Patrycja Kakuba

404899

The role of animacy violation in novel metaphor processing and comprehension

Rola naruszenia żywotności w przetwarzaniu i rozumieniu nowych metafor

Rozprawa doktorska napisana w Szkole Doktorskiej
Nauk o Języku i Literaturze
na Wydziale Anglistyki
Uniwersytetu im. Adama Mickiewicza w Poznaniu
pod kierunkiem dr hab. Bogusławy Whyatt, prof. UAM
i dr Karoliny Rataj

Poznań, 2025

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my PhD supervisors, Dr hab. Bogusława Whyatt, Prof. UAM, and Dr Karolina Rataj, for their continued support and scientific guidance throughout this challenging process. I am truly thankful for sharing your expertise with me, for your constructive feedback, and your faith in my ability to complete my PhD. Your dedication and enthusiasm for research have continually inspired me on this scientific journey.

I would like to wholeheartedly thank Dr Marcin Naranowicz, who has supported me from the very beginning. Thank you for all the fun and memorable moments we shared while collecting data in the lab and for sharing your knowledge with me. Most importantly, thank you for your words of encouragement when I doubted the most.

I would like to extend my heartfelt thanks to my friends: to those who walked alongside me through the PhD journey, offering companionship on a path they know all too well, and to those who, even without fully understanding what I was working on, believed in me and kept me motivated. Thank you.

No words can fully express my gratitude to my family, who have always believed in me. I am endlessly grateful to my parents for their unconditional love and constant support during the most difficult times. Your unwavering faith in me is my greatest source of strength. A special thank you goes to my grandmother, Maria, who has been my greatest cheerleader since the very beginning of my education.

A heartfelt thank you goes to my partner, Artur Buja. You kept me grounded when I couldn't see the finish line and when I questioned everything. Your encouragement and belief in me gave me the strength I needed. Thank you for being my constant.

Last but not least, I want to express my deepest gratitude to all the participants, who quite literally made this research possible.

OŚWIADCZENIE

Ja, niżej podpisana

Patrycja Kakuba

przedkładam rozprawę doktorską

pt. „The role of animacy violation in novel metaphor processing and comprehension"

na Uniwersytecie im. Adama Mickiewicza w Poznaniu
i oświadczam,

że napisałam ją samodzielnie.

Oznacza to, że przy pisaniu pracy, poza niezbędnymi konsultacjami, nie korzystałam z pomocy innych osób, a w szczególności nie zlecałam opracowania rozprawy lub jej istotnych części innym osobom, ani nie odpisywałam tej rozprawy lub jej istotnych części od innych osób.

Jednocześnie przyjmuję do wiadomości, że gdyby powyższe oświadczenie okazało się nieprawdziwe, decyzja o wydaniu mi dyplomu zostanie cofnięta.

....

(miejscowość, data) (czytelny podpis)

Table of contents

List of tables

List of figures

Introduction

Metaphors have been found to occur, on average, once every 7.5 lexical units in discourse (Steen et al. 2010, analysis of the British National Corpus). A frequency that illustrates how metaphors are woven into everyday communication. Often used with little conscious reflection, metaphors serve as powerful tools that allow us to convey abstract meanings by drawing similarities with concrete concepts (Lakoff and Johnson 2011). In recent years there has been a noticeable increase in research on both conventional and novel metaphors. Studies have shown that metaphors can require different cognitive resources to derive meaning than literal sentences, especially if they are creative and novel (e.g., Lai et al. 2009; Rataj et al. 2018; Jankowiak et al. 2021). While much research has focused on processing demands of metaphors, figurative language can also involve animacy violation, which can be additionally cognitively taxing. Such violations occur through a combination of an inanimate noun and a verb or adjective that suggests animacy e.g., *healthy attitude*. Importantly, animacy has been observed to be a fundamental feature of human perception that emerged from evolutionary need for survival (Nairne et al. 2013) and plays a significant role in attention (e.g., Abrams and Christ 2003) and memory (e.g., Daley et al. 2020). Furthermore, the aspect of animacy has been observed to influence language processing (e.g., Kuperberg et al. 2007; Szewczyk and Schriefers 2011; Wang et al. 2020).

Animacy has recently gained attention in studies on language and cognition, however there is still scarcity of findings on the role of animacy and animacy violations in novel metaphor processing and comprehension, as a majority of studies on animacy have focused on literal language. Moreover, some inconsistent results have been reported in the few studies on the role of animacy violations in metaphor comprehension, and these studies were conducted on Mandarin Chinese, which has distinctive features. Therefore, the present study aims to examine the role of animacy in metaphor processing and comprehension by comparing novel metaphors with animacy violations (e.g., *To jest kolejna zmęczona książka 'This is another tired book'*) and novel metaphors without animacy violation (e.g., *To jest kolejna zatruta książka 'This is another contaminated book'*) in Polish using electrophysiology measures (EEG). The obtained results will broaden the perspective on the role of animacy violations in metaphoric language processing and comprehension.

