

Mareike Keller

Compositionality: Evidence from Code-Switching

Abstract: The storage and processing of phrasemes has been discussed many times over the past decades, with varying results. Researchers still disagree as to the degree to which phrasemes are stored and processed holistically or compositionally. This paper approaches the topic of compositionality through bilingual data, which is rarely discussed in theoretical work on phraseology. It provides a qualitative analysis of verb-based phrasemes, highlighting the structural and semantic features of code-switching patterns in and around phrasemes which serve as clues to underlying production processes. The study is based on recordings of German-English informal conversation. The language-mixing patterns are presented in the framework of the MLF model (Myers-Scotton 2002; Myers-Scotton and Jake 2017). The mixing patterns inside collocations and the resistance to mixing of more idiomatic phrasemes suggest that the surface realization of phrasemes in bilingual speech is determined both by morphosyntactic code-switching constraints and by the semantic impact of nominal and verbal phraseme components on the meaning of the phraseme as a whole. The findings support both the Superlemma Theory of phraseme processing (Sprenger et al. 2006) and the MLF model of code-switching, as they provide empirical evidence for the unitary storage of phrasemes at the conceptual level as well as for their compositional assembly in accordance with structural code-switching constraints during language production.

1 Introduction

One of the much-discussed but still unresolved questions related to multi-word sequences like idioms, semi-idioms and collocations (henceforth referred to as *phrasemes*) concerns the way they are stored in the mental lexicon: Are they stored and retrieved holistically or are they assembled compositionally from individual words each time they are produced? A traditional approach to compositionality is the investigation of variation and modification in monolingual canonical data (Moon 1998; Langlotz 2006: 175–224). Further insights have been drawn

from the analysis of non-canonical data like language acquisition, aphasia, attrition, or slips of the tongue (Häcki-Buhofer 2007; Paradis 2004; Kuiper et al. 2007). More recently, psycholinguistic experiments have been conducted measuring processing speed, mostly in comprehension, but also in production (Havrila 2009; Wray 2012). In this paper the subject of compositionality is approached via a largely unexplored type of data: phrasemes in naturally occurring code-switching produced by balanced bilinguals.¹ The approach builds on the assumption that language switching or mixing alongside or within phrasemes can be employed as an indicator for chunking or parsing during language processing (Backus 2003; Wray and Namba 2003; Namba 2012). As differences between monolingual and bilingual language processing concern areas like speed of access or executive control rather than basic processing mechanisms (Paradis 2004; Bialystok and Craik 2010), the conclusions are not restricted to bilingual contexts but could also provide explanations for monolingual storage and processing of complex lexical items.

Phrasemes are a highly heterogeneous group of lexicalized word-strings and the information a phraseme can reveal with respect to language processing depends on the lexical category of its syntactic head as well as on its internal syntactic structure. This paper is devoted exclusively to phrasemes in the form of syntactic constituents with a verb as syntactic head.² Verb-based phrasemes were chosen because their comparatively complex argument structure provides more opportunities for internal language mixing than e.g. nominal phrasemes. All examples were extracted manually from a 50-hour corpus of German-English spontaneous speech.³ The paper provides empirical evidence of the mixing patterns in and around phrasemes and explores the ways in which language contact phenomena can be related to syntactic and semantic properties of phrasemes. The findings provide clues to the mental representation of phrasemes, including the

1 *Balanced bilingualism* is defined here as a native-like level of proficiency in both languages.

2 The phraseological terminology used in this paper is based primarily on Burger (2015). The term *phraseme* will be used as a cover term for idioms, semi-idioms and collocations. Verb-based phrasemes, which are the focus of the study, fall into Burger's category of referential phrasemes in the form of syntactic constituents (*nominative referentielle Phraseme*, Burger 2015: 32).

3 The data were collected between 1999 and 2005 as part of the project "Sprachkontakt Deutsch-Englisch: Code-switching, Crossover & Co." funded by the Deutsche Forschungsgesellschaft (DFG) and headed by Rosemarie Tracy (University of Mannheim) and Elsa Lattey (University of Tübingen). Further details on the speakers and data collection process are given in Tracy and Lattey (2010). My sincere thanks go to Rosemarie Tracy for access to the recordings and the transcripts.

level at which language selection takes place and will hopefully inspire further research into a complex but highly promising type of data.

The structure of the paper is as follows. As most readers will be more familiar with phraseology than code-switching, basic assumptions concerning storage and processing of phrasemes are outlined only very briefly (section 2), before the structural approach to code-switching is introduced in more detail (section 3). Then the empirical data are presented and analysed (section 4). In section 5 the findings are discussed with respect to theoretical issues concerning the compositionality of phrasemes. Section 6 concludes the paper with a short summary and suggestions for future research.

2 Conceptual Unity – Compositional Processing

The observation that different types of phraseme exhibit different degrees of fix-
edness or compositionality has been widely discussed among phraseologists, and over the past decades various taxonomic approaches placing different types of phraseme along a continuum have been proposed and refuted (Wray and Perkins 2000). The one characteristic uniting all types of phraseme, from true idioms to collocations, seems to be that they are recurrently co-occurrent sequences of lexemes which appear to be reproduced rather than creatively assembled. Some of them express meaning beyond the sum of the meaning of their individual components, some are peculiar in their syntactic make-up – but the vast majority do not show any semantic or syntactic characteristics which clearly set them apart from free combinations of words. How, then, can we tell that one string of words is a phraseme and another one is not? One indispensable precondition for recognizing phrasemes as such in actual discourse seems to be their representation as conceptual units at some level in the mental lexicon (Backus 2003: 92). However, unitary representation does not necessarily entail holistic storage and processing all the way from the conceptual level to actual phonological realization. The question that remains is: Which aspects or components of a phraseme are stored in long-term memory, and what can be assembled online during production (see Jackendoff 2002, 152–195)?

It is widely assumed that phrasemes have their own entries in the mental lexicon (Levelt 1989: 186–187; de Bot 1992: 10), but there is no agreement on what

this entry actually looks like.⁴ To describe the representation of a phraseme in the mental lexicon, Levelt and Meyer (2000: 442) introduce the term *superlemma*, which “represents the idiom’s restricted syntax and points to a set of simple lemmas.” This idea is expanded by Sprenger, Levelt and Kempen (2006) into their Superlemma Theory. The theory supports a hybrid view of phraseme processing (Cutting and Bock 1997) and claims that “[f]ixed expressions and idioms and literal language only differ with respect to the *source* of word activation: while the words of a literal phrase are activated by their own lexical concepts, the words of a fixed expression will benefit from a common idiom node” (Sprenger et al. 2006: 167). This means that in a phraseme, the individual lexemes are selected from the lexicon via the superlemma entry for the phraseme. The Superlemma Theory is attractive because it treats the production of phrasemes similarly to the production of free combinations, and it elegantly aligns production and comprehension. In addition, the contradiction between conceptual unity on the one hand and syntactic compositionality on the other is resolved by postulating the superlemma as a conceptual unit and the component lexemes as syntactically related but individually accessed pieces.

