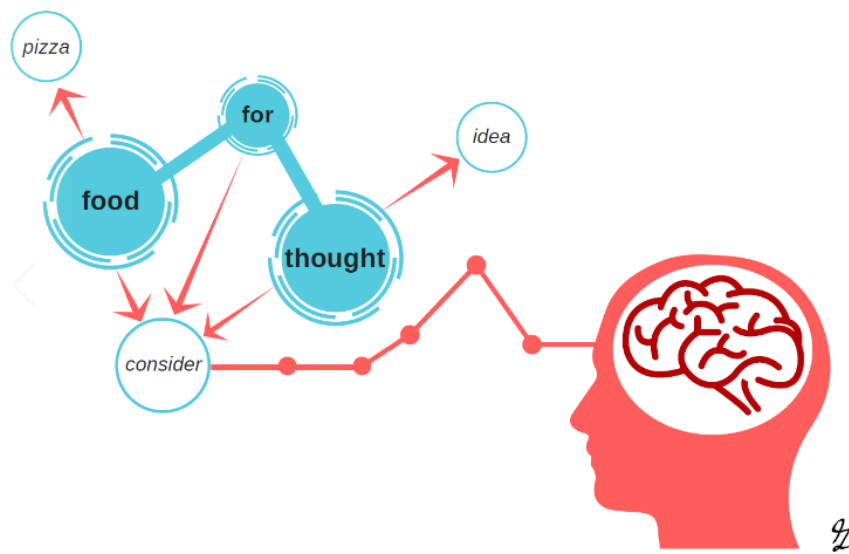




Workshop on Psycholinguistic and Computational Perspectives on Non-Compositional Meaning in Phrases

Hosted by the SFB 833 (Projects A3 and B9) and associates at the University of Tübingen



29TH-30TH NOVEMBER 2018, TÜBINGEN (FÜRSTENZIMMER, HOHENTÜBINGEN)

ORGANIZING COMMITTEE

Sara D. Beck, B9

Patricia Fischer, A3

Ruth Keßler, B9

Yana Strakatova, associated faculty

Workshop on Psycholinguistic and Computational Perspectives on Non-Compositional Meaning in Phrases

Both computational and psycholinguistic approaches attempt to help us better model the bridge between form and meaning, and this workshop provides a platform for resolving interdisciplinary differences and encouraging cross-talk between junior and senior researchers. Particularly, we aim to ask how psycholinguistic modeling of non-compositional meaning can inform computational linguistic models and vice versa.

PROJECTS INVOLVED:

SFB 833 (Projects A3 and B9)

PROGRAM CHAIRS:

Daniël de Kok, A3
Claudia Friedrich, B9
Erhard Hinrichs, A3
Andrea Weber, B9

Table of Contents

▪ Program	1
▪ Keynote Abstracts	3
▪ Participant Abstracts	10
▪ Oral Presentation Abstracts	10
▪ Poster Presentation Abstracts	28
▪ Workshop Dinner Location	51

Program

Thursday, November 29th, 2018: Fürstenzimmer, Castle Hohentübingen

8:30-9:00 Registration

9:00-9:15 Welcome from the organizers

9:15-10:15 **Keynote: Cristina Cacciari (University of Modena)**
Idioms “Electrified”: What ERP Revealed about the Comprehension of Multi-word Expressions

10:15-10:30 Coffee break

Oral Presentations

10:30-11:00 **Eva Smolka and Carsten Eulitz (University of Konstanz)**
Can You Reach for the Planets or Grasp at the Stars? – Modified Noun, Verb, or Preposition
Constituents in Idiom Processing

11:00-11:30 **Amélie la Roi, Simone Sprenger, and Petra Hendriks (University of Groningen)**
Language in the Aging Brain: Using ERPS to Study Idiom Processing across the Life Span

11:30-13:30 Lunch break

13:30-14:30 **Keynote: Aurelie Herbelot (University of Trento)**
Colorless Green Ideas in Formal Distributional Semantics

Oral Presentations

14:30-15:00 **Petr Rosyaykin (Lomonosov Moscow State University)**
Using Context Information for Russian MWE Extraction

15:00-15:30 **Patricia Fischer (University of Tübingen)**
Improvements on Subject-Object Disambiguation in Parsing

15:30-16:00 Coffee break and stretching

16:00-17:00 **Keynote: Inbal Arnon (Hebrew University of Jerusalem)**
Parallels in the Processing of Compositional and Non-Compositional Phrases: Implications for
the Lexicon

19:00-21:00 **Workshop dinner: Restaurant Maugeneschtle**
Burgsteige 18,
72074 Tübingen
<https://www.hotelamschloss.de/restaurant-maugeneschtle/>

****Bolded** authors in attendance.**

Friday, November 30th, 2018: Fürstenzimmer & Rittersaal

Fürstenzimmer

8:30-9:00 Registration

9:00-10:00 Keynote: Johan Bos (University of Groningen)
Multilingual Meaning Banking and Compositionality

10:00-10:30 Coffee break

Oral Presentations

10:30-11:00 Irene Fioravanti (Università per Stranieri di Siena), Marco Silvio Giuseppe Senaldi (Scuola Normale Superiore di Pisa), Alessandro Lenci (Università di Pisa) and Anna Siyanova-Chanturi (Victoria University of Wellington)
Lexical Fixedness and Compositionality in Native and Non-native Speakers' Intuitions about Italian Word Combinations: Psycholinguistic and Computational Perspectives

11:00-11:30 Sara Beck and Ruth Keßler (University of Tübingen)
L1 and L2 Learners Keep Their Eyes on the Prize: Eye-tracking Evidence during Idiom Recognition

11:30-12:30 Keynote: Gareth Carrol (University of Birmingham)
Into the Unknown: What Eye-tracking Data Can Tell Us about How Language Users Process and Resolve Unfamiliar Idioms

12:30-12:40 Poster slam (all posters)

Rittersaal

12:45-14:00 Poster lunch
Refreshments and poster presentations

Fürstenzimmer

14:15-15:15 Keynote: Melanie Bell (Anglia Ruskin University)
The Role of Constituent Disambiguation in the Interpretation of Novel Noun-noun Constructions

15:15-15:30 Concluding remarks, student awards, coffee, and discussion

Keynote Abstracts

1. Cristina Cacciari (University of Modena) 4
2. Aurelie Herbelot (University of Groningen) 5
3. Inbal Arnon (Hebrew University of Jerusalem) 6
4. Johan Bos (University of Groningen) 7
5. Gareth Carol (University of Birmingham) 8
6. Melanie Bell (Anglia Ruskin University) 9

Idioms “*Electrified*”: What Event-related Brain Potentials Reveal about the Comprehension of Multi-word Expressions

Cristina Cacciari

Department of Biomedical, Metabolic and Neural Sciences, Modena

Human brains are a repository of a variety of semantic knowledge that includes word meanings, concepts etc. and familiar word strings that people learn and remember (e.g., movie and book titles, song titles and lyrics, poetry, idioms, proverbs, clichés). Idiomatic expressions (e.g., kick the bucket) are part of this knowledge stored in semantic memory. They are extremely widespread in natural language and are part of the vast family of multiword expressions (MWEs) whose meaning does not derive from the composition of the constituent word. To investigate the neural activity underlying the comprehension of these non-literal expressions, many studies measured the electrophysiological activity in the brain as reflected by Event-related brain potentials (ERPs). The sophisticated time resolution of ERP measures allows to track the timing of comprehension processes with millisecond-level resolution. In this presentation, I discuss the state of the art of the ERP literature on MWEs with specific reference to the role of predictive processing.

Colorless Green Ideas in Formal Distributional Semantics

Aurelie Herbelot

Centre for Mind/Brain Sciences, University of Trento

Phrases such as 'colorless green ideas' are usually considered semantically infelicitous on grounds of non-compositionality. In this talk, I inspect what exactly is meant by 'non-compositionality' and argue that the notion is valid in formal representations of meaning whilst being at best questionable in distributional formalisms. Drawing on existing experimental results in Formal Distributional Semantics, as well as current work on the meaning of colors, I tentatively propose that non-compositionality is a graded notion over a lexical / truth-theoretic continuum.

Parallels in the Processing of Compositional and Non-compositional Phrases: Implications for the Lexicon

Inbal Arnon

Department of Psychology, Hebrew University of Jerusalem, Israel

The representation and processing of non-compositional phrases has been the focus of much research. Such forms (e.g., idioms, multiword expressions) are often treated as distinct from compositional phrases, and used as evidence for dual-system models of language, where compositional and non-compositional forms are represented and learned differently. In this talk, I draw on developmental and psycholinguistic evidence to argue that the distinction between compositional and non-compositional phrases is one of degree and not kind, and present a model of the lexicon where the difference between them reflects differences in distributional properties and semantic coherence.

Multilingual Meaning Banking and Compositionality

Johan Bos

Computational Semantics – Computational Linguistics, University of Groningen

In this talk, I will present ongoing work on creating a parallel meaning bank (a large corpus of texts annotated with formal meaning representations) for four languages. I will show how the idea of compositionality is used, and how it is abused, to produce meaningful analyses of non-compositional phenomena.

Into the Unknown: What Eye-tracking Data Can Tell Us about how Language Users Process and Resolve Unfamiliar Idioms

Gareth Carol

Department of English Language and Applied Linguistics, University of Birmingham

Idiom research widely reports on the importance of familiarity in how native and non-native speakers understand non-compositional expressions (e.g. Cieślicka, 2006, 2013; Libben & Titone, 2008;). But familiarity also masks a number of other processes, in that native speakers can often simply recognise and retrieve a known expression without the need for semantic analysis, inferencing from context, or any of the other processes that we assume are required to understand figurative meaning (Carrol, Littlemore & Dowens, 2018). By using less familiar items – idioms translated from other languages – as our stimuli, English native speakers effectively become “non-natives by proxy”, and must apply the same strategies as language learners when they encounter unknown figurative expressions. I discuss the interplay of factors such as relative transparency (and what contributes to this) and strength of context, and present data from eye-tracking to show how language users attempt to make sense of phrases they haven’t heard before. This includes new data from an ongoing study (Carrol & Littlemore, in prep.) where we compare known idioms, unknown idioms and metaphors, focusing on the differences in how these are recognised, processed and integrated into the surrounding text.

I will also discuss the use of eye-tracking more broadly as a tool for the study of figurative and formulaic language. This includes comparing data from a number of studies that have looked at different kinds of compositional and non-compositional expression, to establish the important differences in how these are tackled during online language processing.

References

- Carrol, G., Littlemore, J., & Gillon Dowens, M. (2018). Of false friends and familiar foes: Comparing native and non-native understanding of figurative phrases. *Lingua*, 204, 21-44.
- Cieślicka, A. B. (2006). Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research*, 22(2), 115-144.
- Cieślicka, A. B. (2013). Do nonnative language speakers chew the fat and spill the beans with different brain hemispheres? Investigating idiom decomposability with the divided visual field paradigm. *Journal of Psycholinguistic Research*, 42(6), 475-503.
- Libben, M., & Titone, D. (2008). The multidetermined nature of idiom processing. *Memory & Cognition*, 36(6), 1103-1121.

The Role of Constituent Disambiguation in the Interpretation of Novel Noun-noun Constructions

Melanie Bell

Department of English and Media, Anglia Ruskin University

In English, constructions consisting of two orthographically separate nouns, in which the first is the semantic modifier of the second, have been referred to both as phrases and as compound nouns. But regardless of which label is preferred, such constructions are inherently non-compositional, because the semantic relation connecting the two constituent words is not overtly expressed. For example, nothing in the string bike rack tells us whether this is a rack for a bike – e.g. for carrying a bike on a vehicle or for parking a bike – or a rack on a bike – e.g. for attaching bags – or any other conceivable interpretation. The ambiguity demonstrated by this lexicalised combination becomes even more problematic when a noun-noun construction is encountered for the first time.

Since Gagné and Shoben (1997)¹, evidence has accumulated that humans use statistical semantic preferences associated with constituent words in the interpretation of noun-noun constructions. Most of this work has focused on the semantic relation between constituents, rather than the interpretation of the constituents themselves. Yet due to the ubiquity of polysemy and homonymy, there are typically several concepts to choose from for any given constituent, and it is not clear at the outset which concept is to be used. This holds for both modifiers and heads, and for concepts that are unrelated (e.g. plant denoting a factory or organism) as well as those that are metaphorically linked (e.g. gold denoting a material or colour). In a series of studies using attested novel constructions, I will show that diversity of interpretation is much greater than suggested by studies that focus only on semantic relations and that the ambiguity of the constituents, especially the head, plays an important role in determining both difficulty and diversity of interpretation.

Overall, there is evidence that, even with the help of linguistic context, humans do not necessarily converge on the same interpretation of novel noun-noun constructions. This may be especially true when the intended reading of either constituent is specialised or infrequent. Assuming that speakers who coin new combinations do have some specific interpretation in mind, how then does communication proceed? A plausible explanation is that a gist interpretation involving significant underspecification is often sufficient, with more detailed – and costly – elaboration occurring only when strictly required.

¹ Gagné, C. L., & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of modifier–noun combinations. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 23(1), 71–87.

Participant Talk Abstracts

1. Eva Smolka and Carsten Eulitz (University of Konstanz)	11
2. Heiner Drenhaus and Vera Demberg (Saarland University) (<i>cancelled</i>)	13
3. Amélie la Roi, Simone Sprenger (University of Groningen)	16
4. Petr Rosseyaykin (Lomonosov Moscow State University)	18
5. Patricia Fischer (University of Tübingen)	21
6. Irene Fioravanti and Marco Silvio Giuseppe Senaldi (Università per Stranieri di Siena), (Scuola Normale Superiore di Pisa)	23
7. Sara Beck and Ruth Keßler (University of Tübingen)	26

Can You Reach for the Planets or Grasp at the Stars? – Modified Noun, Verb, or Preposition Constituents in Idiom Processing

Eva Smolka and Carsten Eulitz

Department of Linguistics, University of Konstanz

Idioms are a special case of complex word formations in that their figurative meaning cannot be constructed compositionally from the meaning of the single constituents. Hence, psycholinguistic research remains puzzled by the question of how the figurative meaning is assembled. For example, some approaches assume some kind of a ‘unitary entry’ that defines the single constituents of an idiom that will activate the figurative meaning (e.g., Cacciari & Tabossi, 1988; Holsinger, & Kaiser, 2013; Sprenger, Levelt, & Kempen, 2006). Other approaches assume that the meaning of the single constituents is activated, and their joint co-activation activates the figurative meaning (e.g. Smolka, Rabanus, & Rösler, 2007).

