

SIMONE CICCOLONE  
(Libera Università di Bolzano)

## *Classificare il code mixing: una reinterpretazione dei parametri di constituency del modello di Muysken*

*Aim of this paper is to discuss the problem of classification and analysis of code mixing, in particular by observing the different conditions of constituency in the three classes of phenomena described by Muysken (2000): insertion, alternation and congruent lexicalization. After a close reading of Muysken's parameters concerning constituency (as detailed in Bilingual Speech), the paper will propose a partial revision of his classification model, by introducing two scales combined in a matrix for the evaluation of the conditions of constituency. This partial revision will then be tested on different cases of code mixing already analyzed by Muysken (2000) and on selected examples from an Italian-German corpus collected in South Tyrol. The proposed model aims at observing code mixing phenomena in relation to the sociolinguistic context in which they occur, in order to achieve a better understanding on the process and factors involved in the interaction between grammars in language contact situations.*

### *1. Problemi di classificazione: cenni introduttivi*

Numerosi studi nel campo della linguistica del contatto si sono confrontati con la necessità di classificare i fenomeni di interazione tra codici che si presentano nel parlato bilingue, spesso proponendo modelli e tassonomie nuovi o confutando, sulla base dei propri dati, modelli proposti da studi precedenti.

Per quanto si possa cercare di adottare criteri apparentemente molto chiari (almeno in astratto) per la categorizzazione di tali fenomeni, inevitabilmente ogni studioso finisce per imbattersi in almeno un esempio capace di invalidare, o quantomeno mettere in dubbio, il modello teorico adottato.

Per fare un esempio, la ricerca di restrizioni sintattiche al *code mixing* è stata immediatamente seguita, e poi addirittura affiancata dalla ricerca di eccezioni che ne negassero o restringessero la validità: l'*equivalence constraint* e il *free morpheme constraint* (Poplack 1980) vengono nel giro di pochi anni testati, confermati e confutati (cfr. Clyne 1987;

Myers-Scotton 1993; Treffers-Daller 1994; Bentahila/Davies 1995), così come le diverse versioni del *Matrix Language Frame* (Myers-Scotton 1993, 2002; Myers-Scotton/Jake 2000) si alternano a studi che ne mettono in dubbio la validità (cfr. Poplack/Meechan 1995; MacSwan 1999, 2005; nonché Berruto 2000, 2001, 2004).

Paradossalmente, i tentativi di individuare i confini strutturali dell'enunciazione mistilingue hanno messo ulteriormente in rilievo l'alto grado di libertà e creatività linguistica presente nel discorso bilingue, tanto da portare la discussione scientifica a negare in più di un'occasione l'esistenza di *constraints* universali (cfr. Clyne 1987; Gardner-Chloros 2009), oppure ad estendere tali confini oltre quelli di una singola lingua (cfr. MacSwan 2005; Berruto 2011<sup>1</sup>).

Anche il modello tripartito proposto da Pieter Muysken, al quale egli giunge unificando le riflessioni teoriche e le analisi empiriche delle principali linee di ricerca nel campo del *code mixing*<sup>2</sup> (i modelli di Myers-Scotton, Poplack e Clyne), sembra incontrare difficoltà a categorizzare univocamente proprio i casi più ambigui, che risultano maggiormente distanti dai prototipi di *insertion*, *alternation* o *congruent lexicalization*.

Muysken stesso, d'altronde, parla di aree di transizione tra le tre tipologie di *code mixing*, individuando poi una lista di parametri binari (o 'pseudo'-binari, vista la possibilità di valori neutri nello schema) per i

<sup>1</sup> Berruto (2011) giunge in realtà a conclusioni ben diverse da quelle di MacSwan (2005), per il quale di fatto anche nel *code mixing* agisce (localmente) sempre un solo sistema grammaticale (il sistema A per gli elementi di A e il sistema B per gli elementi di B): il fatto che elementi di A o di B possano essere indiscriminatamente composti a formare un enunciato sia secondo i vincoli strutturali del codice A che del codice B (la "somma delle rispettive grammatiche", Berruto 2011: 65) pone il *code mixing* (e di conseguenza il parlato bilingue) in una condizione di maggiore libertà strutturale rispetto al parlato monolingue.

<sup>2</sup> Nonostante molti degli studi citati e delle analisi riportate tendano a usare l'etichetta di *code switching* in senso lato anche per i fenomeni di cui si intende discutere in questo contributo (cfr. Myers-Scotton 2002, Poplack 1987), si è deciso di utilizzare dappertutto la denominazione *code mixing*, o enunciazione mistilingue, per indicare fenomeni di contatto nel discorso bilingue nei quali i codici siano compresenti all'interno dello stesso enunciato, mentre si è limitato l'uso di *code switching*, o commutazione di codice, ai casi di cambio di codice interfrasale (cfr. Berruto 2001, Dal Negro/Guerini 2007). Occorre tuttavia segnalare che alcuni degli esempi riportati possono includere più enunciati; in particolare, quelli dal corpus di italiano-tedesco a contatto al quale faremo riferimento (cfr. § 6) costituiscono di solito interi turni conversazionali. Benché il focus del contributo sia l'interazione tra grammatiche nel *code mixing*, verranno inclusi nella discussione anche casi di *code switching* interfrasale occorrenti in uno stesso turno conversazionale, analizzandoli dal punto di vista strutturale e dell'organizzazione in costituenti così come per i casi di *code mixing*.

quali egli fornisce le combinazioni di valori corrispondenti alle tre classi di fenomeni (cfr. Muysken 2000: 230).

Tale lista di parametri (riprodotta in tab. 1), di là da potersi considerare un mero elenco di caratteristiche per la descrizione delle tre categorie, pone lo studioso di fronte a due ordini di problemi: il primo riguarda la relazione tra questi tratti, ovvero, se tutte le possibili combinazioni di valori non sono equidistribuite, quali tratti tendono a co-occorrere, quali combinazioni sono più frequenti e perché; il secondo, strettamente dipendente dal primo, riguarda la natura stessa di tali tratti, ovvero se vi siano relazioni gerarchiche tra i singoli parametri o, infine, se alcuni di questi siano ridondanti e possano quindi essere omessi<sup>3</sup>.

In questo contributo si propone una reinterpretazione di tale modello, limitatamente alla parte relativa ai parametri di *constituency* (con inevitabili sconfinamenti cursori verso altre parti dello schema), nel tentativo di individuare le correlazioni tra tali parametri e di ridurre conseguentemente le dimensioni del modello al minimo.

Ci si concentra qui sulla *constituency* perché sembra rappresentare uno ‘zoccolo duro’ del modello di Muysken, in cui è unicamente l’analisi strutturale dell’enunciato a permettere la classificazione dei fenomeni. Inoltre, come verrà discusso di seguito, l’analisi della *constituency* è indipendente dalla direzionalità dello *switch*, diversamente da quanto avviene per altre sezioni del modello: ad esempio, per poter classificare un caso di *code mixing* in base agli elementi di cui è costituito lo *switch*, occorre in qualche modo aver già deciso quale sia la lingua base nell’enunciato e quali siano gli elementi dell’altro codice che in esso si inseriscono. Successivi lavori completeranno tale tentativo di riformulazione del modello di Muysken considerando anche le parti qui escluse, allo scopo di costruire uno schema ridotto di parametri scalari computabili operativamente applicabile ai casi reali di *code mixing*.

Occorre precisare che non si pretende qui di formulare un modello nuovo, autonomo e completo, che superi in termini di capacità esplicativa e predittiva quello di Muysken, quanto piuttosto riflettere analiticamente su quest’ultimo, in prospettiva teorica e prima dell’applicazione a casi concreti, allo scopo di evidenziare possibili incongruenze, ridondanze o ambiguità strutturali e di proporre infine una riformulazione del

<sup>3</sup> Si vedano a tal riguardo le osservazioni conclusive di Deuchar/Muysken/Wang (2007: 336-337).

		insertion	alternation	congr. lex.
<b>Constituency</b>	single constituent	+	○	○
	several constituents	-	+	○
	non-constituent	-	-	+
	nested a b a	+	-	○
	non-nested a b a	-	+	+
<b>Element switched</b>	diverse switches	-	○	+
	long constituent	-	+	-
	complex constituent	-	+	-
	content word	+	-	-
	function word	-	-	+
	adverb, conjunction	-	+	-
	selected element	+	-	+
	emblematic or tag	-	+	○
<b>Switch site</b>	major clause boundary	○	+	○
	peripheral	○	+	○
	embedding in discourse	○	+	○
	flagging	-	+	-
	dummy word insertion	+	○	-
	bidirectional switching	-	+	+
<b>Properties</b>	linear equivalence	○	+	+
	telegraphic mixing	+	-	-
	morphological integration	+	-	+
	doubling	-	+	-
	homophonous diamorphs	○	-	+
	triggering	○	○	+
	mixed collocations	○	-	+
	self-corrections	-	+	-

Tabella 1:

*Schema dei parametri (pseudo-)binari proposti da Muysken (2000: 230)*

modello stesso. È quindi prevedibile che tale modello riformulato possa riprodurre (per suo difetto congenito) le stesse debolezze e le stesse limitazioni del modello di partenza. L'operazione di riformulazione e ottimizzazione rimane tuttavia utile, proprio perché rende più evidenti gli errori nel modello e più direttamente individuabili le cause.

In quest'ottica, per la discussione di questa prima parte relativa alla riformulazione dei parametri di *constituency*, si utilizzeranno gli stessi esempi prodotti da Muysken (propri o ripresi da altre fonti precedenti) nella descrizione dei parametri e delle tre classi di fenomeni (cfr. Muysken 2000, in particolare capp. 1, 3, 4, 5 e 8), ai quali si aggiungeranno, a titolo esemplificativo, alcuni casi estratti da un corpus di parlato bilingue italiano-tedesco, raccolto nel corso di un progetto di ricerca dedicato allo studio del contatto linguistico in Alto Adige<sup>4</sup>.

L'obiettivo ultimo di questo lavoro, al di là di un'osservazione analitica del modello di Muysken in vista di una sua possibile ottimizzazione, è quello di contribuire alla riflessione scientifica sul funzionamento strutturale del *code mixing*, indipendentemente dalla classificazione dei fenomeni e dalla tassonomia adottata, in modo da indagarne più da vicino i processi in atto e individuare quali fattori intervengano in tali processi – primo fra tutti, la situazione sociolinguistica in cui tali fenomeni si realizzano – e di proporre, in ultima analisi, possibili generalizzazioni sul contatto linguistico e sulle forme e le dinamiche di interazione tra i codici in contatto.

## 2. *Vincoli sintattici al code mixing: la sintesi di Muysken*

Come già accennato sopra, fin dai primi lavori di Poplack (1980) e Myers-Scotton (1993) la discussione scientifica sul *code mixing* si è concentrata particolarmente sul problema di individuare delle costanti nell'uso combinato dei codici che in qualche modo arginassero o vinco-

<sup>4</sup> Il progetto, dal titolo "Italiano-tedesco: aree storiche di contatto in Sudtirolo e in Trentino", è stato finanziato dalla Provincia Autonoma di Bolzano per il triennio 2011-2014 e coordinato da Silvia Dal Negro. Il corpus raccolto finora nell'ambito del progetto consiste di oltre 16 ore di parlato (quasi integralmente già trascritte e con annotazioni per parte del discorso e codice, per un totale di oltre 120.000 parole). Le registrazioni includono tre tipologie di dati, raccolti in condizioni diverse: (1) dati elicitati (attraverso la tecnica del *map task*, che permette di raccogliere parlato dialogico semi-spontaneo); (2) interazioni guidate (interviste e *focus group*); (3) parlato spontaneo non elicitato (interazioni interne alla famiglia e a gruppi di amici).

lassero la libertà combinatoria del parlante bilingue. Questi studi hanno prodotto esiti diversi, a volte anche difficilmente conciliabili, mostrando quanto il campo d'azione del contatto nel discorso fosse eterogeneo e articolato al suo interno.