As long as we are dealing with monolingual data, we can use e.g. speed of access in experimental settings, or performance errors in spontaneous and elicited speech as indicators of chunking or parsing of phraseological units. When we look at bilingual data, the contrast between the two languages involved offers an additional clue to the way in which conceptual units are assembled into actual phonetic strings. One might assume that a string of words which appears as a unit on the level of conceptual representation should be barred from internal language mixing in order to preserve the exact meaning or pragmatic function of the unit. The relevance of phrasemes in contrast to simplex lexemes for the study of code-switching patterns was already noticed in a very early study by Hasselmo (1970: 196), who writes about the data he analysed: “Purely lexical conditioning of switching is obviously an important factor, but throughout this discourse it appears that larger preformulated segments play a role as well.” Later code-switching research has mentioned in passing that phrasemes are often inserted as whole constituents (e.g. Myers-Scotton 2006: 263), supporting the view that phrasemes are processed as units all the way from the mental lexicon/phrasicon to the phonetic level. However, this blanket view does not hold for all types of

4 One problem with previous research on the topic is the definition of the target structures. Earlier works focus mainly on pure idioms, or idioms in a narrow sense. The authors cited in the following paragraphs may not all have had phrasemes in the wider sense in mind, but their findings are applicable nevertheless.

phraseme. Backus (2003) cites examples of phraseme-internal language mixing, which suggest that under specific conditions phrasemes can be broken up into their sub-components at some point in the production pipeline.⁵ Furthermore, it shows that not all phraseme components are language-specific on all levels of language production. I believe that the propensity or resistance of a phraseme to internal language mixing can be used as a clue to how phrasemes are assembled during language production. The following section provides a short introduction to the structural study of code-switching. This will serve as the theoretical background against which the behaviour of phrasemes in code-switching is analysed.

3 Bilingual Code-Switching

Over the past decades, code-switching has been studied from various angles, such as sociolinguistics, psycholinguistics or syntax, with the goal of finding out which factors influence or constrain mixed utterances. Contrary to early beliefs, mixing languages is neither a sign of incompetence, nor does it occur randomly. Instead, it seems to be governed by social as well as syntactic constraints, the nature of which has not been fully understood. What we can say for sure is that code-switching constraints, just like any other grammar rule, are probabilistic rather than absolute.⁶ This paper focuses on the morphosyntactic aspect of language mixing and attempts to link it to semantic factors influencing the surface form of complex lexical items which are the object of investigation of phraseological research. Myers-Scotton's Matrix Language Frame Model (MLF model) (1993 and following) serves as the theoretical framework for the account of bilingual phraseme processing developed in the following pages. For reasons of space the account must remain somewhat superficial. Readers new to the topic are referred to Myers-Scotton and Jake (2009) for a concise but detailed overview.

The MLF model is cognitively based and lexically driven, which means it is focused on processes originating in the mental lexicon. It was devised in accordance with basic assumptions of generative grammar and aims to explain how language production is linked to linguistic competence (Myers-Scotton 2002: 14). At

⁵ Namba (2012) also deals with the topic of mixed phrasemes in code-switching. However, as his analysis is based on bilingual acquisition data from two young children, his examples are very few and cover only a small section of frequent phraseological structures.

⁶ See Mindt (2002: 210–211) who argues that any descriptive grammatical rule will have about 5% exceptions due to online processing errors, idiosyncrasies or variation/language change.

the core of the MLF model lies the claim that the distribution of languages in bilingual clauses is asymmetrical. One language, the *matrix language (ML)*, provides the morphosyntactic frame of a bilingual clause.⁷ Into this ML frame, elements from a second language, called the *embedded language (EL)*, can be inserted. The MLF model's unit of reference is the bilingual clause (Myers-Scotton and Jake 2017: 3).⁸ This means that the two principles restricting the surface realization of morphemes in code-switching are only applicable to bilingual clauses. They are not aimed at syntactic units bigger than one clause. Also, the terms ML and EL only refer to one clause at a time. According to the MLF model, the surface realization of morphemes⁹ in a bilingual clause is constrained by two principles. These two principles state that in mixed-language constituents, word order and particular grammatical morphemes (e.g. morphemes transporting information on agreement or case) have to come from the ML:

The Morpheme-Order Principle: In ML+EL constituents consisting of singly occurring EL lexemes and any number of ML morphemes, surface morpheme order (reflecting surface syntactic relations) will be that of the ML.

(Myers-Scotton 1997: 83)

The System Morpheme Principle: In ML+EL constituents, all system morphemes which have grammatical relations external to their head constituent (i.e. which participate in the sentence's thematic role grid) will come from the ML.

(Myers-Scotton 1997: 83)

The MLF model has been revised several times in order to make its predictions more precise. One crucial step in clarifying which morphemes are affected by the System Morpheme Principle was the introduction of the 4-M-Model (Myers-Scotton and Jake 2000). Myers-Scotton and Jake (2017: 2) state explicitly that the 4-M-Model is not itself a model of code-switching but a general model of morpheme processing, applicable equally well to other types of data. It relates to the MLF

⁷ This assumption is made only about so-called classic code-switching “in which empirical evidence shows that abstract grammatical structure within a clause comes from only one of the participating languages” (Myers-Scotton and Jake 2009: 337). For mixed languages, Myers-Scotton (2002: 100) proposes a composite matrix as the grammatical basis. The term ML should not be mistaken for or confused with the dominant language of a speaker or a discourse. It is a grammatical abstraction, applicable only within one clause (Myers-Scotton 2002: 58).

⁸ In terms of generative syntax: “Our unit of analysis is the clause, or CP, the projection of complementizer, or COMP” (Myers-Scotton and Jake 2015: 418).

⁹ The term *morpheme* is used for surface realizations (phonetic form in the actual utterance) as well as for the underlying lemma entry (abstract form in the speaker's mind) (Myers-Scotton 2002: 106).

model only insofar as it can help to explain the morphosyntactic regularities observed in bilingual clauses. The model assumes four different types of morpheme (hence the name, 4-M[orpheme]-Model): 1. content morphemes, 2. early system morphemes (e.g. plural affixes), 3. bridges (e.g. possessive markers) and 4. outsiders (e.g. case and agreement markers).¹⁰ Content and early system morphemes together transport the meaning of an utterance. They are accessed at the conceptual level. The two types of late system morpheme, bridges and outsiders, make the utterance grammatical in terms of the morphosyntactic structure projected by the matrix language. Their exact phonological form is selected only at the level of the formulator, once the thematic grid of the utterance has been laid out.¹¹ In short, “[c]ontent morphemes and early SMs satisfy the speaker’s intentions, while late SMs provide grammatical structure” (Myers-Scotton and Jake 2017: 3). According to the System Morpheme Principle, only the late outsider system morphemes must be supplied by the ML in bilingual constituents. Their function in a clause lies in “disambiguating grammatical roles and providing argument struc-

10 Explanatory note: The 4-M-Model assumes an asymmetry between content and system morphemes, which is crucial for language processing. In crude terms, content morphemes are conceptually activated lexical items (which assign or receive theta-roles; Myers-Scotton and Jake 2015: 425), whereas system morphemes are structurally assigned functional elements. The system morphemes are subdivided into early and late. Early system morphemes are accessed along with content morphemes from the mental lexicon. They are functional affixes which add to the semantic content but do not affect the grammaticality of the sentence. The sentences *Paul likes Anna’s sister* and *Paul likes Anna’s sisters* are equally grammatical, but the plural affix on the word *sister* in the second one changes the meaning of the proposition. The late system morphemes are subdivided again, into bridges and outsiders. Both help to make the sentence grammatical. A bridge establishes a grammatical relation between lexical items within the systactic constituent in which it occurs. In *Paul likes Anna’s sister* the possessive marker expresses the grammatical relation between Anna and the sister, which in the given sentence are both components of the same object NP. An outsider establishes a grammatical relation with a lexical item outside the systactic constituent in which it occurs. In *Paul likes Anna’s sister* the agreement marker on the verb expresses the grammatical relation between the subject NP and the verb under INFL (Myers-Scotton and Jake 2017: 7). A particular grammatical morpheme is not necessarily assigned to the same group crosslinguistically (Myers-Scotton and Jake 2017: 4). Its type depends on the kind of grammatical information the morpheme carries. In Modern English, the article only carries information about definiteness and is classified as an early system morpheme. In Modern German the article also carries information about case and is thus a late outsider system morpheme.

