Multiword Units in Multilingual Speakers

University of Tübingen, 7th-8th June 2021

To participate, please, send a short message to marlene.wessel@student.uni-tuebingen.de and we will send you the zoom-link.

More about the workshop here:

# Workshop on Multiword Units in Multilingual Speakers, Schedule

## Monday, 7th June 2021

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:30-12:45</td>
<td>Opening</td>
</tr>
<tr>
<td>12:45-13:45</td>
<td><strong>Albert Backus</strong>&lt;br&gt;Not always written in stone: A usage-based account of cross-linguistic influence on word combination</td>
</tr>
<tr>
<td>13:45-14:00</td>
<td>Break</td>
</tr>
<tr>
<td>14:00-14:45</td>
<td><strong>Mikhail Kopotev</strong>&lt;br&gt;Collocations and beyond: constructional generalization over Russian collocations.</td>
</tr>
<tr>
<td>14:45-15:30</td>
<td><strong>Gaëtanelle Gilquin</strong>&lt;br&gt;Phrasal verbs in L2 English: EFL and ESL compared</td>
</tr>
<tr>
<td>15:30-15:45</td>
<td>Break</td>
</tr>
<tr>
<td>15:45-16:30</td>
<td><strong>Albert Backus, Marie Barking &amp; Maria Mos</strong>&lt;br&gt;Similarity in language transfer</td>
</tr>
<tr>
<td>16:30-17:15</td>
<td><strong>Johanna Wolf</strong>&lt;br&gt;In chunks we trust...why multilingual speakers need more than one feature for successful processing</td>
</tr>
<tr>
<td>17:15-17:30</td>
<td>Break</td>
</tr>
<tr>
<td>17:30-18:15</td>
<td><strong>Sandra Deshors</strong>&lt;br&gt;Phrasal verbs in English as a Lingua Franca: a case for corpus-based multifactorial analysis</td>
</tr>
<tr>
<td>18:15-18:45</td>
<td>Open Table + Reception</td>
</tr>
</tbody>
</table>

## Tuesday, 8th June 2021

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:30-12:45</td>
<td>Opening</td>
</tr>
<tr>
<td>12:45-13:45</td>
<td><strong>Jeanine Treffers-Daller</strong>&lt;br&gt;Single words are the odd ones out: interdisciplinary perspectives on multiword units</td>
</tr>
<tr>
<td>13:45-14:00</td>
<td>Break</td>
</tr>
<tr>
<td>14:00-14:45</td>
<td><strong>Ruth Kessler &amp; Tatiana Perevozchikova</strong>&lt;br&gt;What is behind the transfer? A staring look at multi-word units of Russian heritage speakers in Germany</td>
</tr>
<tr>
<td>14:45-15:30</td>
<td><strong>Inga Hennecke &amp; Evelyn Wiesinger</strong>&lt;br&gt;Language contact phenomena in multiword units</td>
</tr>
<tr>
<td>15:30-15:45</td>
<td>Break</td>
</tr>
<tr>
<td>15:45-16:30</td>
<td><strong>Rundi (Wendy) Guo</strong>&lt;br&gt;Processing collocations in first and second languages</td>
</tr>
<tr>
<td>16:30-17:30</td>
<td><strong>Eva Smolka</strong>&lt;br&gt;Psycho- and neurolinguistic evidence for the effects of (native) language structure on the processing of multiword units</td>
</tr>
<tr>
<td>17:30-18:15</td>
<td>Closing</td>
</tr>
</tbody>
</table>
**Not always written in stone: A usage-based account of cross-linguistic influence on word combination**