The aim of the present study was to examine the concept of a ‘unitary entry’ for idioms in more detail: (a) Will the figurative meaning of an idiom be retained, if an idiomatic constituent, such as the noun, verb, or preposition, is modified? (b) Will the word category of an idiomatic constituent (noun, verb, preposition) affect whether a modification will preserve the figurative meaning? To this end, we conducted three paraphrase experiments, in which each canonical idiomatic sentence (e.g., “*She reached for the stars*”) was presented in three versions: (1) with its canonical constituent, (2) with the canonical constituent replaced by a closely associated word, or (3) with the canonical constituent replaced by an unrelated word. We manipulated the noun constituent in Experiment 1 (e.g., canonical/modified/unrelated: “*She reached for the stars/planets/sweets*”), the verb constituent in Experiment 2 (e.g., “*She reached/grasped/looked for the stars*”), and the preposition in Experiment 3 (e.g., “*She reached for/at the stars/for the sweets*”). Each sentence was paired with the paraphrase of the idiomatic sentence, “*She always strived for something unreachable*”), and participants rated on a scale from 1 (“not at all”) to 7 (“completely”) how well the meanings of two sentences mirrored each other. We tested between 33-39 idiom triplets in each experiment, all of which consisted of 7 words, had the same sentence structure (verb + prepositional phrase), and had high cloze probabilities (on average 90%). Close noun or verb associates of the critical idiomatic noun or verb constituents were selected in pre-tests.

We hypothesized that if ‘unitary entries’ define the idiomatic constituents, then sentences whose idiomatic constituents are replaced by close associates will not be considered to reflect the figurative meaning and will yield paraphrase ratings similar to sentences with unrelated constituents. If, however, the assumptions hold (a) that each idiomatic constituent activates its literal meaning, (b) that a close associate of an idiomatic constituent will activate a similar literal meaning and (c) will thus contribute to the joint co-activation of the figurative meaning, sentences holding close associates of an idiomatic constituent will be rated as higher in reflecting the figurative meaning than those with unrelated constituents.

Linear mixed-effect models on paraphrase ratings yielded strong effects of sentence type (canonical/modified/unrelated) and constituent type (noun/verb/preposition), with highest ratings for canonical idiomatic phrases, lowest ratings for control sentences, and ratings in between for idioms with modified constituents (see Figure 1). As Figure 1 further indicates, idioms with modified verbs were rated higher in matching the figurative meaning than idioms with modified prepositions or nouns.

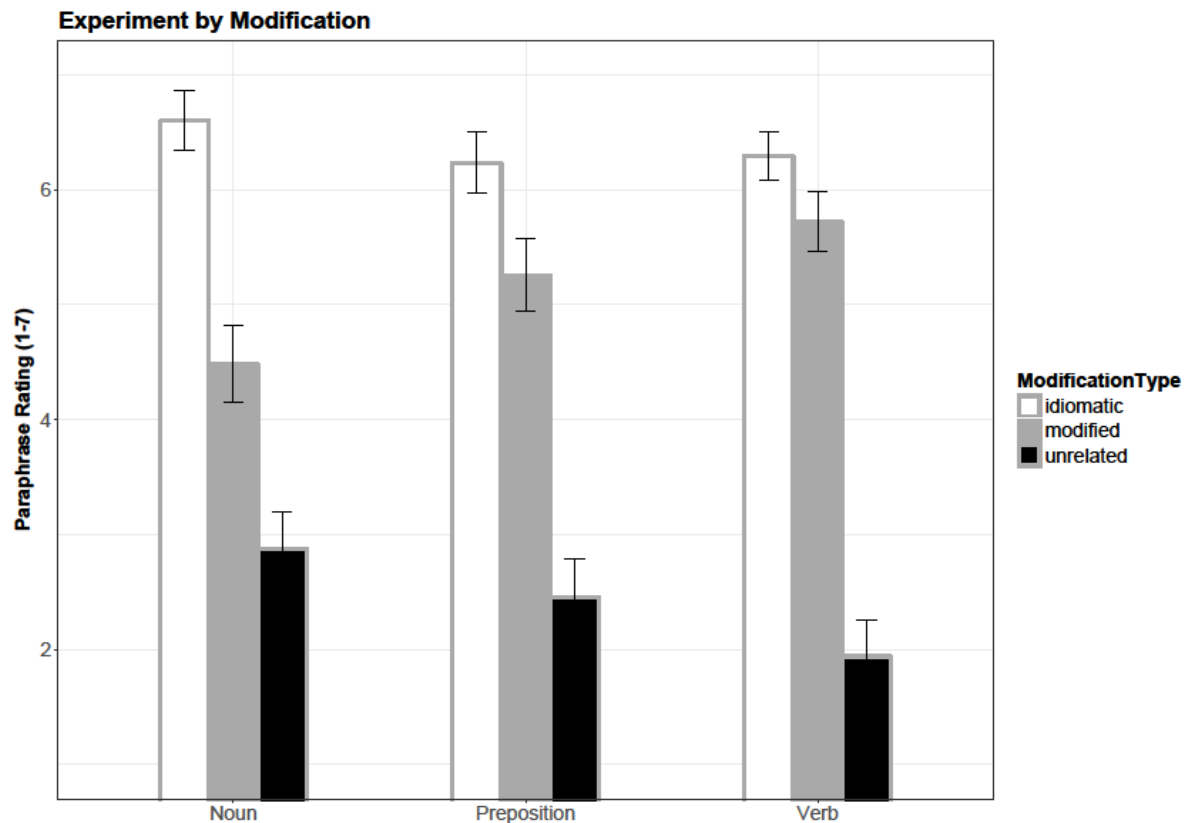


Figure 1. Paraphrase ratings on a scale from 1-7 for idiomatic sentences holding canonical, modified, or unrelated constituents; manipulated noun constituents (left panel), prepositions (mid panel), and verb constituents (right panel); Y-bars indicate standard errors of the mean.

Overall these findings indicate that the figurative meaning was assembled in spite of the modifications. We conclude that idioms are not fully ‘semantically fixed’ and do not possess fixed ‘unitary entries’ but allow for some flexibility in the processing of idioms. Modified constituents that activate meanings similar to those of the canonical constituents will co-activate the figurative meaning of the idiom together with the other idiomatic constituents. We discuss psycholinguistic models on idiom comprehension.

References

- Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, 27, 668-683.
- Holsinger, E., & Kaiser, E. (2013). Processing (non)compositional expressions: mistakes and recovery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 866-878.
- Smolka, E., Rabanus, S., & Rösler, F. (2007). Processing verbs in German idioms: Evidence against the Configuration Hypothesis. *Metaphor and Symbol*, 22(3), 213-231.
- Sprenger, S., Levelt, W., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language*, 54, 161-184.

CANCELLED

The Time Course of Processing Idioms as a Function of Context: An ERP-study

Heiner Drenhaus and Vera Demberg

Saarland University

Existing literature on idioms has argued that certain idioms can be recognized before the end of the idiomatic expression, and then lead to qualitatively different types of prediction effects than non-idiomatic constructions [1,2]. The point in processing after which an idiom is identified is referred to as the idiom's "recognition point" (RP), at which the idiom is retrieved from the lexicon, and the rest of the idiom is then very highly expected. [1] suggest that after idiom recognition, processing might be subject to a categorial template matching process (in line with a proposal from [3], who distinguishes a categorial matching mechanism eliciting a P3, and a probabilistic mechanism of search in working memory eliciting an N400).

The goal of the current study is two-fold: Firstly, we want to test whether we can replicate the finding of a P300 effect for idiom recognition at or after the RP. Secondly, we want to address a gap in the literature by testing whether the location of an idiom's RP is context-dependent. We designed an EEG study with 24 items consisting of idioms with an early RP, for example "*riecht den Braten*", as shown in (2); we manipulated the preceding context (idiomatic 1a vs. literal 1b). Idioms were selected such that cloze probability on the determiner > 0.8 in the idiomatic context.

We tested 34 participants in a 2(literal vs. idiomatic context) x 2(literal vs. idiomatic expression) between subject design. First, participants read a context sentence (1a) or (1b) followed by the idiomatic (2a) or literal (2b) version of sentence (word-by-word presentation with 500ms presentation rate, plus 100ms interstimulus interval). Some of the sentences were followed by a question regarding the plausibility of the second sentence, which participants answered by button press. Plausibility of the conditions 1a-2a, 1b-2a and 1b-2b was matched.

Idiomatic Context: (1a) Petra fährt nach Paris und wird von einem Straßenhändler angesprochen, der ihr gefälschte Uhren verkaufen will.

'On her trip to Paris, Petra is approached by a merchant who wants to sell her fake watches.'

Literal Context: (1b) Petra macht einen Spaziergang und kommt an einem Gasthaus mit Biergarten vorbei.

'Petra is going for a walk and is passing by a restaurant with a beer garden.'

Target sentence: (2) Sie riecht [den] (a) Braten/ (b) Auflauf und geht weiter.

'She smells the roast (idiomatic meaning: smell the rat)/casserole and moves on.'

Verbal Region. For the onset of the idiom (*riecht*), we expect that the idiom might already be recognized in the context that biases for an idiomatic reading (such a very early effect is possible, because the verb's semantic literal is not compatible with that context, while its idiomatic reading is compatible with preceding context). We therefore compare ERPs on the verb *riecht* in literal context vs. idiomatic context. If context does not affect idiom recognition this early on in the idiom, we would expect to see an N400 effect for the idiomatic context; if the idiomatic context however allows for very early recognition of the idiom, we expect to see an early positivity (the P3 observed in [1]) in the literal context. We fitted linear mixed models

with ERP values averaged over the critical items for each participant as dependent measure. Our analysis revealed a significant centro-posterior positivity (200-400ms) on the verbal region following an idiomatic context (1b) compared to a literal context; compare Figure 1.

This effect is in line with Kok's 2001 [3] proposition that the P3 amplitude reflects activation of elements in an event categorization network, i.e. that the P3 reflects the recognition of the encountered verb *riechen* being compatible with the context (confirming this prediction, e.g. [8]). The amplitude of the P3 is larger in the idiomatic context, which may reflect the categorial recognition of the idiom.

Determiner Region. We expected to also see an effect of idiom recognition as a function of context at the determiner: At this point in processing, the target noun has not yet been perceived, and the only difference between conditions lies in context type: the idiomatic context (1a) biases for the exact completion *den Braten*, while the literal context is compatible with several dishes, many of which have same grammatical gender as the idiomatic target.

If the idiom has already been recognized in the idiomatic (in line with the P300 effect found on that region), we predict an N400 effect on the determiner in the literal context condition compared to the idiomatic context (see also [4,5,6]; but [7]).

We fitted linear mixed models with ERP values averaged over the critical items for each participant as dependent measure. Our analysis revealed a significant fronto-centrally distributed negativity (400-600ms) on the determiner region following a non-idiomatic context (1b) compared to a context which assigned an idiomatic reading (1a); compare Figure 2.

The reported negativity on the determiner in our study can be seen as an effect induced by predictive processing (e.g. [5,6]). These results show that effects of context constraint are measurable on the determiner.

Noun Region. On the noun region, we compared the idiomatic completion *Braten* with the literal completion *Auflauf*. In the idiomatic context, we expect to see some signature of processing difficulty or reanalysis on the literal completion.

We again fitted linear mixed models with ERP values averaged over the critical items for each participant as dependent measure. Our analysis revealed a significant long-lasting fronto-centrally distributed negativity (450-1000ms) for the violation (2b) compared to the idiom-consistent completion (2a); compare Figure 3. This effect can be attributed to the non-idiomatic noun eliciting an N400 effect compared to the predicted idiomatic completion, followed by a long lasting negativity which is indicative of higher working memory load, which may be caused by the attempt to process the idiom literally – highlighting the effort of computing a new mental model for an unexpected event in a given context (e.g. [8,9])

In the literal context condition, we found an early positivity for the idiomatic completion, which indicates that the idiom was recognized at this final word in the non-supportive context. Taken together, our data is supportive for a view where an idiom's recognition point is context-dependent, and where idiom recognition leads to strong, potentially categorial predictions of the rest of the idiom. When the expectation was violated, we observed a long-lasting negativity, which may be indicative of retrieval from 'semantic' memory. We will discuss the implications for our understanding of prediction in language processing with respect to the early positivity and the fronto-central negativity observed in these experiments.

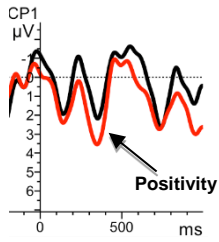


Figure 1: ERP effects on the verbal region following a idomatic context (black line) and following a literal context (red line) .

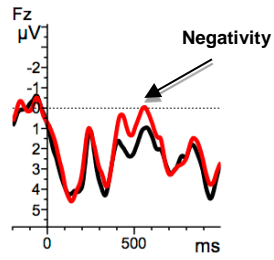


Figure2: ERP effects on the determiner region following a idomatic context (black line) and following a literal context (red line).

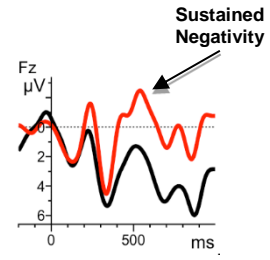


Figure 3: ERP effects on the noun region following a idomatic context , for the idiom-consistent noun (black line) vs. violation (red line) .

References

- [1] Vespignani, Canal, Molinaro, Fonda, Cacciari 2009; [2] Tabossi Fanari Wolf, 2005; [3] Kok, 2001; [4] Van Berkum, Hagoort, & Brown, 1998; [5] DeLong, Urbach, Groppe, & Kutas, 2011; [6] DeLong, Urbach, & Kutas, 2005; [7] Nieuwland et al, 2018; [8] Van Petten & Luka, 2012; [9] Pulvermüller et al. 2009

Language in the Aging Brain: Using ERPs to Study Idiom Processing across the Life Span

Amélie La Roi^{1,2}, Simone Sprenger^{1,2}, and Petra Hendriks¹

¹ Center for Language and Cognition Groningen, University of Groningen

² Neuroimaging Center, University of Groningen

As the population of elderly adults continues to grow worldwide, understanding the cognitive consequences of aging is becoming increasingly important. In younger adults, cognitive functions have been shown to be closely tied to successful language processing and comprehension^{1,2}. However, it remains largely unknown how age-related decline in cognitive functions affects language abilities. Our study examines this question by investigating the processing of idioms, such as the Dutch *tegen de lamp lopen* ('to walk against the lamp'). Idioms are ambiguous in that they carry a literal meaning ('to walk against the lamp'), as well as a figurative meaning (here: 'to get caught'). To activate an idiom's figurative meaning, the reader or listener has to inhibit the literal meanings of the idiom constituents ('to walk' and 'lamp'). Since inhibition skills typically decline with age³, increased age is hypothesized to reduce the suppression, thus increasing the activation, of the literal meanings of idiom constituents. To test this hypothesis, we record the event-related potentials of 60 right-handed native speakers of Dutch (aged 18-80) while they read sentences containing an idiom. We measure the amplitude of the N400 evoked by the target word, which is either the idiom's original final noun (*lamp* 'lamp'), a semantically related word (*kaars* 'candle'), or a semantically unrelated word (*vis* 'fish'). In half of the sentences the target word is the idiom's last word (1), while in the other sentences the target word precedes the idiom's verb (2).