Chapter 1 focuses on most prominent theories of metaphor processing and comprehension. It centres on the discussion of both early accounts (e.g., Standard Pragmatic View, Grice 1975) and more recent accounts (e.g., Giora 1997; Fauconnier and Turner 2003; Bowdle and Gentner 2005). It aims to showcase the intrinsic nature of metaphors and address observed changes in approaches and theoretical framing of metaphoric language. Moreover, the aspect of conventionality is discussed as one of the key factors influencing metaphor processing. Aspects reviewed in this chapter are crucial for the studies reported and discussed in Chapters 4 and 5.

Chapter 2 is devoted to a review of research methods frequently implemented in research on metaphors. The chapter discusses questionnaires, behavioural and electrophysiological methods, underlining the role of triangulation in language research. This discussion provides a background for the following section of this chapter, in which event-related potential (ERP) studies on metaphoric language and animacy are presented in detail. Moreover, this chapter focuses on the N400 and sustained negativity, components which have most frequently been examined in research on animacy and metaphor. This chapter also reviews relevant behavioural findings on metaphors. The hypotheses put forward in the present study were grounded in the discussed studies.

Chapter 3 discusses the role of animacy in language and cognition. This chapter highlights the gradient and context-dependent nature of animacy by reviewing studies that investigated the influence of animacy on cognitive functions (attention and memory) and language processing. This chapter starts with discussing the definition of animacy. Then an overview of electrophysiological research on the aspect of animacy in language processing is provided. At last, this chapter addresses the few available ERP studies on the role of animacy in metaphor comprehension. Similarly to Chapter 3, this chapter provides relevant background for hypotheses proposed in the present study.

Chapter 4 is devoted to the normative studies conducted prior to the ERP study discussed in Chapter 5. Studies have shown that many factors, such as conventionality, familiarity, and frequency influence metaphor processing and comprehension (e.g., Pynte et al. 1996; Arzouan et al. 2007; Goldstein et al. 2012; Rutter et al. 2012; Rataj et al. 2018; Jankowiak et al. 2021). Based on the findings and procedures described in the studies, the normative studies included ratings of meaningfulness, metaphoricity, familiarity and cloze probability of the created and selected novel metaphors with animacy violations, novel metaphors without animacy violations, literal sentences and anomalous sentences. This chapter begins with presenting the research hypotheses, followed by a discussion of the material selection, procedures, and obtained results.

Chapter 5 describes the ERP study conducted to investigate ERP responses to novel metaphors with and without animacy violations. The chapter begins with a presentation of hypotheses formulated based on previous ERP studies on metaphor processing and comprehension (Chapter 2), and animacy violation (Chapter 3). Moreover, these hypotheses were rooted in theoretical foundations (Chapter 1), which also provided a framework for a detailed discussion of the collected results. Furthermore, Chapter 5 reports on the materials, procedures, and both the behavioural (reaction times and accuracy rates) and electrophysiological results. The chapter finished with a detailed discussion of results, addresses potential limitations, and provides suggestions for further research on the role of animacy violations in metaphor processing and comprehension.

Overall, the present study aims to provide valuable insights into novel metaphor research, especially in relation to animacy violations. The methodological approach is situated within the framework of prior psycholinguistic research on metaphor processing and comprehension. This study contributes to the existing body of research as the role of animacy is underinvestigated in research on figurative language comprehension.

Chapter 1: Metaphor comprehension: a review of selected theories

1.1. Introduction

Aristotle first highlighted the unique nature of metaphors, stating that using metaphors is "the token of genius. For the right use of metaphor means an eye for resemblances" (Aristotle 1932: 1459a). Such a perspective extends the understanding of metaphors beyond their merely decorative function, linking them directly to processes of cognition. Over time, metaphors have attracted sustained scientific interest. This chapter focuses on providing theoretical background relevant to the studies discussed in Chapter 4 and 5. It is therefore devoted to discussing theoretical accounts of metaphor processing and comprehension. It also addresses key factors which have been observed to influence figurative meaning processing, such as conventionality. The aim of this chapter is providing a relevant account of metaphor theory and key claims.

1.2. Literal and figurative language

We are surrounded by figurative language in various means of communication without focusing much on its variety or complexity. For instance, we might say that a kind person has a 'heart of gold' or complain that something 'cost us an arm and leg' without focusing on the figurativeness of our statements. However, figurative language takes multiple forms, which may require different cognitive processing effort than literal language (Gibbs and Colston 2012: 19). To properly describe the distinction between literal and figurative language we should first address the challenge of defining these terms, which has proven difficult in language research.