Uno dei principali meriti del lavoro di Muysken (2000) è stato quello di riuscire a riunire i principali filoni di ricerca sul *code mixing* e integrarli in un modello unitario. Per far ciò, era tuttavia necessario ridurre la validità dei diversi modelli (e delle restrizioni sintattiche da questi proposti) a sottogruppi distinti di fenomeni. Muysken parla infatti di tre processi diversi, “constrained by different structural conditions, and [...] operant to a different extent and in different ways in specific bilingual settings. [...] The three processes correspond to dominant models for code mixing that have been proposed” (Muysken 2000: 3).

Ai tre processi di *insertion*, *alternation* e *congruent lexicalization* corrispondono quindi modelli diversi, che Muysken distingue sulla base del tipo di vincoli strutturali che questi formulano. I modelli *insertional* “view the constraints in terms of the structural properties of some base or matrix language” (*ibid.*: 3); ovvero, come avviene ad esempio nel modello del *Matrix Language Frame*, una delle due lingue fornisce le strutture (tutti i *system morphemes* per Myers-Scotton 1993, solo gli *outsider late system morphemes* nella riformulazione del modello proposta da Myers-Scotton/Jake 2000), all'interno delle quali possono inserirsi anche elementi dell'altra lingua.

Nei modelli *alternational*, invece, i vincoli sintattici al *code mixing* sono determinati dalla presenza di strutture equivalenti nelle due lingue; ovvero, il passaggio da un codice all'altro è reso possibile, in questo tipo di modelli, solo nei punti in cui le strutture delle due lingue risultino equivalenti, permettendo così di alternare costituenti ben formati nel codice A a costituenti ben formati nel codice B. Sono in linea con questi modelli l'*equivalence constraint* e il *free morpheme constraint*: per il primo, il *code mixing* può avvenire solo nei punti in cui l'ordine dei costituenti nell'enunciato espresso è equivalente nelle due lingue coinvolte; per il secondo, “codes may be switched after any constituent in discourse provided that constituent is not a bound morpheme” (Poplack 1980: 585-586).

Infine, la *congruent lexicalization* è descritta da Muysken come “a situation where the two languages share a grammatical structure which can be filled lexically with elements from either language” (Muysken

2000: 6), e associa a questo tipo di fenomeni di *code mixing* la variazione stilistica e l'interazione tra varietà di una stessa lingua – ambiti di studio che si sono mostrati particolarmente produttivi anche in Italia, mettendo tra l'altro in crisi i modelli precedenti: sia quelli *insertional*, come il *Matrix Language Frame* (cfr. Berruto 2001, 2004, 2011; Cerruti/Regis 2005), sia quelli *alternational*, come l'*equivalence constraint* e il *free morpheme constraint* di Poplack (cfr. Regis 2005).

Agli studi sulla variazione (per i quali cita i nomi di Labov e Trudgill) aggiunge poi, come unica eccezione nel campo della ricerca sul parlato bilingue, il lavoro di Clyne sugli immigrati tedeschi in Australia. Uno degli aspetti più rilevanti delle riflessioni di Clyne, che si discostano notevolmente da altri studi in questo campo, è l'osservazione di processi di facilitazione del *code mixing* da parte di specifici elementi lessicali (nomi propri, prestiti non integrati e *bilingual homophones*) che lui chiama *trigger words* (cfr. Clyne 2003).

In sostanza, alle tre diverse manifestazioni di interazione tra codici a livello intrafrasale sottendono diversi tipi di vincoli sintattici: corrispondenti a quelli di una delle lingue coinvolte, che funge così da lingua 'base' o matrice (*insertion*); legati alla presenza di punti di equivalenza strutturale tra le due lingue, che permettono il cambio di codice nella forma di alternanza tra parti di enunciato autonome nelle due lingue, formate secondo le rispettive regole sintattiche (*alternation*); corrispondenti a una "largely (but not necessarily completely) shared structure, lexicalized by elements from either language" (Muysken 2000: 4).

I vincoli sintattici dell'*insertion* non si applicano quindi ai casi di *alternation*, così come questi non possono applicarsi ai primi, né tantomeno i primi e i secondi possono realizzarsi nei casi di *congruent lexicalization*: la presenza di una struttura largamente condivisa esclude per definizione la possibilità di individuare una lingua base o matrice, così come perde di senso il vincolo sull'equivalenza strutturale dell'*alternation*, visto che essa si realizza in più punti non permettendo, in ultima analisi, una netta distinzione tra parti dell'enunciato formate secondo le regole del codice A e parti formate secondo le regole del codice B<sup>5</sup>.

<sup>5</sup> Riprendendo le osservazioni di Regis (2003: 152-153), si potrebbe interpretare l'*equivalence constraint* ipotizzando che operi anche nei casi di *congruent lexicalization*, in virtù della struttura condivisa che ne è presupposto; tuttavia, l'*alternation* risulterebbe comunque maggiormente vincolata, dovendo rispettare le regole morfosintattiche delle singole lingue nelle parti di enunciato ad esse afferenti, limitando quindi il passaggio di codice ai soli punti di confine tra costituenti autonomi.

Già da questo veloce *excursus* dei modelli del *code mixing* risulta evidente il ruolo centrale della *constituency* nella distinzione tra i diversi processi (e, quindi, i vincoli ad essi applicabili): mentre la prospettiva *insertional* prevede l'inserimento di un costituente della lingua A in un enunciato formato secondo le strutture della lingua B, la prospettiva *alternational* vincola il passaggio di codice alla giustapposizione tra costituenti ben formati nelle due lingue che si alternano in punti in cui le rispettive strutture corrispondono; quando invece si parte dal presupposto che esista una “largely [...] shared structure”, non è più possibile distinguere in modo univoco tra costituenti formati secondo le regole dell'uno o dell'altro codice.

### 3. I parametri della *constituency* nel modello di Muysken

A questo punto, per poter distinguere tra le tre categorie di fenomeni descritte da Muysken, e di conseguenza individuare i vincoli sintattici che ne regolano il comportamento, risulta fondamentale capire in che modo venga definita la *constituency*. A tal proposito, lo studioso formula un principio di base: l'*Adjacency Principle*.

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 (Muysken 2000: 61).

Un costituente è quindi rappresentato da “any syntactic unit, either a lexical item (e.g. a noun) or a phrase (e.g. a prepositional phrase)” che può essere presente in un enunciato mistilingue (*ibid.*: 61).

#### 3.1. *Constituency ed estensione dello switch*

Sulla base di questa definizione della *constituency* è possibile applicare i parametri del modello di Muysken ad essa afferenti. La tabella seguente (tab. 2) riproduce i primi tre parametri del gruppo, con i rispettivi valori assegnati alle tre classi di fenomeni.



Parametri	insertion	alternation	congr. lex.
single constituent	+	○	○
several constituents	–	+	○
non-constituent	–	–	+

Tabella 2: I parametri legati all'estensione del costituente

La classificazione di singoli casi di enunciazione mistilingue sulla base di questi tre parametri è possibile applicando l'*Adjacency Principle*, per il quale il costituente da prendere in esame è quello maggiormente esteso. Nell'esempio seguente, quindi,

- (1) du muasch a *storia tragica* dârzeeln<sup>6</sup>  
'devi raccontare una storia tragica'

l'unità da prendere in considerazione non sarebbe il solo nome *storia*, ma *storia tragica* come blocco unico (che rappresenta tra l'altro una collocazione, ovvero un'unità coesa anche sul piano semantico). In questo caso, visto che l'intera sequenza nel codice B (l'italiano) è un singolo costituente, sarebbe possibile interpretare questo esempio come un fenomeno di *insertion* (almeno in relazione a questi parametri) all'interno di un enunciato in dialetto sudtirolese.

Tuttavia, osservando solo i valori riprodotti nel modello il parametro *single constituent* non ci permette di escludere che si tratti di un caso di *alternation* o di *congruent lexicalization* (che hanno qui un valore neutro, secondo lo schema di Muysken, cfr. tab. 2). D'altro canto, sia nel caso vi fossero più costituenti adiacenti (*several constituents*) sia che gli elementi adiacenti nel codice B non formassero dei costituenti isolabili (*non-constituent*), dovremmo escludere la possibilità che si tratti di una *insertion*.

Ciò significa che qualsiasi tipo di *constituency* diversa da quella di

<sup>6</sup> Negli esempi riprodotti in questo contributo si seguirà la formattazione del testo d'origine quando riportato da altra fonte, segnalando in corsivo le parti considerate dall'autore come *switch* e inserendo le glossature tra apici. Si useranno le stesse convenzioni per gli esempi, come questo, provenienti dal corpus di italiano-tedesco raccolto in Alto Adige (cfr. § 6), indicando in corsivo le parti nel codice meno attivo nel punto dell'interazione preso in esame e segnalando così la direzionalità dello *switch* dal codice A (in tondo) al codice B (in corsivo).

un singolo costituente nel codice B (o singoli costituenti isolati e non adiacenti) è condizione sufficiente ad escludere l'interpretazione come *insertion*; al contrario, la presenza di un singolo costituente nel codice B è condizione necessaria ma non sufficiente per dichiarare che si tratti di *insertion*.

Allo stesso modo, la presenza di *several constituents* adiacenti non sembra essere una condizione necessaria per classificare un fenomeno come *alternation* (visto che questa può occasionalmente riguardare anche un singolo costituente) né tantomeno esclude che si possa trattare di *congruent lexicalization* (che anche qui ha valore neutro). Infine, solo il parametro *non-constituent*, che sostanzialmente attesta la presenza di elementi di entrambi i codici non assegnabili a costituenti diversi e nettamente separabili (cfr. § 3.3), permette di distinguere chiaramente i casi di *congruent lexicalization* dalle altre due classi di fenomeni, rappresentando quindi una condizione necessaria e sufficiente per definire la *congruent lexicalization* in relazione alla *constituency*.

Da tali osservazioni è evidente come questi tre parametri siano strettamente interdipendenti e richiedano perciò di essere analizzati congiuntamente.

### 3.2. Ridondanza del nesting e sua riformulazione

Risultano invece meno chiari e ridondanti i due parametri successivi considerati da Muysken, inclusi nella sezione relativa alla *constituency*, che fanno riferimento alla disposizione lineare del costituente o dei costituenti nel codice B: *nested* e *non-nested*.

La condizione di *nesting* è esemplificata da Muysken già nella denominazione stessa dei due parametri: se il costituente (o la sequenza di costituenti) nel codice B è sia preceduto che seguito da elementi del codice A “grammatically related” (Muysken 2000: 63), allora siamo di fronte a un *nested switch*; in caso contrario, tale costituente o sequenza nel codice B sarà *non-nested*.

Parametri	insertion	alternation	congr. lex.
nested a b a	+	-	○
non-nested a b a	-	+	+

Tabella 3: I parametri nested e non-nested

Vi sono due ordini di problemi riguardo a questi due parametri. Il primo è che, nonostante per definizione sembrano rappresentare uno l'opposto dell'altro, sono considerati separatamente; se poi osserviamo le coppie di valori sulle tre colonne, troviamo segni opposti solo per *insertion* e *alternation*, mentre la *congruent lexicalization* mostra valori ambigui.