(1) Groningen zal geen last meer hebben van die pyromaan, want hij liep tegen de **lamp/kaars/vis** dinsdagavond.

*Groningen will not be bothered anymore by that pyromaniac, because he walked against the **lamp/candle/fish** Tuesday night.*

(2) Groningen zal geen last meer hebben van die pyromaan, omdat hij tegen de **lamp/kaars/vis** liep dinsdagavond.

*Groningen will not be bothered anymore by that pyromaniac, because he against the **lamp/candle/fish** walked Tuesday night.*

We predict age to correlate with the degree of activation of the target word's literal meaning. If the literal meaning of the idiom constituent has been suppressed, the amplitude of the N400 will not significantly differ in response to semantically related target words compared to semantically unrelated target words (Figure 1, top left), consistent with previous findings⁴. However, if the literal meaning of the idiom constituent is still activated, the amplitude of the N400 in response to the target word is expected to significantly reduce for semantically related target words compared to semantically unrelated target words (Figure 1, bottom left). Furthermore, this reduction in N400 amplitude is most likely to be found for sentences in which the target word precedes the idiom's verb (2) compared to sentences in which the target word is the last word of the idiom (1) (Figure 1, top right and bottom right). The reason for this is that activation of the idiom's figurative meaning relative to the activation of the literal meanings of the idiom constituents is expected to be less strong when the reader has not yet encountered all idiom constituents. By studying idiom processing across the life span we aim to increase our understanding of the relation between cognitive functions and the processing of ambiguous expressions in which one of the meanings is constructed non-compositionally, such as idioms.

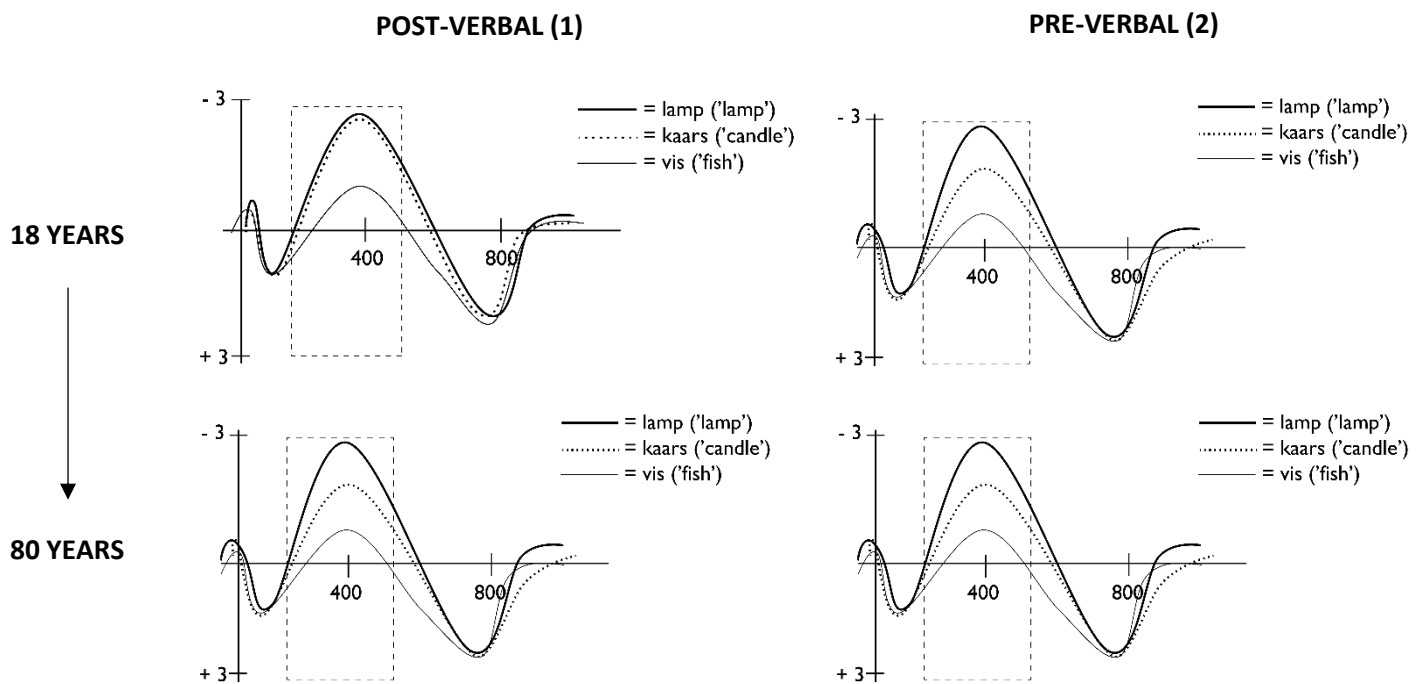


Figure 1. Predictions for the amplitude of the N400 in response to target words that are the idiom's original noun (e.g., lamp 'lamp', solid thick line), a semantically related word (e.g., kaars 'candle', dashed line), or a semantically unrelated word (e.g., vis 'fish', solid thin line)

References

- ¹Friedman, N. P., & Miyake, A. (2004). The reading span test and its predictive power for reading comprehension ability. *Journal of Memory and Language*, 51(1), 136-158.
- ²Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, 5(3), 263-281.
- ³Craik, F. I., & Salthouse, T. A. (Eds.). (2011). *The Handbook of Aging and Cognition*. Psychology Press.
- ⁴Rommers, J., Dijkstra, T., & Bastiaansen, M. (2013). Context-dependent semantic processing in the human brain: Evidence from idiom comprehension. *Journal of Cognitive Neuroscience*, 25(5), 762-776.

Using Context Information for Russian MWE Extraction

Petr Rossyaykin

Lomonosov Moscow State University, Moscow

Multiword expressions (MWEs) are combinations of two or more words that display lexical, syntactic, semantic, pragmatic and/or statistical idiomaticity, or idiosyncrasy (Baldwin & Kim 2010). Semantic idiosyncrasy, or non-compositionality, is the property of MWEs which presents the main interest for linguistics and main difficulty for NLP. For example, many MWEs do not allow for direct translation (English *to hold a lecture* vs Russian *читать лекцию* ‘to read a lecture’ or *вести лекцию* ‘to conduct a lecture’) and their components cannot be substituted with synonyms (*#powerful tea*) due to non-compositionality. Thus, performance of many applications is highly dependent on MWE-awareness. However, manual extraction of MWEs is extremely inefficient because their number in language is comparable to that of single-word lexical items and grows constantly (Sag et al. 2002).

As a result, automatic MWE extraction (or discovery) remains a relevant task which generally includes extraction of a list of candidates from text corpora and candidate ranking. An automatically generated ranked list of MWE candidates can be used in machine translation, paraphrase generation, error detection, thesaurus building, etc.

In this study my goal is to elaborate a method for automatic supplementation of Russian-language thesaurus with new MWEs and thus I focus on extraction of Russian nominal constructions of Adj-N (e.g. *социальная сеть* ‘social network’) and N-N (e.g. *точка зрения* ‘point of view’) patterns. From semantic perspective expressions relevant for thesauri include idioms (*круглый стол* ‘round table’), proper names (*Московская область* ‘the Moscow region’), terms (*гипертонический криз* ‘hypertensive emergency’), etc.

The corpus used for MWE extraction is composed of Russian news’ texts from the Internet published in 2011 and concatenated in a single file. The resulting corpus of 446M tokens was lemmatized and PoS-tagged in order to extract the initial list of 37767 candidate Adj-N and N-N bigrams with the observed frequency of more than 200. No further processing was used since I aim at finding a totally unsupervised and possibly language-independent approach.

9837 of the candidate expressions mostly corresponding to coherent notions on the ontological level (e.g. *фунт стерлингов* ‘pound sterling’) and thus present in the Russian language thesaurus RuThes (Loukachevitch et al. 2014) were regarded as actual MWEs. Precision at K was measured as a ratio of these expressions to all ranked in top K.

Most common methods for MWE extraction/ranking either require external resources (parallel corpora, WordNet, automatic translation) or are based entirely on frequency distribution of MWEs (PMI, t-score, etc.). However, the former lack in coverage while the opportunities of the latter are restricted due to a tiny amount of information they take advantage of (Evert 2007). In this study I compare more than 20 widely used statistical association measures with an equally language- and processing-independent but much less elaborated context-based approach.

Context measures detect MWEs drawing on non-compositionality rather than statistical distribution. The idea behind is to take into account the number of unique words which occur as components or neighbors of MWEs. I propose a modification of GM measure used by Nakagawa & Mori (2003) for automatic term extraction (1). My variant (denoted as type-LR) models lexical rigidity (non-substitutability) of MWE components:

$$(1) \quad \text{typeLR}(x, y) = \sqrt{|r(x)| * |l(y)|}$$

Where $r(x)$ and $l(y)$ are sets of unique words which occur in a corpus immediately to the right from the word x and to the left from the word y respectively.

Taking into account that usual statistical association measures and context measures use different properties of MWEs (statistical idiosyncrasy and non-compositionality) I also combined type-LR with the observed frequency of word pairs being classified (2). This modification gives a significant increase in average precision (see Table 1).

$$(2) \quad \text{typeFLR}(x, y) = \frac{f(x, y)}{\text{typeLR}(x, y)}$$

Where $f(x, y)$ is the observed frequency of the bigram 'x y' in a given corpus.

Recently, Riedl & Biemann (2015) achieved high precision in English and French MWE extraction with their context-based DRUID measure assuming that contexts of MWEs of different lengths are similar to that of single words due to the fact that MWEs are used to denote single entities. I take another perspective on the context data, comparing immediate contexts of MWEs with those of their components. In particular, I propose the two following measures:

$$(3) \quad \text{context intersection}(x, y) = \frac{|l(x, y) \cap l(x)|}{|l(x)|} * \frac{|r(y) \cap r(x, y)|}{|r(y)|}$$

$$(4) \quad \text{independent CI}(x, y) = \frac{|l(x, y) \cap l(x, \bar{y})|}{|l(x, \bar{y})|} * \frac{|r(\bar{x}, y) \cap r(x, y)|}{|r(\bar{x}, y)|}$$

Where $l(x, y)$ is a set of unique words to the left from the bigram 'x y', \bar{x} denotes all words except for x and \bar{y} denotes all words except for y .

Following Riedl & Biemann (2015) and my previous assumption about interplay between non-compositionality and statistical idiosyncrasy I also combined CI and ICI with the observed frequency. As a result, I achieved average precision much higher than that of any association measure. The results for best statistical and context measures are presented in Table 1.

In this report I am going to provide a vast comparison of statistical, distributional and context methods proposed by me for Russian MWE extraction. However, the methods under consideration are in no way bound to Russian and easily applicable to other languages.

Measure	AP@100	AP@500	AP@1000	AP@2500
PMI ³	0,907	0,821	0,795	0,726
log-likelihood ratio	0,778	0,802	0,780	0,705
type-LR	0,521	0,563	0,553	0,529
type-FLR	0,818	0,825	0,796	0,74
log(freq)*ICI	0,915	0,879	0,855	0,789

Table 1. Comparison of statistical and context measures with the highest average precision

References

- Baldwin & Kim 2010 – Timothy Baldwin and Su Nam Kim. Multiword expressions. In Nitin Indurkha and Fred J. Damerau, editors, *Handbook of Natural Language Processing*, pages 267–292. CRC Press, Taylor and Francis Group, Boca Raton, FL, USA, 2 edition.
- Evert 2007 – Evert, S.: *Corpora and collocations*. Extended Manuscript of Chapter 58 of A. Lüdeling and M. Kytö, 2008, *Corpus Linguistics. An International Handbook*, Mouton de Gruyter, Berlin.
- Loukachevitch 2014 – Loukachevitch, Natalia, Dobrov, Boris and Ilia Chetviorkin. "Ruthes-lite, a publicly available version of thesaurus of russian language ruthes." *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference "Dialogue"*, Bekasovo, Russia. 2014.
- Nakagawa & Mori 2003 – Nakagawa H., Mori T. Automatic Term Recognition based on Statistics of Compound Nouns and their Components // *Terminology*. – 2003. – Vol. 9, №2. – P. 201-219.
- Riedl & Biemann 2015 – Riedl, Martin and Chris Biemann. 2015. A single word is not enough: Ranking multiword expressions using distributional semantics. In *Proceedings of EMNLP 2015*, pages 2430–2440, Lisbon.
- Sag et al. 2002 – Sag, Ivan A.; Baldwin, Timothy; Bond, Francis; Copestake, Ann; Flickinger, Dan. Multiword expressions: A pain in the neck for NLP. In *Proceedings of the Third International Conference on Intelligent Text Processing and Computational Linguistics (CICLing 2002)*, pages 1–15, Mexico City, Mexico.

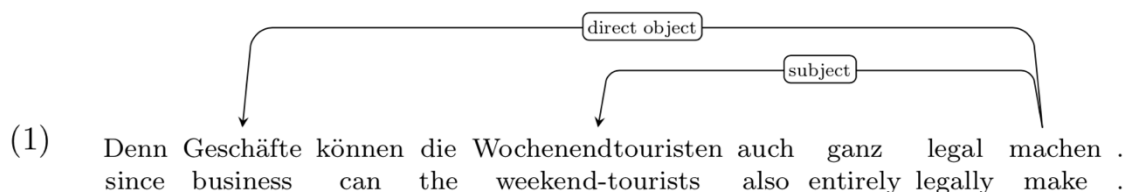
Improvements on Subject-Object Disambiguation in Parsing

Patricia Fischer

Department of General Linguistics, Tübingen

Solving subject-object ambiguities is a challenging task in dependency parsing of German. This is still true for modern state-of-the-art neural network parsers that use word embeddings to capture semantic similarities between words. In the the neural parser of De Kok and Hinrichs [1], 12 percent of all attachment errors concern the subject and direct object relations.