As part of a growing research tradition in contact linguistics, I have been involved in several projects that apply a usage-based approach to the investigation and explanation of language contact phenomena. In this talk, I will first summarize some of that recent work, and relate it to other traditions and trends in the study of bilingualism. I will defend the position that much of this work either uses a usage-based perspective or is compatible with it. Finally, interpreting this as suggesting there is a need for the integration of disciplinary perspectives, I will sketch what such integration could look like, and to what degree this is relevant for research questions about multiword combinations. Work on language contact has taken place in various sub-disciplines and often focuses on separate phenomena that nevertheless all occur together in the everyday lived reality and in the minds of bilingual speakers. These phenomena include language mixing (or codeswitching), contact-induced structural change, the processing of bilingual speech, and the social factors that determine language choice in multilingual settings. I will illustrate these various strands, where possible, with work on Turkish spoken as an immigrant language in the Netherlands and other parts of Western Europe. I will then argue that adopting a usage-based perspective entails the conclusion that these phenomena are not as separable as it seems, and that better, or more encompassing, explanations of linguistic knowledge and of language change could be forthcoming if we manage to integrate these research traditions more. The key, I will argue, is the closer integration of sociality and cognition as dimensions that both need to be taken into account jointly. While both cognitive sciences and social sciences, and their linguistic manifestations psycholinguistics and sociolinguistics, have, of course, valuable things to say about language, the direct association between the way language is used, as determined by the needs of human sociality, and the way it is processed, as determined by the universals of human cognition, makes it necessary to address their interaction as well. In this talk I will explore this integration to account for empirical findings on Immigrant Turkish and Heritage Languages in general. What I hope to contribute to is the contours of a model that explains why languages change, with specific attention to the way this affects the fate of multiword combinations.

**Collocations and beyond: constructional generalization over Russian collocations**

A speech act is produced linearly: after saying A, we may be more likely to say B rather than Y. Syntagmatically, many word combinations can be identified as not idiomatic in a narrow sense, but rather as sharing a common property—a stable co-occurrence. At one extreme, the multiword units are idioms, where no generalization is possible at all, for example, lo and behold. At another dimension, the fixed features are entire morphological, for example, [I]’m
+ V-ing; these cases are colligations. Yet another multiword units appear to be stable on a more abstract level of generalization. They are called collostructions (after Gries & Stefanowitsch 2004), where lexical items are interchangeable but belong to a certain semantic and grammatical class, for example, sleight of [hand/mouth/mind]. The full range of such associations includes grammatical and lexical features—without drawing ad hoc borders between them—that can be determined by context and based on corpus. In my presentation, I show that prior to being coined into more fixed items, co-occurrences exist as a set of almost-ready-to-use prefabricated units.

References

Gaëtanelle Gilquin
(University of Louvain)

Phrasal verbs in L2 English: EFL and ESL compared

Although phrasal verbs in English are widely recognized as forming a single unit, as evidenced among others by the fact that they often have a one-word equivalent (e.g. ‘give up’ = ‘abandon’, see Siyanova and Schmitt 2007), they are made up of two elements, a verb and a particle, which are sometimes separated from each other (e.g. ‘He filled it up’). This special feature makes the phrasal verb a particularly interesting multiword unit to investigate, especially among non-native speakers of English, since phrasal verbs are said to be “one of the most notoriously challenging aspects of English language instruction” (Gardner and Davies 2007: 339).

This presentation will centre around phrasal verbs with ‘up’, which will be studied in two types of non-native English, namely English as a foreign language (EFL) and English as a second language (ESL). The approach will be corpus-based, relying on the phrasal verbs extracted from written and spoken corpora of EFL (the International Corpus of Learner English and the Louvain International Database of Spoken English Interlanguage) and ESL (different sections of the International Corpus of English), as well as comparable corpora of English as a native language (ENL), namely the Louvain Corpus of Native English Essays, the Louvain Corpus of Native English Conversation and the British component of the International Corpus of English.

The frequency of phrasal verbs will be compared across the EFL, ESL and ENL corpora, with special focus on the possible differences between speech and writing, since phrasal verbs are often thought to be more typical of speech than of writing (Marks 2005). Preferences in terms of construction will also be compared, and in particular the distinction between V-Obj-Prt and V-Prt-Obj (see also Gilquin 2015). These comparisons should provide insights into the use and possibly processing of phrasal verbs by two types of non-native speakers who both use English as an additional language, but who differ in the way they have acquired it.
References


Albert Backus, Marie Barking & Maria Mos
(Tilburg University)

