

22

Structural Outcomes of Language Contact

Yaron Matras

1. Introduction: Contact vs. Repertoire Management

Since the works of Haugen (1950, 1969 [1953]) and Weinreich (1953), language contact has been depicted, much as the metaphor CONTACT suggests, as the interplay of two (or more) separate linguistic systems, which “interfere” with one another during speech production; this leads to changes in their respective structures, as components of one system are incorporated into another. Both authors identified the locus of contact as being the bilingual individual. In this way, CONTACT as a metaphor equally captures the difficulties that bilinguals encounter when trying to keep components of their repertoire of linguistic structures apart during speech production.

In the half-century that followed these pioneering works in contact linguistics, the idea of SYSTEMS INTERACTING WITH ONE ANOTHER can be said to have dominated research. Considerable attention had been devoted to identifying structural constraints on language mixing in conversation (Poplack 1980, Berk-Seligson 1986, Clyne 1967, Myers-Scotton 1993b, Muysken 2000), driven by the assumption that there is a “grammar of code-switching” and that the structural outcomes of spontaneous language mixing are to some extent predictable. In typology, Moravcsik (1978) was first to apply the concept of IMPLICATIONAL HIERARCHIES to structural borrowing, describing the outcomes of contact in terms of a gradational ordering of categories and scalar presence of category features (such as morpheme independence, or semantic autonomy and referential transparency; see also Campbell 1993, Field 2002). In a much-cited work on the outcomes of contact, Thomason & Kaufman (1988) put forward the notion that the borrowing of structural categories is gradational and linked to the duration and intensity of social and cultural contacts. Category status and social processes such as identity flagging play a key role also in Bakker’s (1997, 2003) ideas on the predictability of structural INTERTWINING in mixed languages, while Heine & Kuteva (2005) frame contact-induced

change in grammaticalization theory and discuss how cross-system analogies trigger category expansion and re-distribution.

In this contribution I draw partly on my earlier work on language contact (Matras 1998, 2009), which seeks to re-position the study of contact-induced language change in the context of the individual user's management of a complex repertoire of linguistic structures. This approach links structural outcomes of contact with the inherent functions that structural categories have in triggering information processing in communication. I take issue with the notion of languages as demarcated systems, and regard contact instead as a constant negotiation of the selection and deployment of repertoire components. My approach aligns itself with a growing tendency to critique the integrity of systems when it comes to multilingual users and their repertoires. This is reflected in work on bilingual conversation that identifies MIXING as the default rather than the exception in bilingual settings (Meeuwis & Blommaert 1998, Grosjean 1989). It is also found in work in neurolinguistics that questions the separation of linguistic systems during speech production and assumes that, rather than de-activate systems on a wholesale basis, speakers rely on a mental "executive control" mechanism in order to select and de-select individual structural items (Green 1998, Bialystok, Craik, & Luk 2008). And it agrees with more recent work in post-structural sociolinguistics that adopts a holistic view of the linguistic repertoires that users have at their disposal and attributes choices and combinations of structures to localized practice and routines (Blommaert & Backus 2013, Pennycook & Otsuji 2015, Li 2018).

The approach that I adopt here assumes that speakers have access to an integral repertoire of linguistic resources on which they draw in order to communicate, and that boundaries among "languages" are therefore essentially permeable and subject to users' creativity. While not every idiosyncratic innovation in discourse will lead to language change, every structural change is initiated in the form of a local innovation. For that reason, an understanding of the motivations behind diachronic change must rest on an understanding of individuals' usage in actual linguistic interaction. My assumption is that innovations (hence also change) do not affect elements of language in a random way but are linked to the function of structures and structural categories, viz., to their role in instigating mental processing tasks in communication.

Sociolinguistic conditions enable the structural outcomes of contact inasmuch as it is societal conditions that give rise to and sustain multilingualism in the first place, and it is societal norms that allow or constrain the propagation of innovations. But sociolinguistic conditions cannot, on their own, explain the motivation to innovate around certain structures rather than others; duration and intensity of social and cultural contacts on their own fail to explain why some categories are more contact-prone than others. The key to an explanatory model is, I maintain, users' motivation to alter their communicative habits, affecting the way in which interaction

routines trigger the selection of certain structures within the repertoire rather than others. In order to understand the structural outcomes of language contact, we therefore require an integrated model that links users' communicative motivations with communicative tasks and the structural categories that represent them.

2. Code-Switching (Codemixing)

The terms *code-switching* and *codemixing* are often used interchangeably to denote the alternation of languages in discourse. They capture the structural effects of language contact and bilingualism/multilingualism at the level of discourse production rather than the language "system." Pioneer studies showed how language choice in multilingual settings is sensitive to changes in the interaction setting, including the participant constellation, place, and thematic context of the discourse, reflecting the diglossic distribution of languages in the repertoire of individuals and communities.

That approach led Myers-Scotton (1993b) to apply the notion of MARKEDNESS to the selection of languages in multilingual communities, suggesting that speakers make "rational choices" (Myers-Scotton & Bolonyai 2001). Such choices are tailored to participants' expectations and linked to the social values that are associated with languages in a given setting. An "unmarked" choice is one that complies with expectations, while a "marked" choice contradicts expectations and thereby draws attention and emphasis to the utterance or segment, signalling social distance between participants or toward portions of the content of the conversation.

A parallel research trajectory focuses on conversation-internal aspects of code selection. Here a distinction is made between "insertional" switches, which are generally single word or single phrase units, and "alternational" switching, which covers an entire utterance and often the remaining conversation flow (Muysken 2000). Language alternation can be triggered by a variety of factors, including changes to the conversational setting. Single word insertions such as names of institutions, procedures, or even persons, which are not translatable, or words that are shared by both languages, often act as discourse-internal triggers of alternational code-switching (Clyne 1967).

Code-switching itself can constitute a stylistic device that helps structure discourse. Gumperz (1982) argues that code-switching acts as a cue that helps speakers to contextualize information by juxtaposing languages in a way that is meaningful and metaphorical of the contrast between the content of single portions of the conversation. Auer (1984, 1995) lists a range of typical functions of such meaningful contrast of languages that includes reported speech, side comments, emphasis and reiteration, change of mode, topicalization as well as language play. Applying the tools of Conversation Analysis, code-switching can be investigated as a device that

is used to organize the sequentiality of discourse, for instance as a way of marking participants' level of cooperation, expression of disagreement, refusal of an offer, or initiation of a repair (Li 2005).

From the outset, research on code-switching also addressed the question of structural patterns, searching for general constraints on language mixing in conversation, or what some referred to as a "grammar of code-switching" (Sankoff & Poplack 1981, Sankoff 1998, Muysken 2000). It has been observed that speakers tend to maintain the grammaticality of structures in both or all of the languages used in "mixed" utterances, that they avoid violating the word order rules of either of the participating languages, and that switches around inflectional morphemes are also avoided (see, e.g., Pfaff 1979). The latter two generalizations have been identified as the "equivalence constraint" and "free morpheme constraint" respectively by Poplack (1980).

