8
Neurophysiological Markers of Phrasal Verb Processing: Evidence From L1 and L2 Speakers

Silke Paulmann
Zainab Ghareeb-Ali
University of Essex, UK

Claudia Felser
University of Potsdam, Germany

Abstract

Language is often ambiguous. For instance, verb-preposition strings such as *look up* can be interpreted either as a single *verb-plus-preposition* combination leading to a literal interpretation (e.g., *to look up the chimney*), or can be interpreted as a so-called phrasal verb which requires a figurative interpretation (e.g., *to look up the number*). Past research has primarily used behavioural methodologies to investigate how first (L1) and second language (L2) learners deal with this phenomenon. However, *Event-related brain potentials* (ERPs) are highly time sensitive and may shed additional light on this issue. In this chapter, we will first provide an overview of evidence on phrasal verb processing in L1 and L2 speakers. We will then present some of our own ERP data exploring phrasal verb processing in native speakers of English and native Arabic-speaking L2 learners of English. We will conclude with directions for future ERP research in this domain.

Keywords: ambiguity resolution, Arabic, bilingualism, EEG, ERPs, figurative language processing, multi-word expressions, N400, non-native language processing, phrasal verbs, prepositional verbs, second language acquisition, sentence processing

Introduction

Psycholinguistic research has a long tradition in exploring how native speakers successfully master the complexities encountered in everyday language. Lexical and structural ambiguities form a vital part of this complexity and are a frequent feature of language. For instance, multi-word expressions such as phrasal verbs (e.g., *run into*) which can have a figurative interpretation (e.g., *to meet*) have been argued to form about one third of the English verb vocabulary (Li, Zhang, Niu, Jiang, & Srihari, 2003). That is, these verbs are commonly used and language users will have to distinguish them from *prepositional verbs* (i.e., single verb + preposition

combinations) which may contain the same words but require a literal interpretation (e.g., *to enter somewhere running*).

The vast majority of research on this kind of multi-word expression has concentrated on determining how native speakers of English (easily) distinguish between different meanings that can be generated from sentences such as *Peter ran into Zara on Oxford Street* which can either be interpreted figuratively (*A man called Peter met his friend called Zara when walking along Oxford Street*) or literally (*A man called Peter went into the Zara store which is located on Oxford Street*). Underlying each of these two possible interpretations is a different syntactic structure. While *run into* in its figurative sense is often taken to be a single lexical unit (a compound verb (e.g., Di Sciullo & Williams, 1987), the preposition *into* functions as the head of a post verbal prepositional phrase in the 'literal' case.

Little is known to date about how non-native, or second language, speakers deal with the fact that sometimes a lexical form such as *run into* has to be interpreted as a phrasal verb (i.e., figuratively) and sometimes as a single *verb + preposition combination* (i.e., literally). How do non-native speakers of English overcome the problem that it is often not sufficient to simply know each individual word alone to reach the correct interpretation of a sentence? In the first part of this chapter, we will review recent evidence on phrasal verb processing in first (L1) and second language (L2) learners of English. This is followed by presenting data from our own lab in which we investigated the cognitive mechanisms underlying phrasal verb processing in native speakers of English and native Arabic-speaking L2 learners of English by means of *Event-Related Brain Potentials* (ERPs). ERPs possess the sensitivity to assess how meaning is processed (accessed) while words unfold in real-time. Our discussion of the data will be followed by suggestions for future directions of research in this field.

Representation and Access of Phrasal Verbs in the Mental Lexicon

As mentioned above, investigations on how native speakers end up with the correct interpretation of ambiguous multi-word strings such as *run into* or *kick the bucket* have played an important role in psycholinguistic research. Some of this research has focused on how idioms such as *kick the bucket* or phrasal verbs are stored and accessed in the mental lexicon. For instance, while some propose that these expressions are processed as whole lexical chunks (e.g., Estill & Kemper, 1982; Swinney & Cutler, 1979), others suggest that they may be processed compositionally, similarly to any other word string (e.g., Cacciari & Tabossi, 1988). Of related interest is the question of whether specific meanings or interpretations of multi-word expressions are accessible and activated in the lexicon before others. That is, what happens when language users encounter phrases that cannot be interpreted by relying solely on the meaning of each single word of the phrase? For example, is there a temporal processing advantage for one specific interpretation?

Various hypotheses have been proposed to explain how we process figurative language. For instance, when initially introducing the *idiom list hypothesis*, Bobrow and Bell (1973) proposed that literal meanings are accessed first. A few years later, Gibbs (1980) in his *direct access hypothesis*, proposed the opposite, namely that figurative interpretations of multi-word expressions are accessed before literal interpretations (i.e. the figurative meaning of a phrase is preferentially processed). Around the same time, *the lexical representation hypothesis* favored by

Swinney and Cutler (1979) put forward the view that both meanings are processed in parallel (i.e. there is no processing advantage for one specific interpretation). In this case, when encountering expressions which can be interpreted literally or non-literally, two different processing strategies are applied. On the one hand, activated words undergo a structural analysis necessary for a literal interpretation, while simultaneously whole units get activated in order to access the figurative interpretation. Similarly, the *configuration hypothesis* by Cacciari and Tabossi (1988) proposes that initially, both figurative and literal meaning interpretations are considered; however, once idiomatic expressions are recognized as units, only figurative interpretations receive further activation. Literal interpretations of these expressions are no longer pursued once this recognition point has been reached. It is worthwhile to note that the processing mechanisms can be influenced by biasing context. Specifically, if one interpretation is favored over another, for example if the sentence context biases towards a figurativerather than a literal interpretation, then the processing mechanisms described by the *configuration hypothesis* are altered so that an expression can be recognized even before arriving at its uniqueness point. For example, in an appropriate context the expression "build castles in the air", which lacks a literal reading, may be recognized as an idiom before the final noun "air" has been processed.