From traditional theorists, such as Grice (1975), we would gather that literal meaning could be interpreted as what the speaker says, while what the speaker implicates could be defined as figurative meaning. However, Gibbs et al. (1993) point to at least five possible definitions of how to interpret literality, thereby underlining the complexity of language. Moreover, Gibbs (2002) emphasises that contrary to what was initially assumed about literal language that it is “coded, compositional, context-invariant, sentential, and truth-conditional” (Ariel 2002: 361), now has been challenged. More recently, studies show that comprehension of what the speaker says, which was in the past defined as literal, also depends on pragmatic knowledge and shares inferential processes with processing figurative meaning (Gibbs and Colston 2012: 23).

As for figurative language, there have been works devoted to distinguishing and defining varieties of figurative language into types, e.g. metaphors, idioms, irony, proverbs or metonymy (Barlow et al. 1971; Kreuz et al. 1996). However, Gibbs and Colston (2012) point out that the differences between different types of figurative language are not clear cut. Figurative language often blends multiple forms—like metaphor with irony or metonymy—making it difficult to classify each into separate and complete category. They propose a dimensional approach towards identifying and studying figurative language, rather than forcing figurative expressions into definite categories. The authors state that with this approach the researchers will need “to be more particular as to how each dimension of figurativeness contributes to the cognitive effort employed to interpret particular cognitive effects” (Gibbs and Colston 2012: 57). With the proved difficulty of providing clear, concise, and distinguishable definitions of literal and figurative language, we can state that there is a need to look at the meanings as a continuum rather than two different categories. Therefore, we can try to establish a continuum of meaning with placing literal meaning on one end and figurative meaning on the opposite end of the spectrum. In such a case, we can acknowledge that various instances of figurative language will hold different placements on the spectrum. For instance, more novel or poetic metaphors (e.g., *Her kiss is the sting of a bee*) would be considered more non-literal in comparison to highly conventionalised expressions (e.g., *John kicked the bucket*)(Gibbs and Colston 2012: 26). In support of this view, the variations on the literality/figurativeness continuum are reflected and justified by studies on figurative language where we observe the differences in processing effort, e.g. more conventional figurative utterances are processed faster (closer to the literal end of the spectrum) than more novel instances of figurative language (closer to the figurative end of the spectrum), and sufficiently supporting context can facilitate processing of even novel utterances.

Still, Gibbs and Colston (2012: 27) stress that there is “no single dimension along which all instances of literal and figurative meanings nicely align” and we need to acknowledge the complexity and variety of figurative language in our research. Within this diversity, metaphors are particularly noteworthy as they allow us to understand concepts through establishing connections between two conceptually different ideas. Acknowledging the difficulty to provide a concise definition for figurative language types, for the purpose of the research conducted and discussed in this thesis, in which metaphors are central, I would like to henceforth adapt a commonly used interpretation of metaphors as linguistic utterances whose intended meaning differs from its literal interpretation and can be derived from establishing links between two seemingly distinct concepts through cross-domain mappings (e.g., Lakoff and Johnson 2011; Gibbs and Colston 2012; Jankowiak et al. 2021). Recognizing the intricacy of figurative language, I will discuss aspects of metaphor comprehension to an extent that is relevant for this dissertation.

1.3. Theories of metaphor processing and comprehension

To provide a comprehensive overview of the evolving approaches to figurative language, it is essential to examine two predominant theoretical trends that formed over the years. I will begin with discussing the indirect access approach, and then challenge the view with more recently proposed direct access theories. This shift has significantly transformed the understanding of how figurative language is interpreted and processed and gives a comprehensive perspective on the complexity of figurative language processing.

1.3.1. Indirect Access theories on figurative language comprehension

Several theories have been proposed to account for mechanisms involved in figurative language processing. In the early stages of research on figurative meaning and its differences from literal meaning in language comprehension, it was postulated that literal meaning took precedence over figurative meaning during language processing, being easier to access and requiring less processing time. The Standard Pragmatic view (Grice 1975), which gave rise to a group of models called the Indirect Access Models of figurative language comprehension that described figurative language processing as a three-stage process. This view suggested an extensive complexity of figurative language comprehension. Within the approach when one encounters a figurative sentence, they must first compute the utterance, decide if the literal meaning of the utterance was the intended one considering the context, and then reject the literal meaning and compute the figurative meaning (Grice 1975). This theory has its foundation in Grice's conversational maxims provided as guidelines for effective conversation. These maxims most importantly state that speakers should be informative, truthful, relevant, clear, and brief in their communication. Based on Grice's maxims, figurative language violates the primary assumption of effective conversation. According to the theory, upon hearing figurative expressions such as metaphors, similes or irony that go beyond the literal meaning and include ambiguity in meaning, the hearer needs to compare the expression to the conversational context and upon rejecting the literal meaning find the meaning suited for the context (Grice 1975).

