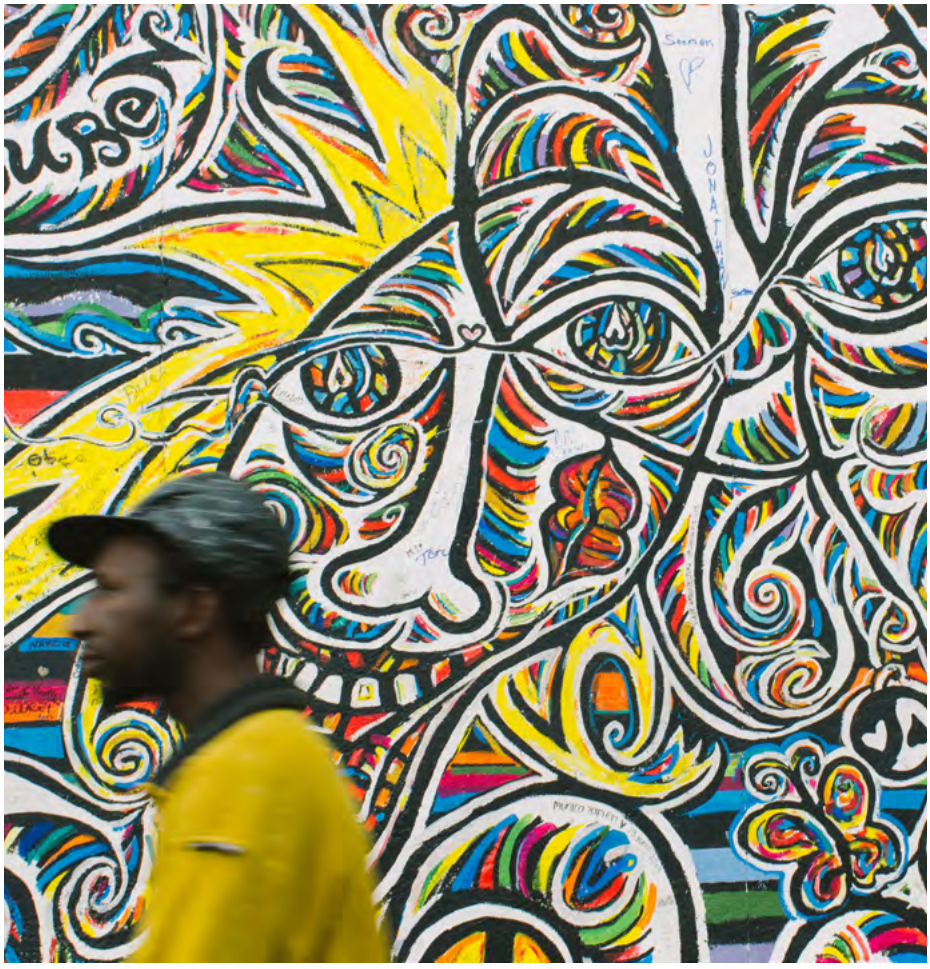
Seemen
JONATHAN

Multilingualism and code-switching

TONJES VEENSTRA & ARTEMIS ALEXIADOU

The global transformations of the age we live in, such as forced displacement of people by war and armed conflicts, economic migration from south to north, population movement due to effects of global warming, as well as rapid urbanization have led to our societies being increasingly multilingual. Although this still may be a common conception in Western societies, it suffers from what sometimes has been called the »monolingual bias«, the idea that human beings, and the societies they constitute and participate in, are essentially monolingual, an idea originally rooted in the nation state ideology of the late nineteenth Century. Contexts like Africa, Asia and India, however, clearly challenge this conceived Western notion of »one-nation-one-language«, as has been pointed out by many, and more recently by Coulmas (2018). For example, given that there are about seven thousand languages spoken in the world but only 193 countries, it follows that most, if not all, societies are multilingual in nature. The top five, based on Ethnologue 2016's rankings, are Papua New Guinea (839), Indonesia (707), Nigeria (526), India (454) and the United States of America (422). But one has to be careful and distinguish between societal and personal multilingualism, because it is not necessarily the case that citizens in highly multilingual societies are multilingual themselves. Thus, the greater majority of citizens of the United States of America are monolingual speakers of English. Nevertheless, it seems safe to say that on a world-wide scale being bi- or multilingual is the default. There are many more reasons to leave this bias behind. One of them, closer to home, is the multilingual character of metropolitan areas such as Berlin. Also, in this case, it is not so much a recent but a century-old development, almost right from the start as a Slavic settlement. From the trade contacts of Jewish merchants during the Middle Ages, Dutch engineers drying up the swamps in the 17th Century, Huguenots after their expulsion from France, Russians coming in after the October Revolution, to the more recent influx of migrant workers from Southern Europe, Turkey and Vietnam, people with different linguistic backgrounds have come to this city and stayed there. In this sense Berlin has always been a multilingual city and is even becoming more and more multilingual nowadays, partly due to some of the global transformations mentioned above. Almost every third citizen of Berlin has by now foreign roots (either as immigrants or as (grand)children of immigrants). It is estimated that more