11 The processing components referred to by Myers-Scotton are based on Levelt’s model of language processing (Levelt 1989; Levelt et al. 1999). The model was adapted to bilingual speech by de Bot (1992) and Wei (2009); see also Myers-Scotton (2005).

ture” (Myers-Scotton and Jake 2017: 7). The intuition that different types of morpheme are accessed at different levels during language production will prove crucial for understanding which types of phraseme, or which phraseme components, are uttered in which language in code-switching discourse.

The MLF model does not make any predictions about the processing of phrasemes. Nevertheless, the following comment shows how the model relates to the question of phraseme storage and processing:

I also see my work as recognizing that explanations lie in linking a theory of language with a theory of language processing in a manner similar to the views expressed in Jackendoff (2002). Jackendoff stresses the need to consider what aspects of an utterance are in long-term memory (content morphemes in my framework) and what aspects can be constructed online with working memory.

(Myers-Scotton 2002: 310)

Myers-Scotton has not published any work focusing on phrasemes specifically but in her discussion of so-called EL-islands (full syntactic constituents from the EL inserted into an ML clause) she mentions that these islands often show phraseological characteristics:

Many of the Embedded Language islands can be considered collocations, combinations of words that often appear together as a single phrase.

(Myers-Scotton 2006: 263)

[M]any Embedded Language islands are either formulaic or routine collocations, perhaps making them similar to the activation required to access singly occurring forms.

(Myers-Scotton 2002: 162)

These comments suggest that phrasemes are likely to be inserted as chunks of lexemes from only one language into bilingual utterances. At first glance my data seem to confirm this. Most phrasemes are inserted as EL chunks and do not show internal mixing at all, among them the vast majority of adverbial and nominal phrasemes:

- (1) KL: *Not anymore.* And **{in} einer Hinsicht** *it's-äh I think it is it's more hygienic.*

[K6:730]¹²

¹² Most of the examples used in this paper are taken from the database compiled for the research project presented in Keller (2014). Transcription conventions: “In order to ensure readability of examples, we added punctuation marks and adopted the following conventions: German items are roman, English are italic; a slash signals a word- or sentence-break, a dash con-

- (2) AS: ...jetzt is des Tor net zugangen. Und **all of a sudden** hat der g'schrien ja geh halt rei', du Depp!
[AIM11b:722]
- (3) LK: Weil wenn ma', wenn ma' sieben Jahr' lang nicht nicht redet/ **first of all**, damals war's no' net so wie heut', dass du...
[L2:401]
- (4) KL: *That keeps me going. I'm pretty sure. And **der gute Wille**. And that's about it.*
[K9:467]
- (5) KL: *Things were then better over here, too, you know? Naja. But **der liebe Gott**, h-he-he evened it out.*
[K16:58]
- (6) TG: ...da ham se bloß das Essen gekriegt und 'n- **place to stay!**
[T29:949]

However, especially among verb-based phrasemes there is a significant number of items that do show internal language mixing:

- (7) TG: ...because- *he **made himself** {so} **wichtig**, you know.*
[T29:278]

Here, the German phraseme *sich wichtig machen* (Engl. *act the big shot*, literally 'make oneself important') is rendered partly in English and partly in German. Examples like (7) suggest that at least verb-based phrasemes are not necessarily accessed as completely prefabricated, language-specific strings of lexemes. Maybe they are accessed as strings of lemmas, or as superlemmas, at the conceptual level – but somewhere along the production process the superlemma must be decomposed and reassembled drawing on lexemes/morphemes from two different languages. This raises the question of which elements of a mixed phraseme appear in which language in a bilingual utterance. Or more precisely: Which elements of a mixed phraseme are realized in the language the phraseme is drawn from and which elements are translated, or calqued?

nects iterated items. Curly brackets mark ambiguous language affiliation. Round brackets indicate incomprehensible sections, square brackets set off meta-linguistic comments and indicate passages left out; [...] Note that we consistently – even within English utterances – employed German orthography for hesitation expressions, i.e. *äh(m)* (Tracy and Lattey 2010: 57). I added curly brackets to mark homophonous diamorphs, i.e. elements which could be English or German. Square brackets following examples contain file and line identification.

In the pursuit of possible constraints regulating the language distribution in the surface realization of phraseme components, Backus (2003: 92) suggests that “ML morphemes will have semantically basic meanings.” Unfortunately, Backus (2003) leaves it to the reader to decide what does and what does not qualify as semantically basic meaning. Unrelated to the topic of phraseological units, Wei (2009: 280–283) regards lemma congruence, i.e. the degree of similarity between word forms from different languages expressing the same lemma, as the organizational principle guiding the production of mixed utterances. For him the main reason for inserting EL content morphemes into an ML frame seems to be insufficient semantic or pragmatic congruence between lemmas. Likewise, Myers-Scotton (2002: 20) suggests that “lack of sufficient congruence may explain why certain structures are avoided or impossible in switching between specific language pairs.” However, what is sufficient and what is insufficient congruence remains unclear. Nevertheless, studying code-switching data might shed more light on the question of which elements are central and which peripheral in lexical entries, simplex or complex ones:

[H]ow an EL content morpheme is accommodated by an ML frame tells us something about which features characterizing that morpheme (ultimately characterizing its supporting lemma) are critical and which may be peripheral in lexical entries. At this stage, we only aim to have shown the effects on CS [=codeswitching] of different aspects of lexical structure, but we do think it is clear how studying congruence in CS has implications far beyond the nature of CS itself.

(Myers-Scotton and Jake 1995: 1019)

This is to say that the study of morphosyntactic details in code-switching data and its implications is more than just a source for understanding more about the possible compositionality of phrasemes. It holds valuable clues to the make-up of entries in the mental lexicon. With this theoretical introduction in mind, the goal of the study presented in the following section is to show how balanced bilinguals integrate verb-based phrasemes in their everyday conversations.

4 A Study of Verb-Based Phrasemes in German-English Code-Switching

The examples presented in this paper are based on 732 utterances containing various types of phraseme of the size of a syntactic constituent, extracted manually from 50 hours of informal interviews with seven German Americans (see footnote

3). Six of the interviewees emigrated to the US as adults, one at the age of fourteen. At the time of recording they were 65–87 years of age and had lived in an English-speaking environment for 42–66 years. After their emigration from Germany some of the speakers continued to use their variety of German on a regular basis, others experienced phases with no or hardly any interaction with other native speakers of German.

