

ILARIA FIORENTINI, MARCO FORLANO

Duj bers et' u vavar tredici anni. Il continuum tra insertion e alternation nel parlato bilingue italiano-sinto lombardo¹

Sulla base di un corpus di parlato bilingue italiano-sinto lombardo, il presente contributo intende analizzare alcuni casi di *code-mixing* intermedi rispetto ai prototipi di *insertion* e *alternation* (Muysken 2000, 2013). A tal fine, all'appoggio di Muysken vengono affiancati quelli di stampo *usage-based* introdotti da Backus (1996, 2003). Dall'analisi emerge come, indipendentemente dal loro rapporto di dipendenza sintattica rispetto al resto della frase nel codice più attivo, gli *switch* presenti nel corpus possano ricondursi a meccanismi di produzione simili. Inoltre, viene mostrato come i rapporti tra *insertion* e *alternation* debbano essere intesi come più continui che discreti in quanto, a partire da inserzioni di singole parole, fenomeni di *triggering* possono attivare in maniera incrementale il materiale del codice meno attivo, fino ad attivare *switch* completi verso una nuova lingua di base dell'interazione.

Parole chiave: *code-mixing*, *switch* estesi di tipo insertivo, *insertion*, *alternation*, sinto lombardo

1. Introduzione

La descrizione e l'analisi dei fenomeni di contatto che emergono nel parlato bilingue rappresenta tradizionalmente un campo molto dibattuto: in particolare, gli studi si sono concentrati (e scontrati) sulla proposta di modelli di classificazione che fossero in grado di dare complessivamente conto dei fenomeni di interazione tra codici. Fin dai primi lavori (cfr. p. es. Poplack 1980, Myers-Scotton 1993), la discussione sul *code-mixing* si è concentrata sulla problematica individuazione di “costanti nell’uso combinato dei codici che in qualche

¹ Il contributo è frutto del lavoro congiunto dei due autori. Ai soli fini accademici, si attribuiranno i §§1, 2 e 3 a Ilaria Fiorentini e 4, 5 e 6 a Marco Forlano.

modo arginassero o vincolassero la libertà combinatoria del parlante bilingue” (Ciccolone 2014: 99-100). Tra le diverse proposte teoriche, quella di Muysken (2000, 2013) ha avuto il merito di conciliare gli esiti a cui si è pervenuti, integrando i diversi modelli (con le rispettive restrizioni sintattiche) in un modello unitario, che riconduce le formulazioni precedenti a tre sottogruppi distinti. I tre processi individuati da Muysken (*insertion*, *alternation*, *congruent lexicalization*) presentano diversi tipi di vincoli strutturali, e possono essere riassunti come segue (Ciccolone 2014: 100):

- l’*insertion* prevede l’inserimento di singole parole o costituenti di una lingua A in una lingua B (la *matrix language*; cfr. Myers-Scotton 1993); l’attivazione del codice meno attivo è prevalentemente lessicale;
- l’*alternation*, invece, consiste in un vero e proprio *switch* da un codice A a un codice B in punti di equivalenza strutturale tra le due lingue; è proprio la presenza di strutture equivalenti a determinare il passaggio di codice, in linea con l’*Equivalence Constraint* teorizzato da Poplack (1980);
- infine, con *congruent lexicalization* si intende quella situazione in cui “the two languages share a grammatical structure which can be filled lexically with elements from either language” (Muysken 2000: 6); si può verificare in contesti bilingui dove le lingue in contatto sono strutturalmente molto vicine, o per vicinanza genetica o a seguito di processi di convergenza.

Per determinare se un fenomeno vada considerato come prevalentemente insertivo o alternante si può applicare una serie di criteri diaagnostici di natura sintattico/posizionale; l’*insertion* riguarda singoli costituenti (parametro *single constituent*), parole contenuto (*content word*), integrazione morfologica (*morphological integration*) e sintattica (*selected element*); l’*alternation* si caratterizza, tra gli altri, per una maggiore lunghezza e complessità (*length and complexity*), più costituenti (*several constituents*), autonomia sintattica (*absence of selection*; Muysken 2000: 102). Tra i due tipi esistono tuttavia dei casi intermedi, non prototipici, per cui diventa necessario prevedere zone di transizione (cfr. Backus 1996, 2003; Demirçay & Backus 2014).

Il nostro studio prenderà dunque in considerazione casi di *code-mixing* intermedi rispetto ai prototipi di *insertion* e *alternation*,

definiti “*switch* estesi di tipo insertivo”, ovvero *switch* di più parole che non superano la dimensione del singolo costituente sintattico, inseriti in una frase prodotta in un altro codice più attivo (Ciccolone & Dal Negro 2021: 99). Tra i parametri sopra elencati, maggior peso sarà dato alla dipendenza vs. autonomia sintattica (*selected* vs. *non selected elements*). L’obiettivo principale è analizzare le loro forme e funzioni in un corpus (descritto al par. 3) di parlato bilingue italiano-sinto lombardo (una varietà di romaní dell’Italia settentrionale), con un duplice fine:

- verificare la loro possibile natura di “unità” nel lessico mentale dei parlanti bilingui (cfr. Backus 2003);
- ipotizzare un possibile meccanismo unitario alla base della produzione di tali *switch*, suggerendo un’unica tipologia per casi spesso tenuti distinti negli approcci di natura più sintattica al *code-mixing*, come quello di Muysken (2000).

Il contributo è strutturato come segue: il background teorico su cui si basa l’analisi sarà presentato al §2, mentre il §3 sarà dedicato alla descrizione dell’area di ricerca e della raccolta dati. Il §4 illustrerà la metodologia adottata, mentre nel §5 saranno discussi e analizzati i dati raccolti. Alcune conclusioni saranno tratte al §6.

2. *Background teorico*

Come anticipato, l’analisi si concentrerà sugli *switch* estesi di tipo insertivo (Ciccolone & Dal Negro 2021: 99), ovvero inserzioni di più parole di un codice A meno attivo in una frase prodotta in un codice B più attivo, dall’estensione coincidente con un singolo costituente sintattico. Seguendo l’*Adjacency Principle* teorizzato da Muysken (2000: 61), la costituenza va intesa in questi termini: “if in a code-mixed sentence two adjacent elements are drawn from the same language, an analysis is preferred in which at some level of representation (syntax, processing) these elements also form a unit”.