than 120 languages are currently being spoken in the Berlin area with a total population of three-and-a-half million people. The five major groups of immigrants and their descendants are Turks, Poles, Italians, Bulgarians and people from the Russian Federation. Many of the languages spoken by these groups can by now be considered as an integral part of the multilingual mosaic of the city. At the same time members of these communities also speak German, although to a varying degree. A similar pattern has also been observed for other metropolitan areas in the (Western) world, such as London, Paris, New York and Toronto. The bottom line is that in general inhabitants of such areas are best characterized as bi- or multilingual.

The monolingual bias has also been pervasive in linguistics, partly due to the following famous quote by Chomsky (1965: 3):

> »*Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance.*«

One of the main goals of linguistic theory is to adequately characterize and explain the structures of the grammar that each individual has constructed in their mind, thereby providing formal models of the linguistic competence of human beings. Most approaches, however, have been solely focused on the monolingual speaker. But what happens if speakers have two or more grammars at their disposal?

Research at the Leibniz-Centre General Linguistics has been dealing with several aspects of multilingual competence. Gagarina and colleagues have been investigating the linguistic development of multilingual children showing that having two languages at one's disposal is an advantage. In this paper, we will focus on code-switching. Other terms that have been used to describe this phenomenon are code-mixing or language mixing, which we will use interchangeably in the remainder. Code-switching, then, is when people use more than one language in a stretch of discourse, even within a sentence. Many bilinguals produce such mixed sentences in ordinary conversations, and these are produced with great ease and complete fluidity:

(1) a. I lose my temper *porque a mi [sic] me da mucho coraje*. English/*Spanish*
›I lose my temper because it makes me so mad‹

b. Les femmes et le vin, *ne ponimayu*. French/*Russian*
›Women and wine, I don't understand.‹

c. *Bericht*-i *schreiben* yapiyor. Turkish/*German*
›He is writing the report.‹

d. *telegrama*-o *manda* shimashita. Japanese/*Braz. Portuguese*
›He/she sent the telegram.‹

This is what bilinguals do often and repeatedly, although no one really likes it – not teachers, not parents, perhaps not even the speakers themselves. They think it reflects a poor command of one or both of the languages. They believe it weakens the language and could even lead to a whole new hybrid language. These are some of the most persisting prejudices that surround this phenomenon. Researchers, on the other hand, have found time and again that taking over foreign words into your own language does not weaken or destroy the integrity of your language, nor that language mixing is unstructured and unsystematic, nor does it reflect limited proficiency in either one of the languages. Also, people don't mix languages for the reasons we tend to think they do, namely out of laziness or because they can't think of the native word. Although most of these findings fly in the face of received wisdom, they are supported by research done over the last four decades, ever since the seminal paper of Carol Pfaff on Spanish/English language mixing in the late seventies. Actually, speakers who code-mix fluently and easily tend to be proficient bilinguals!