La condizione di *nesting* sarebbe quindi sufficiente ad escludere l'*alternation*, ma non la *congruent lexicalization*, e d'altronde non sarebbe sufficiente, per questo stesso motivo, a classificare il fenomeno come *insertion*. L'assenza di *nesting*, benché sufficiente ad escludere l'*insertion*, non permette di distinguere le altre due classi di fenomeni.

Di conseguenza questa coppia di parametri (che dovrebbe essere dicotomica ma che, per i valori che assume nello schema di Muysken e per il fatto stesso che essi vengano tenuti separati, di fatto non lo è) non contribuisce in modo chiaro alla classificazione dei fenomeni di *code mixing*.

Vi è però anche un secondo ordine di problemi, legato alla pertinenza del *nesting* nell'osservazione di tali fenomeni. Si prendano i seguenti esempi, sempre tratti dal corpus di italiano-tedesco raccolto in Alto Adige:

- (2) wâasch wos dâr *focol* isch?  
'sai cos'è il focol?'
- (3) sâl isch dâr *focol*  
'quello è il focol'
- (4) *el focol* # isch sâl # bail sog mân äm  
'il focol – è quella – ascia, come si dice...'

Tutti riguardano lo stesso lemma del codice B (trentino *focol*) inserito in un enunciato nel codice A (tedesco sudtirolese). Se si osserva l'ordine lineare degli elementi, solo nel primo caso si può parlare di reale condizione di *nesting* (nei termini sopra espressi di elemento del codice B inserito tra elementi del codice A strutturalmente legati). Tuttavia, può questa caratteristica di superficie essere una condizione sufficiente a valutare i tre esempi come fenomeni sensibilmente diversi?<sup>7</sup>

<sup>7</sup> A ciò si aggiunga che già Deuchar/Muysken/Wang (2007) avevano osservato come questi parametri non si applichino a tutti i casi di *code mixing*: “‘nested aba’ and ‘non-nested aba’, only apply to switches that have other or matrix language material both before and afterwards. They do not apply to interclausal switches” (Deuchar/Muysken/Wang 2007: 315).

Il (4) ha visibilmente qualcosa di diverso: è in trentino anche il determinante (*el*), mentre negli altri due casi il sostantivo è retto dall'articolo in sudtirolese. Ma questa differenza non ha nulla a che fare col *nesting* così com'è stato definito: se in (2) l'articolo fosse stato in trentino, comunque si sarebbe trattato di un *nested switch* (candidato ad essere classificato come *insertion*), mentre il (3) e il (4), sia che il determinante fosse stato in trentino o in sudtirolese, sarebbero risultati *non-nested*. Questa diversa valutazione però dipende esclusivamente dalle condizioni specifiche dell'evento osservato, che nel primo caso, per via della struttura sintattica utilizzata, porta a produrre la condizione di *nesting*, mentre negli altri due no.

Ora, se il caso (2) è una *insertion*, e i casi (3) e (4) non sono in altro modo diversi se non per la mera condizione accidentale di occorrere a inizio o fine enunciato, perché questi ultimi non dovrebbero essere classificati allo stesso modo?

La condizione di *nesting* risulta eccessivamente sensibile alla mera osservazione lineare, e di conseguenza ai difetti prospettici che tale osservazione può generare. Pertanto il *nesting*, così com'è formulato, risulta una condizione non solo non sufficiente, ma neanche necessaria ad individuare i fenomeni di *insertion*, né tantomeno a distinguerli dalle altre due categorie.

### 3.3. *Costituenti autonomi vs. non autonomi*

Risulta invece più rilevante, anche in relazione alla terna di parametri precedente, un altro aspetto strutturale, che si intende qui introdurre: il grado di 'autonomia' dei costituenti nel codice B rispetto al resto dell'enunciato, ovvero se è possibile distinguere nettamente tra costituenti dell'enunciato interamente in un codice o nell'altro, e se (o quali di) questi siano strutturalmente indipendenti tra loro. I costituenti *nested*, in tal senso, sarebbero da considerare come un sotto-insieme dei costituenti non autonomi.

In primo luogo, proprio in relazione agli esempi riportati precedentemente (che evocano ovviamente la spinosa questione del ruolo del determinante nel *code mixing*, per la quale rimando alle osservazioni di Berruto 2004: 56-61), occorre distinguere tra casi in cui i costituenti dell'enunciato sono composti unicamente da elementi appartenenti a un solo codice e quelli in cui invece si possono trovare costituenti 'misti',

nei quali uno o più elementi del codice B (che formino o meno un'unità) sono inseriti in un costituente con elementi del codice A.

Questa prima distinzione sembra particolarmente pertinente per distinguere casi potenzialmente ambigui di *alternation* e *congruent lexicalization*. Infatti, se andiamo ad analizzare più da vicino gli esempi proposti da Muysken per chiarire i parametri di *constituency*, quello che sembra distinguere più chiaramente le due classi è proprio la presenza di costituenti misti:

- (5) Je dois je dois glisser [*daan veinger*] [*hier*]  
'I have to insert/my finger here' (Treffers-Daller 1994: 213)
- (6) Je téléphone à Chantal, he, [*meestal*] [*voor commiesked te doen*]  
[*en eten*]  
'I call Chantal, hm,/mostly to go shopping and eat' (Treffers-Daller 1994: 213)
- (7) dus Pom yu *begrijp*, san ben lob(i) tak(i) Kip e kon [*op z'n schoot*]  
[*hield zich koest*]  
'Thus Pom you/understand,/who liked [it] that Kip came/on his lap kept quiet' (Bolle 1994: 73)
- (8) èn eeh precies [so wan *geval*] *bijna*  
'And uh exactly such a / case almost' (Bolle 1994: 81)
- (9) *ondanks* [*ellende nanga angri*] in(i) Sranan  
'in spite of misery / and hunger in Surinam' (Bolle 1994: 81)

Gli esempi 5-7 sono riportati da Muysken come rappresentativi della presenza di *several constituents*, mentre 8 e 9 sono esempi di *non-constituent mixing* (cfr. Muysken 2000: 96-97 e 140-141). Nei primi vi sono almeno due costituenti consecutivi interamente in un codice diverso da quello in cui inizia ad essere strutturato l'enunciato, non riconducibili a un costituente unico (i.e. un unico sintagma). Negli ultimi lo *switch* si estende su più costituenti, ma in uno di questi sono presenti anche elementi dell'altro codice; l'elemento nel codice B non è quindi autonomo, senza tuttavia essere necessariamente *nested*, ovvero racchiuso tra elementi del codice A strutturalmente connessi. La presenza poi di altri costituenti contigui con elementi del codice B (autonomi o non) porta ad escludere, per via del principio di adiacenza, l'interpretazione di tali casi come *insertions* (le quali, invece, riguardano di norma costituenti

isolati). Si può invece considerare *insertion* il primo *switch* dell'esempio 7 (*begrijp*): si tratta di un costituente isolato e non autonomo, inserito in un costituente più esteso con elementi del codice A; questa è la situazione tipica dei casi di *nonce borrowing*<sup>8</sup>.

La valutazione della 'autonomia' dei costituenti dovrebbe permettere anche di distinguere casi limite potenzialmente a cavallo tra *insertion* e *alternation*: dato che quest'ultima, nel modello di parametri elaborato da Muysken, non esclude la possibilità di riguardare un singolo costituente, è necessario che vi sia un altro parametro che permetta di distinguere questi casi da quelli di *insertion*; avendo escluso il *nesting* quale parametro pertinente, occorre trovare un tratto discriminante che, indipendentemente dalla sequenza lineare degli elementi nell'enunciato, contrapponga casi di singoli costituenti nel codice B inseriti nella struttura frasale di un enunciato in codice A a casi di *alternation* legata a singoli costituenti isolati all'interno del turno conversazionale.

Si confrontino i seguenti esempi (tratti dal corpus di italiano-tedesco a contatto):

- (10) war schun *na figada*  
'era proprio una figata!'
- (11) e quand che l'è venü # *zum wool*  
'e quando è venuto – salute!'
- (12) perché *wail år so schâi isch* # *POVERINO*  
'perché – perché è così timido – poverino!'

È chiaro che, per distinguere quali di questi esempi siano da classificare come *insertions* e quali come *alternations*, non è più sufficiente considerare la presenza o meno di costituenti 'misti', bensì è necessario incorporare nell'analisi della *constituency* la relazione strutturale tra costituenti singoli isolati nel codice meno attivo e il resto dell'enunciato – ovvero, se tali costituenti assumano un ruolo centrale nella struttura fra-

<sup>8</sup> Non è negli obiettivi di questo lavoro affrontare lo spinoso problema del rapporto tra *code mixing* e prestito, per il quale si rimanda, tra gli altri, a Regis (2003), Matras (2009), Poplack/Dion (2012). Basti qui accennare al fatto che, secondo Myers-Scotton, "from a synchronic point of view, there is no need to make the borrowing vs. codeswitching distinction" (Myers-Scotton 2002: 153), mentre Poplack usa il concetto di *nonce borrowing* per giustificare le eccezioni al *free morpheme constraint* (cf. Poplack 1987: 69; Sankoff/Poplack/Vanniarajan 1990: 74-75); Muysken, infine, inserisce il *nonce borrowing* nella categoria dell'*insertion* (cfr. Muysken 2000: 32).

sale (come elementi argomentali o parte della predicazione) o quantomeno ‘accessorio’ (come elementi circostanziali), o se si trovino piuttosto in posizione marginale, di confine tra due proposizioni o tra un enunciato e il successivo, al punto da poter essere considerati come *switch* esterni alla frase o interfrasali.

Questo tipo di valutazione dell’autonomia dei costituenti ci porta a coinvolgere in realtà altri parametri del modello di Muysken, esterni alla sezione relativa alla *constituency*, e in particolare i parametri *selected element*, *emblematic or tag*, *major clause boundary* e *peripheral*, riportati nelle sezioni *element switched* e *switch site* (cfr. tab. 1).

Parametri	I	A	CL	Autonomia
selected element	+	-	+	-
emblematic or tag	-	+	○	+
major clause boundary	○	+	○	+
peripheral	○	+	○	+

Tabella 4: Parametri interagenti col principio di autonomia dei costituenti

Come si può notare dalla tabella 4, gli ultimi due parametri hanno valori corrispondenti per le tre classi di fenomeni; *emblematic or tag* è quasi corrispondente a questi, con l’unica differenza che l’*insertion*, invece di avere valore neutro, ha valore negativo. Questi tre parametri caratterizzano quindi l’*alternation*, senza tuttavia rappresentare condizioni sufficienti, da sole, ad escludere casi limite di *congruent lexicalization* o *insertion*. Quest’ultima è esclusa solo nel caso di *emblematic switching* (o, più esplicitamente, *extrasentential switching*, cfr. Muysken 2000: 99), ovvero nei casi in cui si può ormai pacificamente parlare di commutazione di codice piuttosto che di vero *code mixing*.

Al contrario, il parametro *selected element* da solo permette di escludere che si tratti di *alternation*; tuttavia, non riesce a distinguere tra casi di *insertion* o di *congruent lexicalization*: mentre l’*alternation* riguarda esclusivamente costituenti autonomi, in cui gli elementi del codice B non sono dipendenti da elementi del codice A (siano essi all’interno dello stesso sintagma, come nel caso dei costituenti ‘misti’, siano essi all’esterno, nella forma di costituenti sovraordinati nei quali il costituente nel codice B rappresenta un elemento *selected*, ovvero un elemento ar-

gomentale nella frase nel codice A), *insertion* e *congruent lexicalization* si realizzano proprio con la presenza di costituenti non autonomi, ‘misti’ o *selected*.