Many of such constructions are easily resolved by humans, even when case markers are absent or ambiguous, since incorrect readings are semantically implausible. For instance, both noun phrases *Geschäfte* `business' and *Wochenendtouristen* `weekend-tourists' in Example 1 are syntactically possible subjects or direct objects of the verb *machen* `to-do'. However, the reading of *Geschäfte* as subject is implausible given that *machen* takes an animate subject and an inanimate object. Knowledge of such selectional preferences can therefore help the parser in making more well-informed attachment decisions.



The use of compatibility metrics has been shown to be beneficial for improving attachment scores for specific constructions such as PP attachment [5, 8], but also for parsing in general [7]. These earlier works use association strength metrics such as pointwise mutual information (PMI). However, these approaches typically estimate the compatibility of rare and unseen events incorrectly.

The shift to word embeddings as the word representation in computational linguistics has opened the possibility to model the compatibility of vector subspaces rather than concrete words. Some preliminary work in this direction has been done by De Kok et al. [2] who introduce a classifier that predicts the compatibilities between a PP and its possible candidate attachment sites using (among other features) word embeddings. Schoener [6] further extends this line of work by using a similar architecture to predict compatibility scores of adjective-noun and verb-(direct object) pairs. Schoener [6] uses cluster-informed sampling to introduce negative examples.

- (2) Und das Erfolgskonzept von heute ist der kalte Kaffee von morgen .
and the success-concept of today is the cold coffee of tomorrow .

The challenges of such compatibility models lie in recognizing and handling non-compositional phrases. For example, such a model would interpret the phrase *kalter Kaffee* `old hat' in Example 2 to be highly compatible with the verb *trinken* `to drink' even though this is not the case. In work on the composition of compounds and adjective-noun combinations, Dima [3, 4] has equally recognized this challenge that non-compositional phrases pose. Therefore, current experiments by Dima on composition functions are also targeted at improving performance on non-compositional compounds and phrases.

A compatibility model could further enrich the parser in obtaining higher accuracy on compositional (subject, verb, direct object) triples. A successful composition model additionally included in a parser could then improve attachment for non-compositional verb phrases.

References

- [1] Daniël de Kok and Erhard Hinrichs (2016). “Transition-Based Dependency Parsing with Topological Fields.” In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics: Volume 2, Short Papers*, p. 1–7.
- [2] Daniël de Kok, Jianqiang Ma, Corina Dima, and Erhard Hinrichs (2017). “PP Attachment: Where do We Stand?” In *Proceedings of the 15th Conference of the EACL: Volume 2, Short Papers*, p. 311–317.
- [3] Corina Dima (2016). “On the Compositionality and Semantic Interpretation of English Noun Compounds.” In *Proceedings of the 1st Workshop on Representation Learning for NLP*, p. 27–39.
- [4] Corina Dima (2019). *Compositional Models for the Representation and Interpretation of Nominal Compounds*. PhD Thesis.
- [5] Donald Hindle and Mats Rooth (1993). “Structural Ambiguity and Lexical Relations.” In *Computational Linguistics*, 19(1), p. 103–120.
- [6] Peter Schoener (2018). *Embedding-Based Approaches to Prediction of Semantic Compatibility*. B.A. thesis under supervision of Daniël de Kok.
- [7] Gertjan van Noord (2007). “Using Self-Trained Bilexical Preferences to Improve Disambiguation Accuracy.” In *Proceedings of the 10th Conference on Parsing Technologies*, p. 1–10.
- [8] Martin Volk (2000). “Scaling up. Using the WWW to Resolve PP Attachment Ambiguities. In *Proceedings of KONVENS-2000*, p. 151–155.

Lexical Fixedness and Compositionality in Native and Non-native Speakers' Intuitions about Italian Word Combinations: Psycholinguistic and Computational Perspectives

Irene Fioravanti¹, Marco S. G. Senaldi², Alessandro Lenci³, and Anna Siyanova-Chanturia⁴

¹ Università per Stranieri di Siena

² Scuola Normale Superiore di Pisa

³ Università di Pisa

⁴ Victoria University of Wellington

The present research focuses on native and non-native speakers' intuitions about Italian verb-noun word combinations. The phraseological approach distinguishes word combinations in terms of lexical restriction and compositionality. The former refers to whether or not a word within a phrase can be substituted with a synonym; while the latter has to do with how much the meanings of individual words contribute to the overall meaning of a phrase. In the phraseological models, word combinations are usually placed on a continuum from free combinations to idiomatic expressions, with collocations being somewhere in the middle. Free combinations are fully compositional in that their constituents can be easily substituted, while idioms are described as least decomposable and highly fixed word combinations. Collocations are distinct from free combinations and idioms. On the one hand, they are characterized by lexical fixedness, allowing a relatively limited substitution of the members. On the other hand, they are fairly transparent with their individual components contributing to the overall meaning of a phrase.

Research suggests that free combinations and idiomatic expressions may be less challenging for second language (L2) learners compared to collocations (Laufer and Waldman, 2011; Nesselhauf, 2005; Gyllstad and Wolter, 2016). Might this be due to their nature? Do L2 learners perceive the three types of word combinations differently? Does the degree of restriction play a role in learners' vocabulary choices? To what extent do L2 speakers' intuitions differ from those of native speakers?

Two studies were carried out to answer these questions. In Study 1, 120 Italian verb-noun combinations were embedded in sentential context; an identical set of 120 sentences was further used in which the verb of the target combination was substituted with a synonym. Native speakers of Italian and English learners of L2 Italian were asked to rate acceptability of the sentences. In Study 2, the same materials were used. However, the verb was removed from all verb-noun combinations. Each sentence was presented with two options: the original verb and a synonym. Native and non-native speakers' task was to choose the verb they felt was the most acceptable given the context.

Mixed-effects modelling showed that the three types of word combinations were perceived differently by the learners and native speakers, and that both groups were sensitive to lexical fixedness. L2 speakers' intuitions were accurate about free combinations and idioms, but not about collocations. Learners judged the use of a synonym in collocations as more acceptable compared to free combinations and idioms. On the contrary, native speakers tended to judge the use of a synonym as less acceptable in the case of collocations. Moreover, they judged the

use of a synonym in free combinations as acceptable, while in idiomatic expressions they were judged as not acceptable. Native speakers judged the substitutions of the verb with a synonym as less acceptable in the case of idioms compared to collocations. In addition, learners' intuitions were found to vary as a function of proficiency; they were further affected by both phrasal frequency and frequency bands.

Considering that word combinations also differ in terms of compositionality, the following analyses were conducted in order to answer these questions: Does compositionality affect native and non-native speakers' intuitions about the restriction of word combinations? To what extent does compositionality interact with lexical fixedness in native speakers' and learners' intuitions about vocabulary use?

Firstly, we collected compositionality and lexical fixedness judgments for free combinations, collocations and idioms from native speakers of Italian. They judged free combinations as the most compositional and least fixed expressions, and idioms as least compositional and most restricted combinations. Collocations appeared in-between free combinations and idioms.

Following this, native speakers' judgments about lexical fixedness and compositionality were compared to corpus-based indices of compositionality. To this aim, we used the variant-based distributional measures proposed by Senaldi, Leboni and Lenci (2016, 2017), who showed them to be reliable in defining lexical variability and compositionality of idiomatic phrases. We extended these measures to free combinations and collocations. In particular, we focused on two types of measures. The first measure is an additive-based index, which computes the cosine similarity between the vector of a combination and the vector resulting from the sum of the vectors of the word combination elements. On the contrary, the second measure is a variant-based measure, the centroid measure, that computes the cosine similarity between the vector of a target combination and the centroid of the vectors of its variants. Native speakers' judgments of lexical fixedness and distributional measures of compositionality were then added to the mixed effects models of behavioural data collected in Study 1 and 2 above. We found that native speakers' intuitions were affected also by compositionality and that compositionality interacted with lexical fixedness in native speakers' intuitions about the use of word combinations. On the contrary, the results showed that learners' intuitions were not affected by compositionality.

Taken together, our findings provide new insights into how native and non-native speakers of Italian perceive word combinations that vary along the continua of lexical fixedness and compositionality, suggesting interesting differences between the two participant groups.

References

- Cowie, A. P. (1994). Phraseology. In Asher, R. E. (ed.), *The Encyclopedia of Language and Linguistics*, Oxford, Oxford University Press, 3168-3171.
- Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations?. *International Review of Applied Linguistics*, 47 (2), 157-177.
- Gyllstad, H., & Wolter, B. (2016). Collocational processing in light of the phraseological continuum model: Does semantic transparency matter?. *Language Learning*, 66, 296-323.
- Ježek, E. (2005). *Lessico. Classi di parole, strutture, combinazioni*. Bologna, Il Mulino.
- Laufer, B., & Waldman, T. (2011). Verb-Noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61 (2), 647-672.
- Lenci, A., Montegmani, S., & Pirrelli, V. (2005). *Testo e computer: Elementi di linguistica computazionale*. Roma: Carocci.

- Masini, F. (2009). Combinazioni di parole e parole sintagmatiche. In Lombardi Vallauri, E., & Mereu, L. (eds), *Spazi linguistici. Studi in onore di Raffaele Simone*. Roma: Bulzoni, 191-209.
- Mitchell, J., & Mirella, L. (2010). Composition in Distributional Models of Semantics. *Cognitive Science*, 34 (8), 1388-1429.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24 (2), 223-242.
- Senaldi, M. S. G., Lebani, G., & Lenci, A. (2017). Determining the Compositionality of Noun-Adjective Pairs with Lexical Variants and Distributional Semantics. *Italian Journal of Computational Linguistics*, 3 (1), 43-58.
- Senaldi, M. S. G., Lebani, G., & Lenci, A. (2016). Lexical Variability and Compositionality: Investigating Idiomaticity with Distributional Semantics Model. *Proceedings of 12th Workshop on Multiword Expressions (MWE 2016)*: 21-31.
- Siyanova-Chanturia, A., & Spina, S. (2015). Investigation of native speaker and second language intuition of collocation frequency. *Language Learning*, 65 (3), 533-562.
- Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom-processing by native and non-native speakers. *Second Language Research*, 27, 251-272.

L1 and L2 Learners Keep Their Eyes on the Prize: Eye-tracking Evidence during Idiom Recognition

Sara D. Beck and Ruth Keßler

University of Tübingen

Idioms present native and non-native models of language processing with the challenge of addressing meaning that appears to be both non-compositional and compositional at the same time. Considering the idiom *to let the cat out of the bag*, it is possible to access the figurative meaning (to reveal a secret or surprise) in addition to the well-formed literal meaning. Idiomatic processing models suggest that once an idiom is recognized, its idiomatic meaning becomes available, but the role of literal language at this point needs further investigation. Although literal word meanings are often not relevant for the understanding of the figurative meaning, there has been evidence of literal word activation in varying types of idioms at idiom offset (e.g., Smolka, Rabanus & Rösler, 2007). To make matters more complex, there is evidence that native speakers are able to recognize and therefore predict the final word of the idiom, a factor that affects the activation of literal constituent meanings (e.g., Titone & Connine, 1994). However, this access has rarely been looked at in online processing in tandem with idiom recognition.

Keßler, Weber and Friedrich (in preparation) found ERP and eye-tracking evidence that adult native speakers both predict an idiom's final word during listening in addition to activating individual constituent meanings automatically during this process. These results are in line with previous evidence showing single word meaning activation (e.g., Sprenger, Levelt & Kempen, 2006). However, how might this process differ in language learners? Research on L1 language learners (children) is lacking in both research on prediction and literal constituent activation during processing. For L2 language learners, there is ample evidence for literal constituent activation (e.g., Beck & Weber, 2016.), but so far no positive evidence for prediction in recognition (e.g. Siyanova-Chanturia, Conklin & Schmitt, 2011).

In this study, we asked to what extent do language learners, namely L1 children (aged 13) and L2 adults, anticipate idioms' final constituents and activate their literal meanings.

We tested this question using a visual world eye-tracking paradigm on 26 (12;8 years) children with German L1 and 26 adults with proficient German L2 (English L1). Participants listened to neutral sentences containing idioms without their final word (*Hannes let the cat out of the ...*) and had to choose the correct idiom completion from one of four displayed words by giving a verbal response. Displayed words were (a) correct completions (*BAG*), (b) distractors semantically related to the correct completion (*BASKET*), and (c) two unrelated distractors (*ARM, STOMACH*). Only familiar German idioms were included in which the phrase-final words were highly predictable. The time-course and number of fixations toward all displayed words was measured, but only trials receiving correct responses were included in the analyses.

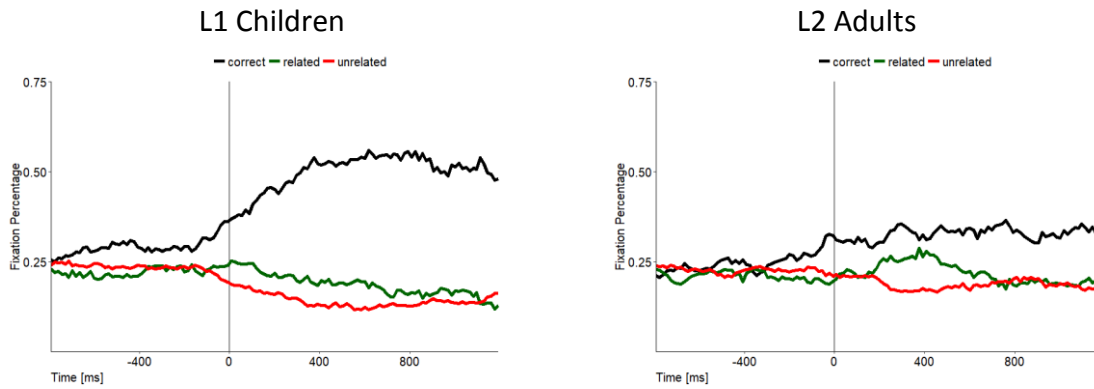


Figure 1. Fixation Percentage to correct completions (black), related distractors (green) and unrelated distractors (red). Eye fixations are aligned to the offset of the acoustic stimulus (0ms)

Like Keßler, Weber and Friedrich (in preparation), we found that both L1 children and L2 adults anticipate final constituent words and activate their literal meanings. However, there are differences in the time-course and overall fixation patterns between the groups tested. Namely, children are faster than L2 adults, but slower than L1 adults, in prediction and literal activation. Additionally, children show a literal activation pattern similar to adults, while L2 adults show overall more uncertainty and stronger literal activation.

