratterizzate da frequenti commutazioni di codice (in croato). La prima parte analitica implica l'analisi delle caratteristiche grammaticali della commutazione di codice, ossia le concordanze evidenti dei modelli ripetitivi che rappresentano le preferenze di combinazioni lessicali specifiche che coinvolgono l'istroveneto e il croato. La seconda parte, invece, si basa sull'analisi degli esempi di commutazione di codice che riflettono le funzioni conversazionali individuate da Gumperz (1982) e Baker (2010), come tipiche della commutazione di codice nelle comunità bilingui in seno alla modalità bilingue.

Parole chiave: bilinguismo, commutazione di codice, modello *usage-based*, dialetto istroveneto, corpus istroveneto-croato, corpus del parlato, CHAT, CLAN.

18. Summary

The first section of this work highlights the most important diachronic moments that have contributed to the formation of today's sociolinguistic situation in the Istrian territory. Some of the most relevant approaches to the study of code switching are explained, such as the matrix model, the markedness model, and the usage-based model. The analytical part consists in the presentation of an experimental examination focused on the study of the phenomenon of code switching in the context of the usage-based model within the Istrovenetian-Croatian corpus, including transcriptions of spontaneous conversations, characterized by frequent language alternations where Istrovenetian dialect presents the matrix language. The first part of the analysis presents the grammatical characteristics of code switching explaining the evident concordances of the repetitive patterns that show the preferences of specific lexical combinations involving Istrovenetian dialect and Croatian language. The second part is based on the analysis of the examples that reflect the conversational functions identified by Gumperz (1982) and Baker (2010) as typical of code switching in bilingual communities within the bilingual mode.

Key words: bilingualism, code switching, *usage-based* model, Istrovenetian dialect, Istrovenetian-Croatian corpus, spoken language corpora, CHAT, CLAN.

19. Sažetak

U prvom dijelu ovoga rada ističu se najvažniji povijesni događaji koji su pridonijeli formiranju današnje sociolingvističke situacije na istarskom području. Objasnjeni su neki od najvažnijih pristupa u proučavanju fenomena preključivanja kodova, kao što su model *matrix*, model *markedness* i model *usage-based*. Analitički dio temelji se na prikazu rezultata eksperimentalnog istraživanja usmjerenog na proučavanje fenomena kodnog preključivanja u kontekstu *usage-based* modela unutar istromletačko-hrvatskog korpusa, uključujući transkripcije spontanih razgovora u kojima je istromletački dijalekt obilježenim čestim preključivanjem koda (na hrvatski). U prvom dijelu analize prikazane su gramatičke karakteristike kodnog preključivanja objašnjavajući repetitivno slaganje riječi koji pokazuju sklonosti ka uporabi specifičnih leksičkih kombinacija koje uključuju istromletački dijalekt i hrvatski jezik. Drugi dio se temelji na analizi primjera preključivanja kodova koji odražavaju razgovorne funkcije koje su Gumperz (1982) i Baker (2010) identificirali kao tipične u dvojezičnim zajednicama unutar dvojezičnog načina komunikacije.

Ključne riječi: dvojezičnost, kodno preključivanje, model *usage-based*, istromletački dijalekt, istromletačko-hrvatski korpus, korpus govornog jezika, CHAT, CLAN.