Myers-Scotton's (1993a, 2002) frequently cited model of code-switching assumes a hierarchical relationship between the languages, identifying one as the Matrix and the other as the Embedded language. It identifies constraints that are derived from that relationship, anticipating that grammatical morphemes and word order rules will be provided only by the Matrix language. A later development of the model (Myers-Scotton & Jake 2000) hypothesizes a hierarchy within System morphemes, making predictions about the likelihood of types of morphemes that are subjected to language alternation. Determiners such as definite articles, possessive expressions, and plural marking are referred to as "early system morphemes." They are considered to be closest to content morphemes in that they are conceptually activated, i.e., directly elected by their head content morpheme as part of its representation in the mental lexicon. These morphemes are more likely to be the subject of language mixing or cross-language adoption.

So-called "late system morphemes" are, by contrast, not conceptually activated. They include morphemes that connect content morphemes with each other without reference to the properties of those morphemes, for example English *of* or French *de*, termed "bridge morphemes," and those inflectional morphemes that depend on information outside the head, such as subject-agreement morphemes and case affixes, termed "outsider morphemes" (Myers-Scotton 2002: 75). These latter morphemes are least likely to be points of code-switching and so are also least likely to be adopted from one language into another.

The hierarchical arrangement of structural categories in relation to their likelihood of being the point of code-switching in a sentence has attracted wider attention. Case studies that deal with different language pairs vary in their observations (e.g., Pfaff 1979, Berk-Seligson 1986, Nortier 1990, Gardner-Chloros 1991, Backus 1996), but some generalizations can be made across studies. Not surprisingly, nouns appear at the top of the list as the most frequent point of code-switching, followed by verbs, adjectives, and adverbs, while pronouns and adpositions are low on the hierarchy.

Conjunctions and discourse markers are often listed in a lower position on the scale of frequent points of switching.

What the statistics fail to uncover, however, is the particular functionality of switches around discourse markers and related elements, including conjunctions, tags, fillers, and interjections. Poplack (1980) attributed the frequency of inserted English discourse markers in the Spanish of some Puerto Rican speakers in New York to speakers' lower degree of fluency in English, and classed it as "emblematic," aimed at flagging bilingual competence by relying on items that are more easily integrated into Spanish utterances and require lower proficiency in English. Maschler (1994), on the other hand, taking a Conversation Analysis approach, interprets switched discourse markers as meta-linguaging strategies that are used to mark out discourse boundaries thereby contributing to the sequential organization of the discourse.

Example (1) from a conversation among Kuwaiti girls who attend an English-medium secondary school in Kuwait (Mahsain 2014) shows such meta-linguaging strategies in operation. The language of choice within the peer group is English, including in casual conversation, but they all share Arabic as a language of the family and the principal language of the surrounding society, including institutions other than school:

(1) Kuwaiti schoolgirl (Mahsain 2014: 69–97):

It's better, *bas yaʕni* [=but I-mean] there are disadvantages, *yaʕni* [=I-mean],
ḥarām, [=shame] they're too young, and it's true, there will be more accidents, like there's no focus,
and *yaʕni* [=I-mean] the boys [H laughs] they are sixteen and what they do,
and there are still, like, younger kids that do drive and their parents don't know *yaʕni*, [=I-mean]
what if something happens in the road?

According to Mahsain, the choice of Arabic discourse markers at the boundaries of utterances grounds the narrative in the participants' shared environment of cultural values, and so it serves to reiterate a shared cultural perspective. In this way, the Arabic discourse markers convey a meaning that English translation equivalents could not convey. In the case of *ḥarām*, which means 'shame' but also 'religious prohibition', the choice of Arabic captures explicitly a culture-specific value system.

In selecting resources from the repertoire of linguistic structures, language users are guided by a "selection and inhibition" mechanism (Matras 2009, chapter 5). Researchers in neurolinguistics have referred to this as the "executive control" (Green 1998, Bialystok, Craik, & Luk 2008). We assume that engaging the selection and inhibition mechanism places a burden on the user. At times, there are lapses in the inhibition of forms that are

functionally purposeful (in terms of their intrinsic meaning) but not contextually permissible (as they are not part of the inventory of structures that might be expected or even understood by the interlocutor in the particular interaction context; i.e., they are in the “wrong language”). Such lapses are frequent around discourse markers and related expressions (utterance modifiers such as tags, fillers, focus particles, and phrasal adverbs; Matras 1998):

(2) Polish-German bilingual in England (Matras 2009: 97):

... *bis auf/ bis auf die Tischdecken*, because/ *eh weil sie* ...

‘...except/ except for the tablecloth, because/ uh because it ...’

In example (2), the speaker is a Polish native speaker residing in Germany. She is speaking German to two friends with whom she is meeting in London, during her stay there on a three-week language course. The selection of English *because* during a portion of German conversation targets the language toward which the speaker has been directing her uppermost intellectual attention during the past weeks. We are dealing with an argumentative connector, one that is inserted in order to intervene with and influence the hearer’s course of processing propositions and deriving conclusions from them. At the same time the connector operates at the interactional level, announcing the speaker’s justification of a preceding statement; thus, *because* captures the speaker in a position of potential vulnerability on the interaction plain. The failure of the inhibition mechanism is not intentional, but a form of “interference,” as proven by the speaker’s own immediate action to self-repair. It is noteworthy that English in this example is not the speaker’s native language, or even her stronger language, but rather the language to which her attention is directed during this particular short phase in her life – or the “pragmatically dominant” language (Matras 1998).

Examples (1)–(2) show that there are distinct motivations for speakers to “mix” components of their repertoire of linguistic resources. Such ad hoc, idiosyncratic innovations in certain areas of structure may become frequent and, subject to the social norms of the community, may become conventionalized, leading to permanent structural outcomes.

3. Lexical Borrowing

The term *borrowing* has been widely used in linguistic literature since the works of Haugen (1950) and Weinreich (1953) to refer to the adoption of a form or structure from one language within the framework or system of another. The language that is the source of the form or structure is usually referred to as the “donor,” while the language that has adopted it is referred to as the “recipient.” Two frequently cited motivations for borrowing are

“gaps” in the recipient system and the “prestige” of the donor system. Gaps may explain the borrowing of terms for new artefacts, products, and cultural innovations (such as English *banana*, *sushi*, or *parliament*), including social and ideological concepts (such as English *majesty* or *redemption*). Comparative data on loanwords in a sample of languages (Haspelmath & Tadmor 2009)¹ show some clear hierarchical trends, with semantic categories such as modern world artefacts leading in proneness to borrowing, followed by religion and belief, clothing and grooming, household items, and the law.

Prestige, by contrast, is a somewhat vague notion. It tends to refer to the dominant status of a language within a particular domain of social interaction. Domain-specific borrowings, such as the frequently cited English words *beef*, *mutton*, *poultry*, and *pork*, which complement inherited terms like *ox*, *sheep*, *chicken*, and *pig*, might be explained as reflecting the role of French as the preferred language of the medieval English elite and a source of imitation in connection with culinary customs. Similarly, the English slang expressions *pal* ‘friend’, *kushty* ‘good’, *chav* ‘guy’, or *minge* ‘vagina’, all derived from Romani, reflect the association of Romani in eighteenth- and nineteenth-century England with defiance of the authority of the establishment. The motivation behind such borrowings is thus connected with the specialized roles of languages in different social interaction domains. Similarly, Khuzistani Arabic (as spoken in Iran) has *bās* ‘bus’, a pan-Arabic loan from English, in the singular, but *otobus-hā* ‘buses’ from Persian (which in turn borrows the singular form from French), in the plural. This reflects the contemporary dominance of Persian in the domain of formal-institutional communication, where plural reference to buses is more likely to occur than in casual speech. Hebrew, by contrast, has the European loanword *oto* ‘car’ in the singular, but the neologism *mexoniyot* ‘cars’ in the plural, a reflection of the impact of language-engineering and revivalist ideologies on formal-institutional communication domains.