References

- Beck, S. D., & Weber, A. (2016). Bilingual and Monolingual Idiom Processing Is Cut from the Same Cloth: The Role of the L1 in Literal and Figurative Meaning Activation. *Frontiers in psychology*, 7, 1350.
- Keßler, R. & Friedrich, C. K. (2018). Spoken idioms are decomposed: evidence from ERP and eye-tracking. Poster, AMLaP in Berlin, 6. September 2018.
- Holsinger, E. (2013). Representing idioms: Syntactic and contextual effects on idiom processing. *Language and speech*, 56(3), 373-394.
- Reuterskiöld, C., & Van Lancker Sidtis, D. (2013). Retention of idioms following one-time exposure. *Child Language Teaching and Therapy*, 29(2), 219-231.
- Sprenger, S. A., Levelt, W. J., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of memory and language*, 54(2), 161-184.
- Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research*, 27(2), 251-272.
- Titone, D. A., & Connine, C. M. (1994). Comprehension of idiomatic expressions: Effects of predictability and literality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(5), 1126.

Poster Talk Abstracts

1. Caitlyn Antal & Roberto G. de Almeida (Yale University), (Concordia University) 29
2. Patrick Brandt (Institut für Deutsche Sprache, Mannheim) 31
3. Lucia Busso, Alessandro Lenci, Florent Perek (ColingLab-Università di Pisa), (University of Birmingham) 33
4. Danielle dos Santos Wisintainer & Mailce Borges Mota (Federal University of Santa Catarina) 36
5. Elena. Kulkova, Martin Fischer (University of Potsdam) 39
6. Evelyn Milburn & Tessa Warren (Norwegian University of Science and Technology), (University of Pittsburgh) 40
7. Dario Paape & Shravan Vasishth (University of Potsdam) 42
8. Ali Safari (Hazrat-e Masoumeh University) 45
9. Yana Strakatova (University of Tübingen) 48
10. Ching Chu Sun (University of Tübingen) 49

*Number indicates poster location.

Composing the Meaning of Indeterminate Sentences in Context

Caitlyn Antal¹ & Roberto G. de Almeida²

¹ Department of Linguistics, Yale University, New Haven, CT, USA

² Department of Psychology, Concordia University, Montreal, QC, Canada

A remarkable feature of language comprehension is that we are able to build unambiguous propositional representations from often incomplete or indeterminate utterances. Much of what we represent seems to be warranted by what is actually said, but often we enrich what is said with what we assume to be intended by the speaker. This classical distinction—between what is actually said and what we represent as the intention of the speaker—runs deep in contemporary cognitive science (e.g., Cappelen & Lepore, 2005; Lepore & Stone, 2015; Borg, 2004; Carston, 2002; Grice, 1967). We address this long-held distinction by focusing on so-called “indeterminate” sentences and the kinds of propositions they yield. For instance, when a student says, “It took me all night to finish a book”, we seem to understand the event as referring to reading the book rather than writing, although the nature of the event is not specified and it is left to the hearer to determine what the speaker intended to convey by the utterance. But how are interpretations actually achieved? Are these sentences necessarily enriched semantically—i.e., by some default mechanism? And if multiple alternative interpretations are possible, what is the role of co-text (or context) in constraining the range of possibilities? As these questions indicate, the ability to comprehend indeterminate sentences seems to require the integration of multiple cognitive systems. These involve linguistic processing (syntax and lexical, compositional meaning) and those responsible for interfacing linguistic meaning with the immediate context and experiential knowledge (pragmatics).

We will discuss a project on indeterminate language processing designed to understand how these various linguistic and cognitive systems interface during natural language comprehension, how they might be implemented in the brain, and their implications for our understanding of compositionality (or lack thereof). We will discuss, in particular, two experimental studies bearing on how sentences with aspectual verbs, such as *x began y*, are interpreted in isolation and in supporting contexts. The experiments employed different techniques, manipulating online and offline measures of indeterminate sentence interpretation.

Thus far, several experimental studies investigating indeterminacy have used sentences in isolation and have relied on behavioural differences that arise at the complement noun phrase—e.g., longer reading times (e.g., McElree, et al., 2001; Pickering et al., 2005) in cases such as *x began the book* compared to *x read the book*. These studies have favoured a view of semantic enrichment known as “coercion” or “type coercion” which takes the noun complement (e.g., *book*) to provide semantic information to the resulting composition—thus claiming that a sentence such as *x began the book* yields an *enriched* composition like *x began reading the book*. We raise three issues against these studies. The first is the lack of consistent replication across laboratories (e.g., de Almeida, 2004; de Almeida et al., 2016). Second, these studies do not experimentally address the issue of how propositions might be enriched—limiting their position to attributing processing time differences between conditions to “coercion”. And third, studies have not investigated the role that rich contexts may have over time in the process of indeterminate sentence interpretation. This is important because context can bias for potential ‘events’ for the enriched composition we are supposed to build over time.

In our first study, we investigated whether aurally presented indeterminate sentences embedded in rich biasing contexts would trigger enriched event representations. We used a long-term memory (LTM) paradigm that relies on recovering the propositional content of sentences (). Subjects ($N=72$) heard 24 passages of continuous discourse and, after each one, responded to visual probe sentences representing either an indeterminate target (*Lisa began the book*), a contextually supported foil (...*began reading the book*) or a contextually unsupported control (...*began writing the book*). As in Sachs (1967), probes were presented either immediately after the original occurrence of the indeterminate sentence in context or following additional neutral context (25s). We found that accuracy (target acceptance, foil rejection) was high (~95%) at the immediate probe position. However, at the delayed (25s) probe position, indeterminate targets and contextually supported foils yielded the same low accuracy rates (~50%). These results were supported by a second, eye-tracking experiment in which subjects ($N=36$) read the same contextual passages and responded to the same visual probe sentences. First-pass reading times were the same for targets and contextually-supported foils at immediate and delayed probe positions. Second-pass reading times, however, were significantly longer for indeterminate targets compared to both foils at both probe positions.

Taken together, the studies suggest that enriched meaning may not be built into the sentential representation per se, but instead might occur beyond sentence composition. Interpretations of what a sentence means versus what it implies are computed at distinct levels of representation—the former at the syntactic-semantic interface, and the latter at the level of thought, driven by abductive-interpretative processes.

References

- Borg, Emma. 2004. *Minimal Semantics*. Cambridge: Cambridge University Press.
- Cappelen, H., Lepore, E. (2004). *Insensitive semantics*. Oxford: Blackwell Publishers.
- Carston, R. (2002). *Thoughts and utterances*. Oxford: Blackwell Publishers.
- de Almeida, R. G. (2004). The effect of context on the processing of type-shifting verbs. *Brain and Language*, 90, 249–61.
- de Almeida, R. G., Riven, L., Manouilidou, C., Lungu, O., Dwivedi, V., Jarema, G., Gillon, B. (2016). The neuronal correlates of indeterminate sentence interpretation: An fMRI study. *Frontiers in Human Neuroscience*, 10, 614.
- Grice, Paul. 1989. Logic and conversation. 1967/1987. In *Studies in the way of words*, Paul Grice. Cambridge, MA: Harvard University Press.
- Lepore, E. & Stone, M. (2015). *Imagination and convention: Distinguishing grammar and inference in language*. Oxford: Oxford University Press
- McElree, B., Traxler, M. J., Pickering, M. J., Seely, R. E., and Jackendoff, R. (2001). Reading time evidence for enriched composition. *Cognition*, 78, B17–B25.
- Pickering, M. J., McElree, B., and Traxler, M. J. (2005). The difficulty of coercion: a response to de Almeida. *Brain and Language*, 93, 1–9.
- Sachs, J. S. 1967. Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics* 2(9): 437-442.
- Sachs, J. S. 1974. Memory in reading and listening to discourse. *Memory and Cognition* 2(1A): 95-100.

Compositional after all: Evidence for Hidden Meaning as Suspended Interpretation

Patrick Brandt

Institut für Deutsche Sprache, Mannheim

(joint work with Petra B. Schumacher, Universität zu Köln)

The success of modern linguistic theory originates in the insight that the meanings of complex expressions can be derived from the meanings of their parts and the way these parts are composed. However, the currently dominating theories of the syntax-semantics interface hastily relegate important aspects of meaning which cannot readily be traced to visible structure to empty projecting heads non-reductively (mainstream Generative Grammar), to syntactic constructions holistically (Construction Grammar), or to lexical ambiguities (constraint-based approaches). We present an alternative, compositional analysis of the hidden aspectual-temporal, modal and comparative meanings of a range of productive constructions in German and other languages, specifically, pseudoreflexive, excessive and directional complement constructions. According to our proposal, the hidden meanings result from a repair mechanism that renders interpretable structures which violate certain basic requirements on the semantic side of the interface, in particular, the requirement that truth-conditional representations be non-contradictory.

The repair mechanism locally suspends the interpretation of a certain part of the literally coded meaning, which, interpreted locally, would give rise to contradiction. The meaning is left to be interpreted with respect to alternative parts of structure, predominantly, higher functional structure that accomplishes the indexical anchoring of propositional meanings in terms of times, worlds and standards of comparison. In contrast to putative coercion operations, then, nothing is forgotten or added, but a local surplus of meaning is pushed elsewhere in order to satisfy interface constraints, to note, exhaustivity of interpretation and non-contradictoriness. The dislocated meaning in question -- known as existential negative, viz. *something is not P* -- is part of a logical form of generalized DIFFERENCE, carried, among other, by the grammatical formatives *sich*, *-er* or *zu* as well as by syntactically transitive (nominative--accusative) configurations. E.g., adopting the "A not A" approach to comparatives (Ross 1969, Lewis 1972 and many since), excessive *zu* is presented as expressing the illegal reflexivization of genuinely asymmetric comparative structures, leading to dislocation of the 'negative' part of the comparative semantics to the purpose clause as sketched in (1).

- (1) Otto ist zu schwer, um Jockey zu sein
Otto is too heavy for jockey to be
'Otto is too heavy to be a jockey.'
≈ Otto is d-heavy and Otto is **not** d-heavy and Otto is a Jockey
→ Otto is d-heavy (and Otto is d-heavy) and Otto is **not** a Jockey
'Otto is so-and-so heavy and Otto is not a jockey in the actual world.'

Two ERP studies show that the processing of *zu*-excessive structures yields a positivity in the event-related-potential signature as characteristic of conceptual reorganization. However, the positivity occurs significantly earlier in the case of *zu*-excessives than in the case of privative predicates (*falsch* 'fake', *gefälscht* 'forged') or animal-for statue constructions (*wooden dove*) the processing of which arguably involves (intermediately) contradictory semantics as well (Schumacher et al. 2018). We propose that this difference is due to *zu* just representing the source of the problem rather than the repair itself: the purpose clause that in the case of

excessives forms the target for the dislocated meaning is structurally further away from its source than in the other cases. Regarding privative adjectives, the repair consists in manipulating the semantics of the nominal head that is directly adjacent to the modifying adjective (e.g., a *fake professor* is a professor in some but not in other respects); in the case of animal-for-statue constructions, the nominal head is both the source of the problem and the target of the repair (e.g., the animacy property of *dove* must be negated in construction with *wooden*).

In sum, we offer a reductive, compositional analysis of a range of apparent syntax-semantics mismatches that finds experimental support and that collects a range of constructions not hitherto grouped together. Interestingly from a theoretical perspective, our results argue in favor of the derivational (vs. representational) nature of the grammar engine as only under a derivational approach, the indexically potent interpretation of the locally uninterpreted logical form (viz. meaning) would seem a natural result.

References

- Brandt, Patrick. 2016. Fehlkonstruktion und Reparatur in der Bedeutungskomposition. *Linguistische Berichte* 248:395-433.
- Lewis, David. 1972. General semantics. In Donald Davidson & Gilbert Harman (eds.), *Semantics of Natural Language*, 169-218. Dordrecht: Reidel.
- Ross, John Robert. 1969. The deep structure of comparatives. Paper presented at *The First and Last Annual Harvard Spring Semantics Festival*.
- Schumacher, Petra B., Patrick Brandt & Hannah Weiland-Breckle. 2018. Online processing of “real” and “fake”: The cost of being too strong. In Elena Castroviejo Miro, Louise McNally & Galit Weidmann Sassoon (eds.), *The semantics of gradability, vagueness, and scale structure: Experimental perspectives*. Basel: Springer.

Coercing Italian: A Priming Experiment on Italian Valency Coercion

Lucia Busso¹, Alessandro Lenci¹, and Florent Perek²

¹ColingLab-Università di Pisa

²University of Birmingham

In Construction Grammar the basic units of language are constructions (henceforth: CXN), form-meaning pairs generally defined as “[a]ny linguistic pattern [...] as long as some aspect of its form or function is not strictly predictable from its component parts or from other CXNs recognized to exist.” (Goldberg,2006:5). In other words, CXNs carry an independent and non-compositional semantic content, i.e. which cannot be (entirely) derived from the meaning of its parts. In addition, also regular and frequent formulations are entrenched and stored as CXNs. A very compelling argument for the idea that CXNs are symbolic units with independent and non-predictable meanings comes from the phenomenon of *coercion* (Michaelis,2004; Lauwers&Willems,2011). This concept refers to a semantic incompatibility or mismatch between a lexical element and the constructional context in which it is embedded. The interpretation of such novel combination is said to be resolved with the CXN “coercing” the mismatching element into a new meaning, in line with the general semantic content of the general CXN. Hence, CXNs can change or override selected semantic features of lexical items, creating new non-compositional constructional meanings in the process.

In this study we investigate only a specific type of coercion phenomena: *valency coercion* effects, i.e. the ability of an argument structure CXN to combine freely with non-prototypical verbs, as in (1).

1. *John sneezed the napkin off the table*

In such cases, the deviant verb (e.g. “to sneeze”) is forced by the general CXN (e.g. Caused Motion) into a new interpretation (e.g. “to cause movement by means of sneezing”).

Over the past few years, an increasing number of studies have investigated coercion effects from experimental perspectives, such as psycholinguistic, neurolinguistic and corpus-based (*inter alia* Baggio et al.,2010; Gries et al,2010; Jezek&Lenci,2007; Lukasek et al,2017; Piñango et al.,2006). However, despite the pivotal role of coercion in Construction Grammar, not many languages apart from English have been investigated with respect to coercion phenomena (e.g. Gonzalez-Garcia,2007; Audring&Booji,2016).

In particular, no attempt has been previously made – to our knowledge –to address the psycholinguistic nature of valency coercion in Italian. This study aims to fill this gap in the literature and focuses on the processing of valency coercion structures in Italian. Romance languages are said to be *valency driven*, i.e. rely more heavily on the lexical meaning of verbs (Perek&Hilpert,2014). Hence, if we adopt a gradable approach to coercion rather than binary, we can conceive coercion as the result of a dynamic relation between lexical and constructional meaning. The degree of semantic (in)compatibility between verb and CXN determines the degree of perceived naturalness of the new formulation (Kemmer&Yoon,2013; Yoon,2016). Starting from this assumption, we present a priming experiment on Italian valency coercion, which aims to address the processing of the new, coerced meaning.