The close typological relatedness of English and German, which might pose an obstacle to some areas of linguistic research, is a definite advantage for an investigation of mixing patterns targeting phraseological material, because the high number of cognates and (near-)homophones along with the large overlap on the morphosyntactic level provokes a variety of mixing phenomena less likely to be found in bilingual data based on typologically more distant languages.

Phrasemes are notoriously hard to define, and the decision as to whether or not a combination of words is phraseological or not is always to a certain degree a subjective one (see Howarth 1998: 29). I cannot guarantee that I did not miss items that another phraseologist would have wanted to include. To confer a certain degree of objectivity, I included only phrasemes listed in major printed and online dictionaries of idioms and collocations.

Out of a total of 732 utterances containing phrasemes in the form of a syntactic constituent (verb-based and other), 146 (i.e. about 20%) exhibit obvious traces of the speaker's bilingualism, either in the form of code-switching in the vicinity of the phraseme or as phraseme-internal language mixing (table 1).

Tab. 1: The frequency of mixing vs. switching (N=146)

	mixing		switching	
verb-based phrasemes	59	75%	18	27%
other	20	25%	49	73%
total	79	100%	67	100%

My argumentation builds on the hypothesis that language mixing inside a phraseme is suggestive of a compositional process. Phraseme-internal language mixing can be observed primarily inside verb-based phrasemes. Therefore, the present paper focuses on verb-based phrasemes (N=451) and refers to other syntactic types of phraseme only for comparative reasons. Very early on during the research process it became clear that the semantic impact of the verb itself appears to be a crucial factor in determining the mixing patterns in utterances containing verb-based phrasemes. Consequently, the target utterances were divided

into two groups. The first group consists of 236 utterances, each containing a phraseme headed by a verb that adds a clearly discernible semantic component to the overall meaning of the phraseme (Example: *live in the lap of luxury*). These phrasemes will be referred to as *VPhr*. The second group consists of 215 utterances, each containing a phraseme headed by a light verb. These phrasemes will be referred to as *vPhr*. In a *vPhr*, the semantic core of the phraseme is carried by the nominal component (Example: *be sorry*).¹³ The verb does not add clearly discernible meaning to the overall meaning of the utterance but rather serves the syntactic function of turning the expression into a predicate (Pottelberge 2007; see also Allerton 2001; Butt 2003, 2010; Winhart 2005).¹⁴ For the present paper I included the verbs *be*, *have*, *make*, *get* from English and *sein*, *haben*, *machen* from German as heads of light-verb phrasemes. The choice is undoubtedly arbitrary, and more verbs could be included in this group.

4.1 Phrasemes with a Semantically Salient Verbal Head

All seven informants produce phrasemes with a semantically salient verbal head (*VPhr*) in both their languages with equal ease and there are hardly any cases of transfer or interlanguage forms of the kind found in contexts of foreign language acquisition. In monolingual English utterances the speakers use idiomatic *VPhr* that do not have a word-for-word translation (8) as well as idioms which can be expressed using the same image in German (9, *Sterne sehen*). In monolingual German utterances the speakers use a wide variety of standard and dialect idioms, some of which they may not have encountered anymore at all after settling in the United States (10 and 11). This shows that all speakers have a well-developed active repertoire of idiomatic expressions in both their languages.

13 The German tradition uses the term *Funktionsverbgefüge* mostly for combinations of light verb + noun. As I could not find a difference in mixing behavior between light verb + noun and light verb + adjective combinations, I have decided to treat them as one group, focusing on the semantic lightness of the verb instead of on the syntactic category of the nominal complement. I have also included more complex combinations like *be close with s.o.*, containing a light verb, an adjective, a preposition and an external valency slot.

14 The light verb constructions discussed in this paper should not be confused with the dummy verb constructions frequently mentioned in works on language contact (Myers-Scotton and Jake 2015: 428; González-Vilbazo and Lopez 2011). Light verb constructions are lexicalized phraseological units listed in monolingual dictionaries. Dummy verb constructions are a type of contact phenomenon where a light verb is used to integrate foreign lexical material from one language into another.

- (8) KL: *I'm- you know, I'm **keeping my fingers crossed**.* [K3:16]
- (9) KL: *I walked right into that door, and fast, because I was in a hurry. I **saw stars**.* [K22:523]
- (10) TG: En Abstecher hier und da und/ **den Rahm** überall **abschöpfen**, ne? [T20:954]
- (11) TG: San aa die die, wo die arme Leit alle/ ois abnehme und dann **leben wie Gott in Frankr-Frankreich**. [T1:890]

In addition to phrasemes in a monolingual context the speakers produce various forms of overt and covert language mixing in and around VPhr. One form of covert language mixing is spontaneous or idiosyncratic calquing, where a phraseme (mostly a collocation rather than a true idiom) is rendered as a word-by-word translation.

- (12) TG: Wenn ma nach *California* g'flogen san, des **hat** ja aa **lang g'numme**. [T16:1144]

In (12) the Bavarian German *hat lang g'numma* is a calque of the English collocation *take long*. Spontaneous calques are unidiomatic in monolingual standard usage and are not listed in idiomatic dictionaries.¹⁵ The calques that are produced by the speakers are limited to a few recurring items which seem to have become established within the speaker community. Apart from those few established calques, the speakers seem to notice their own spontaneous calques and make an effort to repair them:

- (13) TG: *Because, for the children's sake you have **to bring a- a little sa**/ you have to sacrifice something.* [Tel:1121]

In (13) the speaker first begins to translate the German phraseme *ein Opfer bringen* (lit. *bring a sacrifice*). The attempt is abandoned, and the speaker starts the sentence over, using the simplex verb *sacrifice*, thus achieving a non-phraseological but native-like wording.

¹⁵ Traditionally the term *calque* refers to lexicalized items (English *skyscraper* → German *Wolkenkratzer*). The spontaneous word-for-word translations described here are mostly idiosyncratic non-lexical formations.

Especially if a VPhr has no translation equivalent, we could assume that it would most likely be embedded as a whole into a clause from another language. In her code-switching studies Myers-Scotton refers to the insertion of a full EL constituent into an ML frame as an *EL island*. She assumes that phrasemes are frequent triggers for EL islands (2002: 157, 162 and 263). From studies of lexical borrowing we know that noun phrases and adverbials are borrowed quite easily. The data confirm this kind of insertion for phrasemes with a noun head or in the function of an adverbial (see examples (1)–(6) above). The borrowing of verbs is more complex, as it usually requires the borrowed item to be adapted to the morphosyntactic requirements of the recipient language (tense, word-order, etc.). With simplex verbs, borrowing along with morphosyntactic adaptation is still quite common (*to google s.th.* < *etw. googeln* < *Er hat etwas gegoogelt*). Yet, when a verb can express its meaning only in combination with at least one lexically predetermined argument, insertion in the form of an EL island does not occur. What we do find is a number of code-switches which in all likelihood are anticipational and triggered by a VPhr (14–15):

- (14) TG: Aber ich bin froh. *They **keep an eye on her**, too.* Wenn wie/ irgendwie was wär, die würden ihr helfen.