In base a questa formulazione estesa della ‘autonomia’ dei costituenti, possiamo distinguere i costituenti isolati nel codice meno attivo in: (a) non autonomi, quando direttamente dipendenti da costituenti sovraordinati con elementi del codice maggiormente attivo (includendo tra questi i casi di *selected elements*, ovvero costituenti nel codice meno attivo che rappresentano un elemento argomentale in una predicazione nel codice maggiormente attivo); (b) costituenti autonomi, non direttamente dipendenti da costituenti sovraordinati con elementi del codice maggiormente attivo, e che possono perciò trovarsi in posizione marginale (*peripheral switches*, come modificatori avverbiali o elementi dislocati), extrafrasale (*emblematic or tag switching*) o interfrasale (*major clause boundary*).

Questa estensione del concetto di ‘autonomia’ del costituente ci permette di distinguere casi di *insertion* come (10) da casi di *alternation* come (11): in (10), il singolo costituente presente nel codice B (il dialetto trentino) rappresenta un elemento selezionato e isolato da altri possibili costituenti nello stesso codice, per cui non autonomo in relazione alla struttura frasale; in (11), il singolo costituente presente nel codice B (tedesco sudtirolese) rappresenta un elemento extrafrasale (si noti la presenza di una pausa prima e, soprattutto, il cambio di *topic*, dalla narrazione precedente in trentino al brindisi in tedesco sudtirolese che la interrompe), quindi pienamente autonomo rispetto ai costituenti nel codice A benché isolato da altri possibili costituenti nello stesso codice. Nell’esempio (12), infine, sono presenti due costituenti isolati in italiano, separati da una frase in tedesco sudtirolese autonoma rispetto agli elementi nell’altro codice presenti nell’intero turno conversazionale; i due costituenti in italiano sono a loro volta autonomi rispetto agli elementi in tedesco sudtirolese (la presenza di *doubling* è sufficiente a giustificare quest’interpretazione per il primo elemento, così come per il secondo la pausa e il valore connotativo, di valutazione emotiva di quanto espresso nella frase precedente). Secondo un’interpretazione categorica del modello di Muysken, avremmo quindi una prima *alternation* da italiano a tedesco sudtirolese (segnalata dal *doubling*) e una seconda *alternation*, col ritorno all’italiano, in posizione extrafrasale.



#### 4. *La constituency come matrice bidimensionale*

Ora, se rianalizziamo con attenzione i criteri tramite i quali sono state distinte le tre classi di fenomeni presenti negli esempi precedenti, possiamo notare come i parametri di *constituency* presentati da Muysken possano essere ridotti a due dimensioni scalari: il numero di costituenti consecutivi nel codice B (che permette di distinguere tra casi in cui vi sia un *single constituent* e casi in cui vi siano *several constituents* contigui) e la presenza di costituenti non autonomi (che permette di distinguere tra casi in cui lo *switch* sia ben distinguibile e indipendente dal resto dell'enunciato e casi in cui lo *switch* avviene in un punto centrale dell'enunciato o a livello sub-sintagmatico).

Questo secondo parametro, in particolare, permette di estendere il concetto di *nesting* e di rapportarlo con quello di *non-constituent mixing*: qualsiasi *nested switch* comprenderà, in ultima analisi, un costituente non autonomo; se vi sono più costituenti non autonomi consecutivi, sarà difficile parlare di *nesting* e più plausibile l'interpretazione come *congruent lexicalization*.

Il seguente passaggio ripreso da Muysken sintetizza molto bene queste osservazioni:

In insertional code-mixing what is inserted is a constituent. Sometimes several constituents in a row do *not* form a *unique* constituent. When these are switched insertion is not plausible, and alternation or congruent lexicalization are serious possibilities. A number of elements form a unique constituent if that constituent contains no other elements. [...] [W]hen the switched element is a single, well-defined constituent insertion is a plausible option: this holds *a fortiori* for single words. When several words are switched which do not form one or more constituents together, congruent lexicalization is plausible [...] (Muysken 2000: 62).

L'albero decisionale fin qui delineato per la classificazione dei fenomeni di *code mixing* si sviluppa in due passaggi:

1. Quanti costituenti consecutivi nel codice B vi sono?
  - a. Singolo/i costituente/i isolato/i: probabile *insertion*;
  - b. Più costituenti consecutivi: *alternation* o *congruent lexicalization*;
2. I costituenti nel codice B sono autonomi (“*well-defined*”), indipendenti da elementi del codice A?

- a. Costituenti (quasi) esclusivamente autonomi: probabile *alternation*;
- b. (Anche) costituenti non autonomi: *insertion* o *congruent lexicalization*.

Tali criteri di classificazione non sono da interpretare come strettamente binari: nonostante essi sembrino tali sul piano teorico, la realtà osservata mostra non di rado situazioni molto più complesse, che risulterebbero più oscure e difficili da analizzare tramite valori dicotomici.

Occorre poi aggiungere che fino a qui si è parlato di codice A e codice B, intendendo il secondo come il codice in cui si realizza lo *switch* all'interno di un enunciato o di un discorso nel codice A. Tuttavia, anche questa rischia di essere un'astrazione teorica, dato che non è sempre possibile, soprattutto nei casi di *code mixing* intrafrasale più intenso, individuare la direzionalità dello *switch*, ovvero quale dei due codici è da considerare maggiormente attivo in un determinato punto dell'interazione<sup>9</sup>.

Per costruire un modello di classificazione dei fenomeni di *code mixing* sulla base della *constituency* che tenga conto anche dei casi più estremi e complessi, occorre quindi utilizzare dei parametri discreti, non binari ma scalari, che facciano riferimento al numero di costituenti consecutivi nello stesso codice e di costituenti non autonomi presenti nell'enunciato, indipendentemente dal codice e dalla direzionalità dello *switch*, quand'anche chiaramente individuabile. In sintesi, la *constituency* può essere rappresentata tramite i seguenti due valori:

*m*: numero minimo di costituenti consecutivi nell'enunciato in cui compaiono elementi dello stesso codice (esclusi singoli costituenti non autonomi isolati);

<sup>9</sup> Nel modello di Myers-Scotton si parlerebbe di *matrix language*; ma anche fuori da questo modello si tende a ricorrere a concetti simili, benché non corrispondenti sul piano strutturale, tramite i quali poter distinguere il codice atteso in un'interazione monolingue da quello in cui si realizza il passaggio di codice, sia che esso sia inserito tra elementi del codice atteso sia che continui fino alla fine del turno conversazionale (cfr. l'uso di "pragmatically dominant language" in Matras 2009, o di "base language" in Muysken 2000, che tuttavia ne limita l'uso nella descrizione dell'*insertion*). Anche in Berruto (2004) si fa ricorso alla locuzione 'lingua base': "può esistere, o meglio è certo riconoscibile in una data enunciazione particolare, una 'lingua base' per una singola frase o proposizione, ma questo non va confuso con una lingua matrice che governi tutto il discorso commutato, fornendogli gli schemi. Le due lingue agiscono paritariamente" (Berruto 2004: 60). Qui si fa riferimento a quest'idea di 'lingua base' per l'individuazione della direzionalità dello *switch* in un determinato enunciato.

$n$ : numero di costituenti non autonomi nell'enunciato ('misti' o *selected elements* che formano singoli costituenti isolati).

Più nel dettaglio, i valori assunti dalle due scale sono da interpretare come riportato nella tabella 5:

<b>m</b>	<b>costituenti consecutivi</b>
<b>0</b>	uno dei due codici è presente solo in costituenti non autonomi isolati
<b>1</b>	uno dei due codici è presente in un solo costituente autonomo, oltre che in eventuali costituenti non autonomi isolati
<b>2+</b>	uno dei due codici è presente al massimo in due (o più) costituenti consecutivi, oltre che in eventuali costituenti non autonomi isolati

<b>n</b>	<b>costituenti non autonomi</b>
<b>0</b>	uno dei due codici è presente solo in costituenti autonomi
<b>1</b>	uno dei due codici è presente in un solo costituente non autonomo, oltre che in eventuali costituenti autonomi
<b>2+</b>	uno dei due codici è presente al massimo in due (o più) costituenti non autonomi, oltre che in eventuali costituenti autonomi

*Tabella 5: Significato dei valori nelle due scale della constituency*

Come anticipato, le due scale sono indipendenti dal codice e possono essere calcolate prima di decidere la direzionalità dello *switch*: la scala  $m$  considera il minore tra i due valori corrispondenti al numero di costituenti consecutivi nei due codici, mentre la scala  $n$  considera il numero di costituenti non autonomi indipendentemente dalla loro composizione interna.

Dal rapporto tra questi due valori, che rappresentano scale discrete in uno spazio vettoriale a due dimensioni, è possibile classificare i fenomeni in relazione alla *constituency*. Nel diagramma seguente lo spazio della *constituency* viene rappresentato su una scacchiera: sull'asse delle ascisse sono disposti i valori di  $n$ , mentre su quello delle ordinate i valori di  $m$ ; potenzialmente, la scacchiera può estendersi verso valori più alti su entrambe le scale.

		<b>n</b>			
		0	1	2	3
<b>m</b>	0		I		
	1	A	CL		
	2				
	3				

Figura 1: Condizioni minime di constituency per le tre classi di fenomeni

Trattandosi di uno spazio discreto, la distanza tra due punti è pari alla differenza di valori su entrambe le scale (cosiddetta distanza Manhattan<sup>10</sup>). A livello visivo, possiamo rappresentarla col numero di spostamenti di una casella (in orizzontale e in verticale) necessari per raggiungere un altro punto della scacchiera. Ad esempio, la distanza tra *insertion* e *alternation* nel diagramma è pari a 2, mentre tra entrambe queste classi e la *congruent lexicalization* è pari a 1.

Se applichiamo il modello di classificazione descritto sopra, in base al primo criterio possiamo separare la prima riga ( $m=0$ ), che comprende tutti i casi di *mixing* con singoli costituenti non autonomi isolati, dal resto del diagramma, in cui sono presenti almeno un costituente autonomo o due costituenti consecutivi (autonomi o non) in entrambi i codici coinvolti.