We test this question by means of a lexical decision task inspired by Johnson and Goldberg (2013), in which we present different target verbs preceded by coercion sentences, which serves as primes. Participants are asked to read the sentence on screen, and – after a brief fixation cross – to decide if the target verbs are existing words of Italian. The prime sentences belong to 4 CXNs: Caused Motion (CM), Dative (DT), Intransitive Motion (IM) and Sentential

or *Verba Dicendi* (VD). Stimuli were constructed using non-prototypical yet semantically compatible verbs for the target CXNs. For example, for the VD CXN the stimuli include Sound Emission verbs rather than prototypical verbs of Saying or Telling, as in (2). Stimuli were normalized for character length, i.e. the character lengths for the CXNs do not statistically differ ($p > 0.1$). They were evaluated for naturalness on a Likert scale (1-7) by 30 native Italian speakers (mean= 4.7, sd= 1.08). 15 sentences for each CXN were included in the test. Three types of target verbs were matched to the primes: lexical associate (LA), construction associate (CA), and unrelated (U) verbs. LAs are semantically similar to the main verb of the prime sentence, whereas CAs are prototypical verbs occurring in the prime CXN. U verbs serves as a mean of comparison for the two categories of interest. For example, the target verbs for (2) are *canticchiare* (to hum, LA), *dire* (to say, CA) and *invecchiare* (to age, U). Verbs too were normalized for character length ($p > 0.5$).

2. *Giovanni fischietta che forse verrà domani*
(Giovanni whistles that he will maybe arrive tomorrow)

LAs were selected with a production task by 35 native Italian speakers; Us were found using cosine (dis)similarity (cosine distance $> .75$) on *WEISS* (Word-Embeddings Italian Semantic Spaces, Marelli, 2017). CAs were retrieved on *ItWac* (Baroni et al. 2009). To justify the lexical decision task, a set of 15 filler sentences paired with 15 nonce verbs were included. The experiment is between 3 groups of subjects, of 13 participants each. Each group sees all the primes, but in a different combination with the targets. Only filler sentences and non-words are constant. Table 1 summarises the experiment structure.

EXPERIMENT STRUCTURE			
I GROUP	5 sentences + LA	5 sentences + CA	5 sentences + U
II GROUP	5 sentences + CA	5 sentences + U	5 sentences + LA
III GROUP	5 sentences + U	5 sentences + LA	5 sentences + CA

Table 1: Structure of the experiment for each CXN in each group. The data set is composed of 60 primes + 15 fillers

We analysed the data with linear mixed modelling. Results confirmed our hypotheses; namely, reaction times (RTs) of congruent associations of prime-target (LA or CA) resulted significantly faster than RTs of U verbs. Additionally, CA RTs are even faster than LA, signalling a primary role of CXN in the interpretation process. We further refined the analysis by fitting a second mixed model correlating RTs with Likert scores of prime sentences (see above). It emerged that RTs of LA targets are inversely correlated with sentence grammaticality, whereas CA targets are not affected by it.

We interpret these findings as compelling evidence of the importance of both verb and CXN semantics, and of the primary role of CXN in the processing and interpretation of the new holistic meaning of valency coercion phenomena in Italian.

Selected references

Audring, J., & Booij, G. (2016) Cooperation and coercion. *Linguistics*, 54(4), 617–637.
 Baggio, G., et al. (2010) Coercion and compositionality. *Journal of Cognitive Neuroscience* 22(9):2131-40.
 Baroni, M., et al. (2009) The WaCky Wide Web: A Collection of Very Large Linguistically Processed Web-Crawled Corpora. *Language Resources and Evaluation* 43(3): 209-226
 Goldberg, A. (2006) *CXNs at Work: the nature of generalization in language*. Oxford University Press: Oxford.
 Gonzalez-Garcia, F. (2007) ‘Saved by the reflexive’: evidence from coercion via reflexives in verbless complement clauses in English and Spanish. *Annual Review of Cognitive Linguistics* 5, pp. 193–238.
 Gries, S., et al. (2010) Converging evidence II: More on the association of verbs and CXNs. In Sally Rice and John Newman (eds.), *Empirical and experimental methods in cognitive/functional research*, 59-72. Stanford, CA: CSLI Publications
 Jezek, E., & Lenci, A. (2007) When GL meets the corpus: a data-driven investigation of semantic types and coercion phenomena. *Proceedings of GL*, 10–11.

- Johnson, M. A., & Goldberg, A. E. (2013) Evidence for automatic accessing of CXNal meaning: Jabberwocky sentences prime associated verbs. *Language and Cognitive Processes*, 28(10), 1439–52.
- Lauwers, P., & Willems, D. (2011) New reflections on coercion. *Linguistics*. De Gruyter Mouton:Berlin.
- Lukasek, J., et al. (2017) The Semantic Processing of Motion Verbs: Coercion or Underspecification? *Journal of Psycholinguistic Research* 46(4), 805–825.
- Marelli, M. (2017) Word-embeddings Italian Semantic Spaces: A semantic model for psycholinguistic research. *Psihologija*, 50(4), 503-20.
- Michaelis, L., (2004) Type Shifting in CXN Grammar: An Integrated Approach to Aspectual Coercion". *Cognitive Linguistics* 15, pp. 1-67.
- Perek, F & Hilpert, M., (2014) CXNal tolerance: Are argument structure CXNs equally powerful across languages?". In *CXNs and Frames* 6/2, pp. 266-304.
- Piñango, M. M., et al. (2006) Time-Course of Semantic Composition: The Case of Aspectual Coercion. *Journal of Psycholinguistic Research*, 35(3), pp 233–244.
- Pylikkanen, L., et al., (2009) The Anterior midline field: Coercion or decision making? *Brain and Language* 108: 184-190.

Figurative Phrasal Verb Processing: Evidence from Bilingual Sentence Reading

Danielle dos Santos Wisintainer and Mailce Borges Mota

Federal University of Santa Catarina - UFSC, Brazil

Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, Brazil

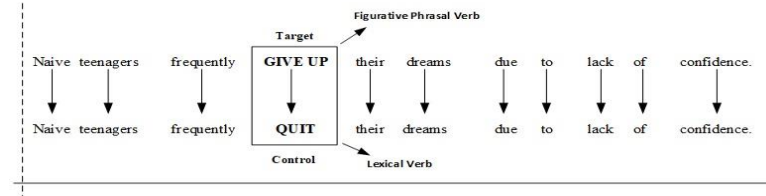
The aim of our study is to investigate the online processing of figurative phrasal verbs (e.g. break up) by nonnative and native speakers of English. Kecskes (2015) claims that phrasal verbs can be considered examples of formulaic language which tend to convey holistic meanings operating as a single semantic unit. Thim (2012) defines phrasal verbs as a verb and a particle which is typically homonymous with an adverb or a preposition. Phrasal verbs can be divided into two types: those with compositional meanings and those with non-compositional meanings. In phrasal verbs with compositional meanings, the verb combines with a particle and the whole construction is transparent from the meaning of its constituents; moreover, the particle can introduce the concept of a goal or an endpoint to durative situations (e.g., finish up). In phrasal verbs with non-compositional meanings, it is not possible to infer the meaning of the construction from the meaning of their separate elements, that is their meaning is non-transparent (e.g., figure out).

Evidence suggests that nonnative and native speakers of English process phrasal verbs differently. For instance, using an online reading task (reaction time), Matlock and Heredia (2002) investigated the processing of figurative phrasal verbs (e.g., Paul went over the exam with his students) and their identical verb-preposition combinations used literally (e.g., Paul went over the bridge with his bicycle). The authors found that, for native speakers and early bilinguals, the figurative meaning is a highly familiar meaning always activated before the literal meaning. However, for the late bilingual group, the literal meaning (verb-preposition combinations) was processed first. These results are in line with Littlemore and Low (2006, p.3 and 4), who explain that learners may approach figurative language analytically. They call this approach “figurative thinking”. The authors suggest that nonnative speakers take more time processing figurative language due to the fact that they try to analyze each component of the figurative multiword item (e.g. to figure out) and this slows down their processing, mainly, in those figurative items which are seen for the first time by nonnative speakers. However, Paulmann, Ghareeb-Ali and Felser’s study (2015) favors the figurative meaning first hypothesis (Gibbs, 1980). The authors investigated the cognitive mechanisms underlying the processing of phrasal verbs by monolingual (native English) and bilingual (native Arabic) speakers, in an event related potential study (ERPs). They compared ERPs elicited in response to when and how figurative (I heard that Mr. Smith ran over the old farmer early this morning) and literal meanings (I heard that Mr. Smith ran over the old bridge early this morning) are accessed. Their results showed that monolinguals and bilinguals used similar processing mechanisms when processing phrasal verbs. In addition, figurative sentence interpretations were favored by bilinguals.

Given the controversial results obtained in the bilingual literature with regard to the processing of phrasal verbs and the scarcity of this kind of studies, we propose our investigation. In our eye-tracking study, we delved into how nonnative speakers and native speakers of English process figurative phrasal verbs and lexical verbs (one-word verb) during the reading of sentences. Twenty-four volunteers participated in the present study and were divided into two groups, one consisting of twelve advanced learners of English as L2, native speakers of Brazilian Portuguese (the BP Group) and the other consisting of twelve native speakers of English (the NE Group). The sentence processing task consisted of 96 sentences in

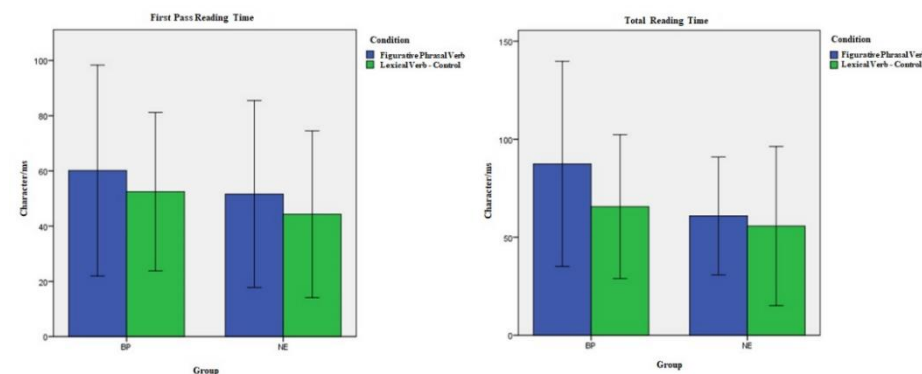
English: 16 sentences contained figurative phrasal verbs, 16 sentences contained lexical verbs, and 64 sentences consisted of filler sentences. Figure 1 presents an example of a sentence with a figurative phrasal verb and its control lexical verb showing their position in the sentence.

Figure 1- Sentences with figurative phrasal verb and its lexical verb show the position of the target and the control verb.



A Repeated Measures ANOVA revealed that nonnative speakers read all the conditions (figurative phrasal verbs and lexical verbs) slower than native speakers of English did on total reading time ($p < .05$). A linear mixed effects model was fitted with 12 nonnative speakers of English and condition type (figurative phrasal verb vs. control) as fixed effects. There was significant effect of condition type ($p < .05$) on total reading time. As predicted, our results show that the participants of the BP group reread and reanalyzed more figurative phrasal verbs than lexical verbs (Figure 2).

Figure 2 – Results of the eye-tracking study on figurative phrasal verbs for nonnative speakers (BP) and native speakers of English (NE).



A possible explanation of the results is that nonnative speakers of English engender greater cognitive effort in processing figurative phrasal verbs. This might be an evidence that figurative phrasal verbs are not processed as a whole chunk by nonnative speakers of English (Littlemore and Low, 2006). Our study provides additional support to the view put forward by Siyanova-Chanturia, Conklin and Schmitt's study (2011) that figurative meaning causes processing difficulty after the recognition point has been reached.

References

- Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition*, 8(2), 149-156.
- Kecskes, I. (2015). Is the idiom Principle Blocked in Bilingual L2 Production? In Heredia, Roberto R., and Anna B. Cieślicka (Eds.), *Bilingual Figurative Language Processing* (pp. 28-52). 1st ed. New York: Cambridge University Press.
- Littlemore, J & Low, G. (2006). What is "Figurative Thinking"? In J. Littlemore & G. Low (Eds.), *Figurative Thinking and Foreign Language Learning* (pp. 3-22). Palgrave Macmillan.
- Matlock, T. & Heredia, R. (2002). Understanding phrasal verbs in monolinguals and bilinguals. In Heredia, R.R.; Altarriba, J., (Eds.), *Bilingual sentence processing* (pp. 251-74). Elsevier.

- Paulmann, S., Ghareeb-Ali, Z. & Felser, C. (2015). Neurophysiological Markers of Phrasal Verb Processing: Evidence from L1 and L2 Speakers. In Heredia, Roberto R., and Anna B. Cieślicka (Eds.), *Bilingual Figurative Language Processing* (pp. 245-267). 1st ed. New York: Cambridge University Press.
- Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and nonnative speakers. *Second Language Research*, 27, 251-272.
- Thim, S. (2012). *Phrasal Verbs: The English Verb-Particle Construction and its History* (Topics in English Linguistics 78). Berlin and New York: De Gruyter Mouton.

Processing Non-Literal Meaning: A Grip Force Study

Elena Kulkova^{1*} and Martin H. Fischer²

¹ University of Potsdam

² University of Potsdam

*Corresponding author

According to an embodied cognition view, conceptual knowledge is grounded in sensorimotor experience. This argument extends to action-associated language processing – sensorimotor cortical areas that are active during a concept’s acquisition establish interconnections with the brain areas that underpin processing language that is semantically related to the concept. However, the evidence of the extent to which sensorimotor systems are involved in *abstract* and *non-literal* language processing is still inconclusive. For example, activation in motor regions was demonstrated for isolated action verbs and, partially, for literal action sentences, but not for idiomatic expressions (Raposo et al., 2009), for literal and metaphoric, but not idiomatic action sentences (Cacciari et al., 2011). Contrarily, Boulenger et al. (2009; 2012) showed that processing of both literal and idiomatic action sentences involving leg- and arm-related verbs leads to a somatotopic activation in the primary motor cortex, thus providing evidence for a partially compositional nature of idioms. In our study, we use grip force sensors, devices that allow bi-manual measurement of subtle fluctuations in the force of gripping, to investigate involvement of motor areas in the left and the right hemispheres in idiom and metaphor processing (Aravena et al., 2012). We measure spontaneous changes in grip force triggered by the semantic features of hand-related and hand-unrelated action verbs and state verbs, embedded in literal, idiomatic, and metaphoric sentences. Grip strength is expected to be related to the amount of “literal” action-related semantics preserved in a sentence context that may influence the intensity of a mental simulation of action properties during meaning processing. Thus, literal meaning, which is more grounded compared with metaphoric, idiomatic or abstract meaning, should lead to a larger grip force. Compared to literal sentences, we hypothesize a decrease in grip force elicited by metaphoric expressions because they use motion verbs at a higher level of abstraction. Even more pronounced decrease is hypothesized for idioms which establish an arbitrary relationship between the literal and the global figurative meaning. Overall, the study aims to enhance our understanding of embodied and simulation-based aspects of figurative meaning comprehension.