Da un punto di vista interpretativo, gli *switch* estesi di tipo insertivo risultano intermedi rispetto, da un lato, alle *insertions* di singole

parole, e, dall'altro, agli *switch* di tipo alternante² (Muysken 2000). Consideriamo l'esempio (1)³:

- (1) *dživena=lo* *inglan di generazione in generazione*
 portare.PRS.3PL=OBJ.3SG.M avanti
 'lo portano avanti di generazione in generazione'

Sebbene il segmento in italiano non rappresenti un'inserzione prototipica (poiché non è composto da una singola parola), l'esempio non è considerabile nemmeno un'alternanza prototipica, perché ha una qualche dipendenza sintattica dalla porzione di frase nel codice più attivo e non è composto da "several constituents", un parametro considerato molto importante da Muysken (2000: 96) per la definizione di uno *switch* come *alternation*.

L'ambiguità nell'interpretazione di *switch* coincidenti con singoli costituenti secondo i parametri di Muysken risiede nel fatto che il parametro *single constituent* da solo non è sufficiente a distinguere tra *insertion* e *alternation* (cfr. Hakimov 2021: 5). Pertanto, tale parametro deve essere considerato almeno in rapporto a un altro parametro sintattico, in particolare il parametro *selected element*, che dà conto dell'autonomia dello *switch* nell'enunciato (cfr. Deuchar *et al.* 2007; Ciccolone 2014; Ciccolone & Dal Negro 2021: 80-81). In base a quest'ultimo, un caso di *code-mixing* tra sinto lombardo e italiano come quello in (2) si qualificherebbe come *insertion*, mentre un caso come (3) come *alternation*:

- (2) *karavas* *i* mediatrice culturale
 fare.IMP.F.1SG ART.F
 'facevo la mediatrice culturale'

- (3) *i=lu* *tacadu* *par*
 essere.PRS.3SG=SBJ.M attaccato.M.SG su
tute ventiquattro su ventiquattro
 2SG.LOC
 'ti sta attaccato ventiquattro su ventiquattro'

² Va precisato che, nel modello di Muysken, sono intesi come casi di *alternation* anche *switch* di singoli elementi periferici e sintatticamente isolati, per esempio gli *switch* di avverbi o segnali discorsivi; tuttavia, questa interpretazione è controversa (cfr. p. es. Ciccolone & Dal Negro 2021: 81).

³ In questo esempio e nei successivi, le forme in sinto lombardo sono indicate in corsivo e quelle italiane in tondo.

In casi del genere, dunque, il criterio strutturale migliore per distinguere tra le due sembra essere quello dell'autonomia rispetto alla dipendenza (cfr. Ciccolone 2014; Ciccolone & Dal Negro 2021): l'*insertion* riguarderebbe *selected elements*, ovvero costituenti non autonomi nel codice meno attivo “che rappresentano un elemento argomentale in una predicazione nel codice maggiormente attivo” (Ciccolone 2014: 110), (“objects or complements” in Muysken 2000: 63); l'*alternation* riguarderebbe invece *non-selected elements* (ovvero, *adjuncts*), come vedremo meglio al §5.

Una visione diversa sulla questione arriva dagli approcci *usage-based* al *code-mixing*, che pongono l'accento sul radicamento (*entrenchment*) dell'elemento commutato nel lessico mentale dei parlanti bilingui. In particolare, Backus (2003) propone la *Unit Hypothesis*, per la quale sequenze di più morfemi di una lingua A inserite in una frase in una lingua B più attiva sono spesso “unità” nel lessico mentale dei parlanti. Seguendo la Grammatica Cognitiva di Langacker, con unità si intende una “thoroughly mastered structure, i.e., one that a speaker can activate as a preassembled whole without attending to the specifics of its internal composition” (Langacker 1990: 15), ovvero senza attivare schemi morfosintattici produttivi. Le unità possono essere di qualsiasi lunghezza e complessità e includere costruzioni parzialmente specifiche (con *slot* fissi e altri liberi); come esempio di costruzione parzialmente specifica, Backus (2003: 88) cita il caso di [*a ... ago*], in cui lo *slot* vacante può essere riempito da un numero piuttosto limitato di varianti lessicali, come per esempio *year*, *month* o *day*. I criteri diagnostici che permettono di individuare un'unità sono il suo significato non composito (ovvero idiomatico o altamente convenzionale), l'irregolarità morfologica, o una frequenza di co-occorrenza particolarmente elevata dei suoi elementi costitutivi; lo statuto di unità o meno non è comunque direttamente falsificabile (Backus 2003: 87).

Secondo la *Unit Hypothesis*, nel *code-mixing* gli *switch* estesi di tipo insertivo costituirebbero potenzialmente delle unità. Un punto centrale per il presente lavoro è che le unità possono essere autonome (*non-selected*) rispetto alla porzione di frase nella lingua più attiva, ed essere inoltre extra-frasali (Demircay & Backus 2014, Goria 2021); proprio perché si tratta di unità, il loro meccanismo di produzione sarà probabilmente soltanto lessicale e dunque più vicino all'*insertion*, contrariamente all'in-

terpretazione di Muysken. Tuttavia, anche negli approcci *usage-based* la dipendenza sintattica rimane importante: secondo Backus (2003: 129), è probabile che la maggior parte dei contro-esempi a questa ipotesi “will turn out to be EL [embedded language] locative and temporal PPs [propositional phrases], because these are often relatively independent of the rest of the clause (...) and may thus be more like alternational than like insertional CS [code-switching]”.

In conclusione, criteri lineari e di dipendenza sintattica non sembrerebbero sempre sufficienti per distinguere tra *insertion* e *alternation*; *switch* più lunghi, complessi e sintatticamente autonomi possono risultare da un’attivazione soltanto lessicale, purché rappresentino unità e non schemi morfosintattici produttivi del codice meno attivo. La domanda di ricerca a cui si tenterà di rispondere è se sia possibile ricondurre gli *switch* estesi di tipo insertivo a un’unica tipologia, indipendentemente da criteri sintattici (*selected* vs. *non-selected elements*), confermando la *Unit Hypothesis*. Si intende dunque verificare fino a che punto i due processi abbiano una tendenza comune, ovvero la possibile natura di “unità”.

3. *Contesto della ricerca e raccolta dati*

L’analisi si concentra sul sinto parlato in Lombardia, in un contesto di contatto prolungato e intenso con l’italiano e (oggi in misura minore) con le varietà italoromanze circostanti. La comunità rom e sinta è presente in Italia dalla fine del XV secolo; la presenza di sinti lombardi è attestata continuativamente dall’Età Moderna (XVI-XVII secolo; Piasere 2009). Tali popolazioni sono diffuse su tutto il territorio nazionale, e sono particolarmente numerose in Piemonte e Lombardia per quanto riguarda i gruppi sinti, mentre i gruppi rom sono stanziate in Italia centrale e meridionale (soprattutto tra Abruzzo e Molise, ma anche in Puglia, Basilicata, Calabria e Campania; Scala 2020: 86). Nonostante si tratti di una comunità nel suo complesso piuttosto consistente, non è facile stabilirne con certezza l’entità: le stime disponibili variano tra gli 80.000 e i 200.000 individui⁴ (cfr. Soravia 2011); di questi, circa 70.000 sarebbero cittadini italiani appartenenti a comunità di antico insediamento (Scala 2020: 87).