Il secondo criterio, invece, divide lo spazio in relazione all'assenza o alla presenza di costituenti non autonomi: la prima colonna ( $n=0$ ) rappresenta casi in cui sono assenti costituenti non autonomi, escludendo così che si possa trattare di fenomeni di *insertion* o *congruent lexicalization*. Il punto d'origine ( $m=0$ ,  $n=0$ ) rappresenta enunciati monolingui, o nei

<sup>10</sup> La matrice della *constituency* rappresenta uno spazio bidimensionale non euclideo, discreto, per il quale vigono le regole della cosiddetta "geometria del taxi". Per calcolare le distanze tra due punti in uno spazio di questo tipo occorre utilizzare la distanza Manhattan, la quale si ottiene sommando il numero di gradi di differenza in una serie di scale discrete. La formula per uno spazio a  $n$  dimensioni è:  $\sum_{i=1, \dots, n} |a_i - b_i|$  (dove  $a$  e  $b$  rappresentano i valori di due punti nella dimensione  $i$ ). In uno spazio bidimensionale, la formula corrisponde a:  $|x_1 - x_2| + |y_1 - y_2|$  (con  $x$  e  $y$  che rappresentano i valori sulle due dimensioni per i due punti considerati). Cfr. Han/Kamber (2006: 387-389).

quali non è possibile individuare con certezza elementi univocamente attribuibili a entrambi i codici. Rientrerebbero in questa casella anche i casi in cui siano presenti singoli *transfer* lessicali ormai assimilati o in via di assimilazione nel sistema del codice maggiormente attivo:

- (13) gânz a grosă polenta  
'proprio una grossa polenta'

I parametri proposti da Muysken possono quindi essere computati all'interno di questo modello assegnando loro dei valori sulle due misure discrete  $m$  e  $n$ , e individuando così le condizioni minime di *constituency* richieste dalle tre classi di fenomeni. Nella tabella seguente i primi tre parametri di Muysken sono integrati con un quarto relativo alla presenza di costituenti non autonomi, che permette una distinzione più chiara e 'ortogonale' rispetto al numero di costituenti consecutivi tra *several constituents* e *non-constituent*:

Parametri	insertion	alternation	congr. lex.
single constituent	$m = 0$	$m = 1$	$m = 1$
several constituents		$m \geq 2$	$m \geq 2$
non-constituent			$m \geq 1$
costituenti non autonomi	$n \geq 1$	$n = 0$	$n \geq 1$
<b>Condizioni minime</b>	<b>(1, 0)</b>	<b>(0, 1)</b>	<b>(1, 1)</b>

Tabella 6: Parametri di constituency tradotti in coordinate in uno spazio discreto bidimensionale

Nella figura 1 sono indicate le condizioni minime di *constituency* per le classi di *insertion*, *alternation* e *congruent lexicalization*, ottenute tramite questa trasposizione scalare dei parametri di Muysken. Sulla base di queste formulazioni è poi possibile procedere alla classificazione dei fenomeni suddividendo lo spazio in aree di pertinenza per le tre classi.

La figura 2 mostra, a sinistra (diagramma *a*), la suddivisione categorica dello spazio in quattro aree: una di pertinenza per ognuna delle tre classi, e una casella che rappresenta il discorso monolingue ( $m=0, n=0$ ). Questa suddivisione rappresenta un'interpretazione 'forte' delle due misure scalari, che potrebbe risultare limitante nelle aree di confine tra

un'area e l'altra, per le quali sarebbe invece opportuno lasciare margine di intervento anche ad altri fattori oltre alla *constituency*.

Se proviamo a calcolare le distanze delle singole caselle dai punti centrali di queste aree (diagramma *b*), otteniamo una matrice con aree di sovrapposizione tra coppie di fenomeni ('I/CL' o 'A/CL') o tra tutte e tre le categorie (segnata con '?'). Questa interpretazione 'debole' del modello permette di mantenere un certo grado di interpretabilità del fenomeno, lasciando spazio di intervento anche a valutazioni connesse ad altri parametri qui ancora non incorporati o analizzati.

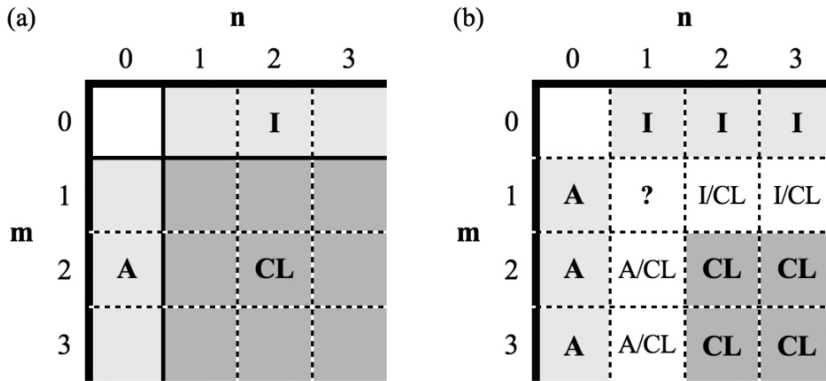


Figura 2: Matrice della constituency per la classificazione dei fenomeni di code mixing

Secondo questa reinterpretazione del modello di Muysken, quindi, benché si riesca a ridurre i parametri per la valutazione della *constituency* a soli due (qui chiamati provvisoriamente *m* e *n*), questi non sono da considerare sufficienti, come non lo erano nel modello originale, a categorizzare in modo univoco i fenomeni di *code mixing* assegnandoli ad una delle tre tipologie.

### 5. Applicazione della matrice della constituency agli esempi di Muysken

Si cercherà ora di applicare tale riformulazione del modello su alcuni esempi classici della letteratura scientifica in quest'ambito, riportando

in particolare quelli utilizzati da Muysken (2000) nella descrizione dei tre fenomeni di *insertion*, *alternation* e *congruent lexicalization*.

In tal modo sarà possibile verificare se tale modello ridotto rispetti due condizioni essenziali: (a) permetta una classificazione pertinente e coerente con la classificazione secondo il modello originario; (b) permetta di classificare con la stessa precisione tutti i casi classificabili tramite il modello originario. Nel caso queste due condizioni vengano rispettate, avremo dimostrato che il modello ridotto raggiunge la stessa efficacia del modello originario con un minor numero di parametri e può quindi sostituirlo.

### 5.1. Esempi di *insertion*

Nella parte iniziale del suo volume *Bilingual Speech*, Muysken presenta alcuni esempi ripresi da precedenti pubblicazioni per illustrare brevemente le tre tipologie di *code mixing*<sup>11</sup>. Vengono di seguito riprodotti gli esempi relativi all'*insertion*, così come sono presentati da Muysken (2000: 4-6).

- (14) kalau dong tukang bikin dong tukang bikin *voor acht personen* dek orang cuma nganga dong makan  
when they always make they always make for eight persons and then people only look they eat  
'When they [cook], it is always for eight people, and then they only look at it, they eat...' (Huwaë 1992)
- (15) na'iish-*crash* lá  
1sg:pass out-crash EMPH  
'I am about to pass out' (Canfield 1980: 219)
- (16) Yo anduve *in a state of shock* por dos dias  
'Ho camminato in stato di shock per due giorni' (Pfaff 1979: 296)

L'esempio (14) mostra un unico costituente in neerlandese inserito in un discorso strutturato in malese delle Molucche. Il costituente non è 'misto', ovvero non ha al suo interno elementi nell'altra lingua, ma è un

<sup>11</sup> Per tali esempi riportati da altre fonti, si indica il riferimento bibliografico tra parentesi come segnalato in Muysken (2000), dando per sottinteso che si tratti di citazioni riprese da quest'ultima fonte. Essendo lo scopo di questo contributo la rivisitazione del modello di Muysken, vengono qui riprodotti solo gli esempi prodotti nella descrizione dei parametri e delle tre classi di fenomeni.

elemento *selected*; rappresenta quindi un costituente non autonomo. I valori di  $m$  e  $n$  sono quindi, rispettivamente, 0 e 1.

Nell'esempio (15) abbiamo un morfema inglese inserito nella struttura verbale in navaho: anche qui i valori sono quindi  $m=0$  e  $n=1$ . In (16), benché lo *switch* sia più esteso, abbiamo comunque un unico costituente in inglese, non autonomo ma dipendente dalla struttura verbale in spagnolo (di nuovo,  $m=0$  e  $n=1$ ).

- (17) Chay-ta *las dos de la noche*-ta chaya-mu-yku  
that-AC the two of the night-AC arrive-CIS-1pl  
'There at two in the morning we arrive' (Muysken 2000: 63)

Muysken presenta l'esempio (17) per esemplificare il *nesting*. In base alla riformulazione del modello, quello che interessa osservare qui non è il fatto che il costituente spagnolo sia inserito sequenzialmente tra due elementi in quechua strutturalmente connessi tra loro, ma che, in termini più generali e indipendenti dalla sequenza lineare, sia da considerare come non autonomo, sia perché elemento *selected* sia perché, ancor più chiaramente, si trova in un costituente 'misto' in cui compare un morfema grammaticale quechua. Si tratta quindi di un singolo costituente non autonomo ( $m=0$ ,  $n=1$ ).

- (18) Paga-wa-y uj qolqe *duro*-wan willa-sqa-yki-taj  
pay-IO-1M one silver hard-with say-PST-2-EMPH  
Se habia comprometido pagarle con plata dura  
'You had promised to pay him with hard cash' (Muysken 2000: 63)

Qui abbiamo più costituenti in spagnolo, tuttavia isolati e per di più non autonomi all'interno della struttura frasale quechua. Se quindi abbiamo un valore diverso per quanto riguarda il numero di costituenti non autonomi ( $n=2$ ), non abbiamo costituenti consecutivi né autonomi ( $m=0$ ). Come si può notare, tutti questi casi rientrano pienamente, secondo lo schema ridotto, nella categoria dell'*insertion*.

## 5.2. Esempi di alternation

Gli esempi di *alternation* che Muysken propone nel capitolo introduttivo sono i seguenti (cfr. Muysken 2000: 5):



- (19) maar ‘t hoeft niet li- ‘*anna ida šeft ana...*  
 but it need not for when I-see I  
 ‘but it need not to be, for when I see, I...’ (Nortier 1990: 126)
- (20) Les femmes et le vin, *ne ponimayu*  
 ‘Women and wine, I don’t understand’ (Timm 1978: 312)
- (21) Andale pues *and do come again*  
 ‘That’s all right then, and do come again’  
 (Gumperz/Hernández-Chavez 1971: 118)

Nell’esempio (19) lo *switch* avviene in un punto di confine tra frasi; sulla base dei parametri di *constituency*, abbiamo una sequenza di più costituenti in entrambi i codici (neerlandese e arabo marocchino), autonomi rispetto ai costituenti nell’altro codice. Volendo escludere la ripetizione del soggetto alla fine dell’esempio, possiamo considerare 3 come valore minimo di costituenti consecutivi nello stesso codice e 0 il numero di costituenti non autonomi ( $m=3$ ,  $n=0$ ).

In (20), pur volendo considerare tutta la parte in francese come un costituente unico, ci troveremmo comunque di fronte a singoli costituenti autonomi nelle due lingue ( $m=1$ ,  $n=0$ ). Lo stesso si potrebbe dire per l’esempio (21), per il quale tuttavia, anche considerando le costruzioni idiomatiche come blocco unico, dovremmo separare il connettivo *and* (ciò nonostante, il numero minimo di costituenti consecutivi nello stesso codice sarebbe rappresentato dal costituente in spagnolo, per cui  $m=1$ ,  $n=0$ ).

Prendiamo ora in considerazione gli esempi 5-7 discussi sopra (§ 2.3): ignorando per un momento la presenza di *begryip*, i tre esempi presentano almeno due costituenti consecutivi nello stesso codice e autonomi; nell’esempio (7), però, è presente anche un costituente non autonomo e isolato. Per (5) e (6) avremmo valori simili a quelli degli esempi precedenti ( $m=2$ ,  $n=0$ ), rientrando così insieme a questi nell’area di pertinenza dell’*alternation*.

Nell’interpretazione debole del modello, l’esempio (7) scivolerebbe invece nell’area di transizione tra *alternation* e *congruent lexicalization* ( $m=2$ ,  $n=1$ ): viene infatti riportato nel capitolo relativo a quest’ultima categoria, benché come esempio della presenza di *several constituents* e non del *non-constituent mixing*. Tuttavia, l’interpretazione come *congruent lexicalization* è legata alla considerazione di altri fattori: innanzi-

tutto, il fatto che l'elemento neerlandese inserito nel sintagma Sranan sia un verbo; in secondo luogo, l'osservazione della distribuzione tra parti del discorso degli elementi neerlandesi in tutto il corpus (cfr. Muysken 2000: 138-139). Entrambe queste osservazioni fanno però riferimento a parametri del modello originario qui non considerati (in particolare, la sezione *element switched*), e che dovranno quindi essere integrati nel modello completo.