**Similarity in Language Transfer**

Bilingual speakers tend to frequently experience language transfer when their languages are typologically closely related, which suggests that similarity between languages is likely to play an important role in the transfer process. In this paper, we explore how three different types of similarity affect transfer of light verb constructions (LVCs, such as take a walk, set an alarm) from Dutch to German, namely (a) similarity to existing constructions, (b) surface similarity based on whether the noun in the LVC is a cognate in Dutch and German, and (c) similarity in the light verb’s usage contexts. The results suggest that more similarity does indeed facilitate transfer, in that speakers experienced more transfer (a) when the LVC was similar to an existing German LVC, (b) when the nouns in Dutch and German were cognates, and (c) when the verbs were used in similar contexts across languages. This is reflected both in the speakers’ language use as well as in their acceptability ratings of transferred constructions. Moreover, the results suggest that having a construction that is highly similar in Dutch and German can even lead to less acceptance of existing alternatives that happen to be less similar. Overall, the results thus show that speakers both add constructions that did not previously exist in German and drop constructions that did previously exist, based on similarity between constructions in Dutch and German, ultimately resulting in more convergence across their languages.

Johanna Wolf
(Paris-Lodron-Universität Salzburg)

**In chunks we trust...why multilingual speakers need more than one feature for successful processing of morphosyntactic information**

Gender assignment in nominal, verbal or adjective phrases is one of the most difficult challenges for foreign language learners. Compared to native speakers or even bilingual children, learners do not seem to be able to process and filter the morphosyntactic rules from the input. They are also often unable to apply the rules in production without committing errors in assigning gender. On the basis of a corpus of authentic learner productions (Romance languages as L3, German as L1), the article investigates where learners have particular problems with assignment and to what extent the errors can be explained by their multilingual
background. In addition, a comparison with data from first language acquisition is used to try to find an explanatory model for the specific problem of gender assignment. On the one hand, the model of feature-driven acquisition will be used in order to explain the difficulty of gender assignment for foreign language learners, and on the other hand, the processing of multi-word units which seems to facilitate the filtering of morphosyntactic rules out of the input.

Sandra Deshors
(Michigan State University)

Phrasal verbs in English as a Lingua Franca: a case for corpus-based multifactorial analysis

This multifactorial corpus-based study on particle placement in English as a Lingua Franca (ELF) focuses on verb-object-particle (VOP) and verb-particle-object (VPO) alternations and identifies the usage patterns that characterize these alternations in ELF multi-participant interactions. Verb-particle constructions (VPC) have been approached differently across Englishes. In native English, studies have shown how all linguistic levels and cognitive processes of language acquisition influence the uses of VOP/VPO constructions. Similarly, in learner English (EFL), the uses of the two constructions have been shown to deviate from those in native English because of their syntactic and semantic complexity, processing demands, input effects, and the typology of speakers’ L1. Methodologically, quantitative studies are yet to handle the complexity of analyzing VPC based on multiple predictors, an important limitation since linguistic and extra-linguistic factors affect, independently, the structure of ELF. In this context, the present study (i) makes a stronger connection between theory and method in ELF research, (ii) identifies the combinations of linguistic and extra-linguistic/conversational factors that influence constructional choices and (iii) aligns current ELF and EFL methodological approaches to VPC. Specifically, I investigate 585 VOP/VPO constructions from the Vienna-Oxford International Corpus of English (VOICE) annotated against individual phrasal verbs, length, determiner, complexity and type of direct object, concreteness of the referent of direct object, semantic use of the verb phrase, sex and age range of the speaker, his/her L1 and conversational role. Statistically, the approach involves a random forests analysis including interactions between linguistic, socio-demographic and conversational factors. I obtained models/trees with significant classification accuracies and C-scores and identified strongest predictors of alternations such as Type, Complex, Det and Length. These factors also participate in several interactions that drive VOP/VPO alternations. Overall, these interactions underscore the importance of integrating extra-linguistic factors to large-scale quantitative linguistic analyses of the structure of ELF. Syntactically, the results indicate that as in native English and EFL, the type of direct object, the type of determiner and the degree of complexity of the direct object all play a significant part in the VOP/VPO alternation, suggesting that VOICE speakers are constrained by the same factors as native/EFL speakers. However, the influence of these factors emerges as a function of other co-occurring extra-linguistic factors such as the speakers’s native language and conversational role. These results revive the question of what
distinguishes ELF from EFL as linguistic systems and open the door for a discussion on bilingual processing in ELF.