⁴ Una stima inferiore è riportata in Videsott (2016: 501), che ipotizza circa 23.000 parlanti.

La documentazione precedente all’arrivo in Italia di questi gruppi è scarsa; per quanto riguarda la popolazione sinta, la forte presenza lessicale tedesca nelle loro varietà testimonia il passaggio da territori germanofoni (Piasere 2009). Per tale popolazione (per cui si stima in circa 30.000 unità; cfr. Forlano 2021), si riconoscono due gruppi principali: sinti italiani, a loro volta suddivisibili su base regionale (piemontesi, lombardi, emiliani, marchigiani, veneti...); e sinti tedeschi, stanziati nell’Italia nord-orientale, che si distinguono in *estrexarja* (‘austriaci’), *eftavagarja* (‘delle sette carovane’), *krasarja* (‘del Carso’). Questi ultimi sono definiti sinti *tejč* (‘tedeschi’) o sinti *gaġkani/e* (‘sinti dei gagi’) dai sinti italiani (cfr. Forlano 2021).

Le varietà parlate da entrambe le comunità derivano da una varietà indoaria centro-settentrionale (Beníšek 2020), ma hanno storie migratorie diverse, con diversi riflessi sulla lingua (Scala 2020): Sorrenti (2014) rileva come quasi il 64% del lessico presente nel sinto lombardo (d’ora in avanti SL) sia costituito da prestiti di origine italoromanza (seguiti dallo strato indiano e da quello tedesco); la varietà, prevalentemente parlata, risulta a oggi priva di una norma ortografica (nonostante esistano tentativi di grafizzazione spontanea; cfr. Desideri 2007; Scala 2015).

Le ricerche di taglio linguistico su queste comunità sono relativamente poco numerose. Come lavori di inquadramento generale si segnalano tra gli altri Soravia (1977, 1998), Scala (2006), Duberti (2010, sui sinti piemontesi), Forlano (2021). Esistono inoltre alcuni studi in prospettiva glottodidattica (Scala 2011, 2020; Sorrenti 2014); non risultano invece lavori relativi al contatto linguistico a livello discorsivo/interazionale. Da un punto di vista sociolinguistico, Scala (2012) rileva, per i sinti lombardi, un repertorio dilalico, che vede l’italiano dominare nei contesti esocomunitari, mentre il SL e l’italiano non sono mantenuti sempre distinti nella comunicazione interna al gruppo. Va inoltre sottolineato come il SL abbia un forte valore identitario, e possa spesso marcare l’appartenenza alla comunità (Scala 2012, 2020). Tuttavia, i parlanti dimostrano anche una certa sfiducia nei confronti della lingua, ritenuta non adatta a usi esterni ai contesti informali e familiari (Scala 2020: 89). In tale situazione, è dunque proprio la condizione di plurilinguismo (più che le singole lingue che compongono il repertorio) a rappresentare “il doppio

e importante ruolo di marker identitario e di modalità negoziata di partecipazione alla realtà sociale italiana” (Scala 2012: 446).

L’analisi si basa su un corpus raccolto tra il 2022 e 2023 presso le comunità di sinti lombardi residenti a Pavia. La città ospita due insediamenti sinti principali (per un totale di poco meno di 400 individui): quello “storico” di piazzale Europa, adiacente al centro cittadino, e quello più recente di via Bramante, nella periferia sud, sorto nei primi anni Duemila. La raccolta dei dati è avvenuta principalmente attraverso interviste semi-strutturate con gli informanti, tutti bilingui italiano-SL; quest’ultimo è stato concordato come lingua di base, ma senza censurare occasionali passaggi verso l’italiano. In totale, il corpus è costituito da circa 7 ore di dati di parlato (interamente trascritti tramite ELAN e corrispondenti a circa 35.000 token), prodotti da un campione di 20 parlanti bilanciato per genere e per fascia d’età.

4. Metodologia

Ai fini della nostra ricerca, abbiamo estratto dal corpus tutte le sequenze di più parole da una lingua meno attiva presenti all’interno di una frase prodotta in un’altra lingua più attiva, dall’estensione non superiore al singolo costituente (definito sulla base dell’*Adjacency Principle*; v. §2).

La lingua più attiva della frase è stata operativamente identificata come quella che fornisce la morfologia flessiva del verbo finito (cfr. Myers-Scotton 2002; Blokzijl *et al.* 2017; Parafita Couto *et al.* 2024). Sono quindi stati esclusi dall’analisi *switch* di singoli costituenti coincidenti con nomi propri (come *Gatto con gli stivali*) o chiaramente ascrivibili all’*alternation* per la presenza di altri tratti diagnostici di questo stile come il *doubling*, il *flagging* o cesure prosodiche nette rispetto al resto della frase (Muysken 2000).

Gli *switch* di singoli costituenti sono stati successivamente annotati secondo alcuni criteri strutturali, tra cui:

- relazione sintattica con il resto della frase, distinguendo tra *selected switches* e *non-selected switches*;
- forma (ad esempio Sintagma Nominale, Sintagma Preposizionale) e composizione interna (ad esempio N-Agg, N-SP).

In aggiunta, sono stati applicati criteri di natura semantica e legati alla frequenza per valutare la possibile natura dell’elemento commutato

come *unit*. A questo proposito, dove possibile è stato riportato il *logDice* (Rychlý 2008) degli elementi costitutivi dello *switch*, ottenuto tramite una ricerca nel corpus italiano itTenTen20. Il *logDice* è un indice statistico che misura la probabilità di due o più parole di co-occorrere piuttosto che di combinarsi con altri elementi in una lingua, e può dunque essere considerato un indicatore della salienza di una data collocazione. La scelta del corpus italiano è motivata in quanto, come vedremo meglio nel prossimo paragrafo, gli *switch* presenti nel corpus si realizzano verso l'italiano. Un valore del *logDice* superiore a zero indica una significatività statistica, ma in diversi studi sono state individuate soglie più alte e specifiche (p. es., Cao & Deignan 2019 considerano salienti le collocazioni con *logDice* superiore a 4).