### 5.3. Esempi di congruent lexicalization

Riporto ora alcuni degli esempi utilizzati da Muysken (2000: 6) per rappresentare la *congruent lexicalization*, escludendo quelli in cui la presenza di *bilingual homophones* renda ambigua l'attribuzione dei costituenti ai diversi codici e, di conseguenza, insufficienti i due parametri scalari di *constituency* proposti, che si basano sul loro conteggio.

- (22) (A) Why make Carol *sentarse atrás* (B) *pa'que* everybody has to sit a the back so that move (C) *pa'que se salga* so that [she] may get out (Poplack 1980: 589)
- (23) Bueno, in other words, el flight [que sale de Chicago *around three o'clock*] 'Good, in other words, the flight that leaves from Chicago around three o'clock' (Pfaff 1976: 250)

Riguardo al (22), Muysken osserva come (B) rappresenti un complemento di (A) e (C) sia a sua volta complemento di (B), e che i costituenti nei singoli codici (rispettivamente quelli inglesi e quelli spagnoli) non siano tra loro strutturalmente connessi. Sul piano della *constituency*, possiamo osservare la presenza di almeno due costituenti consecutivi per entrambi i codici, e di almeno due costituenti non autonomi: *sentarse atrás*, retto dal verbo principale in inglese, e la proposizione (B) presa integralmente.

L'esempio (23) ha anch'esso due costituenti non autonomi: *flight* (inserito in un costituente 'misto') e *around three o'clock* (elemento selezionato). È poi presente almeno un costituente autonomo in entrambe le lingue (*bueno* per lo spagnolo e *in other words*, considerando la locu-

zione come blocco unico). Mentre (22) rientrerebbe a pieno titolo nell'area della *congruent lexicalization* ( $m=2, n=2$ ), l'esempio (23) si posizionerebbe nell'area di transizione tra questa e l'*insertion* ( $m=1, n=2$ ). Se infatti escludessimo la locuzione *in other words*, potremmo considerare *flight* e *around three o'clock* semplicemente come casi di *insertion* di elementi inglesi (il primo tra l'altro facilmente candidabile a diventare un prestito) in un enunciato con lo spagnolo come lingua base. Ma se *around three o'clock* può essere considerato un'*insertion*, perché non potrebbe succedere lo stesso con *in other words*?

La questione rientra solo parzialmente nella discussione dei parametri di *constituency* che si intende qui considerare, ma coinvolge il modello nella sua totalità: mentre per *around three o'clock* possiamo parlare di un elemento selezionato (quindi più simile all'*insertion*), *in other words* rappresenta un costituente in posizione marginale, se non extrafrasale (quindi simile a un caso limite di *alternation*); lo stesso si può dire per *bueno*. Dobbiamo quindi considerare i due costituenti iniziali come casi di *alternation*, e i due costituenti non autonomi successivi come casi di *insertion*, o classificare congiuntamente questi fenomeni come un caso di *congruent lexicalization*?

Per casi come (8) e (9) possiamo tranquillamente parlare di *non-constituent mixing* e, di conseguenza, di *congruent lexicalization*. Tuttavia, rimanendo all'interno di un'interpretazione debole del modello, questi rientrerebbero nell'area di transizione tra *alternation* e *congruent lexicalization* ( $m=2, n=1$ ).

Tutti questi casi, ambigui per il modello nella sua versione debole, sono invece risolti dall'interpretazione forte, rientrando pienamente nell'area della *congruent lexicalization*. Occorrerà quindi verificare l'efficacia delle due versioni del modello scalare della *constituency* in relazione alle parti restanti del modello ancora da incorporare.

## 6. *Classificazione scalare dei fenomeni di code mixing*

Si cercherà ora di applicare lo stesso modello di classificazione scalare ad esempi estratti dal corpus di italiano-tedesco a contatto raccolto in Alto Adige, individuando casi posizionabili nelle aree di transizione tra le tre tipologie di fenomeni. Si confronterà poi un tentativo di classi-

ficazione tipologica, in termini categorici di tutti gli esempi presentati con la classificazione scalare secondo i due parametri di *constituency* proposti in questo contributo.

### 6.1. Esempi dal corpus di italiano-tedesco a contatto

Fino a questo punto del testo sono stati inseriti 8 estratti dal corpus: gli esempi (1-4) e (10-13). Gli esempi (1-4) e (10) presentano singoli costituenti non autonomi e isolati (il primo in italiano, gli altri in trentino) all'interno di enunciati in dialetto sudtirolese. Possono quindi essere classificati come casi di *insertion* ( $m=0$ ,  $n=1$ ). Come già discusso, l'esempio (11) presenta un singolo costituente autonomo, più simile a un caso di *intersentential code switching* che di vero *code mixing*<sup>12</sup> (nello schema occuperebbe una casella nell'area dell'*alternation*, con  $m=1$  e  $n=0$ ), così come (12) ( $m=2$ ,  $n=0$ ). Infine, (13) può essere considerato come un enunciato monolingue ( $m=0$ ,  $n=0$ ), dato che *polenta* è ormai pienamente integrato non solo nel tedesco sudtirolese (che tra l'altro avrebbe la variante locale *Plenten*, cfr. Ammon *et al.* 2004: 580), ma in tutta l'area tedescofona.

Analizziamo invece i seguenti esempi:

- (24) non che adesso mi voglio sparare perché *wail i laivesotto red*  
'... perché/perché io parlo laivesotto'
- (25) *prima di venir äh # befor i do es intärwiu gmächt hân hân i mâr*  
*dainä briafäh widâr durchglesn non chiedermi perché*  
'prima di venire - eh - prima di aver fatto l'intervista qui mi sono  
riletto le tue lettere - non chiedermi perché'
- (26) *cioè i bin rimasta di merda*  
'cioè io [ci] sono...'
- (27) *du muasch dänkn dass äh sankt iakob isch spaccata in due*  
'devi pensare che - eh - San Giacomo è spaccata in due'

Il fenomeno di *doubling* negli esempi (24) e (25) fa propendere per

<sup>12</sup> Come osserva lo stesso Muysken, mentre il processo di *insertion* si può considerare più simile all'interferenza, il processo di *alternation* è più vicino al *code switching stricto sensu*, ovvero interfrasale: "I avoid using the term *code-switching* for the general process of mixing. Switching is only an appropriate term for the alternational type of mixing [...] Alternation is just a special case of code-switching, as it takes place between utterances in a turn or between turns" (Muysken 2000: 4-5).

l'interpretazione del caso come *alternation*; di fatto, tutta la frase *wail i laivesotto red* in (24) può essere considerata costituita da elementi del tedesco sudtirolese, dato che *laivesotto* funziona qui come un nome proprio e non è necessariamente da attribuire ad uno dei due codici in contatto<sup>13</sup>. In (25), i costituenti in italiano sono autonomi rispetto alla parte di enunciato in tedesco sudtirolese, mentre non sono presenti costituenti non autonomi. L'analisi della sola *constituency* conferma quindi l'interpretazione come *alternation* per entrambi i casi.

In (26) abbiamo almeno un costituente autonomo e un costituente non autonomo ( $m=1, n=1$ ). Il caso (27) mostra una notevole somiglianza col precedente, che ha in più solo il segnale discorsivo *cioè* posto all'inizio. Si tratta quindi in entrambi i casi di *insertion*, col primo che presenta anche una *alternation* iniziale, oppure (26) è da considerare come un esempio di *congruent lexicalization*? Di nuovo, l'ambiguità può essere risolta o adottando un'interpretazione 'forte', categorica del modello della *constituency*, oppure integrandolo, come sarebbe auspicabile, con altri parametri non ancora considerati del modello originario di Muysken.

## 6.2. Posizione nella matrice e distanze scalari tra tipi di code mixing

Posizioniamo ora tutti i 27 esempi qui proposti all'interno della matrice della *constituency*, calcolando poi le distanze dalle caselle rappresentanti i casi prototipici delle tre classi di *insertion*, *alternation* e *congruent lexicalization*. La tabella seguente (tab. 7) riporta i valori  $m$  e  $n$  degli esempi, insieme alle distanze per le due versioni del modello: nell'interpretazione 'forte', le distanze vengono calcolate dai valori di *constituency* minimi; nell'interpretazione 'debole', le distanze vengono invece calcolate dal punto centrale delle aree di pertinenza per le tre classi di fenomeni.

<sup>13</sup> Si tratta ovviamente di un lemma facilmente riconoscibile come italiano, ma che, in qualità di etnonimo, funge da nome proprio per un referente specifico. Come osserva Broersma (2009), "[a]lthough proper nouns presumably need to be stored and processed as any other lexical item [...], they form a quite specific subset and there is some evidence that they might sometimes be treated differently than other words" (Broersma 2009: 448). I nomi propri rientrano poi nella categoria delle *trigger words*, che tendono a formare, secondo Broersma/de Bot (2006: 4), un *subset* nel lessico mentale del parlante bilingue indipendente dalla lingua (cfr. anche de Bot *et al.* 2009, Broersma *et al.* 2009; sul *lexical triggering*, cfr. Clyne 2003).

Esempi	Modello forte						Modello debole				Muysken
	m	n	I	A	CL	min.	I	A	CL	min.	
13	0	0	1	1	2	–	2	2	4	–	–
1-4, 10, 14-17, 27	0	1	0	2	1	I	1	3	3	I	I
18	0	2	1	3	2	I	0	4	2	I	I
11, 20-21	1	0	2	0	1	A	3	1	3	A	A
5-6, 12, 25	2	0	3	1	2	A	4	0	2	A	A
19, 24	3	0	4	2	3	A	5	1	3	A	A
22	2	2	3	3	2	CL	2	2	0	CL	CL
23*	1	2	2	2	1	CL	1	3	1	I/CL	CL
7-9*	2	1	2	2	1	CL	3	1	1	A/CL	CL
26*	1	1	1	1	0	CL	2	2	2	?	–

Tabella 7: Posizionamento e classificazione degli esempi nella matrice della constituency

Le colonne I, A e CL riportano la distanza rispettivamente per *insertion*, *alternation* e *congruent lexicalization*. La colonna ‘min.’ riporta l’etichetta della classe (o delle classi) con la distanza minima<sup>14</sup>. L’ultima colonna riporta la classificazione proposta da Muysken per gli esempi da lui riportati, che viene qui estesa agli esempi del corpus di italiano-tedesco a contatto con condizioni di *constituency* corrispondenti.

Come si può notare, vi sono solo 5 casi su 27 in cui i due modelli danno risultati diversi (segnati con un asterisco), e in tutti questi il modello debole non esclude la classe risultante dal modello forte. Inoltre, i valori delle distanze per il modello debole mostrano di essere maggiormente polarizzati. Questo potrebbe aiutare a risolvere, all’interno di un modello più ampio, l’eventuale ambiguità presente in altri parametri esterni alla matrice della *constituency*. Il modello forte, d’altronde, mostra di corrispondere pienamente alla classificazione di Muysken, pro-

<sup>14</sup> Come accennato in § 4, le distanze sono calcolate ipotizzando che la matrice della *constituency* rappresenti uno spazio bidimensionale non euclideo, discreto, per il quale vigono le regole della cosiddetta “geometria del taxi”. I valori riportati nella tabella 7 sono quindi distanze Manhattan, pari alla somma delle differenze di grado sulle due scale considerate.

ducendo lo stesso risultato tramite l'utilizzo di due parametri scalari al posto dei parametri binari del modello originario.