Neither “gaps” nor “prestige,” however, can easily explain borrowings that replace inherited items, or the apparent hierarchical nature of borrowing in some cases. For example, borrowed terms are more likely to target more remote kin – as in English *uncle*, *aunt*, *grandparents*, *niece*, *nephew* (from French), or Maltese *nannu* ‘grandfather’, *ziju* ‘uncle’, *kugin* ‘cousin’, *neputi* ‘nephew’ (from Italian/Sicilian) – than proximate kin like words for parents, siblings, and direct offspring. Languages that borrow numerals, such as Swahili and Palestinian Domari (both from Arabic) and Romani (from Greek), are more likely to borrow numerals above than under ‘five’. There are thus, in addition to the effect of diglossia, also cognitive factors in operation, whereby everyday concepts that are simple, close, intimate, and more frequently used tend to resist borrowing, whereas their paradigm

¹ See also World Loanword Database (<http://wold.cld.org/>).

counterparts that indicate greater complexity, remoteness, formality, and tightly regulated routines are more borrowing-prone. From this, one might draw the conclusion that borrowing can serve, in its incipient stage, as a meaningful way of marking asymmetry between paradigm values (Elšík & Matras 2006: 385).

There is thus interplay between cognition, in the sense of accessibility and routine of information, and the social values and functions that are associated with a contact or donor language. The proneness to borrowing of higher numerals (as well as those associated strictly with formal mathematical routines rather than with everyday counting, namely 'zero' and fractions) reflects their association with non-casual, institutional procedures and transactions, and therefore with the language that dominates such routines. Borrowed kin terms are more likely to relate to those who are more remote, less frequently mentioned, and more likely to be named in the context of formal titles and genealogical inventories.

The very foundations of the comparative method in linguistics rely on recognizing cognates between languages and therefore on the notion of the diachronic stability of at least some parts of the lexicon. The realization that some meanings prove more reliable for this exercise than others has brought forward the idea of a "basic lexicon." Yet the precise definition of "basic lexicon" has always been somewhat impressionistic. The Swadesh (1952) list remains to this day a popular tool and standard measure of lexical stability and so of language genetic relatedness among languages. As work with the list became more ambitious and scholars ventured into the field of "lexico-statistics," claiming to be able to reconstruct the time depth of separation between related languages on the basis of their shared lexicon, the issue of lexical borrowing took on a center-stage position (cf. Embleton 1986, Renfrew, Trask & McMahon 2000).

The collection of case studies presented by Haspelmath & Tadmor (2009) has succeeded in superseding the Swadesh 100-item list while at the same time partly confirming its validity. The material for the volume is based on a project questionnaire: the Loanword Typology Meaning List, with 1460 entries. For entire sets of meanings we find low borrowing rates across the entire corpus of some 40 different languages. These include body parts, universally present natural phenomena, generic actions, basic properties, personal pronouns, and basic interrogatives. Overall, the "Leipzig-Jakarta" list (as it is called by Haspelmath & Tadmor) of 100 basic vocabulary items includes body parts, in particular external organs like 'mouth', 'ear', 'eye', 'arm', 'nose'; universally present natural phenomena like 'water', 'fire', 'rain', 'night', 'star', 'wind', 'rock/stone'; and generic (i.e., geographically widespread) animal terms like 'fish' and 'bird', as well as terms for creatures found wherever there are humans, such as 'louse', 'ant', 'fly', and 'dog'.

Generic actions on the list include motion verbs like 'to go' and 'to come', as well as basic activities such as 'to eat', 'to drink', and 'to laugh', and

perception verbs like ‘to see’ and ‘to hear’. The list also includes basic properties like ‘big’, ‘small’, ‘old’, and ‘new’, as well as the color terms ‘black’ and ‘red’, the singular pronouns ‘I’, ‘you’, and ‘he/she/it’, and the interrogatives ‘what?’, ‘where?’, and ‘which?’. The new list ends up bearing a close resemblance to the Swadesh list, with 62 items of overlap.

The contributions to the volume by Haspelmath & Tadmor (2009) also allow impressions about different rates of lexical borrowing by language, determined on the basis of the percentage of loanwords in the 1460 item list. Two languages are identified as showing “very high borrowing,” amounting to over 50 percent of the list: Selice Romani (a Romani dialect spoken in southern Slovakia) with 62.7 percent and Tarifyt Berber with 51.7 percent. Around half the sample languages are “high borrowers,” with over 25 percent loanwords on the list. They include English, Romanian, Indonesian, Japanese, Swahili, Thai, and Gurindji. The others are “average borrowers,” with anywhere from 10–25 percent loans; examples include Hausa, Dutch, Malagasy, Hawaiian, and Hup. Only four languages are “low borrowers,” with less than 10 percent loans: Old High German, Manange, Ket, and Mandarin Chinese, which has the lowest borrowing rate in the sample – 1.2 percent. Although little is said in the comparative summaries about the sociolinguistic circumstances of the sample languages, selecting the two extreme cases, notably Selice Romani and Mandarin Chinese, one might conclude that high borrowing correlates with the following factors: universal multilingualism, a minority language status, the absence of a written standard, and socio-political marginalization. On the other hand, low borrowers show little or no bilingualism, have a majority-language status and a powerful standard, and are associated with a socio-politically dominant population.

Loanwords are a product of bilingual speakers’ way of adjusting their overall repertoire of lexical words and the constraints on the selective use of words in certain settings or with certain interlocutors. Insights into the hierarchical nature of lexical borrowing provide us with an excellent opportunity to explore how the process of re-negotiating the bilingual lexical repertoire is related to the conceptualization of reality. It seems that shared lexical repertoire (i.e., lexical items that are generalized throughout the bilingual’s repertoire and used irrespective of interaction setting or interlocutor, i.e. “borrowed”) is symbolic of activities that are shared with a neighboring linguistic community: commerce, religion, administration, and technology. By contrast, personal and family experiences (body, emotions, space) remain conceptually protected and individualized, and this is reflected in the enduring compartmentalization of the relevant linguistic expressions in the bilingual repertoire.

There are other types of lexical borrowing that show very different underlying motivations. They include the replacement of taboo vocabulary, as in the case of Haruai, a Papuan (i.e. non-Austronesian) language of Papua New Guinea (Comrie 2000), where there is a taboo against saying the name

of relatives of a certain degree, such as in-laws or cousins. Since most personal names are also ordinary words, the taboo affects the articulation of everyday vocabulary. Other forms of lexical borrowing are the persistence, in certain styles of community-internal speech, of lexical vocabulary from an ancestral language, used to mark out in-group identity. Such is the case in what researchers have called “Angloromani” (Matras 2010), a style of speech that users simply term *Romani*. It is a particular style of English speech that is common among Romani Gypsies in England and Wales in some interaction contexts, as in the following extract from a family interaction describing a sensitive state of affairs:

(3) Angloromani (Matras 2010: 146):

And me Aunt Alice was crying, she said: “Oh it’s mored [= killed] me I can’t lift me sherra [=head] up. Everybody’s rokkerin [=talking]” she said “among our fowki [=people].”