5. Analisi dei dati

5.1 Panoramica generale

Dal corpus sono stati estratti in totale 475 *switch* estesi che soddisfano i parametri delineati. Una prima considerazione è che, come anticipato, tutti gli *switch* si verificano dal SL verso l'italiano. La motivazione principale risiede probabilmente nel fatto che il SL è la lingua di base delle interviste e l'italiano trova spazio nei dati soltanto sotto forma di *switch*. Nondimeno, questa distribuzione sembra in linea con l'ecologia delle pratiche discorsive della comunità, in cui il *code-mixing* si verifica in maniera sistematica quando la lingua di base è il SL, non quando è l'italiano (cfr. Dal Negro 2005 per riflessioni analoghe in altri contesti minoritari).

In secondo luogo, una prima distinzione tra *selected switches* (§5.2) e *non-selected switches* (§5.3) evidenzia che i secondi sono leggermente più frequenti, come mostrato in Tabella 1.

Tabella 1 - Distribuzione tra *selected switches* e *non-selected switches* nel corpus

<i>Tipo di switch</i>	<i>Ocorrenze</i>
Selected switches	219 (46,1%)
Non-selected switches	256 (54,9%)
<i>Totale</i>	475

Da un punto di vista sociolinguistico, la presenza di *switch* estesi di tipo insertivo appare piuttosto equamente distribuita nei dati di tutti i parlanti del campione e non sembrerebbe dunque dipendere dalle loro caratteristiche socio-anagrafiche.

5.2 Selected switches

La maggior parte dei *selected switches* risulta costituita da sintagmi “pieni” (nominali o preposizionali) o da sintagmi nominali “misti”, introdotti cioè da determinanti o quantificatori in SL. Si noti tuttavia che i sintagmi nominali “pieni” sono sempre sprovvisti di determinanti, come nell’esempio in (4); quando invece i determinanti sono espressi, appaiono pressoché sempre in SL, come nell’esempio in (5).

- (4) *karavas* [ginnastica artistica]

fare.IMP.F.1SG
‘facevo ginnastica artistica’

- (5) *is* [*ja* persona normale]

essere.IMP.F.3SG INDF
‘era una persona normale’

Di seguito, indipendentemente dal loro coincidere con costituenti pieni o con strutture sub-sintagmatiche, vengono analizzati i *selected switches* maggiormente rappresentati nei dati, ovvero: (a) sequenze Nome-Aggettivo, Aggettivo-Nome e Nome-Sintagma Preposizionale; (b) sequenze Numerale-Nome; (c) sintagmi preposizionali pieni.

5.2.1 Sequenze [N-AGG], [AGG-N] e [N-SP]

Una prima tipologia di *selected switches* verso l’italiano è rappresentata da combinazioni di nome e aggettivo (N-Agg o Agg-N, quest’ultima più rara) e di nome e sintagma preposizionale (N-SP), per un totale di 47 occorrenze.

Si può notare che un numero non trascurabile di queste sequenze presenta un significato chiaramente idiomatico, come *luna di miele*, *fidanzati in casa* o *miseria nera* (es. 6); in virtù di questa caratteristica, esse non sono prodotte composizionalmente dai parlanti, bensì come unità non analizzate.

Più in generale, molti degli esempi in questo gruppo rappresentano il nome convenzionale per particolari classi di entità, risultando salienti nella loro globalità piuttosto che nei singoli elementi costitutivi. Generalmente si tratta di espressioni legate alla realtà istituzionale italiana, prive di possibili equivalenti in SL, come *mediatrice culturale*, *scuola media*, *quinta superiore*, *datore di lavoro*, *conto corrente*, ecc. (es. 7)⁵.

- (7) *is=lu* *ja* conto corrente che
 essere.IMP.F.3SG.=3SG.OBL INDF
is=lu *pren* *but* *love*
 essere.IMP.F.3SG=SBJ.3SG sopra molto soldi(M.PL)
 ‘aveva un conto corrente che aveva su molti soldi’

L'elevata incidenza, nei dati discussi in questo paragrafo, di sequenze di più parole altamente convenzionali nell'italiano monolingue sembra coerente con la *Unit Hypothesis* formulata da Backus (2003). A supporto di ciò, notiamo che 35 esempi su 47 hanno un *logDice* superiore a 4 in italiano. Come tali, in un modello del *code-mixing*, questi casi possono essere accostati alle *insertions* prototipiche, di singole parole.

Tuttavia, non tutte le giustapposizioni di più parole italiane all'interno di frasi in SL si prestano necessariamente a un'interpretazione unitaria, come, ad esempio, nel caso della sequenza *dado intero* (che peraltro, se ricercata in itTenTen20, mostra un *logDice* negativo). Per questi casi, comunque piuttosto rari, sembra più plausibile ipotizzare due diverse inserzioni sincroniche volte a colmare lacune adiacenti nel discorso, o per l'assenza di equivalenti in SL, o per difficoltà di recupero lessicale (si consideri d'altronde che i sinti lombardi, per quanto bilingui, sono comunque sbilanciati verso l'italiano quanto ad accessibilità lessicale; cfr. Sorrenti 2014). Pertanto, la semplice successione lineare di più elementi di una lingua meno attiva nel parlato bilingue

⁵ L'assenza di equivalenti nella lingua ricevente costituisce certamente un fattore che favorisce l'inserzione di tali elementi in discorsi in SL; sulla particolare propensione al prestito di nomi con referenti unici e semanticamente specifici, cfr. almeno Myers-Scotton (1993), Backus (2001) e Matras (2020).

non risulta sempre condizione sufficiente a ipotizzare la loro attivazione come singole unità.

5.2.2 Costruzioni del tipo [Num-ANNI]

Una seconda tipologia di *switch* di sintagmi nominali *selected* è rappresentata da costruzioni con numerali del tipo [Num-Anni], [Num-Mesi] e [Num-Giorni]. Si riconducono a questa tipologia 52 *switch*, spesso usati in funzione di oggetto diretto del verbo ‘avere’.

La tendenza al trasferimento dalla lingua maggioritaria di contatto delle espressioni indicanti età e date è stata ampiamente osservata in letteratura e motivata dalla tendenza dei parlanti a esprimere questi concetti in contesti istituzionali associati alla lingua dominante (cfr. Matras 2020). Tuttavia, questa spiegazione da sola non sembra pienamente appropriata se si osservano tutte le costruzioni indicanti età nel corpus, alcune delle quali sono espresse in SL. Tale variazione è mostrata in (8), in cui una parlante ricorre prima a un sintagma interamente in SL (*duj bers* ‘due anni’) e subito dopo a un sintagma interamente in italiano (*tredici anni*).