Phrasal Verb Processing in Native and Non-native Speakers

The processing models mentioned above were built on evidence gathered from L1 speakers. Research investigating whether non-native speakers process multi-word expressions in the same way as native speakers do is scarce. This is surprising given that difficulties of phrasal verb usage for L2 learners are well documented. For instance, one of the earlier studies conducted by Dagut and Laufer (1985) explored the avoidance of phrasal verbs using three different tasks in native Hebrew-speaking L2 learners of English. For instance, in one task, participants were required to fill in the blanks of sentences with one of four verb choices, one of them being a phrasal verb (e.g. *We didn't believe that John could ever _____ his friends* [let down, solve, disappoint, carry on]). In another task, the same blanks had to be filled, but instead of providing the learners with different verb options, a Hebrew translation was included (e.g *We didn't believe that John could ever _____ his friends* [leachzev]). Their results showed that Hebrew speakers predominantly preferred single verbs over phrasal verbs. Hebrew does not have phrasal verbs, but the authors suggested that the avoidance of their use is linked to this absence. In other words, L2 learners prefer to use single verbs whose usage they fully comprehend. Other research has since replicated this avoidance phenomenon in native Dutch (Hulstijn & Marchena, 1989), as well as native Swedish and native Finnish-speaking (Sjöholm, 1995) learners of English. However, results from these studies suggest that both proficiency (Hulstijn & Marchena, 1989) and L1-L2 language distance (Sjöholm, 1995) can modulate phrasal verb avoidance. It should be noted that Laufer and Eliasson (1993) found no phrasal verb avoidance in advanced Swedish learners of English, which indirectly supports the assumption that proficiency alters phrasal verb usage. This is in contrast to more recent evidence from Siyanova and Schmitt (2007) who investigated usage of single (e.g., *to train*) and multi-word verbs (e.g., *to work out*) in native English speakers and advanced learners of English. In a questionnaire study, the authors asked both participant groups to indicate their preferred usage of 26 selected single and multi-word verbs on a 6-point Likert scale. Results revealed that non-native speakers were less likely to use multi-word verbs than native speakers. Thus, although they were advanced learners of English,

the non-native speakers in Siyanova and Schmitt's study showed a higher tendency to use one-word verbs as opposed to multi-word verbs.

One of the few studies that explored phrasal verb comprehension (rather than production) in native and non-native speakers of English applied a cross-modal semantic priming paradigm (McPartland-Fairman, 1989). Phrasal verbs (*sign up, carry on*) as well as matched verb + preposition combinations were embedded in sentences that either biased towards the literal (1a) or figurative (1b) interpretation.

1a. The soldier was writing to his girlfriend and he had a lot to tell her that day. When he finished, there wasn't enough space for his name at the bottom of the letter. He didn't have any choice, so he signed up # the side of the paper.

1b. The doctor told the patient he was working too hard and needed to do more exercise or he would get a heart-attack. He didn't have any choice, so he signed up # the next day for an exercise class).

Participants' task was to name visually presented target words that were either related to the figurative or literal interpretations, or unrelated control words, which appeared during auditory sentence presentation at phrasal verb offset (at the position marked # in the example above). In general, naming times were faster for related than for control targets. However, similar naming times were found for target words that were related to literal and figurative interpretations, suggesting that both meaning interpretations were activated during reading, as predicted by the lexical representation hypothesis. This was the case for both language groups, suggesting that on-line comprehension of multi-word expressions is comparable between native and non-native speakers of English (McPartland-Fairman, 1989).

More recently, Matlock and Heredia (2002) revisited phrasal verb processing in native and non-native speakers of English. Specifically, the authors sought to explore differences between native and non-native speakers in phrasal verb processing by means of a sentence completion and an on-line reading comprehension task. The sentence completion task examined whether non-native speakers of English preferentially produced phrasal verb (figurative) or single verb + preposition combinations (literal interpretation) of ambiguous lexical forms. Interestingly, the results from this task revealed that both native and non-native speakers of English produced *more* phrasal verbs than verb + preposition combinations, suggesting that both language groups were equally comfortable using phrasal verbs.

The on-line reading task investigated the time-course of computing literal versus figurative interpretations. Specifically, Matlock and Heredia (2002) explored whether readers from both language groups would activate the figurative meaning of phrasal verbs (e.g., *to go over the exam*) or the literal meaning (*to go over the street*) first. For the second experiment, the L2 learners were divided into participants who had started to learn English before the age of 12 (i.e., early bilinguals), or after 12 (i.e., late bilinguals). The results revealed that L1 speakers and early bilinguals read sentences with phrasal verbs more quickly than sentences that required single verb + preposition interpretations. This suggests that figurative meanings were processed more quickly than literal meanings. In contrast, late bilinguals showed no such effect. This seemed to suggest that late learners of English have difficulties interpreting sentences that do not allow for a one-to-one translation of words, supporting evidence from production studies which showed that non-native speakers suffer from difficulties using phrasal verbs and that avoidance behavior may be modulated by language proficiency (e.g., Dagut & Laufer, 1985; Hulstijn &

Marchena, 1989; Liao & Fukuya, 2004; Siyanova & Schmitt, 2007). Liao & Fukuya (2004), for example, report that intermediate but not advanced Chinese-speaking learners of English avoided using phrasal verbs in comparison to English native speakers.

Although the results from Matlock & Heredia's (2002) second experiment seem to complement findings from the production literature, they should be interpreted with some caution. For instance, both early and late bilinguals came from a variety of language backgrounds; while the early bilingual group was predominately native Spanish speakers (54% of participants) and closely followed by native Chinese speakers (22%), the late bilingual group comprised speakers from a larger pool of backgrounds, with Spanish (23%) and Chinese (30%) native speakers being less dominant. Although unlikely, the difference between early and late bilinguals might thus stem from differences in L1 background. More importantly, group sizes differed (22 early vs. 13 late bilinguals), and given that the late bilinguals showed a numerical difference in reading times for phrasal verbs and verb + preposition combinations that went into the same direction (a 171 ms advantage for reading phrasal verbs) as found in the early bilinguals and native monolinguals, the question arises as to whether the lack of an effect is actually a statistical power problem. This is supported by the observation that the authors' between-subjects analysis did not reveal a significant two-way interaction between group (early vs. late) and verb type (phrasal verb vs. verb phrase), pointing to the possibility that phrasal verb comprehension may not be fundamentally different between early and late bilinguals. Finally, as Matlock and Heredia (2002) point out themselves, sentence reading times lack the temporal resolution needed to assess the time-course underlying on-line phrasal verb processing. Taken together, we believe more research using time-sensitive methodologies is needed to explore differences and similarities between native and non-native speakers of English when processing phrasal verbs.