“Why?” she said.

“Our Debbie” she said, “She’s bori [=pregnant].”

‘bori’ means with child.

Yeah, mhm.

She said. “Oh” she said/ she said: “And she can’t *pukker* [=tell] me”, she said “the *chor* [=poor]”, she said, “who’s the father of the *tiknas* [baby]”. Coz she’d been with that many, she’d had that many different *mushes* [=men].

And me mam used to say: “you know if you take any *izers* [=trousers] off Debbie, *yog* [=burn]’em”.

Angloromani clearly does not involve a consistent lexicon–grammar split, as its structure has sometimes been portrayed (for example by Thomason & Kaufman 1988: 103–4). Nor can it be regarded as a form of code-switching: Users have a choice to integrate Romani-derived lexicon into English utterances, but not the other way around. While users sometimes refer to their distinct style of speech as a “language,” they also describe it as “broken” or “not proper.” From an analytical perspective it is hard to view this as anything but a particular conversation-level mode of speaking English, albeit not a form of English that is easily accessible to those who are not group members.

The dilemma as to whether Angloromani is an independent “language,” a “hybrid” or “mixed” language, a form of English that is not accessible to other speakers of English, or indeed a form of Romani that is not accessible to speakers of Romani who are not fluent in English highlights precisely the limits of conceptualizing repertoires as closed language systems. English Gypsies license themselves to make use of the full expressive potential of their repertoire of resources under certain circumstances, usually defined by the choice of addressee and bystander, and potentially by the choice of topic. They do so while maintaining just a certain degree of compartmentalization

within the repertoire, one that pertains to lexicon (and to some minor stylistic conventions) but not to other structures. In other words, they do so without maintaining a strict separation of languages. The outcome of the underlying historical process is difficult to accommodate within our traditional notion of LANGUAGE, even within our fixed notions of MIXED LANGUAGES (Bakker & Matras 2003).

4. Functional and Grammatical Borrowing

Contact-induced language change is a reorganization of the balance of factors that condition the way in which users manage elements of their repertoire of linguistic resources. The following example from Palestinian Domari shows just how far-reaching the process can be. Domari, the Indo-Aryan language of the peripatetic Dom of the Middle East, has absorbed strong influences from the surrounding languages, in particular Arabic. The Jerusalem variety, now moribund, has adopted several structural categories on a wholesale basis from the neighboring majority language Palestinian Arabic. In a setting where use of Domari has been limited to oral domestic interaction, all Domari speakers are fully bilingual, and even the family setting has become largely Arabic-speaking, as Domari is now confined to the elderly generation. As a result, individual utterances may appear identical in all but a few features to their Arabic counterparts:

(4) Excerpts from Jerusalem Domari conversation (Matras 2012: 383–4):

- a. *aktar min talātīn xamsa ū talātīn sana*
 more from thirty five and thirty year
ma lak-ed-om-is
 NEG see-PAST-1SG-3SG.OBL
 ‘It has been more than thirty, thirty five years since I’ve seen her.’
- b. *hāda kān ūmr-om yimkin sitte snīn*
 this was.3SG.M age-1SG maybe six years
sabſa snīn
 seven years
 ‘This/I was maybe six or seven years old.’

In both examples, speakers unequivocally identify the utterance as Domari rather than Arabic, and have no alternative structure to express the same utterance in Domari. In other words, none of the Arabic-derived items (italicized in the examples) have potential substitutes that are not Arabic-derived or are uniquely or distinctively Domari. In both examples, the element that makes the utterance distinctively Domari is the anchor of the predication: in (4a), it is the verb *lak-ed-om-is* ‘I saw her’, for which the Arabic equivalent would be *šuf-t-hā*, with identical morpheme sequence as in the gloss above, the final segment being specifically feminine in gender.

In (4b), which has a nominal sentence with no verbal predicate, it is an even simpler element, namely the 1SG possessive marker *-om* (for which the Arabic equivalent is *-ī*). In these selected utterances, context-bound separation of languages is thus expressed entirely through the use of a single inflected word, in the first case, and through a single suffix, in the second. Such utterances may not be the only possible pattern in Domari conversation, but they are not at all uncommon. A great part of the resources in the repertoire of Domari speakers can be used indiscriminately regardless of context (in the sense of “language choice”). Despite this fact, speakers are still able to maintain a separation that can be conceptualized upon reflection as the use of distinct “languages.”

Example (4) makes it clear that the structural outcomes of contact can, over time, be very far-reaching and can affect not just the domain of the lexicon but also of the organization of information and grammatical relations at the level of the utterance, as reflected in the domain of functional categories. We find the Arabic-derived past-tense auxiliary *kān* ‘was’, numerals, conjunctions, negation particles, discourse markers and fillers, indefinites, and plural agreement in borrowing nouns that accompany the borrowed numerals.

In searching for the motivations for, predictions of, and constraints on functional borrowing, we can be guided by two theoretical considerations: the postulation of meaningful implicational hierarchies of functional borrowing and the rareness of borrowing around certain functional categories. Notional implicational hierarchies had been proposed by Moravcsik (1978), in relation to specific pairs of typological categories such as the following:

- (5) Implicational hierarchies in structural borrowing (after Moravcsik 1978):
- a. lexical > non-lexical
 - b. nouns > non-nouns
 - c. free morphemes > bound morphemes
 - d. derivation > inflection

The overall theme is one of semantic transparency as a facilitator for borrowing, or non-transparency as a restriction (see also Field 2002). Thomason & Kaufman (1988) presented a frequently cited gradient of likelihood of borrowing, in which, however, some categories such as “function words” or “typological features” remain rather vague and undifferentiated, while the overall theme merely suggests that prolonged and intensive contact between languages is likely to yield more extensive borrowing. More specific results have since been obtained through targeted cross-linguistic sampling (Stolz & Stolz 1997, Matras 1998, Elšík & Matras 2006, Matras 2007). These works all tend to point to a connection between susceptibility to borrowing and the truth- or presupposition value that is

assigned by a category to an element of propositional content. A good example is the borrowing hierarchy for connectives: *but* > *or* > *and* (where 'greater than' indicates greater likelihood of borrowing and the order is implicational; see Matras 1998), a hierarchy that has been widely attested across different samples. Contrast – the unexpected that is beyond the speaker's control and therefore associated with interactional tension as it puts the speaker's assertive authority potentially in jeopardy – is more prone to borrowing than addition, which conveys a continuous and expected inferential chain. Similar hierarchies have been identified for other functions such as indefinites, comparatives, tense and modality, and more:

- (6) Borrowing hierarchies that reflect control and the speaker's assertive authority (following Matras 2007, 2009):
 - a. contrast > disjunction > addition
 - b. modality > aspect/aktionsart > future tense > (other tenses)
 - c. obligation > necessity > possibility > ability > desire
 - d. concessive, conditional, causal, purpose > other subordinators
 - e. factual complementisers > non-factual complementisers
 - f. superlative > comparative > (positive)
 - g. indefinites > interrogatives > (other) deixis, anaphora

Such hierarchies are difficult to explain, especially in cross-linguistic perspective, with strict reference to either formal-structural features or merely to sociolinguistic conditions and prestige. Rather, they capture an interactional dimension, whereby those categories that clash with or cannot be firmly derived from presuppositions are more prone to borrowing. This suggests that the speaker's management and control of the interaction is linked to the speaker's executive control of repertoire components – the selection and inhibition mechanism. That in turn suggests that the diachronic process of borrowing is linked to the processing of language in conversation in bilingual settings.