Jeanine Treffers-Daller
(University of Reading, UK)

Single words are the odd ones out: interdisciplinary perspectives on multiword units

Since the seminal publication on formulaic sequences by Wray (2000), a wealth of studies on first and second language learners as well as multilinguals has appeared which reveal the importance of multiword units of different kinds in everyday communication. In this paper I will summarize what we know about the use of multiword units by multilinguals, as evidenced in corpus data (e.g. Doğruöz & Backus, 2009; Treffers-Daller, 2012; Treffers-Daller et al, 2016), and review recent evidence regarding the processing of these units, both among monolinguals (e.g. Kessler, Weber & Friedrich, 2020) and multilinguals (e.g. Siyanova, Conklin & Schmitt, 2011). Key to the study of multiword units, in my view, is the integration of knowledge from corpus linguistic and experimental approaches, which each provide essential, but complementary information. However, obtaining corpus linguistic evidence about bilinguals’ use of multiword units can be challenging, as bilingual corpora tend to be much smaller than monolingual corpora, which means that frequency patterns may be less informative in the former than in the latter. Experimental evidence is therefore crucially important in studies of multiword units in multilinguals. I will also look at implications of the findings for models of bilingual processing, such as the Revised Hierarchical Model (Kroll & Stewart, 1994) and the Multilink model (Dijkstra & Rekké, 2010), which have so far mainly been informed by studies of processing of single words (often presented in isolation). I will finish by sketching the key questions that researchers interested in this area may wish to address in future work.

References


**Ruth Keßler & Tatiana Perevozchikova**

(University of Tübingen)

**What is behind the transfer? A staring look at multi-word units of Russian heritage speakers in Germany**

Heritage speakers, i.e., bilinguals who have limited access to their native language at the expense of the majority language, have been shown to typically produce many hybrid constructions that are motivated by both transfer from their dominant language and by a similar construction in their heritage language (Kopotev et al. 2020, Rakhilina et al. 2016). However, it is unclear whether they do so because they are unaware of the underlying construction in a monolingual variety of the heritage language or because they just cannot access it in the online speech production.

To test this question, we collected acceptability judgments for adverbial and prepositional constructions produced by heritage speakers (“non-standard”) and their equivalents in the standard variety of Russian (“standard”). Non-standard constructions (e.g., *po čuvstvu* ‘by feeling’) have been extracted from heritage speaker corpus data. Standard constructions (e.g., *po intuicii* ‘by intuition’) were the equivalents of heritage speaker constructions in monolingual Russian. The experiment was conducted as a browser-based survey with acceptability ratings on a 5-point Likert scale and is supposed to throw some light on the mechanisms behind the use of multi-word units by heritage speakers as a particular type of multilinguals.

**References**


Language contact phenomena in multiword units: Code-switching and calquing

Code-switching and calquing are two frequent language contact phenomena in bilingual speech. Code-switching traditionally refers to “[…] the juxtaposition within the same speech exchange of passages of speech belonging to two different grammatical systems or subsystems” (Gumpertz 1982: 59). Muysken (2000) distinguishes different types of code-switching by insertion, alternation and congruent lexicalization. Calquing traditionally includes loan translations/formations at the lexical, phraseological or structural level that are coined with material from one language according to a model in the other language (cf. Backus & Dorleijn 2008).

While both phenomena have been discussed extensively in research on language contact in the last decades, only very few studies explicitly investigate code-switching and calquing within multi-word units. Multi-word units are generally assumed to be more or less prefabricated or chunked units that can be stored and accessed by the speakers as a whole (cf. Christiansen & Arnon 2017). Whereas they have been found to play a central role in language acquisition and use, their role in bilingual speech is still far from clear.

In our contribution, we aim at bringing together recent usage-based psycholinguistic and Construction Grammar theories and modeling (e.g. Arnon & Snider 2010; Masini 2009) with cognitively oriented approaches to language contact and bilingualism (e.g. Backus 2015; Filipović & Hawkins 2018; Verschik 2019) in order to develop a more differentiated account of code-switching and calquing in multi-word units. Against this theoretical backdrop, we will discuss data from 1st, 2nd and 3rd-generation bilingual Spanish-English speakers in Texas and Miami with a special focus on code-switching and calquing within nominal (e.g. N Prep N), verbal (e.g. V (Prep) Adv) and adverbia (e.g. Prep NP) constructions. The analysis highlights the importance of partially lexically filled schemas as an intermediate level between lexical chunks and fully abstract constructions as well as gradual differences between more lexical and more grammatical multi-word units.