Electrophysiological Investigations on Figurative Language Processing

Event-related brain potentials (ERPs) are highly time sensitive and are now frequently used in psycholinguistic research on bilingualism (see Mueller, 2005, for a review). Whereas behavioral methodologies always measure at discrete points of time (e.g., after a decision has been made), ERPs allow for psychological processes underlying language comprehension to be monitored while words or sentences unfold in real time. Briefly, ERPs are small voltage variations in the *electroencephalogram* (EEG) and result from the brain's response to an event (e.g., auditory/visual stimulus). The series of voltage peaks caused by an event (or stimulus) are called ERP components. Over the past 30 years, several language related components have been identified (for a short review see, Friederici, 2004). For instance, the so-called N400 component is a negative ERP peaking at around 400 ms after the onset of a critical event and has been linked to lexical-semantic processes. Specifically, the N400 is elicited in response to words that mismatch preceding sentence context (e.g., the word *socks* elicits a larger N400 than the word *butter* when preceded by sentence contexts such as *He spread the warm bread with…*; Kutas & Hillyard, 1980), making the N400 an ideal candidate when investigating lexical-semantic expectancies.

Several previous studies exploring figurative language processing have reported N400 effects. For instance, Laurent and colleagues investigated idiom processing in a lexical decision task to test Giora's (2003) *graded salience hypothesis*, which claims that salient meanings enjoy a processing advantage over less salient ones. According to Giora (2003: 10), salient

meanings are “coded meanings foremost on our mind due to conventionality, frequency, familiarity, or prototypicality”. Briefly, in the lexical decision tasks, participants are asked to determine if a presented word (e.g., *HOUSE*) is a legal word in English (YES) or a non-legal word (NO) in English (e.g., *HOUST*). Participants were presented with strongly (e.g. *rendre les armes* 'to surrender weapons') and weakly (e.g. *enfoncer le clou* 'to hammer it home') salient idiomatic expressions followed by targets that could be related to the figurative or literal interpretation. The strength of idiomatic saliency of stimuli was determined in a previous study by asking participants to read each expression and then to jot down the first word that struck them. N400 amplitudes measured at the last word of strongly salient idiomatic expression were smaller than amplitudes measured at the last word of weakly salient idiomatic expressions. This suggests that salience (or expectancy) is critical when processing idioms. That is, highly salient expressions are more easily processed than less salient expressions (Laurent, Denhieres, Passerieux, Iakimova, & Hardy-Bayle, 2006). An earlier study exploring idiomatic, literal, and non-sense language processing in schizophrenics and healthy controls also revealed N400 effects. Specifically, participants were presented with stimuli such as *vicious circle* (idiomatic expressions) or *vicious dog* (literal expressions). The authors report stronger N400 amplitudes in response to literal expressions than in response to idiomatic expressions for the healthy control group. This suggests that literal language can be harder to integrate and process than figurative language, especially if figurative language is high in cloze probability as was the case in this study (Strandburg, Marsh, Brown, Asarnow, Guthrie, Harper, Yee, & Nuechterlein, 1997). Other studies report N400 effects in response to metaphors (e.g., Coulson & Van Petten, 2002; 2007; Pynte, Besson, Robichon, & Poli, 1996) or irony (e.g., Cornejo, Simonetti, Aldunate, Ibanez, Lopez, & Melloni, 2007; Regel, Gunter, & Friederici, 2011).

The Present Study

Thus, it seems as if the N400 can be particularly useful when exploring how and when figurative and literal meanings are accessed during phrasal verb processing. In particular, we can look at ERPs elicited in response to disambiguating nouns when reading a phrasal verb embedded in neutral sentence contexts to assess whether one reading is preferred (that is, more expected) over another. The present study explores exactly this in both monolingual (native English) and bilingual (native Arabic) populations. We presented sentences such as (2a-b) which contain temporarily ambiguous verb-preposition strings and compared ERPs elicited in response to the disambiguating noun (e.g., *bridge* vs. *farmer*). Notice that in 2a, *ran over* means to *walk over something* and in 2b, it means to *kill someone by driving*.

2a. *I heard that Mr. Smith ran over the old bridge early this morning.*
2b. *I heard that Mr. Smith ran over the old farmer early this morning.*

If figurative meanings are preferred over literal meanings, we expect nouns that allow for such an interpretation to elicit a smaller N400 than nouns that require a literal sentence interpretation. Conversely, if literal meanings are easier to compute than figurative ones, we might expect nouns which disambiguate towards a literal reading to elicit a smaller N400 component. In short, in the present study, the N400 is used as an indicator of integration difficulty. Specifically, component

amplitudes are used to infer which reading (literal/figurative) of a noun following neutral sentence requires enhanced cognitive effort.

Methods

Participants: Overall, 20 (10 women, *M* = 25.6 years of age) students from the University of Essex volunteered to participate in the experiment. They received a small fee for participation. Ten participants (two women, *M* = 23.2 years of age, range = 18 to 43 years of age) were native speakers of British English and ten participants (eight women, *M* = 28 years of age, range = 24 to 36 years of age) were native speakers of Arabic. Arabic speakers came from Syria (*n* = 1), Saudi Arabia (*n* = 3), Kuwait (*n* = 3), and Libya (*n* = 3). The L2 participants had started to learn English at school around the age of 8 years (range = 4 to 12 years of age) and, on average, had spent 3 years and 9 months in an English-speaking country (range = 4 to 84 months). L2 participants self-assessed their English proficiency on a four-point scale (4 = Excellent, 3 = Very good, 2 = Good, and 1 = Poor) for auditory comprehension (*M* = 3.4, *SD* = 0.69), reading comprehension (*M* = 3.2, *SD* = 0.63), speaking (*M* = 3.2, *SD* = 0.63), and writing skills (*M* = 3.1, *SD* = 0.56). The grammar part of the Oxford Placement Test (OPT; Allan, 2004) was also given to L2 learners prior to taking part in the experiment. Total scores ranged from 70-97% (*M* = 81%, *SD* = 9), placing the L2 participants within the upper intermediate (competent user) to highly proficient (very advanced user) range according to the OPT language scale.