But what is code-switching exactly, and in which type of social settings do we find it? Starting out with the second issue, we basically find it in a wide variety of settings across the globe, ranging from the interaction between two indigenous languages to the interaction between indigenous and non-indigenous languages, the latter especially in colonial and immigrant contexts:

(2) SOCIAL SETTINGS

a. frontiers between languages or language families (e.g. French and Germanic in Brussels and Strasbourg, French and English in Ottawa or Montreal);
b. clusters of multilingual tribal groups, the members of which speak each other's languages (Amazonian Basin, Australia, Papua New Guinea);
c. dialect/standard language relations (e.g. Netherlands, Germany, Italy);
d. minority language islands (Basque in Spain, or Sorbian in Germany)
e. bilingual of native elites (French/Russian in Czaristic Russia, French/German in Prussia);
f. colonial language/dominated indigenous language (e.g. French in Morocco, English in East Africa, Dutch in Surinam);
g. migrant communities (Puerto Ricans and Mexicans in the USA, Turks in Western Europe, minority communities in Europe and the Americas in general)

A number of sociolinguistic differences between these bilingual communities have been identified, either in terms of the degree of acceptance in the community of code-mixing, and attitudes towards bilingualism in general (e.g. the highly normative situation in Belgium), the structure of linguistic domination (French/English in Montreal compared to Ottawa), or whether we are dealing with a transplanted or endogenous bilingual community (e.g. Hindi/English in Great Britain or in India). Another factor is how the patterns of bilingual speech are distributed over different generations. The result is that there is an amazing variety of mixing patterns found in naturalistic conversations in these different bilingual communities and settings, which makes it difficult, or impossible, to come up with a grand overall theory of code mixing. Nevertheless, it has been observed time and again that not everything goes: code-switches are not just randomly distributed in bilingual sentences but occur at specific points. Grammatical constraints on how speakers mix their sentences provide us then with a new window on the speakers' linguistic competence.

Turning now to code-switching proper, a first distinction is whether the switch occurs between sentences or within a sentence. Inter-sentential switching is relatively easy, because in this case the two languages are kept maximally apart, as in (1a) where the main clause is in English, and the dependent clause, which is in a peripheral position, is in Spanish:

(1a) I lose my temper *porque a mi [sic] me da mucho coraje*. English/*Spanish*
›I lose my temper because it makes me so mad‹

In such cases there is no direct grammatical dependency between the clauses which facilitates switching. Also, in switching within sentences peripherality plays a role. Thus, although the French phrase *Les femmes et le vin* is an object of the Russian verb *ponimat'* (понимать), it does not occur in the canonical object position, Russian being a SVO language. Here it is the re-ordering of the constituents that leads to the object appear in a left peripheral position of the sentence that facilitates the switching between the languages:

(1b) Les femmes et le vin, *ne ponimayu*. French/*Russian*
›Women and wine, I don't understand.‹

Peripherality is one factor constraining and/or facilitating language mixing, as it creates a configuration in which the two languages can in some sense retain their grammatical autonomy. The two examples we just discussed are instantiations of a particular subclass called »alternational code-switching«: fragments of different languages can be combined quite independently of the grammars involved, the only major constraint being that the order of sentence constituents immediately adjacent to and on both sides of the switch point must be grammatical with respect to both languages involved.

The other major subclass is called »insertional code-switching«, which involves the embedding of words or constituents from one language into the other language. Here the two grammars do come together and mingle in one sentence. In (1c) above, repeated here for convenience, a German verb phrase *Bericht schreiben* is combined with a light verb *yapiyor* from Turkish, giving rise to a structure that is absent in either of the source languages:

(1c) *Bericht*-i *schreiben* yapiyor. Turkish/*German*
›He is writing the report.‹

These constructions are known in the literature as »bilingual light verb constructions«, because they consist of the combination of a lexical verb from language X (German), and a light verb, or functional verb, from language Y, (Turkish). This is a pattern that we find in many bilingual communities from Asia to the Middle East. In (1d) we find a combination

of a verb and its object *manda telegrama* from Brazilian Portuguese is inserted into a Japanese sentence headed by the inflected light verb *shimashita*:

(1d) *telegrama*-o *manda* shimashita. Japanese/*Braz. Portuguese*
›He/she sent the telegram.‹

This construction has especially attracted a lot of attention from linguists, as it sheds new light on some of the most vexing problems in linguistic theory: (i) word order; (ii) nominal licensing through Case-marking; (iii) what is a verb? We discuss them in turn.