Ovviamente, per avere certezza della validità ed efficienza del modello ridotto occorre testarlo non solo su un campione molto più ampio di esempi, ma applicarlo come strumento di analisi su più corpora di parlato bilingue, che coinvolgono codici e contesti di interazione. Anche la scelta tra un'interpretazione forte o debole del modello è ancillare a quest'operazione di verifica, nonché all'integrazione nel modello dei parametri restanti dello schema di Muysken qui ancora non incorporati.

### 7. *Fuori e oltre la matrice: fenomenologia del contatto e repertori linguistici*

Verificata l'efficacia della matrice della *constituency* come sostitutiva dei corrispondenti parametri del modello di Muysken (almeno limitatamente agli esempi con esso proposti e qui riprodotti), occorre ora tornare a chiedersi cosa questa matrice rappresenti, o meglio, come tale matrice sintetizzi la realtà analizzata al punto da costituirne un modello affidabile per lo studio del contatto linguistico.

Benché la finalità prima di questo contributo sia la riflessione sul funzionamento del modello di Muysken (rivolta quindi, per così dire, verso i suoi meccanismi interni piuttosto che verso la sua validità esterna), qualsiasi tentativo di riformulazione della teoria deve tener conto non solo della sua funzione interpretativa, ma anche della sua capacità esplicativa dei fenomeni osservati sulla base di principi più generali; deve, in ultima analisi, avere come obiettivo finale una migliore rappresentazione della realtà studiata e una maggiore capacità predittiva del tipo di fenomeni riscontrabili in relazione a determinate condizioni di partenza (come rapporto tra i codici nel repertorio, tipologia di situazione di contatto etc.).

L'analisi dei fenomeni di *code mixing* riscontrabili in un determinato contesto sociolinguistico non dovrebbe quindi limitarsi alla classificazione tassonomica di tali fenomeni in categorie prototipiche a cui delegare il compito, in qualità di entità astratte sovraordinate, di rappresentare la realtà fenomenica in modo univoco (col rischio di sostituire le categorie astratte del modello alla molteplicità di fenomeni osservati, schiacciando così la variazione interna alla realtà studiata verso le classi

di fenomeni più frequenti). Piuttosto, l'obiettivo dovrebbe essere quello di osservare più da vicino i processi in atto nei fenomeni di contatto presenti nel parlato bilingue e individuare, tramite tale osservazione, quali fattori intervengano in tali processi e in che modo.

In tal senso, la matrice della *constituency* ha il compito di rappresentare lo spazio di variazione negli usi linguistici dei parlanti bilingui (in particolare, quando è attivo il cosiddetto “bilingual mode” e l'uso combinato dei codici è una scelta non marcata sul piano pragmatico-conversazionale; cfr. Grosjean 2010) in relazione alla presenza di costituenti (autonomi o non) nei due codici; tale spazio di variazione, piuttosto che essere diviso in modo categorico per ottenere delle macro-categorie di fenomeni (*insertion, alternation e congruent lexicalization*), potrebbe essere usato per individuare eventuali aree di addensamento o correlazioni tra scale, adattando così l'analisi ai dati osservati. Inoltre, se il modello sintetizza uno specifico sotto-insieme di fenomeni all'interno dei fatti di lingua, dovrebbe essere possibile ipotizzare cosa vi sia al di fuori di tale modello, o più precisamente lungo i suoi confini, e come esso si inserisca in uno schema teorico più generale.

Come mostrato già in § 4, in cui si è proposta una prima versione della matrice della *constituency*, il polo all'origine degli assi rappresenta sostanzialmente una situazione di monolinguisimo, in cui gli unici possibili fenomeni di contatto riscontrabili sono prestiti ormai integrati nel sistema linguistico.

Al polo opposto della diagonale, verso i valori più alti su entrambe le scale, possiamo invece collocare situazioni in cui si è sviluppata una varietà di contatto, oppure situazioni in cui sono coinvolte varietà della stessa lingua: in entrambi i casi, infatti, i parlanti avrebbero a disposizione una “struttura largamente condivisa” nella quale inserire costituenti, autonomi e non, indipendentemente dal codice di partenza, realizzando così la *congruent lexicalization*.

Occorre ricordare che Muysken associa questo processo a “second generation migrant groups, dialect/standard and post-creole continua, and bilingual speakers of closely related languages with roughly equal prestige and no tradition of overt language separation” (Muysken 2000: 9); inoltre, fa corrispondere la *congruent lexicalization* alla variazione stilistica, incorporandola così all'interno dei fenomeni di contatto<sup>15</sup>. I ca-

<sup>15</sup> Muysken afferma infatti: “The phenomenon of style shifting can be seen as one subtype of



si di cosiddetto bilinguismo endogeno, come quello italo-romanzo (cfr. Berruto 1993), si collocherebbero quindi vicino ai *fused lects*, che occupano lo spazio più esterno della matrice<sup>16</sup>.

È da notare che, nel passaggio riportato sopra, Muysken pone come vincolo tra i codici a bassa distanza strutturale anche un “roughly equal prestige”, che possiamo qui intendere come una sovrapposizione funzionale che permetta una maggiore libertà nel passaggio da una varietà all’altra all’interno del discorso bilingue: in tal senso, la *congruent lexicalization* richiederebbe non solo una condivisione strutturale ma anche una condivisione delle situazioni comunicative da parte dei codici in contatto, entrambi appropriati all’interazione bilingue e non marcati pragmaticamente. Ci troviamo, in sostanza, al punto estremo del *continuum* ipotizzato da Auer (1999) per i fenomeni di contatto, che parte da un polo pragmatico, rappresentato dalla commutazione di codice prototipica (“*conversational code switching*”, considerato “interactionally” o “locally meaningful”, cfr. Auer 1999: 310-313), passa per i fenomeni di *code mixing* (o “language mixing”, nei termini dell’autore) e arriva infine al polo grammaticale, rappresentato appunto dai *fused lects*, in cui “the use of one language or the other for certain constituents is obligatory” (Auer 1999: 321).

D’altronde, mentre la *congruent lexicalization* è associata al polo dei *fused lects*, l’*alternation* è più vicina al *code switching* conversazionale, di cui rappresenta un caso limite, realizzandosi all’interno del turno conversazionale invece che tra turni consecutivi (cfr. Muysken 2000: 4). Inoltre, diversamente dalla *congruent lexicalization*, l’*alternation* si realizza tra codici con notevole distanza (strutturale o percepita, cfr. Muysken 2000: 152), e soprattutto sembra associarsi a situazioni di bilinguismo sociale, in cui i parlanti bilingui di fatto mantengono distinte le due lingue (in tal modo rendendo possibile l’uso pragmatico della commutazione di codice).

code-mixing, namely congruent lexicalization” (Muysken 2000: 123); o ancora: “Given that there is a gradient both in perceived distance and in analytically established distance, the distance argument as such cannot be the basis for a principled distinction between mixing and shifting” (*ibid.*: 152).

<sup>16</sup> Muysken inserisce in uno schema triangolare, con ai vertici le tre classi di fenomeni da lui individuate, le varie situazioni di contatto esemplificate nel volume (cfr. Muysken 2000: 245). In questo schema, pone vicino al polo della *congruent lexicalization*, corrispondente al punto più esterno della matrice della *constituency*, sia il contatto tra frisone e neerlandese che quello tra neerlandese e sue varietà dialettali (cfr. Muysken 2000: 245).

Possiamo quindi prevedere che maggiore sarà sia la distanza (strutturale e/o percepita) tra i codici sia la loro equipollenza nella distribuzione diafasica (svincolando così l'uso di uno o l'altro dei codici dalla situazione comunicativa e rafforzandone la salienza pragmatica, "locally meaningful"), più l'interazione tra i codici nel parlato bilingue tenderà verso strategie esclusivamente *alternational* o di commutazione di codice interfrasale. A questo tipo di bilinguismo equilibrato possiamo quindi associare l'area a sinistra della matrice, nella quale i due codici in contatto sono nettamente separati e autonomi, usati alternativamente e raramente o mai mescolati, se non con specifiche funzioni pragmatiche legate al singolo evento comunicativo. Si avvicina a questa tipologia la situazione di contatto tra francese e neerlandese a Bruxelles (cfr. Treffers-Daller 1994<sup>17</sup>), nella quale emerge infatti un *mixing pattern* alternante; ma anche i dati di Nishimura (1989), benché riguardino la seconda generazione di immigrati giapponesi negli Stati Uniti, mostrano una certa predominanza dello schema *alternational*, forse determinata dalla notevole distanza strutturale tra le due lingue.

Non avviene lo stesso, invece, per situazioni di contatto tra lingue parimenti distanti a livello strutturale ma non equipollenti in termini di domini d'uso, ovvero nelle quali vi sia una chiara gerarchia tra i codici, come nel caso del contatto tra swahili e inglese studiato da Myers-Scotton (1993). Queste sarebbero infine da collocare nell'area superiore del modello, che comprenderebbe situazioni in cui la direzionalità dei fenomeni di contatto sia fortemente sbilanciata, se non univoca, con un codice a svolgere la funzione di lingua base nella maggior parte delle interazioni (o, nei termini di Myers-Scotton, di lingua matrice), nella quale si inseriscono elementi isolati, di natura prevalentemente lessicale o nella forma di locuzioni o costituenti singoli, dell'altro codice.

In questo tipo di contesti di bilinguismo 'sbilanciato', come nel caso del contatto tra lingue locali e lingua coloniale (cfr. Myers-Scotton 1993, Bentahila/Davies 1995, Poplack/Meechan 1995), il rapporto gerarchico tra codici determina la direzionalità dello *switch*, mentre la loro netta distanza (strutturale o solo percepita) ostacola l'emersione di forme di *mixing* più estese (mancando i presupposti di equivalenza strutturale richiesti da *alternation* e *congruent lexicalization*).

<sup>17</sup> A tal proposito, si confronti di nuovo Muysken (2000: 245), che pone i dati di Bruxelles di Treffers-Daller (1994) più vicini al polo *alternational* del suo schema.

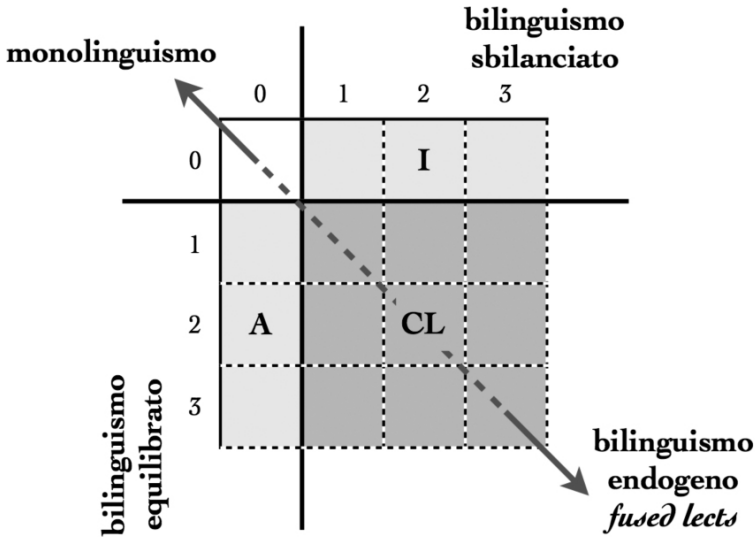


Figura 3: Relazione tra la matrice della constituency e la situazione di contatto

La figura 3 rappresenta tale possibile relazione tra la distribuzione dei fenomeni di *code mixing* nel modello e la situazione di contatto a cui fanno riferimento. L'ipotesi qui avanzata è che l'analisi e la classificazione sistematica dei singoli fenomeni di contatto presenti in un corpus sufficientemente esteso di parlato bilingue permetta di evidenziare delle relazioni coerenti, gradualmente più vicine al modello all'aumentare del numero di osservazioni, con il tipo di situazione sociolinguistica in cui tali fenomeni si realizzano.