As we saw above in example (2), lapses in the application of the selection and inhibition mechanism (executive control) may lead to selection of items from the repertoire that are functionally adequate (i.e., have the correct intended meaning) but contextually inappropriate (i.e., are unexpected and possibly not understandable to the interlocutor, and so not in line with the speaker–hearer contract for the interaction). Such lapses across a variety of structural domains may, depending on the acceptability of multilingualism, become commonplace and conventionalized, leading to long-term structural outcomes.

Invariably, the social power relations between the languages play a role in determining the extent of propagation of what starts off as idiosyncratic innovations in discourse. The conditions are similar to those mentioned above for lexical borrowing: Languages that are spoken by minority

communities, primarily in informal and domestic settings, where bilingualism is ubiquitous and there is little or no institutional support for the heritage language, are more likely to show lax normative control that favors the adoption of functional loans from a more dominant language. Still, the social setting conditions the acceptability and propagation of innovations. Cognitive triggers and universal conditions of interaction management are behind the occurrence of innovations in discourse, and that is confirmed by the hierarchical nature of functional borrowing, which makes the pathways of borrowing to some extent predictable.

The message to take away from such generalizations is not, however, that absolute constraints operate that might rule out the borrowing of certain forms or functional categories. Rather, it is that factors are in operation that motivate borrowing around certain types of categories. Where such motivations are absent, we are less likely to encounter borrowing.

A case in point is the borrowing of pronouns. Wallace (1983) points to the borrowing of personal pronouns in South Asian languages, but notes that in the relevant languages pronouns encode social relations and as such are more akin to lexical titles (such as ‘your honour’ or ‘your majesty’, both, notably, containing borrowed components in English). Thomason & Everett (2001) add to the list cases of “mixed languages,” though here the point is precisely that the processes that lead to the mixture are distinct and involve speakers’ intentional and conscious intervention in re-organizing their repertoire (much like secret languages, which often show camouflage pronouns). The point about genuine deictic and anaphoric forms is that they rely on a harmonious mental referential map that speaker and listener share. They are thus precisely at the opposite end of the structural cline of more easily borrowable forms, compared to the contact-prone items where the speaker’s proposition ventures potentially into the unknown and unconfirmed. Personal pronouns in the sense of actual deictic or anaphoric rather than lexicalized titles of the South Asian type are more rarely borrowed, because speakers are less motivated to generalize them across their repertoire of forms. That is because they pose less of a challenge to the selection and inhibition mechanism, being, as indexical procedures, high on the hierarchy of shared presupposition elements.

Inflectional morphology has also been identified as relatively low on the cline of borrowing. By contrast with derivational morphology (for example, markers of agentivity or diminutives, or word-class changing morphology such as French derived English *-able*), inflectional morphology does not carry meaning in the sense of Myers-Scotton & Jake’s (2000) “early system morphemes.” Here, we can speak of the integrity of processing operations at the level of the predication as an expression of harmonious mapping of information between speaker and listener. This too may not be an absolute constraint on borrowing. Where the speaker needs to make choices when managing complex repertoires and accommodating to the expectations of interactional settings, the drive to compromise on the consistent selection

and inhibition (and so in effect do away with language boundaries around particular functional categories) is lower, and functional categories therefore prove to be more stable and less prone to borrowing.

In the following, I discuss a number of cases where borrowing occurs in areas of function that are seemingly less prone to contact-induced structural change, but the examples are indicative of particular local motivations and their interplay with the structural features of the languages involved. In a sense, these are the exceptions that prove the rule, or the rarities that illustrate what particular conditions need to be met for borrowing to occur that we are less likely to expect.

While the borrowing of definite article forms is relatively rare, it is not uncommon to find definite articles that are borrowed alongside lexical items and re-interpreted as part of the lexical root, as in Spanish *alcalde* 'mayor', from Arabic *al-qāḍī* 'the-magistrate'. Algerian Arabic integrates some French nouns into its inflectional template for nominal plural formation, as in *tumubīl* pl. *tumubilāt* 'automobile(s)', *šumbra* pl. *šnāber* 'room(s)'. But for some nouns, the French definite article is reanalyzed and becomes the marker of plurality: *kādu* pl. *likādu* 'present', *ʔami* pl. *lizami* 'friend(s)'. In both cases, borrowing is enabled as the articles are reanalyzed from determiners with an indexical function, to derivational morphemes that adjust or simply accompany meaning.

The borrowing of overt markers of tense and aspect is also relatively rare. But in the central and eastern European dialects of Romani, so-called Slavic Aspect or aktionsart markers are adopted on a wholesale basis and applied to pre-European Romani verbal roots, as in Polish Romani *za-pindžkirel* 'to introduce' (from *pindžkirel* 'to recognize' – Polish *za-poznać* and *poznać*), *do-resel* 'to obtain' (from *resel* 'to arrive' – Polish *do-stąpić* and *stąpić*), *pše-džal* 'to cross, climb over' (from *džal* 'to go' – Polish *prze-chodzić* and *chodzić*). Here we have a wholesale adoption of the structural procedures of meaning derivation, which in effect leads to a fusion of the two languages in this domain. What is crucial is that what is widely referred to in discussions of Slavonic languages as "aspect" is in fact a meaning-derivational procedure, that is, not one that anchors the event in the interactional context of immediately relevant knowledge shared by speaker and listener (as in the case of the English progressive aspect). ~~It does not pertain~~ to the internal structure of the event independently of the interaction context (hence *aktionsart*).

Formal similarities among functionally related inflectional morphemes may trigger analogies that can facilitate borrowing. In Maltese, adjectives that are borrowed from Italian are generally adopted into the inherited (Semitic) inflection patterns; and they also retain gender and number agreement with borrowed Italian nouns. Maltese adjectives follow the noun, as they normally do in the principal contact language, Italian (and Sicilian). Like Italian, Maltese too has two genders. Borrowed adjectives that end in a consonant take the inherited (Semitic) inflection pattern, in which the M.SG is treated as default and has no identifiable vocalic ending, while the F.SG ends in *-a*. Thus, with

masculine nouns Maltese has *f'kuntest modern-ø* 'in a modern context', corresponding to Italian *in un contesto modern-o*, while with feminine nouns it has *poezija modern-a* 'modern poetry', matching Italian *poesia modern-a*. The coincidental inflectional similarity in the feminine singular triggers an analogy with the plural, where the Italian inflectional ending *-i* (Italian M.PL) is preserved on adjectives that are borrowed from Italian: *binjiet modern-i* 'modern buildings', corresponding to Italian *edifice modern-i*.

Some Romani dialects of the Balkans replicate parts of the person concord set from Turkish (Elšík & Matras 2006: 136). This results from an analogy that is based on chance similarities between the inherited Romani and Turkish conjugations. The inherited Romani past-tense concord markers contain the consonant *-m* in the first person (singular *-om/-em/-im*, depending on dialect, and plural *-am*) and a consonant *-n* in the second person (singular *-an*, plural *-en*). They thus resemble the corresponding Turkish singular forms 1SG *-Vm* and 2SG *-Vn* (with variation subject to vowel harmony). The Turkish plural pronouns are augmented forms of the singular morphemes: 1PL *-VmVz*, 2PL *-VnVz*. By analogy, a number of Romani dialects spoken in Bulgaria (such as the dialects of Sindel, and of Kaspičan in the Shumen region) form a past-tense 1PL concord marker *-amus* and a past-tense 2PL marker *-enus*. Here, the agglutinative marking of plurality in the contact language makes the marker *-us* analyzable. It is replicated in Romani with inherited verbs, replacing the original marker (which is preserved in other dialects of the language).