[T28:22]

- (15) KL: *...but the situation in Osoppo, I think **that really/ that- des is ma sehr nahe gegangen**, I mean I couldn't understand anybody wanting to live like that.*

[K1:164]

In each case the language is switched not only for the VPhr but for the entire clause. In (14) the switch-point coincides with the beginning of a new independent clause. Planning and production difficulties are obvious in (15), where the anaphoric subject *that* of the switched clause is first uttered in the ML, repeated in the ML and then uttered in the EL as *des*. The decision as to whether a language switch was triggered by a phraseme or was due to other factors is undoubtedly subjective. Reasonable cues are hesitation¹⁶, self-correction, hedges or metalinguistic comments, and maybe also the lack of a translation equivalent.

¹⁶ Code-switching per se is not concomitant with an increase in hesitation phenomena compared to monolingual speech (Ehinger 2003). However, in my corpus phrasemes in bilingual utterances show significantly more hesitation than those in monolingual utterances. This is particularly noticeable around verb-based phrasemes (bilingual utterances: VPhr 56% and vPhr 41%; monolingual utterances: VPhr 18% and vPhr 16%). This suggests significantly higher production costs.

The corpus contains a handful of VPhr where the phraseme as a conceptual unit is clearly attributable to language A but some components of it are realized through words or morphemes from language B. The result is a form of overt language mixing which, for lack of an established term, we will for now refer to as *partial calque*. This is rare and produced only by the speaker TG who, compared to other members of her German-American social group, is most at ease with mixing her languages:

- (16) TG: Und dann war's f- für Freudenmädchen. Sin' se ***fin} line gestanden!*** Die Soldaten, die Fl- die die Flotte, die amerikanische, war im Hafen.

[T6:284]

- (17) TG: Und mein Vater, der ging mal zur Bank in {New York} **und hat sich *{i-in}-äh line ge-gestanden***, to- get to the teller...

[T29:259]

In both (16) and (17) the underlying phraseme seems to be the English *stand in line*. The verb is realized in German, the perfect tense is selected in accordance with German colloquial norm. The nominal component, *line*, appears in English. It is preceded by the preposition *in*, which in German-English language mixing cannot be assigned to either of the two languages. Such elements are referred to as *homophonous diamorphs* (following Clyne 1967) and are often found at switch points.

The last example in this section is a rare and curious form of covert language mixing which we can call *bilingual contamination*. Contamination is a well-documented phenomenon affecting phrasemes in monolingual contexts where two phrasemes are merged into one (Cutting and Bock 1997; Burger 2015: 26). In our case, one of the phrasemes comes from English, the other one from German:

- (18) KL: Ah, des is nett, *well*, dann **gibst ihr viele Grüsse**.

[K8:51]

In (18) the verb from the German phraseme *jmdm. viele Grüße sagen* is replaced by a translation of the English *give*, which is most probably a transfer of the verbal component from the English phraseme *give s.o.'s love to s.o.* The surface lexicalization is entirely monolingual. What makes this example interesting in the given context is that just as in the overtly mixed examples it is the verbal component which is calqued.

So far, we have established that the speakers have a well-developed repertoire of VPhr in both their languages. They use them in monolingual as well as in

bilingual turns. If during a turn a speaker wants to use a phraseme from the language which is currently not the ML, he or she switches the language, possibly in anticipation of the phraseme, for the entire clause. The use of VPhr in bilingual clauses in the form of overt mixing is rare and often accompanied by hesitations and repairs. In the following section we will look at verbal phrasemes with a semantically light verb. These show more overt phraseme-internal mixing and thus provide more interesting evidence with respect to the question of compositional processing.

4.2 Phrasemes with a Semantically Light Verbal Head

In this section we zoom in on verb-based phrasemes with *be*, *have*, *make*, *get* from English and *sein*, *haben*, *machen* from German as their syntactic head. Some of the phenomena and findings described in the section on VPhr are also applicable to vPhr. The speakers use them with equal ease in both their languages, in monolingual as well as in bilingual turns. Most vPhr occur in monolingual clauses:

(19) LK: Un' na sag i, *well*, i wollt'- **Mittag** mit dir **make** heut, un' i hab Zeit. [L2.518]

(20) KL: *And afterwards she **was sorry** that she didn't buy it.* [K9.118]

As with VPhr, insertions limited to a vPhr alone do not occur. However, in contrast to VPhr, anticipational switching is not frequent either. There are a few switches following abandoned calques of more idiomatic vPhr:

(21) KL: Na, aber die war nicht/ *She **was not** what we call here **my cup of tea**.* [K16:38]

In (21) the entire clause is repaired and also the phraseme is flagged as an item specific to American culture by the meta-comment *what we call here*. Although there are no obvious complete calques, there are also a few cases of attempted calquing, abandoned mid-sentence. In these cases, it is not the complete clause that is started over; rather, the repair is limited to the nominal component of the vPhr:

(22) TG: Is' die Elsie Eigel noch **in gut/ {in} good shape**, *Elsa?* [T28:200]

In (22) the English *be in good shape*¹⁷ is first translated but abandoned at the point where the speaker would have to assign a German gender-specific adjective ending to *gut*. The repair begins with the homophonous diamorph *in*. In the repaired version the verb still remains in the ML, whereas the complete NP is inserted in the EL.

Overt phraseme-internal language mixing in the form of partial calques is found with significant frequency (17% of 215 targets) and across speakers (6 out of 7 speakers):

- (23) LK: ...und deshalb si/ **bin** i ja {so} **close mit denen**.
[L2:454]
- (24) TG: Hätt's ihn grad' runterschlagen können, *because-* he **made himself** {so} **wichtig**,
you know!
[T29:278]

In (23) the verb of the English vPhr *be close with s.o.* is calqued, as is the preposition *with*. The semantically most salient component, the adjective *close* remains in its original English form. Before the English insertion we have the intensifier *so* as a homophonous diamorph. In (24) the verb of the German vPhr *sich wichtig machen* is calqued, as is the reflexive pronoun *sich*, whereas the adjective *wichtig* remains in its original language. Again, the homophonous diamorph *so* appears between the calqued components and the EL insertion.

All mixed verbal phrasemes appear to follow one consistent mixing pattern: the verb is calqued and the semantic core (mostly a noun or an adjective) appears in the original language of the phraseme. The mixing pattern is not dependent on the language of the phraseme, English or German. The partial calque in (25) will now be discussed in more detail in order to relate this recurrent pattern to the theoretical assumptions about code-switching and language processing outlined in sections 2 and 3.

- (25) LK: ...wie mer unser/ uns die Häuser anschaut ham, da wollte mer **sure mache**, dass mer e Haus kriege, wo mer e Eckbank neistelle kann.
[L1:42]

The underlying phraseme appears to be the English collocation *make sure*. A possible German translation equivalent is *sichergehen* (literally: *go sure*). Thus, a conflict on the level of lexical congruence could be expected with respect to the

¹⁷ Whether or not to include the copula verb in the phraseme is a complex issue which for reasons of space is not addressed in this paper (see Fix 1971: 72 and Keller 2014: 195–198).

semantically non-congruent verb rather than the congruent adjective. However, this is the reverse of what we actually see happening: the speaker chooses to calque the semantically incongruent verb *make* as German *machen* and to leave the semantically congruent adjective *sure* in its original form. This suggests that in partial calques superficial lexical equivalence is not the primary force at work. So, what exactly is motivating lexical selection during the production of mixed vPhr?