What are possible word orders in a language, and in which part of the grammar is this information on linearization wired in? Is it a general rule of grammar (also called a macro-parameter), or is the information present on the words themselves (or classes of words)? Word order variation has traditionally been captured in terms of a macro-parameter (e.g. Chomsky 1981), with its settings largely coinciding with the main typological classes, the major classes being head-initial (Verb-Object) and head-final (Object-Verb) languages. But the overall correlations were shown not to be perfect: there are too many languages that are not uniformly head-initial or head-final, Persian being an example of a language with a head-final verb phrase (Object-Verb) but with a head-initial adpositional phrase (Preposition-Object), the Gbe languages, spoken in West Africa, examples of a language with a head-initial verb phrase (Verb-Object) but with a head-final adpositional phrase (Object-Postposition). Hence, there is no general setting of the parameter for a language, but the parameter is set within subparts of the grammar, the verbal domain, the adpositional domain, the adjectival domain, etc.

When we investigate linearization solutions in code-switching, we have to look at language pairs, in which the source language of the light verb and the source language of the verb phrase have different word order settings: (i) the light verb taken from a Verb-Object language and the verb phrase from an Object-Verb language and (ii) the light verb taken from an Object-Verb language and the verb phrase from a Verb-Object language. The solutions adopted in these apparently conflicting sites are the most interesting for linguistic theory. In short, what we find is that the light verb is the one that determines the linearization of its complement, the verb phrase.

Let us start with the first option, of which Spanish/German code mixing presents a clear example: The light verb is taken from Spanish, a Verb-Object-language, while the verb phrase is drawn from the German lexicon, German being a head-final language. The prediction is

that, since the light verb is drawn from Spanish, the constituents in the verb phrase will have to be head initial. In fact, this is what exactly happens, as shown in example (2):

(2) Juan ha hecho *verkaufen die Bücher.* Spanish/*German*
›Juan has sold the books.‹

In order to exemplify the other option, we draw from Japanese/Portuguese and Korean/English code switching. In both instances, the light verb is drawn from the head final language (Japanese or Korean) while the constituents of the verb phrase are drawn from the head initial language (Portuguese or English). In the examples below (one of them is (1d) above), we see that the light verb determines the order of the constituents in the verb phrase:

(3) a. *telegrama*-o *manda* shimashita. Japanese/*Braz. Portuguese*
›He/she sent the telegram.‹

b. *three cups of cha drink* -ha-go sip-ta Korean/*English*
›want to drink three cups of tea.‹

Additional language pairs that have been documented as displaying this pattern are Bhojpuri/ Mauritian Creole, Japanese/English, Panjabi/English and Tamil/English. Needless to say, we take this to constitute a typologically very robust pattern.

The general conclusion we draw from this discussion is that the solutions bilingual speakers come up with as to how to linearize their utterances when they code-switch points towards the importance of the functional part of the lexicon in determining word order. It is the light verb that determines the linearization of the verb phrase headed by the lexical verb.

The second issue we want to discuss concerns the licensing of nominal constituents through Case-marking. Normally, the verb determines the Case of the object it combines with. Thus, in German the verb *helfen* ›to help‹ selects for a Dative argument and not an accusative argument:

(4) Sie hilft dem Mann. (*den Mann)
›She helps the man.‹

The majority of verbs in German of course select for an accusative argument. That is why the accusative is sometimes called a structural Case, and the dative a lexical one. One of the major questions in linguistic theory is how the arguments of verbs are licensed. Are all Cases (accusative, dative, etc.) lexically specified on the verb, or only a subset of them (e.g. dative) whereas the others are assigned by default in a particular configuration. Also, in this case the syntax of code-mixing can shed some light on this issue. Consider the examples in (1c) and (1d) again, repeated below:

(5) a. *Bericht*-i *schreiben* yapiyor. Turkish/*German*
›He is writing the report.‹

b. *telegrama*-o *manda* shimashita. Japanese/*Braz. Portuguese*
›He/she sent the telegram.‹