Domari shows an interesting example of the outcome of interplay between a motivation to borrow the comparative/superlative function (which signals a modification to presuppositional knowledge), and resistance toward the replication of a complex inflection template from the contact language Arabic. Arabic employs a morphophonological template *áCCaC* to derive comparative/superlative forms from consonantal roots: *kbīr* 'big', *ákbār* 'bigger'; *zǧīr* 'small', *ázǧār* 'smaller'. This template cannot easily be isolated or integrated into the agglutinative-inflectional morphology structure of Domari, nor is it simple or even possible to break down Domari adjectives such as *tilla* 'big' or *kištota* 'small' into tri-consonantal roots for insertion into the Arabic-based derivation template. The solution adopted by Domari speakers is to borrow the full Arabic word-form for all comparative/superlative forms, resulting in complete borrowing-based suppletion of the inventory of adjectives – *tilla* 'big', *ákbār* 'bigger'; *kištota* 'small', *ázǧār* 'smaller' – and thus also a complete fusion of the entire word class of Domari comparative/superlative forms with Arabic.

An interesting case is that of verbs. Various studies have pointed out the complexities of the structural integration of borrowed verbs (see already Moravcsik 1975; cf. Wichmann & Wohlgemuth 2007, Wohlgemuth 2009). Verb borrowing shows a continuum across languages of various integration patterns, which can be summarized as follows:

- (7) Types of loan verb integration (Matras 2009: 176):
- a. No modification of the original form of the verb (“direct insertion”)
 - b. Morphological modification of the original form of the verb (“indirect insertion”)
 - c. Insertion of the original form of the verb into a compound construction where it is accompanied by an inherited verb (“light verb”)
 - d. Import of the original verb along with its original inflection (“paradigm transfer”)

The last strategy, that of paradigm transfer, is rare, but attested for example in Balkan dialects of Romani, where Turkish lexical verbs are adopted along with their Turkish verb inflections, including both tense and person inflections, which they retain even in communities where speakers no longer have active command of Turkish. Both the light verb strategy and that of morphological modification bear certain functional similarities, in that both in effect assign to the borrowed lexical root a “verbalizing” element, often one that attributes to it specific valency. Light verbs, such as those used in Turkic and Indo-Iranian languages to adapt borrowed roots (often from Arabic, but in modern Indo-Aryan languages also from English and other languages), generally distinguish between *BECOME*-verbs, where the subject is an experiencer (typically intransitives), and *DO*-verbs where the subject is an agent (typically transitives). Arabic and Hebrew integrated loan roots by assigning them to an inflectional paradigm that expresses intensification or reiteration.

In many languages, the morphological markers used to integrate loan verbs are those that are also used within the language to derive verbs from non-verbs. It seems then that cross-linguistically there is a frequent need to attribute to lexical roots explicitly a predication-anchoring function, drawing on elements of the recipient language that have the power to assign such a function. At the same time, the difficulty of carrying over inflectional paradigms across languages, especially for verbs, suggests that predication-anchoring morphology is what speakers most clearly perceive as marking out a particular “language.” This happens where they most consistently hesitate to compromise the perceived boundaries between “languages” in the sense of sets of structures that are reserved for particular interaction settings and constellations and not others.

5. Convergence and Grammaticalization

In the previous sections I discussed the replication of what I call linguistic *matter* (Matras 2009) from one language to another. This pertains to the use of actual phonological word forms or affixes and their integration into another language, as we defined “borrowing” above. But cross-linguistic

generalizations can also pertain to *patterns*. Here, it is the form-function or form-meaning mapping that is replicated from one language to another. Weinreich (1953) described pattern replication as a change in the function of morphemes in a “replica language,” inspired by a “model language,” referring to the process as “convergent development.” The phenomenon is also widely known by Haugen’s (1950) label *calque*. Later works refer to change in form-function mapping that is triggered by an external model as “pattern transfer” (Heath 1984), “metatypy” (Ross 1996, 2001), and “partial or selective copies” (Johanson 2002).

Following Weinreich, we might regard convergence as a kind of compromise strategy that allows speakers to continue and flag language loyalty through a more-or-less rigid choice of word-forms and at the same time to reduce the load on the selection and inhibition mechanism by allowing patterns to converge, thus maximizing the efficiency of speech production in a bilingual situation. Consider the following example:

- (8) Trilingual (German, Hebrew, English) child, aged 4:6 (Matras 2009: 26):
ze avál yafe!
this but pretty
'This is very pretty indeed!'

In a Hebrew utterance, addressed to the Hebrew-speaking parent, the trilingual child generates an exclamatory particle *avál*, which does not exist in norm-based, adult monolingual Hebrew. He does this on the basis of the model pattern of German: *Das ist aber schön!* The creative process is motivated by a wish to conform to the expectations of the present context of interaction and select word forms that are licensed for use with the present interlocutor (identified socially as words belonging to “Hebrew”). At the same time the speaker wishes to exploit the full expressive potential of their repertoire as a whole, which includes the attributive-exclamative particle *aber*. Reconciling the two motivations, the speaker searches for a word form that would be permissible for selection in the present context and would at the same time convey the expressive meaning of *aber*, driven by the awareness that *aber* itself is not licensed for the present interaction context. In the outcome, the speaker identifies the polysemy of *aber*, which is also used in German as an adversative conjunction. That conjunction has a translation equivalent in Hebrew *avál*. The speaker then maps the same polysemy onto the Hebrew equivalent.

The replication of patterns depends on the ability to match a new pattern to available word-forms. Speakers of Sinti Romani, for instance, replicate the German lexicalized aktionsart pattern as found in constructions such as *ich mache auf* ‘I open’, literally ‘I make up’. They do so by identifying pivotal features of the pattern and by matching them with corresponding inherited word forms in Sinti. The result is the Sinti verb *kerau pre* ‘I open’, based on a

composition of *kerau* ‘I make’ with the local relations expression *pre* ‘up, above’. A similar replication of the German *ich mache zu* ‘I shut’, literally ‘I make to’, is impeded, however, by the absence of an isolated word-form with a dative-allative meaning. Sinti uses synthetic suffixes to indicate dative-allative case. As a result, German *zu* is borrowed directly, and we get *kerau zu* (or rather: *cu*) ‘I shut’.² There is thus an interplay and mutual conditioning of matter and pattern replication.