If there were no morpho-syntactic constraints governing the production of mixed utterances, one could imagine the following alternative renderings of the phraseme *make sure* in a bilingual clause with German as the ML (note: for ease of explication the dialect from the original is adapted to standard German):

- (a) Da wollten wir *make sure*, dass...
- (b) Da wollten wir *sure make*, dass...
- (c) Da wollten wir sicher *make*, dass...
- (d) Da wollten wir *sure make-en*, dass...
- (e) Da wollten wir *sure machen*, dass...

The MLF model provides arguments for why versions (a)–(c) should be dispreferred by a balanced bilingual. The complete EL insertion of the phraseme as in the hypothetical realization given in (a) violates the morpheme order principle, which states that word order must come from the ML. According to the rules of German word-order, the non-finite verb *make* should be preceded by the adjective *sure*. The EL insertion in (b) fixes this problem and follows ML word order. However, it still violates the system morpheme principle: the non-finite EL verb *make* doesn't carry the ML infinitive suffix *-en*.¹⁸ The same holds for the mixed option in (c), which calques only the semantically congruent adjective and retains the original but non-congruent light verb. Option (d) is in line with both

¹⁸ Myers-Scotton and Jake (2017: 10) refer to French infinitive suffixes as early SMS, based on the observation that in their data French infinitives appear to be inserted along with their French infinitive suffixes. This does not seem to be so for the inserted German and English infinitives in the corpus I used. There are instances where the German infinitive suffix is omitted, e.g. in “Na, let's fahr- \emptyset nach England, wegen deine Geschwister und die alle” (Keller 2014: 219). Conversely, when an English infinitive is adapted to German, an infinitive ending is added, e.g. “Zwei *languages* zusammen-*put-en!*” (Münch and Stolberg 2005: 74). Therefore, I am inclined to assume that the German infinitive suffix is a late outsider, which – as all other outsiders – conveys grammatical rather than semantic information.

MLF principles (word-order and outsiders from the ML) but includes word-internal language mixing.¹⁹ The combination in (e), i.e. the one actually produced by the speaker, is the one that optimally solves or integrates congruence issues on the morphosyntactic as well as on the semantic level: word-order and the infinitive marker on the light verb come from the ML (German), satisfying the MLF constraints. The adjective, which carries the semantically salient core of the phrase, is retained in its original language and inserted as an EL element into the clause. It does not carry any late outsider system morphemes and occurs in a position which does not violate ML syntax.

5 Discussion

The examples provided in section 4 show that the code-switching constraints proposed by Myers-Scotton in her MLF model also hold for phraseological units. So, the study of phrasemes in code-switching lends further support to the model. However, as the subject of this paper is the processing of phrasemes rather than the predictive power of a code-switching model, the crucial question is: what can the behaviour of phrasemes in code-switching tell us about the internal make-up and processing of phrasemes?

Mixed vPhr all show the same distribution of languages: The (light) verb is calqued and produced in the ML of the clause. The nominal component is inserted in its original language.²⁰ The order of the elements follows the syntactic requirements of the ML. This pattern integrates two challenges in an optimal way. First, retaining the nominal element carrying the semantic weight of the phrase in its original language serves as a cue for the language-specific multi-word sequence stored in the mental lexicon and helps to convey the intended propositional content to the hearer. Second, calquing of the semantically light verb allows integration of a phrase from language A into a clausal frame from language B in a manner that does not violate the grammatical rules of language B as

19 Word-internal mixing resulting from the addition of a language-B system morpheme to a language-A content morpheme is commonly observed among early bilinguals during simultaneous acquisition (Lanza 1997). The adult speakers who participated in our study seem to avoid word-internal mixing and use it mainly to achieve a comic effect.

20 This distribution of languages matches findings presented by Marian (2009: 172), who, without reference to phrasemes, writes that in her data verbs tend towards covert mixing (calquing), whereas nouns are more often overtly inserted. She attributes this to the stronger syntactic relations of verbs with other syntactic constituents in a clause.

proposed in the MLF model. The pattern is repeatedly produced by six out of the seven speakers and is thus not an idiosyncratic feature. The mixing pattern leads to the following hypothesis concerning the roles of semantics and syntax in the production of mixed phrasemes in classic code-switching:

The lexeme carrying the semantic core of an EL phraseme needs to be produced in its original language as a cue to the language-specific superlemma stored in the mental lexicon. Semantically lightweight elements can be calqued in order to satisfy ML morphosyntactic requirements.

This hypothesis is an empirically derived synthesis of Myers-Scotton's (2002: 240) assumption that the primary function of an EL is to supply content morphemes in mixed constituents and Backus's (2003: 92 and 123) claim that in mixed constituents based on conceptual units ML elements will have semantically basic meanings. It also supports the claim that "basic vocabulary" tends to be calqued whereas "specific vocabulary" will be inserted as an EL form (Backus and Dorleijn 2009: 92).

With respect to language processing, the question now is: How do we get from a language-specific superlemma entry to a mixed phonological realization? With no explicit reference to phrasemes, De Bot proposes the following – fairly vague – suggestion concerning language-sensitivity or -specificity of the levels of speech production:

[The conceptualizer] is probably partly language-specific and partly language-independent. Further it is hypothesized that there are different formulators for each language, while there is one lexicon where elements from different languages are stored together. The output of the formulator is sent to the articulator, which makes use of a large set of non-language specific speech motor plans.

(De Bot 1992: 1)

Also without reference to phrasemes, Myers-Scotton and Jake (1995: 987) suggest that at the conceptual level, language-specific lemmas are selected and sent to one of the language-specific formulators, which then adds the required predicate-argument structure, word-order and inflections.

Let us assume that the initial²¹ step is the same for lemmas and superlemmas: Guided by the intent of the speaker, a language-specific superlemma is selected

21 We can avoid the unresolved question of relative timing of the sub-processes if we adopt Jackendoff's Parallel Architecture model, according to which lexical/semantic and morphosyntactic processes run in parallel and influence each other (Jackendoff 1998: 39).

at the conceptual level. According to Sprenger et al. (2006: 167), each component of the superlemma is accessed individually from the mental lexicon, but through one common idiom node. This complex of individual but connected lemmas is sent to one of the language-specific formulators. As we want to explain the language distribution in mixed EL phrasemes, we are interested in the case where an EL phraseme is sent to an ML formulator.²² The formulator is supposed to project ML argument structure onto the elements it receives from the conceptualizer and to convert lemmas into lexemes and then word forms which can be sent on to the articulator. Under the current assumption that bilinguals might have two separate grammars but only one joint lexicon, it might not be all that surprising that a lemma from one language could be realized by a word-form from the other language, even if this lemma is part of a phraseme. However, the choice of surface language does not appear to be random. Judging from the mixing patterns we find in the data, the choice of word-forms at the formulator level appears to be subject to two constraints, one conceptual-semantic and one morphosyntactic in nature (Tab. 2).