References

1. Aravena, P., Delevoye-Turrell, Y., Deprez, V., Cheylus, A., Paulignan, Y., Frak, V., & Nazir, T. (2012). Grip force reveals the context sensitivity of language-induced motor activity during “action words” processing: evidence from sentential negation. *PLoS One*, 7(12), e50287.
2. Boulenger, V., Hauk, O., & Pulvermüller, F. (2009). Grasping ideas with the motor system: semantic somatotopy in idiom comprehension. *Cerebral cortex*, 19(8), 1905-1914.
3. Boulenger, V., Shtyrov, Y., & Pulvermüller, F. (2012). When do you grasp the idea? MEG evidence for instantaneous idiom understanding. *Neuroimage*, 59(4), 3502-3513.
4. Cacciari, C., Bolognini, N., Senna, I., Pellicciari, M. C., Miniussi, C., & Papagno, C. (2011). Literal, fictive and metaphorical motion sentences preserve the motion component of the verb: a TMS study. *Brain and language*, 119(3), 149-157.
5. Raposo, A., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2009). Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia*, 47(2), 388-396.

Idioms Show Effects of Meaning Relatedness and Dominance Similar to those Seen for Ambiguous Words

Evelyn Milburn¹ and Tessa Warren²

¹Norwegian University of Science and Technology

²University of Pittsburgh

The tension between word-level and phrase-level processing has been a central focus of idiom research (e.g. Gibbs, 1980). But this tension may not be limited to idiom processing; literal multiword phrases can also be processed as units. Evidence for this comes from the fact that more frequent literal strings (lexical bundles) are processed more quickly (e.g. Arnon & Snider, 2010) and remembered more accurately (Tremblay et al., 2011) than less frequent strings. Indeed, findings of facilitated processing for non-adjacent frequent collocations (Vilkaite, 2016) suggest that the language system may process at the word and bundle levels simultaneously. This can account for classic evidence that idioms are not *only* processed as unitary chunks—in particular, the fact that modifying an idiom does not always impair its processing (e.g. Nordmann et al., 2014). This raises the possibility that the language system may process flexibly across different grain sizes (i.e. single words or multiword chunks) and literality statuses. If this hypothesis is correct, effects that are evident in the processing of ambiguous literal units like words should also appear for ambiguous figurative units like idioms. Indeed, there are hints that this may be the case: higher meaning relatedness facilitates processing for polysemes over homonyms (Klepousniotou, 2002), and for more decomposable idioms (Titone & Connine, 1999).

We tested this hypothesis by investigating whether effects of meaning dominance and meaning relatedness previously found for ambiguous words also appear for idioms. We drew three predictions from the literature on ambiguous words. First, following Foraker and Murphy (2012) and Brocher and colleagues (2018), meaning dominance and meaning relatedness would interact. Second, idioms with more related figurative and literal meanings should be read more quickly and re-read less (c.f. Klepousniotou, 2002). Third, idioms with highly dominant figurative meanings should be read more slowly in literal-biasing contexts (c.f. Brocher et al., 2016).

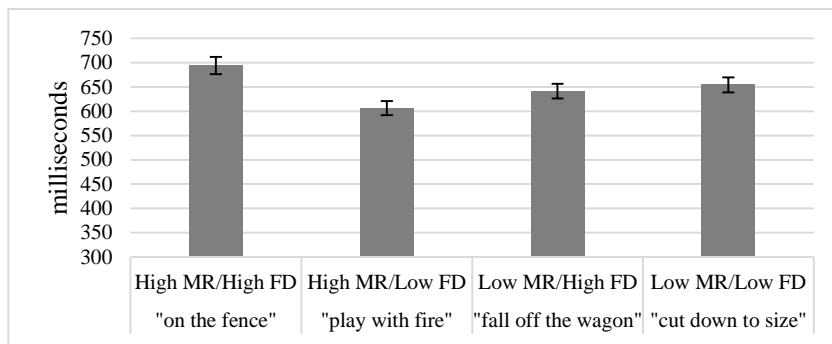
The experimental design was based on Brocher et al. (2016) and Foraker and Murphy (2012). We embedded idioms that varied in meaning relatedness and figurative meaning dominance into contexts biasing either their literal or figurative interpretations. We did not include post-critical regions disambiguating the idioms. This meant that any observed effects, even in re-reading behaviors, would reflect relatively uninterrupted processing rather than revision processes.

College-aged adults (n=36) read 45 three-sentence passages intermixed with 80 fillers while their eyes were tracked. Passages consisted of a biasing context sentence, a sentence containing an idiom, and a short wrap-up sentence (below; idiom is underlined):

(1A) Ben cancels our plans all the time. He always gets cold feet no matter what. It's really frustrating. (figuratively-biased context)

(1B) Ben wears thick socks all the time. He always gets cold feet no matter what. It's really frustrating. (literally-biased context)

We normed idiom figurative meaning dominance and idiom meaning relatedness; these values were treated as continuous factors in LMER analyses. Norms also verified that context sentences biased the intended meaning of the idiom, passages were comparably natural across conditions, and idioms were familiar to participants.



Fixation-based measures on the idiom region showed an interaction between meaning dominance and meaning relatedness: go past time, rereading time, and total time were longest when the idiom had highly

related literal and figurative meanings but a highly dominant figurative meaning (GP: $\beta=76.96$; $p<.05$; RR: $\beta=26.60$; $p<.05$; TT: $\beta=65.65$; $p<.05$, see Figure 1 showing median splits of TT raw data). As predicted, higher meaning relatedness was associated with facilitation on the idiom region in all of the above eye-tracking measures. However, there were no effects of context bias in any region for any eye-movement measure. This is likely because we did not include a disambiguating post-critical region, which is where such effects usually appear.

The current results suggest that ambiguous idioms and single words are processed similarly. The fact that the facilitative effect of meaning relatedness in idioms parallels the polysemy advantage observed for ambiguous words hints that processing facilitation associated with idiom decomposability and transparency (e.g. Titone & Connine, 1999) may be driven by their close relationship to meaning relatedness. If idioms and ambiguous words are indeed processed similarly, the results of the present study open new questions about idiom representation. Research into lexical ambiguity resolution has begun to investigate whether different lexical meanings are stored separately or together, and if together, whether the representation is underspecified or based on core features (Eddington & Tokowicz, 2015). These questions may also be relevant to idiom processing.

References

- Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *J. of Mem. & Lang.*, 62(1), 67–82.; Brocher, A., Chiriacescu, S. I., & von Heusinger, K. (2018). Effects of information status and uniqueness status on referent management in discourse comprehension and planning. *Disc. Pro.*, 55(4), 346–370.; Brocher, A., Foraker, S., & Koenig, J.-P. (2016). Processing of irregular polysemes in sentence reading. *J of Exp. Psy.: Lea., Mem., and Cog.*, 42(11), 1798–1813.; Eddington, C. M., & Tokowicz, N. (2015). How meaning similarity influences ambiguous word processing: the current state of the literature. *Psy. Bull. & Rev.*, 22(1), 13–37.; Foraker, S., & Murphy, G. L. (2012). Polysemy in sentence comprehension: Effects of meaning dominance. *J. of Mem. and Lang.*, 67(4), 407–425.; Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Mem. & Cog.*, 8(2), 149–156.; Klepousniotou, E. (2002). The Processing of Lexical Ambiguity: Homonymy and Polysemy in the Mental Lexicon. *Brain and Lang.*, 81(1–3), 205–223.; Nordmann, E., Cleland, A. A., & Bull, R. (2014). Familiarity breeds dissent: Reliability analyses for British-English idioms on measures of familiarity, meaning, literality, and decomposability. *Acta Psychologica*, 149, 87–95.; Titone, D. A., & Connine, C. M. (1999). On the compositional and noncompositional nature of idiomatic expressions. *J. of Prag.*, 31(12), 1655–1674.; Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing Advantages of Lexical Bundles: Evidence From Self-Paced Reading and Sentence Recall Tasks: Lexical Bundle Processing. *Lang. Lea.*, 61(2), 569–613.; Vilkaite, L. (2016). Are nonadjacent collocations processed faster? *J. of Exp. Psy.: Lea., Mem., and Cog.*, 42(10), 1632–1642.

Meaning Reversal in Multiply Negated Sentences – New Evidence from Two Tasks

Dario Paape and Shravan Vasishth

Department of Linguistics, University of Potsdam

Consider the sentence *No head injury is too trivial to be ignored*. The correct meaning is nonsensical (*Head injuries should be ignored, even if they seem trivial enough to be treated*). However, the sentence is persistently misinterpreted to mean *Treat even seemingly trivial head injuries*, in violation of the compositionality principle [1,2,3,4]. This is the ‘depth charge’ illusion. The standard explanation is in terms of negation overload and subsequent misinterpretation of the final verb *ignore*; call it the *composition-failure account* [1]. Expt. 1 tests this account and shows that the illusion arises even before the verb *ignored* is encountered. Expt. 2 investigates the claim that the illusion is specific to constructions of the type *No X is too Y to ...* [5]. We show that the illusion generalizes to related constructions, and depends on the particle *too* being present.

Experiment 1 – Is the final verb the locus of the illusion?

In order to find out whether the source of the illusion lies before the verb, we conducted a sentence completion study in German with 60 participants and 32 sentences. The stimuli were truncated after *to*, as shown below. The adjectival negation condition is a control condition for which no illusion is expected.

DOUBLE NEGATION	Keine		Kopfverletzung	ist	zu	ungefährlich,	um	...
	no		head injury	is	too	un-dangerous	to	
ADJ. NEGATION	Manch	eine	Kopfverletzung	ist	zu	ungefährlich,	um	...
	some	a	head injury	is	too	un-dangerous	to	

Under a compositional interpretation, ... *be treated* would be a sensible continuation for both sentences. However, if the depth charge illusion occurs, participants should instead supply a continuation like ... *be ignored* for the double negation sentence. If the illusion appears in this task, it is unlikely that the verb is the main culprit.

Because this was a production task, participants supplied a variety of continuations. For analysis, it was necessary to group these responses according to a binary criterion (illusion/no illusion). As there is no established method of diagnosing the appearance of the illusion in the absence of explicit comprehension tests, we devised two different coding schemes. The 21 coders (9/12 per scheme) were blind to the experimental manipulation, given that the preambles were not presented to them (see below). Scheme A asked whether a combination such as *head injury – see a doctor about* (when the latter was the supplied continuation) indicated that the head injury was considered to be ‘of importance’ or not. This scheme was used because potentially low importance of the subject (e.g., *ignore trivial head injuries*) typically accompanies the illusion. For Scheme B, coders judged whether the continuation supplied for the experimental sentence fit with a negation- and quantification-free sentence, as shown below.

NO NEGATION	Diese	Kopfverletzung	ist zu	gefährlich,	um	[behandelt/ignoriert]	zu werden.
	this	head injury	is too	dangerous	to	treated/ignored	to get

Under the illusion, the illicit meaning of the double negation sentence (*Any head injury is too dangerous to be ignored*) is compatible with that of the no negation sentence (*This head injury*

is too dangerous to be ignored). On the other hand, the compositionally correct continuation of the double negation sentence, ... *be treated*, does not fit with the no negation sentence. Inter-coder agreement was higher for Scheme A (Fleiss' κ : 0.77) than for Scheme B (Fleiss' κ : 0.49). According to both coding schemes, participants produced illusion-signaling continuations like ... *be ignored* more often in the double negation than in the adjectival negation condition (Scheme A – mean difference: 77%, Bayesian credible interval (95% CrI): [65%, 86%] ; Scheme B – diff: 67%, CrI: [51%, 78%]). Contrary to previous speculation [1], the appearance of the illusion in the completion task suggests that the origin of the depth charge illusion lies before the verb, whose prediction is corrupted in the presence of too many negations (see also [4]).

Experiment 2 – Is the illusion construction-specific?

Our second study investigated the recent claim by Fortuin [5] that *No X is too Y to ...* is a crystallized construction which assumes the reversed meaning if the larger discourse context implies that meaning. While Fortuin notes that the construction has a number of relatives (e.g., *No X is so Y that ...*), it is possible that these do not show the illusion if the non-compositional meaning of *No X is too Y to ...* is highly idiosyncratic.

In order to study the generalizability of the depth charge phenomenon, we identified two candidate constructions besides *No X is too Y to ...*. The first of these (see below) also involves the particle *zu*, 'too', which carries an implicit negation (*too Y to Z* → *not Z*), while the other one contains *so*, 'so', instead. As *so* does not imply negation, an explicit negation appears in the embedded clause to keep the semantics matched. Besides the depth charge configuration with both global and adjectival negation, we also tested sentences without global negation as a control, so that a 2x3 design was employed.

DEPTH CHARGE	Keine		Kopfverletzung	ist ...					
	no		head injury	is					
CONTROL	Manch	eine	Kopfverletzung	ist ...					
	some	a	head injury	is					
TOO ... TO			... zu	ungefährlich,	um	ignoriert	zu werden.		
			too	un-dangerous	to	ignored	to get		
TOO ... AS THAT			... zu	ungefährlich,	als	dass man sie	ignorieren könnte.		
			too	un-dangerous	as	that one it	ignore could		
SO ... THAT			... so	ungefährlich,	dass	sie nicht	ignoriert werden sollte.		
			so	un-dangerous	that	it notignored	get should		

Sixty participants read 30 sentences of this type in whole-sentence presentation and rated on a 1–7 scale whether they seemed sensible (7 = perfectly sensible). The classical *too ... to* construction and the *too ... as that* construction showed an equally strong illusion, that is, an equal increase in ratings in the depth charge condition. These two constructions also had similar whole-sentence reading times, which showed no evidence of an effect of the additional negation. The *so ... that* construction showed a different picture, such that reading times were increased in the depth charge condition, but the increase in ratings was much smaller.