References


Processing collocations in first and second languages

Collocations are two words that have a tendency to co-occur within a few words’ span, e.g., “drink coffee” and “dark chocolate” in English. Growing empirical evidence suggests that both native (L1) speakers and advanced second language (L2) learners process collocations faster than unconnected word pairs, and that they are sensitive to the frequency distributions of linguistic units beyond individual words. Here we investigate the processing of L1 and L2 collocations with three questions in mind: 1) Are collocations processed like “big words” by native speakers of various languages, i.e., are native speakers sensitive to the frequency distribution of the collocations? 2) Are the collocation facilitatory effects applicable in L2 speakers, and if so, are the effects modulated by language proficiency and dominance, i.e., are second language speakers sensitive to the frequency distribution of the collocations and does the level of sensitivity change as language experience increases? And 3) What are the underlying factors that impact the ease of processing of collocations in speakers varying in their language experience? In dual Lexical Decision Tasks (LDT), 42 English monolinguals and 163 Chinese(L1)-English(L2) sequential bilinguals read 432 pairs of letter strings and judged if both strings in a pair were correctly-spelling words in English. Some of the word pairs are collocations in English but not Chinese (EN incongruent, e.g., honest mistake), some are collocations in Chinese (if translated verbatim) but not English (CH incongruent, e.g., black chocolate), some are collocations in both languages (congruent, e.g., direct flight), and some are unconnected in both languages as the baseline control (e.g., bright hand). Additionally, all...
the bilinguals also completed a Chinese version of the same task where all the items are presented in simplified Chinese. Preliminary results showed facilitated processing of English collocations vs. English non-collocations for English L2 speakers and facilitated processing of Chinese collocation vs. Chinese non-collocations for Chinese L1 speakers. There was also facilitated processing of English collocations for Chinese-English bilingual speakers, where the effect size increases as English experience increases. Additionally, general linear regression models including factors such as word length, word frequency, and collocation frequency in both languages revealed that highly advanced L2 speakers process collocations in a similar way as L1 speakers do in the sense that performance is more impacted by collocation frequency and less impacted by word-level factors in both L1 speakers and highly proficient L2 speakers, while the pattern is reversed for low proficiency L2 speakers. This suggests that as language proficiency grows, the unit for language processing becomes larger – from single words to collocations and potentially other multiword units.

Eva Smolka  
(University of Innsbruck, University of Vienna)

Psycho- and neurolinguistic evidence for the effects of (native) language structure on the processing of multiword units

Humans possess the ability of coining meaning in an almost unlimited number of different ways. The focus of this talk resides on the combinatorial ability to coining multiword units in which the meaning of the whole unit dramatically differs from the meaning of the single constituents. For example, the meaning of mouthwash remains related to the meanings of both mouth and wash and is thus considered to be semantically compositional (i.e. transparent). By contrast, the meaning of hogwash cannot be easily constructed from the separate meanings of hog+wash and is thus considered to be rather noncompositional (i.e. opaque). Also the meaning “to strive for something impossible” only remotely relates to the meaning of the constituents reach+stars in the multiword unit ‘to reach for the stars’. Psycholinguistic and neurolinguistic research thus remain puzzled by the question how multiword units are stored and processed in lexical memory. In particular: Do we access the meaning of the single constituents like wash in the course of processing mouthwash or hogwash? And, does the specific language experience of a native speaker affect how multiword units are stored and processed in lexical memory?

This talk will present the findings of a series of behavioral, EEG and fMRI studies on the processing of different types of multiword units, including complex verbs, compounds, and idioms. Overall, the findings indicate that multiword units in German are processed by referring to the single constituents, indicating that lexical processing and memory structures in German differ from those in other Indo-European languages. These findings are relevant to a key debate about general versus language specific aspects of cognition.