Stimuli: The experimental items were created using 18 different, temporarily ambiguous verb-preposition strings such as *run over.* These were embedded in neutral sentences and were semantically disambiguated either towards a literal (single verb + preposition combination) or a figurative interpretation (phrasal verb construction) by the following noun, as illustrated in (2a-b) above. Each experimental sentence consisted of 13 words in total and pairs of critical/disambiguating nouns were matched on frequency from the British National Corpus, $t(35) = .793$, $p > .433$, and were also approximately matched in length. To increase the number of trials in the ERP experiment, each critical verb-preposition string was used four times in total. Specifically, each verb-preposition string was embedded in two different sentence contexts, each of which came in two conditions (i.e. literal vs. figurative). Thus, 72 experimental sentences were presented to each participant. In addition, 144 filler sentences, of which some contained idioms (*My grandfather is as old as the hills*), binomials (*Lisa and her friend ate some fish and chips)*, collocations (*Angelina likes to drink strong tea*) or compounds (*Last night my daughter did all of her homework*), were presented. We created four differently randomized presentation lists in which the 36 test items (18 for literal and 18 for figurative interpretation) were intermixed with 144 fillers. Using SuperLab Version 4.07b, sentences were presented using *Rapid Serial Visual Presentation* (RSVP) of one word at a time. In order to encourage participants to read the sentences actively for meaning, a set of yes/no comprehension questions were constructed which were randomly included in the experiment. For each list, 18 comprehension questions followed critical experimental sentences and 36 questions followed filler sentences.

Procedure: All participants were tested in a quiet laboratory room. Before the start of the EEG experiment, participants were asked to fill out a short demographic and language questionnaire. For the EEG experiment, participants faced a computer monitor from a distance of approximately one meter. Before the start of the experiment, participants engaged in a practice session of four trials. Each trial started with an eye fixation cross displayed for 450ms, followed by the sentences presented word by word (word presentation duration was set at 450 ms and

words were separated by blank screens of 200 ms), followed by a 1000 ms inter-trial interval. All words were displayed in lower case Arial font (64 point) black letters against a white background in the center of the screen. Comprehension questions were presented randomly between trials to ensure that participants were paying attention to the sentences. Breaks were included after every 54 trials. At the end of the EEG session, participants were asked to complete a plausibility rating questionnaire which consisted of 86 sentences in total and included all 72 experimental sentences and additionally 14 fillers which were either perfectly plausible (e.g., *The researchers who are researching the causes of cancer are making progress*) or totally implausible sentences (e.g., *Diana met her whistle in a blue skirt full of beans*). Participants were asked to rate the sentences according to their plausibility on a scale from one to three (1 to indicate *low plausibility* and 3 to indicate *high plausibility*). The results from this questionnaire can be found in Table 1. The EEG experiment lasted approximately 40-45 minutes, while the plausibility rating questionnaire took approx. ten minutes to fill out. One experimental session (incl. EEG set up) lasted no longer than two hours.

ERP Recording: Sixty-four EEG channels were recorded from the scalp by means of Ag/AgCl electrodes attached to an elastic Quikcap (Neuroscan) according to the international 10-20 system: FP1, FPZ, FP2, AF7, AF3, AFZ, AF4, AF8, F9, F7, F5, F3, FZ, F4, F6, F8, F10, FT9, FT7, FC5, FC3, FCZ, FC4, FC6, FT8, FT10, T9, T7, C5, C3, CZ, C4, C6, T8, T10, TP9,TP7, CP5, CP3, CPZ, CP4, CP6, TP8, TP10, P9, P7, P5, P3, PZ, P4, P6, P8, P10, PO7, PO3, POZ, PO4, PO8, O1, OZ, O2, M1, M2. Each EEG channel was amplified with a band pass from DC to 100 Hz with a digitization rate of 500 Hz. AFz served as a ground electrode. All electrodes were on-line referenced to the left mastoid (M1). Horizontal and vertical electro-oculograms (EOG) were recorded to control for eye movement artifacts. Electrode impedances were kept below 7 kΩ.

ERP Data Analyses: EEG data were processed with EEGLab (Delorme & Makeig, 2004). For each participant, EEG recordings were first re-referenced to the average of both mastoids off-line. Next, recordings were band pass filtered between 0.1 Hz and 40Hz.The continuous EEG was then epoched and baseline corrected using a 200ms pre-stimulus baseline. Epochs were extracted from 200ms before the appearance of the disambiguating critical noun up to 800 ms after noun onset. Data for each participant were scanned for artifacts and epochs contaminated with eye blinks and/or muscle/electrical artifacts were removed for each participant using the *find abnormal values* function in EEGLab. The threshold for this automatic rejection procedure was set at 75μV. Data were also visually inspected for artifact rejection purposes. Following this procedure, 23% of trials had to be rejected for English speakers, whereas 31% of trials were rejected for Arabic speakers. Finally, separate ERPs for each condition at each electrode site were averaged for each participant. For graphical illustration purposes only, grand average ERPs were smoothed with a 7Hz low-pass filter.

In all experiments, the critical group comparisons of the ERP data were quantified for correct responses by calculating amplitudes relative to a 200-ms pre-stimulus baseline. A 2 (Native vs. Non-native participants) × 2 (Literal vs. Figurative meaning) x 7 (Region of Interest, see below) repeated measurements analysis using the PROC GLM function in SAS 9.2 was conducted. The factor Region of Interest (ROI) defined a critical region of seven scalp sites: left frontal (LF): F7 F5F3 FT7 FC5 FC3; right frontal (RF): F8 F6 F4 FT8 FC6FC4; left central (LC): T7 C5 C3 TP7 CP5 CP3; right central (RC): T8 C6 C4 TP8 CP6 CP4; left posterior(LP): P7 P5 P3 PO7 PO3 O1; right posterior (RP): P8 P6P4 PO8 PO4 O2 and the midline (ML): FZ FCZ CZ CPZPZ POZ. The Geissser-Greenhouse correction (Geisser & Greenhouse, 1959) was

applied to all repeated measures with greater than one degree of freedom. Main effects of topographical factors are not of interest for the present investigation and are thus not followed up.

Results

Plausibility Rating: A 2(Language Group: Native vs. Non-native) × 2(Verb Type: Literal vs. Figurative) ANOVA revealed a significant main effect of verb type, $F(1,18) = 19.6$, $p < .001$, reflecting the fact that participants' plausibility rating scores were higher for figurative than for literal sentences.. There was no significant main effect of group, $F(1,18) = .679$, $p > .421$, suggesting that overall, plausibility rating scores did not significantly differ between native and non-native speakers ($M = 2.26$ vs. $M = 2.15$), and no significant interaction between group and verb type was revealed, $F(1,18) = .015$, $p > .904$. Means of the plausibility rating can be found in Table 1 below.