In (5a), the object of the verb *schreiben* ›to write‹ inside the verb phrase from German is marked with the Accusative Case marker *-i* from Turkish, and (5b) the object of the verb *manda* ›to send‹ is marked with the Accusative Case marker *-o* from Japanese. Thus, it seems that the licensing of nominals via Case-marking is not done by the lexical verb, but by the light verb. These examples, however, only show that structural Case (i.e. accusative) is dependent on the light verb. The following examples show that also so-called lexical Cases are dependent on the light verb. In Turkish, verbs of motion may govern one of two cases in their complement. They can govern locative Case, to indicate location, or dative Case, to indicate motion into a place. German has a similar distinction, but in this language the dative indicates location and the accusative indicates motion into a place:

(6) a. Lili fährt im Amazonasgebiet. Dative
›Lili drives (around) in the Amazon.‹

b. Lili fährt ins Amazonasgebiet. Accusative
›Lili drives into the Amazon.‹

Now, have a look at the Case-marking pattern in German/Turkish code mixing:

(7) a. *Parkhaus*-a *fahren* yapıyor — Dative – Turkish/*German*
›He is driving into the parking lot.‹

b. *Parkhaus*-da *fahren* yapıyor — Locative
›He is driving (around) in the parking lot.‹

c. **Parkhaus*-i *fahren* yapıyor — Accusative

In (7a), we have the »drive-into«-reading, and in (7b) the »drive-around«-reading, whereas the accusative marker is not possible in this position, as shown in (7c). Thus, we have the full Case-marking system of Turkish at work, and not the one from German. This means that Case-marking is dependent on the light verb and does not come from the lexical verb. This is not only true for structural Case, like accusative, but also for lexical Case, such as dative or locative.

Recent work in linguistic theory which ZAS researchers have also contributed to indicates that verbs are not monolithic entities, but actually consist of a functional part and a lexical part. The functional part is responsible for the grammatical properties and lexical part for the conceptual meaning of verbs. If we take this partition seriously, we can state that the two verbs in a »bilingual light verb construction« actually spell out the two different parts that make up the verbal category: the light verb spells out the functional part of verbs, whereas the lexical verb spells out the lexical part of verbs (cf. Veenstra & López 2016, Alexiadou 2017, López, Alexiadou & Veenstra 2017)

In conclusion, research on multilingualism has also deep implications for formal theories of linguistic knowledge. The linguistic behavior of multilingual speakers enables us to identify universal patterns and further ours understanding of how language is shaped.

REFERENCES

Alexiadou, Artemis. 2017. Building verbs in language mixing varieties. *Zeitschrift für Sprachwissenschaft* 36(1), 165–192

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT Press.

Chomsky, Noam. 1981. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Coulmas, Florian. 2018. *An Introduction to Multilingualism: Language in a Changing World*. Oxford: Oxford University Press.

López, Luis, Artemis Alexiadou & Tonjes Veenstra. 2017. Code-Switching by Phase. *Languages* 2(3), 1–17.

Pfaff, Carol W. 1979. Constraints on Language Mixing: intrasentential codeswitching and borrowing in Spanish/English. *Language* 55, 291–318.

Veenstra, Tonjes & Luis López. 2016. Little v and cross-linguistic variation: Evidence from code switching and the Surinamese creoles. In Ermenegildo Bidese, Federica Cognola & Manuela Caterina Moroni (eds.) *Theoretical Approaches to Linguistic Variation*, 317–336. Amsterdam: John Benjamins.

Prof. Dr. Dr. h.c.
Artemis Alexiadou

Dr. Tonjes Veenstra

PROF. DR. DR. H.C. ARTEMIS ALEXIADOU'S
research focuses on syntactic and morphological theory and cross-linguistic variation. She is Professor of English Linguistics at the Humboldt University, where she is the leader of the research group Experimental Syntax and Heritage Languages, and one of the Vice-Directors of the Leibniz- ZAS, where she is also the head of the Research Area Syntax and the Lexicon.

DR. TONJES VEENSTRA
is an expert on evolutionary trajectories of language, in particular creole languages, their structure and their genesis. He is a founding member of an European Network on the study of these languages (http://www.pidgins-creoles.cnrs.fr), as well as one of the presidents of International Association for Colonial and Postcolonial Linguistics.