Tramite la rappresentazione dimensionale, ovvero la matrice della *constituency*, della variazione all'interno dei *mixing patterns* in una determinata comunità linguistica, è possibile tener conto non solo dello schema dominante (limiterebbe la classificazione a comunità bilingui *insertional*, alternanti o che usano la *congruent lexicalization*), bensì dell'insieme di fenomeni osservati, e non in base alla loro categorizzazione finale ma mantenendone la variabilità interna.

Si prendano ad esempio le conclusioni di Deuchar/Muysken/Wang (2007) sull'analisi di tre diversi corpora: uno gallese/inglese, uno mandarino/tsou e uno mandarino/taiwanese. Per ognuno di essi sono stati

classificati 100 casi di *code mixing*, calcolando la corrispondenza con i tre schemi prototipici sulla base della media dei valori attribuiti ad ogni singolo caso su tutti i parametri del modello di Muysken<sup>18</sup>. La tabella 8 riporta tali valori, indicando poi lo schema dominante (in questo caso, valori maggiori indicano maggiore corrispondenza).

Schemi di code mixing	gallese / inglese	tsou / mandarino	taiwanese / mandarino
insertion	7,1	5,43	4,94
alternation	-3,85	-5,05	-4,49
congruent lexicalization	4,34	5,17	6,38
<b>Schema dominante</b>	<b>insertion</b>	<b>insertion</b>	<b>congr. lex.</b>

Tabella 8: Punteggi per i diversi schemi di code mixing calcolati su 100 casi (dati ripresi da Deuchar/Muysken/Wang 2007)

Come si può notare, nonostante il responso finale per il corpus gallese/inglese e per il corpus mandarino/tsou sia lo stesso (con *insertion* come schema dominante), i valori assunti dai dati sono sensibilmente diversi, con una dominanza più netta dell'*insertion* per il primo e un rapporto di maggiore equilibrio tra lo schema *insertional* e la *congruent lexicalization* per il secondo. Non a caso, nel momento in cui gli autori, per escludere possibili errori di valutazione determinati da eventuali prestiti inclusi nel calcolo, restringono il campo ai soli *multi-item switches*, la distribuzione tra i valori per il corpus mandarino/tsous cambia, facendo risultare la *congruent lexicalization* come schema di *code mixing* dominante; questo secondo corpus assumerebbe quindi una posizione intermedia tra gli altri due (quello gallese/inglese più chiaramente *insertional*, quello taiwanese/mandarino decisamente legato alla *congruent lexicalization*). Come osservano gli studiosi,

In conclusion, the results of analysing our Welsh-English, Tsou-Mandarin and Taiwanese-Mandarin data suggest that no data set can be exclusively

<sup>18</sup> In realtà, occorre qui segnalare che tre dei 27 parametri del modello non sono applicabili a singoli casi ma solo all'intero corpus di dati o alle situazioni di contatto nel loro complesso: *diverse switches* fa riferimento alla variabilità di tipi di *switch* (minima se questi si concentrano su una o due classi di parole, massima se vi è una distribuzione su tutte le classi da parte di entrambi i codici); *bidirectional switching* fa riferimento alla presenza di casi di *switch* in entrambe le direzioni (dal codice A al codice B e viceversa); *homophonous diamorphs* si riferisce invece alla percentuale di elementi omofoni nelle due lingue all'interno del corpus.

categorised by one of Muysken's (2000) three CS patterns. There is always a more or less strong secondary pattern. However, there are clear preferences, which may be related to grammatical or extralinguistic factors (Deuchar/Muysken/Wang 2007: 335-336).

Un'osservazione non in termini categorici di tali *mixing patterns*, bensì scalare, a partire dalla distribuzione dei fenomeni in una matrice di dimensioni di variazione direttamente analizzabile, rispecchierebbe forse in modo più fedele la variabilità interna dei processi di *code mixing* e la loro correlazione con i fattori strutturali ed extralinguistici che ne possono condizionare la realizzazione all'interno delle diverse situazioni di contatto.

Simone Ciccolone  
c/o Centro Competenza Lingue  
Libera Università di Bolzano  
Piazza Università 1  
39100 Bolzano  
E-mail: s.ciccolone@noam.it

## Bibliografia

- Ammon, Ulrich *et alii*, 2004, *Variantenwörterbuch des Deutschen. Die Standardsprache in Österreich, der Schweiz und Deutschland sowie in Liechtenstein, Luxemburg, Ostbelgien und Südtirol*, Berlin/New York, de Gruyter.
- Auer, Peter, 1999, "From codeswitching via language mixing to fused lects: Toward a dynamic typology of bilingual speech". *International Journal of Bilingualism* 3/4: 309-332.
- Bentahila, Abdelâli / Davies, Eirlys E., 1995, "Patterns of code-switching and patterns of language contact", *Lingua* 96: 75-93.
- Berruto, Gaetano, 1993, "Le varietà del repertorio". In: Sobrero, Alberto A. (ed.), *Introduzione all'italiano contemporaneo. I. La variazione e gli usi*. Roma/Bari, Laterza: 3-36.
- Berruto, Gaetano, 2000, "La sociolinguistique européenne, le substandard et le code switching". *Sociolinguistica* 14: 66-73.

- Berruto, Gaetano, 2001, "Struttura dell'enunciazione mistilingue e contatti linguistici nell'Italia di Nord-Ovest (e altrove). In: Wunderli, Peter / Werlen, Iwar / Grünert, Matthias (Hrsg.), *Italica-Raetica-Gallica. Studia linguarum litterarum artiumque in honorem Ricarda Liver*, Tübingen, Francke: 263-283.
- Berruto, Gaetano, 2004, "Su restrizioni grammaticali nel codemixing e situazioni sociolinguistiche. Annotazioni in margine al modello MLF". *Sociolinguistica* 18: 54-72.
- Berruto, Gaetano, 2011, "Interazione e compenetrabilità di grammatiche nel contatto linguistico. Il caso dell'enunciazione mistilingue". In: Bombi, Raffaella / D'Agostino, Mari / Dal Negro, Silvia / Franceschini, Rita (a cura di), *Lingue e culture in contatto. In ricordo di Roberto Gusmani*. Atti del 10° Congresso della Associazione Italiana di Linguistica Applicata. Perugia, Guerra Edizioni: 47-69.
- Broersma, Mirjam, 2009, "Triggered codeswitching between cognate languages". *Bilingualism: Language and Cognition* 12 (4): 447-462.
- Broersma, Mirjam / de Bot, Kees, 2006, "Triggered codeswitching: A corpus-based evaluation of the original triggering hypothesis and a new alternative". *Bilingualism: Language and Cognition* 9 (1): 1-13.
- Broersma, Mirjam / Isurin, Ludmila / Bultena, Sybrine / de Bot, Kees, 2009, "Triggered code switching: Evidence from Dutch - English and Russian - English bilinguals". In: Isurin, Ludmila / Winford, Donald / de Bot, Kees (eds.), *Multidisciplinary Approaches to Code Switching*. Amsterdam/Philadelphia, John Benjamins: 103-128.
- Cerruti, Massimo / Regis, Riccardo, 2005, "'Code switching' e teoria linguistica: la situazione italo-romanza". *Rivista di Linguistica* 17/1: 179-208.
- Clyne, Michael, 1987, "Constraints on code switching: How universal are they?". *Linguistics* 25: 739-764.
- Clyne, Michael, 2003, *Dynamics of language contact*, Cambridge, Cambridge University Press.
- Dal Negro, Silvia / Guerini, Federica, 2007, *Contatto: Dinamiche ed esiti del plurilinguismo*, Roma, Arcane.
- de Bot, Kees / Broersma, Mirjam / Isurin, Ludmila, 2009, "Sources of triggering in code switching". In: Isurin, Ludmila / Winford, Donald / de Bot, Kees (eds.), *Multidisciplinary Approaches to Code Switching*, Amsterdam/Philadelphia, John Benjamins: 85-102.
- Deuchar, Margaret / Muysken, Pieter / 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.
- Gardner-Chloros, Penelope, 2009, *Code-switching*, Cambridge, Cambridge University Press.
- Grosjean, François, 2010, *Bilingual. Life and Reality*, Cambridge, Harvard University Press.
- Han, Jiawei / Kamber, Micheline, 2006, *Data Mining: Concepts and Techniques*, San Francisco, Morgan Kaufman.
- MacSwan, Jeff, 1999, *A Minimalist approach to intrasentential code switching*, New York, Garland Press.
- MacSwan, Jeff, 2005, “Codeswitching and generative grammar: A critique of the MLF model and some remarks on ‘modified minimalism’”. *Bilingualism: Language and Cognition* 8 (1): 1-22.
- Matras, Yaron, 2009, *Language Contact*, Cambridge, Cambridge University Press.
- Muysken, Pieter, 2000, *Bilingual Speech: a typology of code-mixing*, Cambridge, Cambridge University Press.
- Myers-Scotton, Carol, 1993, *Duelling languages: Grammatical structure in code switching*. Oxford, Clarendon Press.
- Myers-Scotton, Carol, 2002, *Contact Linguistics. Bilingual Encounters and Grammatical Outcomes*. Oxford, Oxford University Press.
- Myers-Scotton, Carol / Jake, Janice L., 2000, “Four types of morpheme: evidence from aphasia, code switching, and second-language acquisition”. *Linguistics* 38/6: 1053-1100.
- Nishimura, Miwa, 1989, “The topic-comment construction in Japanese-English code-switching”. *World Englishes* 8/3: 365-377.
- Poplack, Shana, 1980, “Sometimes I’ll start a sentence in Spanish *y termino en Español*”. *Linguistics* 18: 581-618.
- Poplack, Shana, 1987, “Contrasting Patterns of Code-Switching in Two Communities”. In: Wande, E. *et al.* (eds.), *Aspects of Multilingualism*. Proceedings from the Fourth Nordic Symposium on Bilingualism, Uppsala, Borgströms: 51-77.
- Poplack, Shana / Dion, Nathalie, 2012, “Myths and facts about loanword development”. *Language Variation and Change* 24/3: 279-315.
- Poplack, Shana / Meechan, Marjory, 1995, “Patterns of language mixture: Nominal structure in Wolof-French and Fongbe-French bilingual discourse”. In: Milroy, Leslie / Muysken, Pieter (eds.), *One speaker, two languages*. Cambridge, Cambridge University Press: 199-232.

- Regis, Riccardo, 2003, “Enunciazione mistilingue e prestito: una storia infinita?”. *Plurilinguismo* 10: 127-164.
- Regis, Riccardo, 2005, *Appunti grammaticali sull'enunciazione mistilingue*, Lincom, München.
- Sankoff, David / Poplack, Shana / Vanniarajan, Swathi, 1990, “The case of the nonce loan in Tamil”. *Language Variation and Change* 2: 71-101.
- Treffers-Daller, Jeanine, 1994, *French-Dutch Contact in a Comparative Perspective*. Berlin/New York, de Gruyter.