Tab. 2: Assigning surface language to phraseme components

Conceptualizer	A superlemma representing a complete language-specific phraseme is selected from the mental lexicon.	
Formulator	Morphosyntactic constraint (Myers-Scotton's SMP) Phraseme components which host late outsider system morphemes activated only at the level of the formulator must come from the ML	Conceptual-semantic constraint Semantically salient phraseme components must come from the language with which the phraseme is affiliated in monolingual speech

The morphosyntactic constraint, Myers-Scotton's System Morpheme Principle, holds for classic code-switching in general. The semantic constraint is specifically formulated for phrasemes in code-switching. The two constraints, applicable in parallel rather than consecutively, offer a theoretical explanation for the

²² If an ML phraseme is sent to the ML formulator, we will get a monolingual utterance. If an EL phraseme is sent to the EL formulator, the result will be an EL island. An ML phraseme sent to an EL formulator would not be an option, as it renders the basic idea of having an ML completely mute.

recurring overt language mixing pattern in vPhr. Also, they can help to explain the apparent resistance of VPhr to internal mixing. When an EL vPhr is inserted into an ML clause, the semantically salient element (mostly a noun or an adjective) must be realized in the original language of the phraseme in order to satisfy the semantic constraint. The semantically light verb, which in the actual phonetic string carries a late outsider, can be adapted to ML morphosyntactic requirements by way of calquing (or “literal translation”) to satisfy the syntactic constraint. However, if a speaker wants to make use of a VPhr which is not part of the language he or she is currently using as the ML, the verb is semantically salient and thus needs to be realized in the original language of the phraseme. But it also carries a late outsider. Without word-internal mixing the verb cannot be adapted to ML morphosyntactic well-formedness conditions. Therefore, the only solution appears to be anticipational switching of the entire clause. The observed resistance of VPhr to internal mixing might suggest that more idiomatic phrasemes are processed holistically. I don’t think this is the case. Rather, the “grammar” of classic code-switching (outsiders have to be supplied by the ML) prevents overt mixing of idiomatic VPhr.²³ The few instances where the speakers start with the production of a mixed or calqued VPhr are quite instructive: The observation that these attempts are often abandoned and rephrased indicate that the speakers are aware of the “unlawfulness” of such translations of phrasemes. And it suggests that the bilingual language monitor checks for idiomaticity not necessarily before but rather while assembling a phraseological unit from individual language-specific lexemes. The abandoned calques show that more idiomatic elements of a phraseme can be calqued individually as well but that the result is rejected by the language monitor.

6 Conclusion

Of 451 verb-based phrasemes analysed for the present study, 20% show overt or covert language contact phenomena, either inside the phraseme (language mixing) or in its direct vicinity (language switching). The analysis has shown that phrasemes are subject to the same morphosyntactic constraints as free combinations of words proposed in the MLF model and the 4-M-Model (Myers-Scotton

23 If we look at cases of attrited (or attriting) phrasemes, which for reasons of space have been left out of the discussion, we can observe that in cases where an automatized production route is no longer available, VPhr also appear to be assembled from individual components (Keller 2014: 251–253).

2002; Myers-Scotton and Jake 2017): word order and late outsider system morphemes, i.e. the inflectional morphemes, which only serve a grammatical function, have to be supplied by the ML of the clause. Consequently, the verb, which in German and English carries a late outsider, has to be realized in the ML, at least in speaker groups where word-internal mixing is dispreferred. In contrast to the verb, nominal or adjectival complements can appear in the EL. This distribution of languages based on word-class is reflected in a recurring mixing pattern which is mostly found with inserted EL phrasemes containing a light verb but is also occasionally observable in more idiomatic phrasemes: the noun or adjective which carries the semantic core of the phraseme is realized in the EL, whereas the verb is calqued and produced in the ML of the clause.

The observation that phrasemes in code-switching can be composed of elements from different languages also supports the Superlemma Theory (Sprenger et al. 2006), which claims that the components of a phraseme are accessed individually, but through one common idiom node at the conceptual level. The findings suggest that at the level of the formulator, the production of phrasemes is determined not only by morphosyntactic code-switching constraints but also by phraseme-specific semantic considerations: The semantic core of a phraseme must be produced in the original language of the phraseme, while functional elements, including light verbs, can also be realized in a different language.

A promising next step to test the theoretical modelling of language distribution or language assignment to surface lexemes in mixed phrasemes proposed in this paper would be an extended analysis of more utterances with semantically light verbs as their syntactic head. But of course, there is a lot more to explore in the context of phrasemes and code-switching, for example the status of the copula verb (included in or excluded from the phraseme) or de-automatisation as observable in attrition of phrasemes. Also, the influence of internal and external valency or of semantic compositionality could be analysed in more detail in order to further enhance our understanding of storage and processing of phrasemes.

References

- Allerton, David J. (2001): *Stretched verb constructions in English*. Oxford: Routledge.
- Ayto, John (2010): *Oxford dictionary of English idioms*. Oxford: Oxford University Press.
- Backus, Ad (2003): Units in Codeswitching: evidence for multimorphemic elements in the lexicon. *Linguistics* 41 (1), 83–132.
- Backus, Ad & Margreet Dorleijn (2009): Loan translations versus code-switching. In Barbara Bullock & Almeida Jaqueline Toribio (eds.), *The Cambridge Handbook of Linguistic Code-switching*, 75–93. Cambridge: Cambridge University Press.

- Bialystok, Ellen & Fergus I. M. Craik (2010): Cognitive and linguistic processing in the bilingual mind. *Current Directions in Psychological Science* 19, 19–23.
- Burger, Harald (2015): *Phraseologie: Eine Einführung am Beispiel des Deutschen*. 5., neu bearbeitete Auflage. Berlin: Erich Schmidt.
- Butt, Miriam (2003): The Light Verb Jungle. *Harvard Working Papers in Linguistics* 9, 1–49.
- Butt, Miriam (2010): The Light Verb Jungle: Still Hacking Away. In Mengistu Amberber, Brett Baker & Mark Harvey (eds.), *Complex Predicates: Crosslinguistic Perspectives on Event Structure*, 48–78. Cambridge: Cambridge University Press.
- Clyne, Michael (1987): Constraints on code-switching: how universal are they? *Linguistics* 25 (4), 739–764.
- Cutting, J. Cooper & J. Kathryn Bock (1997): That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiomatic blends. *Memory & Cognition* 25, 57–71.
- De Bot, Kees (1992): A bilingual production model: Levelt's 'Speaking' model adapted. *Applied Linguistics* 13, 1–24.
- Ehinger, Anja (2003): *Simultaneous activation of languages: Investigating co-production and hesitation phenomena in bilingual speech*. Tübingen, Univ., MA Thesis.
- Fix, Ulla (1971): *Das Verhältnis von Syntax und Semantik im Wortgruppenlexem*. Leipzig: Diss., Karl-Marx-Universität.
- González-Vilbazo, Kay & Luis Lopez (2011): Some properties of light verbs in code-switching. *Lingua* 121 (5), 832–850.
- Häcki-Buhofer, Annelies (2007): Phraseme im Erstspracherwerb. In Harald Burger, Dmitrij Dobrovolskij, Peter Kühn & Neal R. Norrick (eds.), *Phraseology: An International Handbook of Contemporary Research*, 854–869. Berlin, New York: De Gruyter.
- Havrila, Marek (2009): Idioms: Production, Storage and Comprehension. *philologica.net*. Retrieved from <http://philologica.net/studia/20091107224500.htm>, accessed September 28, 2019.
- Howarth, Peter (1998): Phraseology and second language proficiency. *Applied Linguistics* 19 (1), 24–44.
- Jackendoff, Ray (2002): *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.
- Keller, Mareike (2014): *Phraseme im bilingualen Diskurs: All of a sudden geht mir ein Licht auf*. Frankfurt am Main: Peter Lang.
- Kuiper, Koenraad, Marie-Elaine van Egmond, Gerard Kempen & Simone Sprenger (2007): Slipping on superlemmas: multi-word lexical items in speech production. *The Mental Lexicon* 2 (3), 313–357.
- Langlotz, Andreas (2006): *Idiomatic creativity: A cognitive-linguistic model of idiom-representation in English*. Amsterdam, Philadelphia: John Benjamins.
- Lanza, Elizabeth (1997): *Language Mixing in Infant Bilingualism: A Sociolinguistic Perspective*. Oxford: Clarendon.
- Levelt, Willem J.M. (1989): *Speaking: From Intention to Articulation*. Cambridge, Mass.: Massachusetts Institute of Technology Press.
- Levelt, Willem J. M. & Antje Meyer (2000): Word for word: Multiple lexical access in speech production. *European Journal of Cognitive Psychology* 12 (4), 433–452.
- Levelt, Willem J. M., Ardi Roelofs & Antje Meyer (1999): A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22 (1), 1–75.