Table 1. Mean Rating Scores (SDs In Parentheses) of the Plausibility Rating Split by Group and Condition

	Verb type	
Group	Literal	Figurative
Native English (n = 10)	2.15 (.585)	2.37(.447)
Native Arabic (n = 10)	2.05 (.473)	2.25(.349)

Visual inspection of ERPs

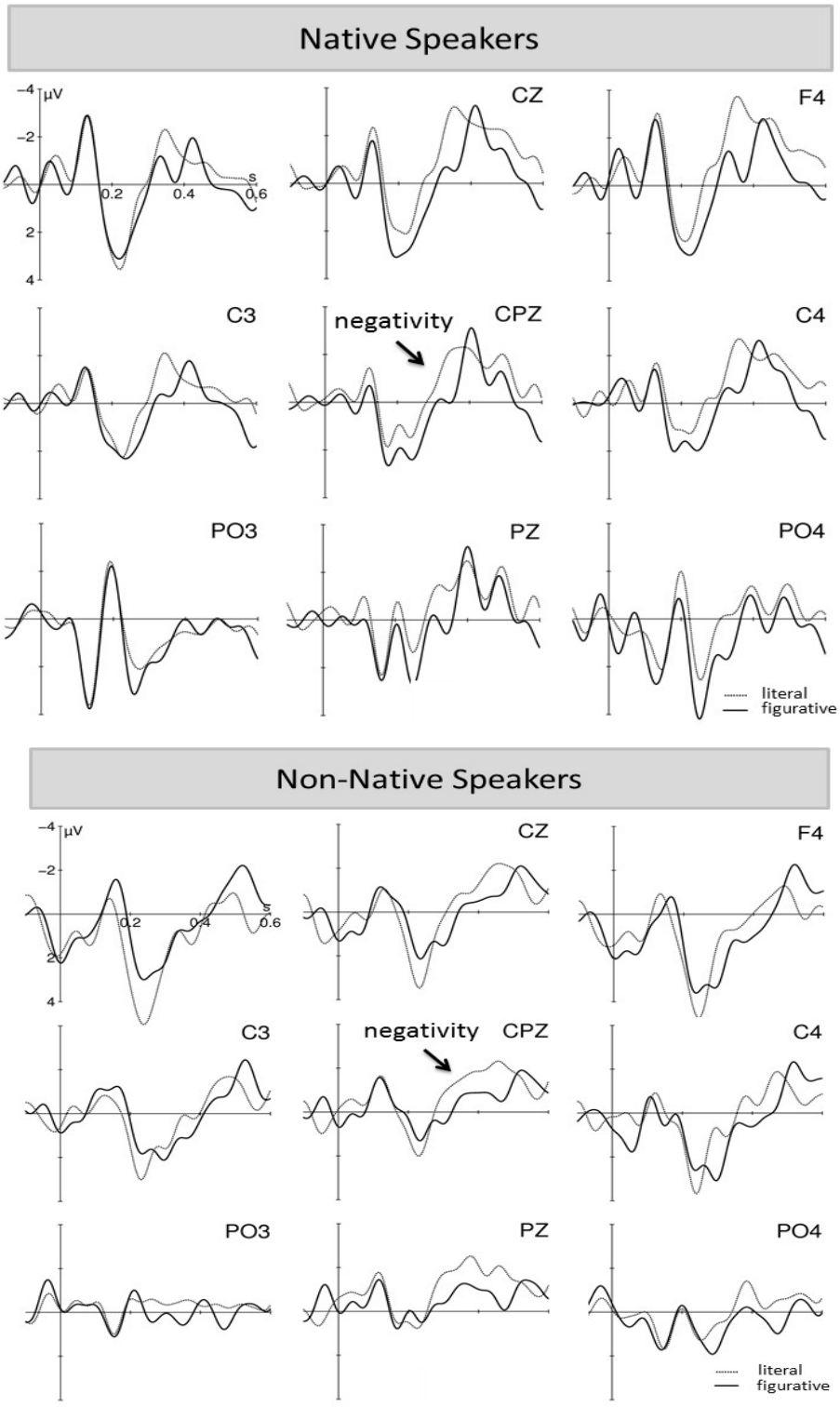
See Figure 1 for visualization of the ERP data.Visual inspection of the data shows an early negativity (N100) peaking at about 100 msec poststimulus, followed by a positivity (P200) that peaks at around 200 msec, followed by a negative-going wave peaking at 400 msec (N400). Modulations dependent on Type are particularly pronounced in the last component . Thus, a classical N400 time window ranging from 300 to 500 ms after noun onset was chosen for analysis after visual inspection of the data. In addition to a mean *amplitude* analysis, a peak *time* analysis was also run as visual inspection of the data revealed slightly later component onsets for non-native speakers (bottom-half of the figure) when compared to native speakers (top half of the figure).

N400: The ERP analysis for mean amplitudes showed no significant group effect, $F(1,18) = 2.32$, $p = .15$. Crucially, the Verb Type effect was significant, $F(1,18) = 4.61$, $p < .05$, showing more negative ERP amplitudes in response to nouns requiring a literal interpretation than nouns requiring a figurative interpretation. There were no significant interactions with the factor group or Verb Type.

To assess whether N400 *peak amplitude times* differ between nouns requiring a literal and those requiring a figurative interpretation, we ran an additional ERP peak time analysis. This analysis helps to establish whether the two groups differed in processing time rather than in the

way they processed the ambiguity. There were no significant main effects. The three-way interaction between ROI x Verb Type x Group reached significance, $F(4,72) = 2.81$, $p = .05$. However, follow-up analyses revealed no further significant effects. Taken together, these results revealed comparable processing mechanisms for native and non-native speakers of English when processing sentences containing phrasal verbs or verb + preposition combinations.

Figure 1: The illustration shows the significant N400 effect at selected electrode sites for native and non-native speakers of English**.**


Native Speakers
−4 μV
−2
2
4
0.2
0.4
s
0.6
CZ
F4
C3
negativity
CPZ
C4
PO3
PZ
PO4
literal
figurative
Non-Native Speakers
−4 μV
−2
2
4
0.2
0.4
s
0.6
CZ
F4
C3
negativity
CPZ
C4
PO3
PZ
PO4
literal
figurative

Discussion

The present study set out to explore the time-course underlying phrasal verb and verb + preposition processing in native and non-native speakers of English by means of ERPs. Specifically, we sought to establish whether native and non-native speakers process sentences containing phrasal verbs and verb + preposition combinations in a similar fashion. The current results suggest that this is indeed the case; we report larger N400 components in response to literal when compared to figurative interpretations for both native speakers of English and proficient L2 learners of English with native Arabic language background.