But the following example pairs from the co-territorial languages Macedonian and Macedonian Turkish (Matras 2011: 146), and Kurdish and Neo-Aramaic (Matras 2009: 261), respectively, show how convergence involves the matching of key features that are selected among key features of inherited structures, resulting in a resemblance of patterns that does not, however, necessarily show complete isomorphism:

- (9) a. Macedonian:
čovек-ot što dojde
 man-the what arrived
- b. Macedonian Turkish
adam ne geldi
 man what arrived
 ‘the man who arrived’
- (10) a. Kurdish (Kurmanji):
ez rabû-m û min derî vekir
 1SG.NOM stood.up-1SG and 1SG.OBL door opened.Ø
- b. Saqqez Neo-Aramaic:
qîm-na, tara-kê plix-li
 stood.up-1SG.ITR door-the opened-1SG.TR
 ‘I stood up and opened the door’

In the Balkan case (Macedonian/Macedonian Turkish), the relative clause is a finite predication that follows the head, unlike the historical and Standard Turkish preposed gerundial *gel-en adam* ‘the man who arrived/arrives’. It is introduced by a relativizer, which in turn is derived by drawing on the semantic-pragmatic properties of the interrogative as a marker of clarification, *ne* ‘what’, which alters its illocutionary force from a question (where clarification is to be provided by the listener) to an assertion (where it is provided by the speaker). Other properties of the two clauses differ, such as the marking of definiteness on the head noun. In the Anatolian/Caucasian case involving Kurdish and Neo-Aramaic, different inflectional paradigms are used to align the subject with intransitive and transitive verbs. This pattern shows consistent ergativity in Kurmanji,

² Common transliteration conventions in Romani linguistics use [ç] for [ts]; but Sinti is frequently written in a system based on or heavily oriented towards German orthography.

where not just verb agreement, but also the case marking of the subject pronoun differs. It also shows different person markers in Neo-Aramaic, derived in the intransitive verb from a copula form *-na*, and in the transitive verb from the prepositional form of the subject *-li* lit. ‘to me’.

Both cases thus involved functionalization of an inherited element in order to adapt to the model of the contiguous language. Heine & Kuteva (2005) therefore speak of contact-induced grammaticalization as a driver of morphosyntactic change in contact situations. They propose that general notions of grammaticalization theory such as SEMANTIC EXTENSION, EXPANSION OF DISTRIBUTION, and INCREASED FREQUENCY can account for and help frame processes of convergence. The issues of “belonging” to an original donor system and “employment” in the receiving system are dealt with by identifying one of the language-specific constructions as a “model” and the other as a “replica.”

Another way of approaching the question of borrowing in cases of morpho-syntactic convergence is the notion of PIVOT-MATCHING (Matras 2009: 240–2). This is compatible in principle both with the ideas of contact-induced grammaticalization and with that of a replication of constructions (a topic dealt with in several of the contributions to Wiemer, Wälchli, & Hansen 2012). However, it is not constrained by the directionality condition of the grammaticalization model: it allows to account both for the loss of categories as a result of contact and for the acquisition of new ones. And it can offer a local account of the actual construction components that are affected by cross-language replication.

Convergence can affect the overall typology of a language, as it drifts for instance from relying on preposed nominalized constructions, as in Turkish, to postposed subordinated clauses, as in the Balkan Turkish dialects. While word order is a natural candidate for convergence, some categories appear to attract pattern replication but to resist the replication of matter. Thus, definite articles often cluster in geographical areas, a well-cited example being the emergence of postposed definite articles in some contiguous languages of the Balkans (Romanian, Albanian, Bulgarian, and Macedonian), whereas matter replication of definite articles is rare.

The driving force behind convergence or pattern replication seems to be in the conventionalization of what we might call *contextually inferred meanings* into regularly attributable meanings. Meanings that are contextually derived from particular configurations, rather than fixed and attributable to the individual word-forms and morphs that compose those configurations, are more likely to be treated as flexible and to be generalized throughout the repertoire. This is related to the flexibility of pragmatic inferences and the tendency to regard them as universal rather than bound to particular contexts and settings, and therefore more open to creative and innovative processes. The organization of clauses seems more susceptible to convergent developments, as speakers seek to reduce the processing burden on organizing entire utterances and at the same time allow themselves

greater flexibility in combining words than at the level of individual word configuration, which is more closely associated with tighter context-bound norms. Other features that show greater susceptibility to convergence are related to loose configurations among constituents that are more likely to be standalone and independent (and whose combination is thus regarded as more ad hoc), such as possessive constructions or attributive constructions, followed by those that appear at the single word-level.

6. Conclusion

A comprehensive overview of grammatical borrowings in the languages of the world is still outstanding, although Matras & Sakel (2007b) include a collection of case studies based on a comprehensive template and Seifart (2017) presents an assessment of a database on morphological borrowing. (See also the other sources cited above.) A complicating factor, in addition to the usual challenges of sampling, is the need for diachronic and etymological information in order to be able to identify borrowing, and for historically grounded sociolinguistic information about the nature of contacts and power relations that have shaped communicative practices in multilingual settings, in the recent and more distant past. For the languages of many parts of the world, such information is missing.

Nonetheless, sampling has so far succeeded in exposing many patterns. The search for constraints on borrowing and for the factors that facilitate borrowing rests on an epistemology that identifies overwhelming trends as worthy of attention even if isolated exceptions exist. It purports to be able to derive explanatory models from trends and sees the cross-linguistic examination of borrowing as leading to the formulation of an explanatory model. If we can identify patterns, then that suggests that borrowing is not random, but that it is a reflection of human communicative behavior. In this way, the study of structural outcomes of language contact can contribute to a better understanding of the language faculty itself, and possibly even of key aspects of the evolution of human language.

References

- Auer, Peter. 1984. *Bilingual conversation*. Amsterdam: John Benjamins.
- Auer, Peter. 1995. The pragmatics of code-switching: A sequential approach. In *One speaker, two languages. Cross-disciplinary perspectives on code-switching*, ed. by L. Milroy & P. Muysken, 115–35. Cambridge: Cambridge University Press.
- Backus, Ad. 1996. *Two in one. Bilingual speech of Turkish immigrants in the Netherlands*. Tilburg: Tilburg University Press.

- Bakker, Peter. 1997. *A language of our own. The genesis of Michif – the mixed Cree-French language of the Canadian Métis*. New York: Oxford University Press.
- Bakker, Peter. 2003. Mixed languages as autonomous systems. In Bakker & Matras 2003, 107–50.
- Bakker, Peter & Yaron Matras (eds.). 2003. *The mixed language debate. Theoretical and empirical advances*. Berlin: Mouton de Gruyter.
- Berk-Seligson, Susan. 1986. Linguistic constraints on intra-sentential code-switching: A study of Spanish/Hebrew bilingualism. *Language in Society* 15.313–48.
- Bialystok, Ellen, F.I.M. Craik, & G. Luk. 2008. Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics* 21.6.522–38.
- Blommaert, Jan & Ad Backus. 2013. Superdiverse repertoires and the individual. In *Multilingualism and multimodality. The future of education research*, ed. by I. Saint-Georges & J.J. Weber, 11–32. Rotterdam: SensePublishers.
- Campbell, Lyle. 1993. On proposed universals of grammatical borrowing. In *Historical linguistics 1989*, ed. by H. Aertsen & R.J. Jeffers, 91–109. Amsterdam: John Benjamins.
- Clyne, Michael. 1967. *Transference and triggering*. The Hague: Nijhoff.
- Comrie, Bernard. 2000. Language contact, lexical borrowing, and semantic fields. In *Languages in contact*, ed. by D.G. Gilbers, J. Nerbonne, & J. Schaeken, 73–86. Amsterdam: Rodopi.
- Elšík, Viktor & Yaron Matras. 2006. *Markedness and language change: The Romani sample*. Berlin: Mouton de Gruyter.
- Embleton, Sheila M. 1986. *Statistics in historical linguistics*. Bochum: Brockmeyer.
- Field, Frederic W. 2002. *Linguistic borrowing in bilingual contexts*. Amsterdam: John Benjamins.
- Gardner-Chloros, Penelope. 1991. *Language selection and switching in Strasbourg*. New York: Oxford University Press.
- Green, David. 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1.67–81.
- Grosjean, François. 1989. Neurolinguists beware! The bilingual is not two monolinguals in one person. *Brain and Language* 36.3–15.
- Gumperz, John. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.
- Haspelmath, Martin & Uri Tadmor (eds.). 2009. *Loanwords in the world's languages. A comparative handbook*. Berlin: Mouton de Gruyter.
- Haugen, Einar. 1950. The analysis of linguistic borrowing. *Language* 26.210–31.
- Haugen, Einar. 1969 [1953]. *The Norwegian language in the Americas: A study in bilingual behavior*. Bloomington, IN: Indiana University Press.
- Heath, Jeffrey. 1984. Language contact and language change. *Annual Review of Anthropology* 13.367–84.