(8)	<i>kek</i>	<i>i=lu</i>	<i>duj</i>	<i>bers</i>
	uno	essere.PRS.3SG=3SG.OBL	due	anni (M.PL)
	<i>et'</i>	<i>u</i>	<i>vavar</i>	<u><i>tredici anni</i></u>
	e	ART.DEF.M	altro	
‘uno ha due anni e l’altro tredici anni’				

Uno sguardo più approfondito ai dati mostra che la variazione nella lingua di tali costruzioni è legata al tipo di numero usato, apparendo interamente in italiano quando il numero è superiore a 6 e interamente in SL quando il numero è inferiore a 6. Costituenti misti con nome italiano e numero SL (o viceversa) non sono attestati. Le ragioni di questa distribuzione si potrebbero spiegare considerando che il sistema dei numeri in SL prevede generalmente numeri di origine romani fino a 6, mentre le forme italoromanze coprono quelli successivi (cfr. anche Scala 2017)⁶. Appare dunque evidente che l’accesso da parte dei parlanti bilingui a un numero prestato dall’italiano co-attiva stabilmente la parola successiva in italiano (almeno limitatamente

⁶ Un revisore anonimo (che ringraziamo) ha osservato che alcuni parlanti SL possono utilizzare i numeri romani anche per 20 e per 100. Nel nostro corpus, tuttavia, sono attestate soltanto forme italiane per questi valori, anche al di fuori delle costruzioni di età analizzate.

alle costruzioni qui discusse); in altre parole, un'inserzione lessicale prototipica porta a uno *switch* esteso. A nostro avviso, si tratta di un meccanismo leggermente diverso sia rispetto all'inserzione di una sequenza di più parole come una singola unità lessicale, sia rispetto al ricorso a inserzioni adiacenti autonome (processi entrambi discussi nel paragrafo precedente). Si può inoltre notare come il ricorso a queste costruzioni in italiano possa talvolta innescare uno *switch* completo, di tipo chiaramente alternante, verso l'italiano (es. 9):

- (9) *mujas* ch' *is=lu*
 morire.PRF.3SG essere.IMPF.3SG=SBJ.3SG.M
 venticinque anni ed è stato un lutto
 'è morto che aveva venticinque ed è stato un lutto'

5.2.3 Sintagmi preposizionali

Una terza categoria di *selected switches* (N=55) è rappresentata da sintagmi preposizionali, principalmente dipendenti da verbi di moto o da verbi con significato di 'essere' / 'stare'.

Più nello specifico, notiamo come tali sintagmi si possano ricondurre a due tipologie: (a) locuzioni fisse in italiano, talvolta idiomatiche (come *in giro*, *a posto*, *a momenti*, *in regola*, *a disposizione*, *nel mondo dei sogni*); (b) sintagmi locativi i cui complementi nominali sono toponimi (del tipo *a Roma*, *a Pavia*, *a Piacenza*, *all'Arena di Milano*). Insieme, questi casi costituiscono la quasi totalità dei sintagmi preposizionali *selected* nel corpus.

Mentre per i casi in (a) un'attivazione come singole unità sembra plausibile, la stessa interpretazione sembrerebbe forse meno adeguata per i casi in (b), più compostionali. A questo proposito, va notato che il SL dispone di preposizioni per esprimere relazioni locative; nondimeno, i toponimi nel corpus risultano sempre preceduti da preposizioni italiane. Al contrario, gli altri sostantivi nel corpus (SL e italiani) mostrano una preferenza per preposizioni locative romaní in contesti dominanti SL.

A nostro avviso, ciò che motiva l'emergere di *switch* estesi contenenti un toponimo è, nuovamente, l'azione di fenomeni di *triggering*; in particolare, è il toponimo a essere selezionato dal parlante bilin-

gue e la preposizione italiana viene co-attivata⁷. Questa ipotesi trova riscontri in letteratura. In particolare, analizzando svariati corpora bilingui, Clyne (2003) nota che questi casi di *triggering* si realizzano proprio attraverso nomi propri, e in particolare attraverso toponimi, poiché essi, portatori di doppi indici linguistici, possono facilitare il *code-mixing*. In alternativa, Backus (1996) nota che i toponimi, caratterizzati da una semantica molto specifica, tendono a imporre restrizioni agli elementi grammaticali che possono occorrere nel loro immediato intorno sintattico. Come risultato, la costruzione nella quale il toponimo è inserito può a sua volta diventare piuttosto specifica e, dunque, potenzialmente accessibile come unità ai parlanti durante la produzione discorsiva.

Al termine di questa panoramica, osserviamo come la maggior parte dei *selected switches* individuati ai fini dell’analisi possa ricondursi a due principali meccanismi di produzione:

- una sequenza di più parole è stata inserita come unità a causa del suo statuto di unità già nella lingua di partenza, l’italiano;
- l’inserzione di una specifica parola presa a prestito dall’italiano (come un numerale o un toponimo) ha co-attivato ulteriore materiale italiano nelle sue vicinanze.

5.3 Non-selected switches

Passiamo ora ad analizzare i *non-selected switches*, che nel modello di Muysken (2000) sono visti come più vicini all’*alternation* in virtù della loro condizione sintatticamente autonoma. Per via di tale autonomia, essi spesso sono anche periferici nella frase, un’altra condizione diagnostica di *alternation* (si vedano i parametri “peripherality” e “major clause boundary”; Muysken 2000: 100).

In sede di analisi ci è parso opportuno, seguendo Ciccolone & Dal Negro (2021), distinguere tra *non-selected switches* caratterizzati da una funzione prevalentemente referenziale e *non-selected switches* con una funzione prevalentemente testuale/interazionale. Mentre i primi introducono elementi referenziali nell’universo del discorso, i secondi hanno la fun-

⁷ La direzionalità di tale fenomeno, si noti, può essere individuata con relativa certezza, in quanto non si dà mai il caso contrario di una preposizione italiana che introduce un nome SL in costruzioni locative.

zione di marcare la struttura informativa dell'enunciato o l'interazione stessa. La discussione procederà considerando in breve le due tipologie.

5.3.1 *Non-selected switches* con funzione referenziale

I *non-selected switches* con funzione referenziale sono rappresentati da sintagmi pieni dalla semantica principalmente temporale, locativa o di maniera.

Nuovamente, molti di questi sembrano prestarsi a un'interpretazione unitaria, trattandosi di sequenze altamente convenzionali, se non idiomatiche, in italiano. Un esempio è lo *switch ventiquattro su ventiquattro*, in cui peraltro l'ellissi della testa nominale *ore* contribuisce a rendere la sequenza meno compositazionale (v. es. 3, già discusso al §2). Altri possibili esempi sono *mille volte*, *mattino pomeriggio e sera*, *metà e metà*, ecc.