The present results challenge the view that phrasal verb processing is *per se* difficult for second language learners. Rather, results emphasize once more that language production and language comprehension mechanisms do not always have to go hand in hand. While second language learners may well avoid using phrasal verbs in everyday speech (e.g., Dagut & Laufer, 1985; Hulstijn & Marchena, 1989; Siyanova & Schmitt, 2007), our results support previous findings from comprehension studies applying cross-modal priming (McPartland-Fairman, 1989) or on-line reading tasks (Matlock & Heredia, 2002), which show that comprehension of phrasal verbs is not necessarily problematic in proficient L2 learners of English.

Moreover, our results complement previous behavioral studies investigating phrasal verb and verb + preposition processing in monolinguals and bilinguals (Matlock & Heredia, 2002) which show *preferred* processing of figurative as opposed to literal meanings for native speakers and early, arguably highly proficient, bilingual speakers. The finding that nouns that require a sentence to be interpreted in a figurative way are more easily integrated into a sentence context than nouns that require the same preceding sentence to be interpreted in a literal way suggests that figurative sentence interpretations are anticipated (i.e. predicted) by readers. Specifically, enhanced N400 components in response to nouns leading to a literal sentence interpretation suggest that these nouns were less expected and hence require enhanced processing effort during sentence integration processes. Note that from an incremental sentence processing perspective, analyzing verb-preposition strings as phrasal verbs would also seem to be the easier option as the structural processor is thought to prefer integrating new upcoming words (here, a preposition immediately following a verb) into the current constituent over postulating a new phrase (Frazier, 1979).

As mentioned in the introduction, previous ERP studies investigating how idiomatic expressions are processed by native speakers have revealed similar N400 results. For instance, Strandburg et al. (1997) compared processing of two-word phrases that were either highly idiomatic (e.g., *pot luck*, *fat chance*, *vicious circle*), literal (e.g., *vicious dog*), or non-sensical phrases (e.g., *square wind*) in schizophrenics and healthy controls. Healthy controls showed reduced N400 amplitudes in response to idiomatic two-word phrases when compared to literal and non-sensical phrases suggesting that the first word of the two-word phrase provides enough context for readers to expect the second word of the phrase (i.e. idiomatic expressions are highly conventionalized). This leads to ease of integration of the second part of the two-word phrase, suggesting preferred processing of figurative interpretations of two-word phrases as opposed to literal interpretations. Similarly, Vespignani, Canal, Molinaro, Fonda, and Cacciari (2010) report reduced N400 amplitudes in response to idioms when compared to literal control conditions, again suggesting that figurative analysis of phrases is highly expected. Finally, Laurent et al. (2006) tried to disentangle *figurativity* (i.e. whether an expression is literal or idiomatic) and saliency effects. Specifically, they presented weakly and strongly salient idiomatic expressions, that is, highly conventionalized idioms such as *rendre les armes* ('surrender the weapons') versus expressions whose idiomatic meaning is less salient such as *enfoncer le clou* ('to hammer it

home'). These were followed by target words that were either related to a figurative or literal interpretation. The authors report smaller N400 amplitudes for highly conventionalized (i.e. familiar) idioms when compared to less salient idioms. They also report a reduced N400 in response to figurative targets (e.g. *abandonner* 'to give up') that followed highly salient idioms such as *rendre les arms* 'surrender the weapons' when compared to (a) literal targets that followed highly salient idioms (e.g. *déposer* 'to put down' following *rendre les arms*) as well as (b) literal (e.g. *fixer* 'to fix') and (c) figurative targets (e.g. *insister* 'to insist') that followed less salient idiomatic expressions such as *enfoncer le clou* 'to hammer it home'. Given that the authors do not report any diminished amplitudes to figurative targets following *less* conventionalized idioms when compared to literal targets following the same idioms, they argue that it may not be *figurativity* per se that drives the effects reported in previous studies, but that the saliency of idiomatic expressions is more crucial. This claim is partly supported by recent eye tracking data. In their study, Siyanova-Chanturia, Conklin, and van Heuven (2011) investigated whether native and non-native speakers of English are sensitive to phrasal frequency of multiword sequences. Specifically, the authors compared processing of phrases such as *king and queen* or *right and wrong* with their reversed form *queen and king* or *wrong and right*. It was found that both native and proficient non-native speakers of English were sensitive to frequency information at the phrasal level. Interestingly, reading times of less proficient learners of English were not influenced by phrasal frequency. Taken together, this suggests that multi-word sequences are subject to learning and that their saliency or frequency can influence readers' processing mechanisms.

Although the present material contained highly conventionalized expressions (e.g., *eat up the fish* vs. *eat up the hill*), we also included less conventionalized items (e.g., *coloring in the picture* vs. *coloring in the garden*). Given that we nevertheless found N400 differences for nouns requiring a verb + preposition or phrasal verb interpretation, our results might suggest that saliency of expressions is not as crucial when processing phrasal verbs as it is when processing idiomatic expressions. Unfortunately, the numbers of items included in the present study do not allow further disentangling this effect (i.e. compare highly conventionalized forms with less conventionalized forms) but future studies of phrasal verb processing could try to control for this potential confound more closely.

Taken together, our results are thus in line with previous electrophysiological findings which support the view that figurative sentence interpretations are often strongly favored by readers. Our off-line plausibility rating task also suggests that figurative sentence interpretations are preferred over literal sentence interpretations. When processing temporally ambiguous sentences the *default* for both L1 and L2 learners might be to go for the figurative interpretation. Literal interpretations would only be considered if a figurative analysis is not successful. Further studies are needed to explore whether this preference is modulated by L2 proficiency or exposure by testing learners of English across a range of proficiencies, and how it is affected by the frequency or salience of items.

The view that figurative interpretations of multi-word expressions are preferentially processed is emphasized in different theoretical frameworks (e.g., Cacciari & Tabossi, 1988; Gibbs, 1980; Swinney & Cutler, 1979;). For instance, Gibbs (1980) suggests that the figurative meaning of a phrase is accessed before its literal counterpart. Thus, when reading a sentence containing a verb such as *look into*, readers automatically activate lexical-semantic meaning for the whole phrase (*investigate, dig, search*) rather than for its individual constituents *look* and *into* (literal interpretation). Gibbs (1980) based his *direct access theory* on the observation that

participants were quicker to rate the meaningfulness of phrases that could be interpreted in an idiomatic way (e.g., *kick the bucket*) than those that could only be interpreted in a literal way (e.g., *lift the bucket*). However, judging the meaningfulness of a phrase is a meta-linguistic task which is carried out *at the end* of an on-line reading process. Here, we applied an on-line sentence reading task and time-locked brain activity to the point in time when readers would first know whether the verb phrase had to be interpreted as a phrasal verb (figurative) or as a verb + preposition combination (literal). Given that we find no differences in the *latency* of the N400 peak amplitude between the two conditions, it can be hypothesized that both literal and figurative interpretations were activated simultaneously, but that phrasal verbs received stronger activation than verb-preposition strings given the processor's preference for late closure (Frazier, 1979).