- Marian, Viorica (2009): Language interaction as a window into bilingual cognitive architecture. In Ludmila Isurin, Don Winford & Kees De Bot (eds.), *Multidisciplinary Approaches to Code Switching*, 161–185. Amsterdam, Philadelphia: John Benjamins.
- Mindt, Dieter (2002): What is a grammatical rule? In Leiv Egil Breivik & Angela Hasselgren (eds.), *From the COLT's mouth ... and others: Language corpora studies - in honour of Anna-Brita Stenström*, 197–212. Amsterdam, New York: Rodopi.
- Moon, Rosamund (1998): *Fixed Expressions and Idioms in English*. Oxford: Oxford University Press.
- Münch, Alexandra & Doris Stolberg (2005): “Zwei languages zusammenputten”: Bilingual ways of expressing bicultural identities. In Bent Preisler, Anne H. Fabricius, Hartmut Haberland, Susanne Kjærbeck & Karen Risager (eds.), *The Consequences of Mobility*, 71–79. Roskilde: Roskilde University: Department of Language and Culture.
- Myers-Scotton, Carol (1997): *Duelling languages: Grammatical structure in codeswitching*. Oxford: Clarendon.
- Myers-Scotton, Carol (2002): *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Myers-Scotton, Carol (2005): Supporting a Differential Access Hypothesis: Code switching and other contact data. In Judith F. Kroll & Annete M. B. deGroot (eds.), *Handbook of bilingualism. Psycholinguistics approaches*, 326–358. New York: Oxford University Press.
- Myers-Scotton, Carol (2006): *Multiple Voices: an introduction to bilingualism*. Oxford: Wiley-Blackwell.
- Myers-Scotton, Carol & Janice Jake (1995): Matching lemmas in a bilingual language competence and production model: evidence from intrasentential code switching. *Linguistics* 33, 981–1024.
- Myers-Scotton, Carol & Janice Jake (2000): Four types of morpheme: evidence from aphasia, code switching, and second-language acquisition. *Linguistics* 38, 1053–1100.
- Myers-Scotton, Carol & Janice Jake (2009): A universal model of code-switching and bilingual processing and production. In Barbara Bullock & Almeida Jaqueline Toribio (eds.), *The Cambridge Handbook of Linguistic Code-switching*, 336–357. Cambridge: Cambridge University Press.
- Myers-Scotton, Carol & Janice Jake (2015): Cross-language asymmetries in code-switching patterns: Implications for bilingual language production. In John W. Schwieter (ed.), *The Cambridge Handbook of Bilingual Processing*, 416–458. Cambridge: Cambridge University Press.
- Myers-Scotton, Carol & Janice Jake (2017): Revisiting the 4-M model: Codeswitching and morpheme election at the abstract level. *International Journal of Bilingualism* 21 (3), 340–366.
- Namba, Kazuhiko (2012): *English-Japanese Code-Switching and Formulaic Language: A Structural Approach to Bilingual Children's Interactions*. Saarbrücken: Lambert Academic Publishing.
- Paradis, Michael (2004): *A neurolinguistic theory of bilingualism*. Amsterdam, Philadelphia: John Benjamins.
- Pottelberge, Jeroen van (2007): Funktionsverbgefüge und verwandte Erscheinungen. In Harald Burger, Dmitrij Dobrovolskij, Peter Kühn & Neal R. Norrick (eds.), *Phraseology: An International Handbook of Contemporary Research*, 436–444. Berlin, New York: De Gruyter.
- Röhrich, Lutz (2010): *Lexikon der sprichwörtlichen Redensarten*. Freiburg: Herder.
- Scholze-Stubenrecht, Werner (2013): *Duden – Redewendungen*. Berlin: Dudenverlag.

- Sinclair, John (2011): *Collins COBUILD idioms dictionary*. Glasgow: Harper Collins.
- Spears, Richard (2005): *McGraw-Hill's dictionary of American idioms and phrasal verbs*. New York: McGraw-Hill.
- Sprenger, Simone A., Willem J. M. Levelt & Gerard Kempen (2006): Lexical access during the production of idiomatic phrases. *Journal of Memory and Language* 54, 161–184.
- Tracy, Rosemarie & Elsa Lattey (2010): 'It wasn't easy but irgendwie äh da hat sich's rentiert, net?': a linguistic profile. In Michaela Albl-Mikasa, Sabine Braun & Sylvia Kalina (eds.), *Dimensionen der Zweitsprachenforschung/Dimensions of Second Language Research. Festschrift für Kurt Kohn*, 53–73. Tübingen: Narr.
- Wei, Longxing (2009): Code-switching and the bilingual mental lexicon. In Barbara Bullock & Almeida Jaqueline Toribio (eds.), *The Cambridge Handbook of Linguistic Code-switching*, 270–288. Cambridge: Cambridge University Press.
- Winhart, Heike (2005): *Funktionsverbgefüge im Deutschen: zur Verbindung von Verben und Nominalisierungen*. Tübingen, Univ., Diss.
- Wray, Alison (2012): What do we (think we) know about formulaic language? An evaluation of the current state of play. *Annual Review of Applied Linguistics* 32 (1), 231–254.
- Wray, Alison & Kazuhiko Namba (2003): Use of formulaic language by a Japanese-English bilingual child: a practical approach to data analysis. *Japan Journal for Multilingualism and Multiculturalism* 9 (1), 24–51.
- Wray, Alison & Michael R. Perkins (2000): The functions of formulaic language: an integrated model. *Language & Communication* 20, 1–28.
- Wörterbuch für Englisch-Deutsch und andere Sprachen* (2002–2018). Retrieved from <https://www.dict.cc>, accessed September 28, 2019.
- Wörterbuch für Redensarten, Redewendungen, idiomatische Ausdrücke, Sprichwörter, Umgangssprache* (2001–2018). Retrieved from <https://www.redensarten-index.de>, accessed September 28, 2019.