Tuttavia, rispetto ai *selected switches*, diversi *non-selected switches* con funzione referenziale sembrano caratterizzarsi per un maggiore grado di compositazionalità nell'italiano monolingue, rendendo meno plausibile un'interpretazione come unità. Inoltre, tendono a includere lessemi dalla semantica meno specifica, per i quali sono disponibili equivalenti in SL altamente frequenti. Si tratta dunque di una situazione in parte diversa rispetto a quella osservata per i *selected switches*, costituiti da unità idiomatiche o prestiti (come numerali e toponimi) che hanno co-attivato ulteriore materiale italiano nelle immediate vicinanze.

Esempi di *non-selected switches* compositizionali sono sequenze quali *tutti i giorni*, *tutte le sere* e *tutta la vita*:

(10)	<i>me</i>	<i>čava</i>	<i>kai</i>	<u><i>tutta la vita</i></u>
	1SG.SBJ	stare.PRS.1SG	qui	
			'io sto qui tutta la vita'	

Al netto della compositazionalità semantica, la loro relativa frequenza nel parlato in SL sembra permettere di individuare uno schema ricorrente del tipo [tutto-Det-X] relativamente *entrenched* nel lessico mentale dei parlanti. Infine, alcuni casi, *hapax* nel corpus, rappresentano esempi di possibili creazioni innovative da parte dei parlanti bilingui, che, nel produrre tali *switch*, hanno verosimilmente attivato schemi morfosintattici produttivi dell'italiano. Esempi possono essere gli *switch* dei singoli costituenti *con mio figlio* in (11) e *in modo antipatico e dispregiativo* in (12). Quanto in particolare a quest'ultima

sequenza, se cercata su un motore di ricerca come Google essa non produce alcun risultato, sostenendo l'idea che il parlante l'abbia prodotta *ex novo* nel discorso, piuttosto che reperita come un'unità già immagazzinata nel suo lessico mentale.

- (11) alla fin *xajom* con mio figlio

mangiare.1SG.PRF

'alla fine ho mangiato con mio figlio'

- (12) *joi* *dikel=ma* in modo antipatico e dispregiativo

3SG.SBJ vedere.PRS.3SG=1SG.OBL

'lei mi guarda in modo antipatico e dispregiativo'

Vale la pena riflettere brevemente sul perché, in posizione avverbiale, aumenti l'incidenza di *switch* di sequenze più compostizionali (che non di rado includono materiale grammaticale italiano che non sarebbe ammesso in posizione argomentale, come i determinanti). A nostro avviso, questa circostanza è favorita dalla maggiore autonomia (sintattica e semantica) delle posizioni avverbiali (*non-selected*) rispetto alle posizioni argomentali (*selected*); queste ultime sono maggiormente vincolate dalle restrizioni grammaticali della lingua più attiva della frase (in nostro caso in SL), che ha il suo centro a livello del verbo finito. La condizione di maggiore autonomia caratterizzante gli *slot* avverbiali permetterebbe ai parlanti bilingui di selezionare più liberamente un nuovo codice (l'italiano), attivando schemi caratterizzati da diversi livelli di *entrenchment*, fino al caso di sequenze create *on-the-spot* che prevedono un pieno accesso agli schemi morfosintattici dell'italiano (cfr. anche Backus 1996). Correlando con una maggiore attivazione dell'italiano, casi come questi sembrano potersi collocare leggermente più vicini al polo dell'*alternation*.

5.3.2 *Non-selected switches* con funzione testuale/interazionale

Di particolare interesse risultano infine i *non-selected switches* con funzione testuale/interazionale. Le funzioni a cui essi si riconducono sono piuttosto varie; le principali sono riportate in Tabella 2, seguendo Ciccolone & Dal Negro (2021: 103).

Tabella 2 - *Principali funzioni dei non-selected switches con funzione testuale/interazionale nel corpus*

<i>Funzione</i>	<i>Esempi</i>
<i>Attenuazione/intensificazione</i>	Un po' (<i>x1</i>)
<i>Funzione conclusiva o di chiusura di topic</i>	In poche parole (<i>x2</i>), e bom (<i>x17</i>), e bona (<i>x8</i>)
<i>Gestione del discorso</i>	Tra parentesi (<i>x2</i>)
<i>Esemplificazione</i>	Per dire (<i>x8</i>), per dirti (<i>x1</i>), per esempio (<i>x2</i>)
<i>Riformulazione</i>	Nel senso (<i>x20</i>), cioè nel senso (<i>x4</i>)
General extenders	'Ste cose (<i>x1</i>)
<i>Modulazione del grado di commitment del parlante</i>	Mi sa (<i>x4</i>), non so (<i>x1</i>)

Un primo dato notevole è che questi *switch* sono molto numerosi in termini di *tokens* (rappresentando la categoria più frequente nel corpus, con 146 occorrenze), ma meno in termini di *types*, riducendosi, dunque, a un numero limitato di forme altamente frequenti. Proprio in virtù di ciò, si tratta verosimilmente di forme molto *entrenched* nel lessico mentale dei parlanti, peraltro stabilmente associate a funzioni pragmatiche già in parte indagate nell'italiano monolingue (cfr. Fiorentini & Sansò 2017 su *nel senso*) o in altri contesti minoritari (cfr. Dal Negro & Fiorentini 2014). Date queste condizioni, una loro attivazione come singole *units* sembra plausibile.

Tra gli altri, risultano particolarmente interessanti i casi di *switch* estesi che coinvolgono verbi italiani flessi (i più frequenti sono le forme dei verbi *dire*, *sapere* e *pensare*; cfr. Dal Negro 2015 per altri esempi simili nel parlato bilingue). Questi verbi ricorrono in forme specifiche del loro paradigma, all'interno di costruzioni più ampie che in italiano mostrano un certo livello di grammaticalizzazione. Nei nostri dati, tali costruzioni possono occorrere in posizione periferica in un enunciato a lingua dominante SL, senza realmente attivare una struttura argomentale (es. 13), oppure possono introdurre una completiva introdotta da *che* (es. 14).