- Heine, Bernd & Tanya Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.
- Johanson, Lars. 2002. *Structural factors in Turkic language contacts*. Richmond: Curzon.
- Li Wei. 2005. How can you tell? Towards a commonsense explanation of conversational code-switching. *Journal of Pragmatics* 37.375–89.
- Li Wei. 2018. Translanguaging as a practical theory of languages. *Applied Linguistics* 39.1.9–30.
- Mahsain, Fatemah. 2014. Motivations behind code-switching among Kuwaiti bilingual schools students. Doctoral dissertation, University of Manchester.
- Maschler, Yael. 1994. Metalanguaging and discourse markers in bilingual conversation. *Language in Society* 23.325–66.
- Matras, Yaron. 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36.281–331.
- Matras, Yaron. 2007. The borrowability of grammatical categories. In *Grammatical borrowing in cross-linguistic perspective*, ed. by Yaron Matras & Jeanette Sakel, 31–74. Berlin: Mouton de Gruyter.
- Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.
- Matras, Yaron. 2010. *Romani Britain. The afterlife of a language*. Edinburgh: Edinburgh University Press.
- Matras, Yaron. 2011. Explaining convergence and the formation of linguistic areas. In *Geographical typology and linguistic areas*, ed. by Osamu Hieda, Christa König, & Hiroshi Nakagawa, 143–60. Amsterdam: Benjamins.
- Matras, Yaron. 2012. *A grammar of Domari*. Berlin: Mouton De Gruyter.
- Matras, Yaron & Jeanette Sakel (eds.). 2007. *Grammatical borrowing in cross-linguistic perspective*. Berlin: Mouton.
- Meeuwis, Michael & Jan Blommaert. 1998. A monolectal view of code-switching: Layered code-switching among Zairians in Belgium. In *Code-switching in conversation*, ed. by Peter Auer, 76–98. London: Routledge.
- Moravcsik, Edith. 1975. Verb borrowing. *Wiener Linguistische Gazette* 8.3–30.
- Moravcsik, Edith. 1978. Universals of language contact. In *Universals of human language*, ed. by Joseph H. Greenberg, 94–122. Stanford, CA: Stanford University Press.
- Muysken, Pieter. 2000. *Bilingual speech. A typology of code-mixing*. Cambridge: Cambridge University Press.
- Myers-Scotton, Carol. 1993a. *Social motivations for code-switching. Evidence from Africa*. Oxford: Oxford University Press.
- Myers-Scotton, Carol. 1993b. *Duelling languages. Grammatical structure in code-switching*. Oxford: Oxford University Press.
- Myers-Scotton, Carol. 2002. *Contact linguistics. Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Myers-Scotton, Carol & Agnes Bolonyai. 2001. Calculating speakers: Code-switching in a rational choice model. *Language in Society* 30.1–28.

- Myers-Scotton, Carol & Janice Jake. 2000. Four types of morpheme: Evidence from aphasia, code-switching, and second language acquisition. *Linguistics* 38.1053–100.
- Nortier, Jacomine. 1990. *Dutch/Moroccan Arabic code-switching among young Moroccans in the Netherlands*. Dordrecht: Foris.
- Pennycook, Alistair & Emi Otsuji. 2015. *Metrolingualism: Language in the city*. London: Routledge.
- Pfaff, Carol. 1979. Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language* 55.291–318.
- Poplack, Shana. 1980. Sometimes I'll start a sentence in Spanish y termino en español. *Linguistics* 18.581–618.
- Renfrew, Colin, R.L. Trask, & April McMahon (eds.). 2000. *Time depth in historical linguistics*. Cambridge: McDonald Institute.
- Ross, Malcolm. 1996. Contact-induced change and the comparative method. In *The comparative method reviewed*, ed. by Mark Durie & Malcolm Ross, 180–217. Oxford: Oxford University Press.
- Ross, Malcolm. 2001. Contact-induced change in Oceanic languages in north-west Melanesia. In *Areal diffusion and genetic inheritance: Problems in comparative linguistics*, ed. by Alexandra Y. Aikhenvald & R.M.W. Dixon, 134–66. Oxford: Oxford University Press.
- Sankoff, David. 1998. A formal production-based explanation of the facts of code-switching. *Bilingualism: Language and Cognition* 1.39–50.
- Sankoff, David & Shana Poplack. 1981. A formal grammar of code-switching. *Papers in Linguistics* 14.3–46.
- Seifart, Frank. 2017. Patterns of affix borrowing in a sample of 100 languages. *Journal of Historical Linguistics* 7.3.389–431.
- Stolz, Christel & Thomas Stolz. 1997. Universelle Hispanismen? Von Manila über Lima bis Mexiko und zurück: Muster bei der Entlehnung spanischer Funktionswörter in die indigenen Sprachen Amerikas und Australiens. *Orbis* 39.1–77.
- Swadesh, Morris. 1952. Lexicostatistic dating of prehistoric ethnic contacts. *Proceedings of the American Philosophical Society* 96.452–63.
- Thomason, Sarah G. & Dan L. Everett. 2001. Pronoun borrowing. In *Proceedings of the twenty seventh Annual Meeting of the Berkeley Linguistics Society*, ed. by Ch. Chang, 301–15. Berkeley, CA: Berkeley Linguistics Society.
- Thomason, Sarah G. & T. Kaufman. 1988. *Language contact, creolization and genetic linguistics*. Berkeley, CA: University of California Press.
- Wallace, S. 1983. Pronouns in contact. In *Essays in honor of Charles F. Hockett*, ed. by B. Agard, G. Kelley, A. Makkai, & V. Becker Makkai, 573–89. Berlin: Mouton de Gruyter.
- Weinreich, Uriel. 1953. *Languages in contact*. The Hague: Mouton.
- Wichmann, Søren & Jan Wohlgemuth. 2007. Loan verbs in a typological perspective. In *Aspects of language contact. New theoretical, methodological and*

- empirical findings with special focus on Romanisation processes*, ed. by Thomas Stolz, R. Palomo, & Dik Bakker, 89–121. Berlin: Mouton de Gruyter.
- Wiemer, Björn, Bernard Wälchli, & Björn Hansen (eds.). 2012. *Grammatical replication and borrowability in language contact*. Berlin: De Gruyter.
- Wohlgemuth, Jan. 2009. *A typology of verbal borrowings*. Berlin: Mouton de Gruyter.

PROOF