Language processing is strongly determined by expectancies and context-based predictions (e.g., Federmeier, 2007); here, nouns that allowed for a figurative sentence interpretation elicited smaller N400 amplitudes than frequency and length matched nouns that enforced a literal sentence interpretation. Cacciari and Tabossi (1988) proposed that both figurative and literal meaning interpretations would be considered initially (i.e. no timing differences) by the reader but that readers would disregard one interpretation as soon as a recognition point has been reached. Given our neutral sentence context, the *recognition point* (i.e. the point in time at which the ambiguity was resolved) must have occurred at the same time for both tested conditions, meaning that at least initially both sentential interpretations were pursued to the *same* degree. However, after finishing reading the two-word phrase, participants anticipated a noun that allowed for a figurative interpretation of the verb phrase. Nouns that did not match this expectancy elicited larger N400 components in both groups. In conclusion, the current results provide support for models that allow for *simultaneous* activation of phrasal verb and verb + preposition interpretations of two-word phrases; however, processing mechanisms (e.g., timing or degree of activation) can be altered by predictability of (upcoming) constituents (c.f. Vespignani et al., 2010).

Future Work

The present study set out to explore how native and non-native speakers of English process phrasal verbs and verb + preposition combinations. Both language groups show an enhanced N400 component in response to nouns that require the two-word phrase to be interpreted in a literal as opposed to figurative way. This is in line with previous ERP studies exploring idiom processing in native speakers as well as behavioral studies testing phrasal verb processing in native and proficient learners of English (e.g., Matlock & Heredia, 2002). Our results suggest that non-native but proficient speakers of English use similar processing mechanisms when processing phrasal verbs. In particular, expectancy seems to favor figurative sentence interpretations over literal ones. Clearly, figurative meanings have to be learned over time and as mentioned previously cannot always be derived based on individual constituents alone. Future research should thus investigate when second language learners start to prefer figurative interpretations over literal ones. Matlock and Heredia (2002) suggest that age of acquisition can influence processing mechanisms; we suggest to also test language proficiency (irrespective of age of acquisition). Testing learners with the same native language background but who master English to a different degree could give rise as to *when* figurative meaning interpretation is considered to be the default interpretation. Moreover, testing learners with different first

language backgrounds allows assessing how far language transfer can influence processing mechanisms.

In addition to exploring the influence of proficiency on phrasal verb processing mechanisms, future studies could also investigate the role of sentence context. In the present investigation, phrasal verbs and verb + preposition strings were embedded in neutral sentence context; however, one might ask whether the apparent preference for figurative interpretations is upheld when introducing biasing contextual information. In other words, will the figurative meaning of phrasal verbs be accessed when the sentence or discourse context is biased towards the literal interpretation?

Acknowledgements

The authors would like to thank Katharina Mursin for help with data collection, Chelsea Harmsworth for help with data analysis, and Roger Deeble for technical assistance.

List of Keywords

Age of acquisition, Ambiguity resolution, Amplitude, Avoidance, Configuration hypothesis, Context-based predictions, Conventionalized idioms, Cross-modal semantic priming paradigm, Direct access hypothesis, Early bilinguals, Electroencephalogram (EEG), Figurative language processing, Idiom list hypothesis, Incremental sentence processing perspective, Late bilinguals, Lexical decision task, Lexical representation hypothesis, Lexical semantic expectancies, Literal interpretation, Mental lexicon, N400, Non-native language processing, On-line comprehension, Phrasal verb, Rapid serial visual presentations (RSVP), Recognition point, Region of interest (ROI), Saliency, Simultaneous activation, Time-locked brain activity, Uniqueness point, Voltage Peaks

Thought Questions:

1. If non-native speakers are native-like in their ability to access phrasal verb meanings during comprehension, then why do they tend to avoid using phrasal verbs in language production?
2. What are the advantages and possible disadvantages, of using ERPs to investigate figurative language processing?
3. Does the similarity of native and non-native speakers' brain responses to figurative vs. literal disambiguation in the study reported mean that the same neural mechanisms and pathways are involved in both populations?
4. Many linguists assume that phrasal verbs such as *look up* in *to look up a number* are mentally stored as lexical units, whereas other verb-preposition combinations (as in *to look up the chimney*) are not. How might this difference help account for comprehenders' apparent preference for the figurative (phrasal verb) interpretation of verb-preposition combinations?

5. What linguistic and non-linguistic factors might influence L2 learners' ability to process phrasal verbs?
6. Should L2 learners of English whose native language also uses verb-particle combinations find phrasal verbs easier to acquire and process in the L2 compared to learners whose native language does not use phrasal verbs?

Suggested Student Research Projects

1. Design an experiment that investigates how native and non-native speakers' interpretation preferences for ambiguous verb-preposition combinations are affected by different types of biasingcontext. This could, for example, be an offline sentence completion task in which sentence fragments such as *Mary ate up* ___ are preceded by a context sentence or paragraph which either biases towards the literal (prepositional verb) or towards the figurative (phrasal verb) interpretation. For examples of suitable sentence fragments, see the materials provided in Matlock & Heredia (2002, exp.1). Do the proportions of phrasal vs. prepositional verb completions differ between the two context conditions, and are native and non-native speakers' completions affected in the same way by contextual biases? Can you suggest a paper or an article? Please be more specific as to how they could do it? Perhaps have them replicate a published paper? Please suggest where to get the stimuli, what stimuli to get, and how to do it.. see of example http://www.tamiu.edu/~rheredia/materials.html for software.. You can also suggest them to use the RSVP task you used. Can you provide it to them? Can they down load it from a site?
2. One diagnostic for identifying phrasal verbs in English is the fact that these normally require definite object pronouns to appear between then verb and the preposition (e.g., *She looked it up* vs. **She looked it into*). Design an acceptability judgment experiment to test whether L2 learners are aware of this grammatical difference between phrasal and prepositional verbs. You could ask participants to make binary (yes/no) judgments on grammatical and ungrammatical stimuli like those above. To help ensure that participants understand whether the figurative (phrasal verb) or literal (prepositional verb) reading is intended, the critical stimuli should be presented within appropriate contexts (e.g. *John could not remember Susan's number. Mary quickly looked it up.*). To verify whether participants are familiar with the phrasal and prepositional verbs used in the acceptability judgment task, you could additionally carry out a brief vocabulary check or paraphrase task, e.g. asking participants to paraphrase sentences such as *Mary looked up the number*. Do learners who are aware of the meaning difference between (figurative) phrasal verbs and prepositional verbs also make the correct corresponding acceptability judgments? If not, then what does this tell us about the mental representation of phrasal verbs in the L2? The same issue as number 1. Please provide some guidance…