- (13) *Na suenas inketane pensa te*
 NEG dormire.IMPF.3PL insieme
 ‘non dormivano insieme, pensa te’

- (14) *Mi sa che i da pal mende*
 essere.PRS di dietro 1PL.LOC
 ‘mi sa che sono dietro di noi’

Alla luce di quanto osservato, secondo una prospettiva *usage-based* sarebbe estremamente improbabile che i parlanti, nel produrre questi *switch*, siano effettivamente ricorsi a schemi morfosintattici produttivi dell’italiano; ad esempio, è difficile che, nel produrre lo *switch pensa te* in (14), il parlante abbia effettivamente coniugato il verbo, mentre è più probabile che abbia reperito la costruzione nel suo insieme come un’unità già pronta. Se dunque, strutturalmente, questi *switch* sembrerebbero rappresentare casi di *alternation*, trattarli come tali porterebbe a ignorare che i parlanti li hanno attivati come singole unità. Un’interpretazione insertiva per casi di questo tipo risulterebbe dunque più appropriata.

6. Conclusioni

In questo contributo abbiamo discusso casi non prototipici di *insertion* e *alternation* in un corpus di italiano-sinto lombardo, mostrando come gli *switch* che si inseriscono in questa “zona grigia”, lungi dal rappresentare un’eccezione, siano frequenti nel parlato bilingue, soprattutto in contesti di contatto intenso e prolungato come quello analizzato (cfr. anche Dal Negro 2024).

La principale domanda di ricerca riguardava la possibilità di ricondurre a un medesimo processo gli *switch* di singoli costituenti *selected* e *non-selected* (che rappresenterebbero rispettivamente casi di *insertion* e *alternation*), eventualmente ridiscutendo l’utilità di questi stessi parametri diagnostici. Abbiamo osservato che, tanto in posizione *selected* quanto in posizione *non-selected*, gli *switch* di singoli costituenti sembrano spesso riconducibili a simili meccanismi di inserzione (sotto forma di singole unità o via *triggering*), indipendentemente dal loro rapporto di dipendenza sintattica con la porzione di frase nel codice più attivo (ovvero indipendentemente dalla loro natura *selected* o *non-selected*). In tal senso, adottare la nozione di *unit* proposta dagli approcci *usage-based* permette di espandere il potenziale esplicativo di

meccanismi di tipo insertivo nel parlato bilingue (cfr. anche Demirçay & Backus 2014), riconducendo a un meccanismo unitario casi spesso tenuti distinti nei modelli più strutturalisti al *code-mixing*.

Nondimeno, abbiamo notato che il parametro della selezione sembra avere un qualche ruolo nell'influenzare la struttura del parlato bilingue. In particolare, come visto soprattutto per i *non-selected switches* con funzione referenziale, in posizione avverbiale i parlanti bilingui tendono ad attivare più di frequente schemi morfosintattici produttivi della lingua meno attiva, realizzando un effettivo *switch* verso l'italiano attraverso il ricorso a schemi morfosintattici produttivi. Ciò si potrebbe probabilmente motivare per la maggiore autonomia (semantica e sintattica) di questa posizione rispetto al centro della frase, su cui agiscono maggiormente le regole della lingua più attiva. In questo senso, i nostri risultati sono coerenti con quanto osservato da Backus (1996), per cui l'*insertion* e l'*alternation* si pongono lungo un *continuum* di autonomia semantica e sintattica dell'elemento commutato.

In conclusione, l'analisi ha confermato come i confini tra i prototipi di *insertion* e *alternation* debbano essere intesi come continui, più che come discreti: l'inserzione di una singola parola da una lingua meno attiva (*insertion* prototipica) può attivare, mediante fenomeni di *triggering*, uno *switch* esteso, che può a sua volta innescare uno *switch* di tipo alternante, determinando un vero e proprio cambio nella lingua di base dell'interazione.

Riferimenti bibliografici

- Backus, Ad. 1996. *Two in one. Bilingual speech of the Turkish immigrants in the Netherlands*. Tillburg: Tillburg University Press.
- Backus, Ad. 2001. The role of semantic specificity in insertional codeswitching: evidence from Dutch-Turkish. In Rodolfo Jacobson (a cura di), *Codeswitching worldwide. II*, 125–154. Berlin/New York: Mouton de Gruyter.
- Backus, Ad. 2003. Units in code-switching: evidence for multimorphemic elements in the lexicon. *Linguistics* 41 (1). 83–132.
- Benísek, Michael. 2020. The historical origins of Romani. In Matras, Yaron & Tenser, Anton (a cura di), *The Palgrave handbook of Romani language and linguistics*, 13–47. London: Palgrave Macmillan.
- Blokzijl, Jeffrey & Deuchar, Margaret & Parafita Couto, M. Carmen. 2017. Determiner asymmetry in mixed nominal constructions: the role of gram-

- matical factors in data from Miami and Nicaragua. *Languages* 2(4). (<https://www.mdpi.com/2226-471X/2/4/20>) (Consultato il 29.11.2024.)
- Cao, Dung & Deignan, Alice. 2019. Using an online collocation dictionary to support learners' L2 writing. In Wright, Clare & Harvey, Lou & Simpson, James (a cura di), *Voices and practices in applied linguistics. Diversifying a discipline*, 233–249. York: White Rose University Press.
- Ciccolone, Simone. 2014. Classificare il code-mixing: una reinterpretazione dei parametri di constituency del modello di Muysken. *Linguistica e Filologia* 34. 95–134.
- Ciccolone, Simone & Dal Negro, Silvia. 2021. *Comunità bilingui e lingue in contatto. Uno studio sul parlato bilingue in Alto Adige*. Bologna: Caissa.
- Clyne, Michael. 2003. *Dynamics of language contact: English and immigrant languages*. Cambridge: Cambridge University Press.
- Dal Negro, Silvia. 2005. Il code-switching in contesti minoritari soggetti a regressione linguistica. *Italian journal of linguistics*, 17 (1). 157–178.
- Dal Negro, Silvia & Fiorentini, Ilaria. 2014. Reformulation in bilingual speech: Italian cioè in German and Ladin. *Journal of Pragmatics* 74. 94–108.
- Dal Negro, Silvia. 2015. Contatto linguistico e organizzazione del discorso: il ruolo dei verbi. In Consani, Carlo (a cura di), *Contatto interlinguistico fra presente e passato*, 83–100. Milano: LED.
- Dal Negro, Silvia. 2024. Sociolinguistic typology and language contact in Northern Italy. In Hans-Bianchi, Barbara & Truppi, Chiara & Vogt, Barbara (a cura di), *Speakers and structures in language contact: pluralistic approaches to change and variation*, 199–224. Berlin/Boston: Mouton de Gruyter.
- Demirçay, Derya & Backus, Ad. 2014. Bilingual constructions: reassessing the typology of code-switching. *Dutch Journal of Applied Linguistics* 1 (3). 30–44.
- Deuchar, Margaret & Muysken, Peter & Wang, Sung-Lan. 2007. Structured variation in codeswitching: towards an empirically based typology of bilingual speech patterns. *International Journal of Bilingual Education and Bilingualism* 10 (3). 298–340.
- Duberti, Nicola. 2010. “Trin kamlé tikné”: studenti sinti a Rocca de’ Baldi (Cuneo). *Bollettino dell’Atlante linguistico italiano* 34. 37–78.
- Fiorentini, Ilaria & Sansò, Andrea. 2017. Reformulation markers and their functions. Two case studies from Italian. *Journal of Pragmatics* 120. 54–72.
- Forlano, Marco. 2021. *La romanì in Italia. Un’indagine sociolinguistica*. Pavia: Università di Pavia (Tesi di laurea).