Expt. 2 suggests that the depth charge illusion is not limited to the *No X is too Y to ...* construction, but can also occur in a similar configuration, as long as the particle *too* is present. We propose to maintain the composition-failure account, and suggest that the illusion arises due to the compositional meaning of *too* [6] being changed in one of two ways: either *too* is essentially interpreted as *enough*, reversing its meaning, or global negation is illicitly applied to both the copula and the final verb, as opposed to only the copula. More broadly, our results

imply that seemingly extreme examples of non-compositionality like the depth charge illusion may be due to comparatively small failures in the normal semantic computation.

References

[1] Wason & Reich (1979). *Q J Exp Psychol*. [2] Natsopoulos (1985). *J Psycholinguist Res*. [3] Kizach, Christensen & Weed (2016). *J Psycholinguist Res*. [4] O'Connor (2015). *USC diss*. [5] Fortuin (2014). *Cogn Linguist*. [6] Meier (2003). *Nat Lang Semant*.

English–Persian Translation: Non-compositional Meaning and Semantic Frames

Ali Safari

Hazrat-e Masoumeh University

As it is observed in natural languages, there are some aspects of meaning for which no formal element is present in the construction. For example, in one of the variants of the locative alternation, there is holistic meaning but no word or phrase in the sentence is responsible for this meaning. Other constructions may have holistic properties that cannot be reduced to properties of their individual constituents. This applies to morphological constructions as well. A clear example is that reduplication structures often indicate some form of increase of a semantic property denoted by the base word. For instance, in Persian the intensifying forms of adjectives is created by adding a full copy of the base adjective connected with Ezafe morpheme (e.g. xub ‘good’ – xub-e xub ‘very good’). The meaning of intensifier does not derive from one of the constituents of xub or -e, but is evoked by the copying construction as such. In the examples above the whole construction is responsible for the holistic meaning. As an example, in English, Double Object Construction contains the meaning of transfer without any explicit constituent bearing such meaning. Again, the meaning of transfer is evoked by the constructional configuration as a whole and the construction itself is responsible for the transfer meaning.

When the target language lacks a construction corresponding to that of the source language, a translator machine like Google Translate fails to translate the construction correctly. This problem arises especially when the construction is not compositional i.e. there are aspects of meaning for which there is no overt element in the construction. In this paper, I will discuss two examples of this problem, i.e., the English Resultative Construction and Double Object Construction when translated into Persian. Both constructions contain some aspects of meaning which cannot be attributed to an element in the construction and the whole construction is responsible for the meaning. Double object construction contains the meaning of transfer and Resultative construction contains the holistic meaning. These constructions are not lexicalized in Persian and should be expressed using paraphrase constructions. For example (1a) is the English resultative construction but Persian lacks such construction. (1b) is the paraphrase to the resultative construction and it is similar to Persian construction. (2a) is the English double object construction which Persian lacks. (2b) is the dative construction which is considered a paraphrase to (2a) and it is similar to Persian construction. Google Translate routinely fails to translate these constructions into Persian (3) but the translation improves when paraphrases are used (4).

- (1) a. John licked his plate clean. (no construction in Persian)
b. John licked his plate and cleaned it. (John bošqabaš ra lisid væ tæmiz kærd)
- (2) a. John gave Mary a book. (no construction in Persian)
b. John gave a book to Mary. (John ye ketab be Mary dad)

(3) a. John licked his plate clean.

Natural translation:

John lisid bošqab-aš ra væ an ra tamiz kærd.

John licked plate-his Om and it OM clean did

John licked his plate and cleaned it.

Google Translate:

John lisd ?u ra bořqab tæmiz konæd.
 John licked his OM plate clean do
 (ungrammatical but comprehensible)

b. John gave Mary a book.

Natural translation:

John yek ketab be Mary dad.
 John one book to Mary gave
 John gave a book to Mary.

Google Translate

John Mary dad yek ketab æst.
 John Mary gave one book is
 (ungrammatical and incomprehensible)

(4)

a. John licked his plate and cleaned it.

Google Translate

John lisd bořqabe xod ra væ an ra tæmiz kærd.
 John licked plate own OM and it OM clean did
 John licked his plate and cleaned it.

b. John gave a book to Mary.

Google Translate

John ketab ra be Mary dad.
 John book OM to Mary gave
 John gave a book to Mary.

A *Frames Approach* to syntax and semantics inherently leads to predictions about what types of situation are amenable to paraphrase. This approach is based on the FrameNet¹ project, an online lexical database built on the principles of *FrameSemantics* (Fillmore 1982, 1985, 1994, Fillmore and Atkins 1992). FrameNet incorporates diverse types of linguistic information, more than any other online resources of a similar kind, and it is particularly utilitarian in investigation of paraphrase (Hasegawa et al. 2010, 2011, forthcoming).

In FrameNet, a *frame* is a schematic understanding of types of events, situations, individuals, and things, including the participants, props, parts, and their relations to each other and to the larger situation. Words are understood with respect to a particular frame as background. FrameNet groups words with the same background into frames, and defines these frames and the parts of the frame in prose. For instance, the words, *auction*, *retail*, *sell*, and *seller* are grouped in the **Commerce_sell** frame, for they all have as background a commercial transaction from the point of view of the seller of goods. Each frame has a number of *frame elements* (FEs), which can be thought of as semantic roles. For instance, the **Attaching** frame involves at least three entities, an AGENT, an ITEM, and a GOAL; the AGENT causes the ITEM to be connected to the GOAL.

The Frames Approach to syntax and semantics predicts that these situations must be paid keen attention in translation when: (i) the target language lacks a construction corresponding to that of the source language; (ii) the target language has such a structure but it is

¹ Information at: <http://framenet.icsi.berkeley.edu/>

dispreferred; (iii) the source language and target language have encoded asemantic domain with different frames; and (iv) those frames may be expressed in very different syntactic ways. In relatively well-defined areas of the lexicon and grammar, it is possible to set out a more-or-less definitive list of principles about which structures are preferred by which languages. Then, given a source sentence, one can generate a large range of typologically-attested grammatical or lexical patterns within the source language (even if some of the sentences are unnatural), and select the paraphrase whose structure best matches the typological profile of the target language for this domain. The end resulting translation, whether produced by MT or not, will likely improve.

This paper has shown that there is a clear place for principled linguistic analysis in modern translation studies. Linguists have established many ways of analyzing the source of similarities and differences between such typologically divergent languages as English and Persian. The Frames Approach is especially powerful as it can recognize deep semantic correspondences and simultaneously handle syntactic and lexical differences. The tractability of such complex examples in this approach shows that there is a rich, untapped opportunity for augmenting MT with linguistic and typological analysis, especially when Frames and Constructions take center stage.

References

- Fillmore, Charles. 1982. Frame semantics. In *Linguistics in the Morning Calm*, ed. Linguistic Society of Korea, 111-137. Seoul: Hanshin.
- Fillmore, Charles. 1985. Frames and the semantics of understanding. *Quaderni di Semantica* 6: 222-254.
- Fillmore, Charles. 1994. The hard road from verbs to nouns. In *In Honor of William S-Y. Wang: Interdisciplinary Studies on Language and Language Change*, eds. Mathew Chen and Ovid Tzeng, 105-129. Taipei: Pyramid Press.
- Fillmore, Charles, and B.T.S. Atkins. 1992. Towards a frame-based organization of the lexicon: The semantics of RISK and its neighbors. In *Frames, Fields, and Contrast: New Essays in Semantics and Lexical Organization*, eds. Adrienne Lehrer and Eva Kittay, 75-102. Hillsdale: Lawrence Erlbaum Associates.
- Hasegawa, Yoko, Russell Lee-Goldman, Kyoko Ohara, Seiko Fujii, and Charles J. Fillmore. 2010. On expressing measurement and comparison in Japanese and English. In *Contrastive Construction Grammars*, ed. Hans C. Boas, 169-200. Amsterdam: John Benjamins.

Automatic Collocation Identification Using Word Embeddings

Yana Strakatova

Department of General Linguistics, University of Tübingen

Collocations such as ‘black coffee’ and ‘French window’ are multiword expressions whose constituents show a high degree of statistical association and whose meaning is not entirely semantically transparent, i.e. one of the constituents carries a special meaning found only in this combination. Collocations are in the grey area between free phrases like ‘black car’, where both of the components are fully transparent, and idiomatic expressions such as ‘black sheep’, where all the constituents are opaque. Over the last decades, there has been a lot of research on the automatic identification of collocations. The majority of studies have concentrated on establishing statistical lexical association measures (AM) best suited for this purpose. The performance of different AMs depends on such factors as the language under investigation, the definition of a collocation, the size and the quality of the corpus, and the frequency threshold.

The present work aims at developing an efficient machine learning algorithm for automatic collocation identification using pre-trained word embeddings and combining them with AM scores. The experiments focus on German adjective-noun collocations using the dataset described in [Evert, 2008]. The database contains 1252 expressions annotated by professional lexicographers, where an expression was evaluated as a ‘true collocation’ if it is useful for a bilingual dictionary: 520 true collocations and 732 non-collocations. The task at hand is the binary classification of the collocation candidates from the Evert dataset. In the previous experiments on the dataset in the MWE 2008 shared task [Pecina, 2008], the highest mean average precision (MAP) of 0.62 was achieved by the AM Piaterky-Shapiro coefficient followed by a machine learning algorithm where multiple AMs were used as features with the MAP of 0.61.

In the proposed model the classical approach of using AM scores is combined with implementing dense word vectors as features in the logistic regression classifier. The word embeddings used in the experiment had been trained on the raw text from the DECOW14ax corpus [Schäfer and Bildhauer, 2012], the vocabulary contains one million words with corresponding 50-, 100-, 200-, and 300-dimensional vectors. Each value of the vector is treated as a separate feature in the classifier as well as the four AM scores: mutual information, dice, chi-squared, and log-likelihood. The algorithm outperforms the previous models achieving the MAP of 0.74 in the classifier with 604 features in the 10-fold cross-validation setting and the MAP of 0.72 on the unseen test data.

References

- [Dima, 2015] Dima, C. (2015). Reverse-engineering language: A study on the semantic compositionality of German compounds. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pages 1637–1642, Lisbon, Portugal. Association for Computational Linguistics.
- [Evert, 2008] Evert, S. (2008). A lexicographic evaluation of German adjective-noun collocations. Proceedings of the LREC Workshop Towards a Shared Task for Multiword Expressions (MWE 2008), Marrakech, Morocco.
- [Pecina, 2008] Pecina, P. (2008). Lexical Association Measures: Collocation Extraction. PhD thesis, Faculty of Mathematics and Physics, Charles University in Prague, Prague, Czech Republic.
- [Schäfer and Bildhauer, 2012] Schäfer, R. and Bildhauer, F. (2012). Building large corpora from the web using a new efficient tool chain. In Chair), N. C. C., Choukri, K., Declerck, T., DoA’Y’an, M. U., Maegaard, B., Mariani, J., Moreno, A., Odijk, J., and Piperidis, S., editors, Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC’12), pages 486–493, Istanbul, Turkey. European Language Resources Association (ELRA).

N-gram Frequency Effects in Mandarin Chinese

Ching Chu Sun and **Peter Hendrix**

University of Tübingen

At the word level, the effect of frequency is perhaps the most well-documented effect in the psycholinguistic literature. Recently, frequency effects above the word level have been observed as well. Arnon and Snider (2010), for instance, observed that phrasal decision latencies were shorter for high frequency phrases such as “all over the place” as compared to low frequency phrases such as “all over the city”. (see also Siyanova-Chanturia, Conklin and Van Heuven, 2011; Bannard & Matthews 2008).

Here, we present the results of two studies that extend previous work on frequency effects for multi-word sequences in two ways. First, the current studies are the first to report frequency effects of multiword sequences for a non-alphabetic language: Mandarin Chinese. Second, whereas previous studies primarily reported frequency effects of multiword sequences in language comprehension, the focus of the current work is on frequency effects of multiword sequences in language production.

First, we carried out a reading experiment in Mandarin Chinese in which participants read word trigrams. Whereas previous studies contrasted low frequency phrases with carefully selected high frequency phrases, the multiword sequences in the current experiment were randomly selected trigrams from a large-scale corpus of Mandarin Chinese: the Simplified Chinese Corpus of Webpages (SCCoW; Shaoul, Sun & Ma, 2016). We analyzed the experimental data analyzed with generalized-additive mixed-effect models (GAMMs; Wood 2006, Wood 2011).

The analyses revealed significant effects of trigram frequency across the predictor range for both naming latencies ($\chi^2 = 22.256$, $p < 0.001$) and acoustic durations ($\chi^2 = 37.520$, $p < 0.001$). Response times were shorter for high frequency multiword sequences, as were acoustic durations. The effects of trigram frequency existed over and above the effects of unigram frequencies, bigram frequencies, and phonological length (i.e., the number of phonological segments in a trigram).

Second, we extracted the acoustic durations of word trigrams from a corpus of spontaneous speech in Mandarin Chinese: the Taiwan Mandarin corpus of spontaneous speech (Fon, 2004). As was the case for the experimental data, GAMM analyses revealed a robust effect of trigram frequency on acoustic durations ($F = 42.667$, $p < 0.001$), independent of the effects of component bigram and unigram frequencies and the effect of phonological length. The effect of trigram frequency was qualitatively similar to the effect of trigram frequency in the experimental data, with shorter pronunciation durations for more frequent word trigrams (cf. Arnon & Cohen Priva, 2013 for comparable findings in English).

The results of the work reported here demonstrate that the frequency of multiword sequences influences lexical processing not only in speech comprehension, but also in speech production. In addition, we established that frequency effects of multiword sequences are not limited to alphabetical languages, but also exist in the character-based language Mandarin Chinese. The observed pattern of results furthermore indicates that language users are aware of the combinatorial properties of words.

At least two possible interpretations of this awareness exist. First, in theories like data-oriented parsing (Bod, 2006) and memory-based learning (Daelemans & Van den Bosch, 2005), at least some multiword sequences are stored as a whole in the mental lexicon. Given the fact that we observed trigram frequency effects across the trigram frequency range, this

would require the storage of hundreds of millions of multiword sequences. Alternatively, knowledge about combinatorial properties of words is stored in association patterns between word-level representations (Baayen, Hendrix & Ramscar, 2013). Such an interpretation, we argue, fits more straightforwardly with the current findings.

Workshop Dinner

Restaurant Mauganeschtle (in Hotel am Schloss)

Burgsteige 18,
72070 Tübingen

<https://www.hotelamenschloss.de/restaurant-mauganeschtle/>



The restaurant is located just a few meters below the entrance to the castle.