Related Internet Sites

Even Related Potentials: https://en.wikipedia.org/wiki/Event-related_potential

Phrasal Verbs: http://en.wikipedia.org/wiki/Phrasal_verb
Phrasal verb dictionary: http://www.usingenglish.com/reference/phrasal-verbs/
Phrasal verb bibliography: http://mwe.stanford.edu/phrasalV.html
The N400 and meaning: http://www.youtube.com/watch?v=5d9DPhGSKVo&feature=relmfu

Suggested Further Readings

Dehé, Nicole (2002). *Particle verbs in English: Syntax, information structure, and intonation*. Amsterdam/Philadelphia: John Benjamins.

Moreno E. M., Rodriguez-Fornells, A., & Laine M. (2008). Event-Related Potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics, 21,* 477-508.

Mueller, J. L. (2005). Electrophysiological correlates of second language processing. *Second Language Research, 21*, 152-174.

References

Bobrow, S., & Bell, S. (1973). On catching on to idiomatic expressions. *Memory & Cognition, 1*, 343-346.

The British National Corpus, version 3 (BNC XML Edition). 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk/Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language, 27*, 668-683.

Cornejo, C., Simonetti, F., Aldunate, N., Ibanez, A., Lopez, V., & Melloni, I. (2007). Electrophysiological evidence of different interpretative strategies in irony comprehension. *Journal of Psycholinguistic Research, 36*, 411-430.

Coulson, S., & Van Petten, C. (2002). Conceptual integration and metaphor: An event-related potential study. *Memory & Cognition, 30*, 958-968.

Coulson, S., & Van Petten, C. (2007). A special role for the right hemisphere in metaphor comprehension? ERP evidence from hemifield presentation. *Brain Research, 1146,* 128-145.

Dagut, M., & Laufer, B. (1985). Avoidance of phrasal verbs: A case for contrastive analysis. *Studies in Second Language Acquisition, 7,* 73-79.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods, 134*, 9-21.

Di Sciullo, A. M., & Williams, E. (1987). *On The Definition of Word*. Cambridge, MA: MIT Press,

Estill, R. B., & Kemper, S. (1982). Interpreting idioms. *Journal of Psycholinguistic Research, 11*, 559-568.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*, 491-505.

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies*. Bloomington, IN: Indiana University Linguistics Club.

Friederici, A. D. (2004). Event-related brain potential studies in language. *Current Neurology and Neuroscience Reports, 4*, 466-470.

Geisser, S., & Greenhouse, S. (1959). On methods in the analysis of profile data. *Psychometrica, 24*, 95-112.

Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition, 8,* 149-156. Giora, R. (2003). *On our mind: Salience, context, and figurative language*. Oxford: Oxford University Press.

Hulstijn, J. H., & Marchena, E. (1989). Avoidance: Grammatical or semantic causes? *Studies in Second Language Acquisition, 11*, 241 -255.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203-208.

Laufer, B., & S. Eliasson. (1993). What causes avoidance in second language learning: L1-L2, difference, L1-L2 similarity, or L2 complexity? *Studies in Second Language Acquisition, 15*, 35-48.

Laurent, J. -P., Denhieres, G., Passerieux, C., Iakimova, G., & Hardy-Bayle, M. -C. (2006). On understanding idiomatic language: The salience hypothesis assessed by ERPs. *Brain Research, 1068*, 151-160.

Li, W., Zhang, X., Niu, C., Jiang, Y., & Srihari, R. (2003). An expert lexicon approach to identifying English phrasal verbs. *In proceedings of the 41st annual meeting of the ACL* (pp. 513- 520). Sapporo, Japan. Liao, Y. D., & Fukuya, Y. J. (2004). Avoidance of phrasal verbs: The case of Chinese learners of English. *Language Learning 54*, 193–226.

Matlock, T., & Heredia, R. R. (2002). Lexical access of phrasal verbs and verb-prepositions by monolinguals and bilinguals. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 251-274). Amsterdam: Elsevier.

McPartland-Fairman, P. (1989). *The processing of phrasal verbs by native and non-native speakers of English* (Unpublished doctoral dissertation). The City University of New York.

Mueller, Jutta L. (2005). Electrophysiological correlates of second language processing. *Second Language Research, 21,* 152-174.

Pynte, J. L., Besson, M., Robichon, F. H., & Poli, J. (1996). The time-course of metaphor comprehension: An event-related potential study. *Brain and Language, 55,* 293-316.

Regel, S., Gunter, T. C., & Friederici, A. D. (2011). Isn't it ironic? An electrophysiological exploration of figurative language processing. *Journal of Cognitive Neuroscience, 23,* 277-293.

Siyanova-Chantuira A., Conklin K., Van Heuven W. J. B. (2011). Seeing a phrase 'time and again' matters: the role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 776–784.

Siyanova, A., & Schmitt, N. (2007). Native and nonnative use of multi-word vs. one-word verbs. *International Review of Applied Linguistics, 45*, 109-139.

Sjöholm, K. (1995). *The influence of crosslinguistic,semantic and input factors on the acquisition of English phrasal verbs*. Turku: Abo Akademi University Press.

Strandburg, R. J., Marsh, J. T., Brown, W. S., Asarnow, R. F., Guthrie, D., Harper, R., Yee, C. M., & Nuechterlein, K. H. (1997). Event-related potential correlates of linguistic information processing in schizophrenics. *Biological Psychiatry, 42,* 596-608.

Swinney, D., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Learning and Verbal Behavior, 18*, 523-534.

Vespignani, F., Canal, P., Molinaro, N., Fonda, S., & Cacciari, C. (2010). Predictive mechanisms in idiom comprehension. *Journal of Cognitive Neuroscience, 22,* 1682-1700.