- Goria, Eugenio. 2021. Complex items and units in extra-Sentential code-Switching. Spanish and English in Gibraltar. *Journal of Language Contact* 13. 540–572.
- Hakimov, Nikolay. 2021. *Explaining Russian-German code-mixing. A usage-based approach*. Berlin: Language Science Press.
- itTenTen20 = <https://www.sketchengine.eu/ittenten-italian-corpus/> (Consultato il 29.11.2024.)
- Langacker, Ronald. 1990. *Concept, image, and symbol: the cognitive basis of grammar*. Berlin/New York: Mouton de Gruyter.
- Matras, Yaron. 2020. *Language contact*. Cambridge: Cambridge University Press.
- Muysken, Peter. 2000. *Bilingual speech. A typology of code-mixing*. Cambridge: Cambridge University Press.
- Muysken, Peter. 2013. Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition*, 16 (4): 709–730.
- Molinelli, Piera. 2014. ‘Sai cosa ti dico? Non lo so, se non me lo dici’. ‘Sapere’ come segnale pragmatico nell’italiano parlato contemporaneo. In Danler, Paul & Konecny, Christine (a cura di), *Dall’architettura della lingua italiana all’architettura linguistica dell’Italia. Saggi in omaggio a Heidi Siller-Runggaldier*, 483–498. Frankfurt am Main: Lang.
- Myers-Scotton, Carol. 1993. *Duellng languages: grammatical structures in code-switching*. Oxford: Clarendon Press.
- Myers-Scotton, Carol. 2002. *Contact linguistics: bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Piasere, Leonardo. 2009. *I rom d’Europa. Una storia moderna*. Roma & Bari: Laterza.
- Poplack, Shana. 1980. Sometimes I’ll start a sentence in Spanish Y TERMINO EN ESPAÑOL: toward a typology of code-switching. *Linguistics*, 18 (7-8): 581–618.
- Parafita Cuoto, M. Carmen, Pouw, Charlotte, Laanen, Rodi & Lopez, Luis. 2024. The role of INFL in code-switching: a study of a Papiamento heritage community in the Netherlands. *Frontiers in Language Sciences* 2. (<https://www.frontiersin.org/journals/language-sciences/articles/10.3389/flang.2023.1288198/full>) (Consultato il 29.11.2024.)
- Rychlý, Pavel. 2008. A lexicographer-friendly association score. In P. Sojka, Petr & Horák, Aleš (a cura di), *Proceedings of recent advances in Slavonic natural language processing (RASLAN)*, 6–9. Brno: Masaryk University.

- Scala Andrea. 2006. La penetrazione della romaní nei gerghi italiani: un approccio geolinguistico. In Banfi, Emanuele & Iannàccaro, Gabriele (a cura di), *Lo spazio linguistico italiano e le "lingue esotiche"*, 493–503. Roma: Bulzoni.
- Scala, Andrea. 2011. Così vicini, così lontani: i parlanti romaní, l’italiano e la scuola. In Bozzone Costa, Rosella & Fumagalli, Luisa & Valentini, Ada (a cura di), *Apprendere l’italiano da lingue lontane: prospettiva linguistica, pragmatica, educativa*, 249–265. Perugia: Guerra.
- Scala, Andrea. 2012. Purché la lingua non sia una sola... Trasformazione dei repertori e conservazione del plurilinguismo presso i Sinti italiani dall’Unità ad oggi. In Telmon, Telmon & Raimondi, Gianmario & Revelli, Luisa (a cura di), *Coesistenze linguistiche nell’Italia pre- e postunitoria*, 393–404. Roma: Bulzoni.
- Scala, Andrea. 2015. Se proprio dobbiamo scrivere, almeno facciamolo come gli altri Italiani. I Sinti dell’Italia settentrionale e la grafizzazione della loro lingua. In Dal Negro, Silvia & Guerini, Federica & Iannàccaro, Gabriele (a cura di), *Elaborazione ortografica delle varietà non standard. Esperienze spontanee in Italia e all'estero*, 97–118. Bergamo: Sestante Edizioni.
- Scala, Andrea. 2017. I numerali da 1 a 10 in sinto lombardo. In Prada, Massimo & Sergio, Giuseppe (a cura di), *Italiani di Milano. Studi in onore di Silvia Morgana*, 789–796. Milano: Ledizioni.
- Scala, Andrea. 2020. La romaní. In Fiorentini, Ilaria, Gianollo, Chiara, Grandi, Nicola (a cura di), *La classe plurilingue*, 85–98. Bologna: Bononia University Press.
- Sorrenti, Elena. 2014. Italiano e sinto lombardo a confronto: somiglianze, divergenze e prospettive didattiche. *Italiano Lingua Due*, 1. 117–147.
- Soravia, Giulio. 1977. *I dialetti degli Zingari Italiani*. Pisa: Pacini.
- Soravia, Giulio. 1998. Zigeunersprachen und Romanisch. Lingue zingaresche e lingue romanze. In Holtus, Günter & Metzeltin, Michael & Schmitt, Christian (a cura di), *Lexikon der Romanistischen Linguistik. VII. Kontakt, Migration und Kunstsprachen. Kontrastivität, Klassifikation und Typologie*, 419–427. Tübingen: Niemeyer.
- Soravia, Giulio. 2011. Comunità zingare. In Simone, Raffaele (a cura di), Enciclopedia dell’italiano. Roma: Istituto della Enciclopedia Italiana. ([https://www.treccani.it/enciclopedia/comunita-zingare_\(Encyclopediadell'Italiano\)/](https://www.treccani.it/enciclopedia/comunita-zingare_(Encyclopediadell'Italiano)/)) (Consultato il 29.11.2024.)
- Videsott, Paul. 2016. Lingue di minoranza, comunità alloglotte. In Lubello, Sergio (a cura di), *Manuale di linguistica italiana*, 484–506. Berlin & Boston: Mouton de Gruyter.