Grice's work was at that time the theoretical basis for research on metaphors and several studies partially supported the claim that comprehending metaphors is more cognitively taxing than literal sentences. However, over the years this approach has been mostly rejected by researchers. The main critique made against the Standard Pragmatic View is based on an argument raised by pragmaticians, psycholinguists and cognitive linguists that the three-stage process is psychologically implausible (Genovesi 2023). Studies conducted using experimental methods to compare processing speed of literal and figurative expressions show that both literal and figurative language can be processed in equal time to some extent. For example, Ortony et al. (1978) tested how context influences figurative language processing. In the self-paced reading task, the participants read ambiguous target sentences preceded by literal context or context allowing metaphorical interpretations of the target sentences. The results showed that in a supporting context there were no differences in time needed to read the sentences between the literal and metaphorical condition. The results suggest that adequate context allows for understanding metaphors without increased cognitive cost. Moreover, Mcelree and Nordlie (1999) conducted a study of which results provide proof that context is not necessary. In the first experiment participants read literal sentences (e.g., *Some temples are stone*), nonsense sentences (e.g., *Some clouds are stone*) and relatively conventional nominal metaphors (e.g., *Some hearts are stone*) in a semantic decision task, deciding if a given sentence was a meaningful or meaningless statement, while their reaction time was recorded. In the second experiment the participants were asked to judge whether the statements were literally true. The results showed that even though figurative expressions were less likely to be evaluated as meaningful, there was no evidence to support the theory that figurative expressions take longer to process. The results showed that even without context there were no differences that would support the Standard Pragmatic Theory. Now, over the years more studies have been published contradicting Grice's view on figurative language comprehension, providing evidence that arrival at the figurative meaning might not be as complex and time consuming as it was previously believed, especially when supported by sufficient context or conventionalized.

1.3.2. Direct Access theories on figurative language comprehension

The Standard Pragmatic View has been mostly rejected in favour of the Direct Access Models (e.g., Gibbs et al. 1993; Gibbs Jr. 2002), which suggest that upon encountering a figurative sentence, one does not have to analyse the literal meaning first, but rather is able to understand the figurative meaning immediately (Gibbs Jr. 2002). In contrast to the Standard Pragmatic View, the Direct Access Models do not assume that metaphor processing requires more complex cognitive processing and is more time consuming. Furthermore, the Direct Access approach shows that literal meaning does not hold priority over non-literal meaning in language processing. To this point, the Direct Access Models have gained most evidence. The fundamental assumption that grounds most of the research on metaphor comprehension is founded on the Conceptual Metaphor theory that proposes that metaphors can be analysed based on their construction and they convey meaning through establishing a relationship between two domains: one abstract (topic/target domain) and the other concrete (base/source/vehicle domain)(Lakoff and Johnson 2011). This can be illustrated by analysing a metaphor *Debt is an anchor* (Thibodeau and Durgin 2011). Comprehension of this metaphor is possible by drawing similarities and shared properties of the target domain *debt* and the source domain *anchor.* As a result, we can understand that debt in a burden and has a constraining effect limiting a person's freedom and progress, much like an anchor is heavy and keeps the ship in place. This section focuses on the most prominent theories on metaphor processing highlighting relevant evidence to the conducted study reported in Chapter 5.

Another theory was proposed by Fauconnier and Turner 2003. Broadly, the Conceptual Blending theory refers to "a basic mental operation that leads to new meaning, global insight, and conceptual compressions useful for memory and manipulation of otherwise diffuse ranges of meaning" (Fauconnier and Turner 2003: 1-2). More specifically it proposes that meaning is constructed by mapping features from numerous mental spaces into a new blended space forming a new meaning from combination of the input spaces. Gibbs and Colston (2012) explain the theory using the metaphor "*Surgeons are butchers*". Extending the Conceptual Metaphor theory (Lakoff and Johnson 2011), which proposes mapping features from the source domain (*butchery*) into the target domain (*surgery*), the Conceptual Blending theory accounts for the intended interpretation of the surgeon as incompetent, which is not provided by the Conceptual Metaphor theory. In this example, the blended space includes the elements from the target domain (e.g., process of surgery, the person conducting surgery, the place where surgeries are conducted) and the source domain (e.g., the butcher's tasks, using sharp objects, cutting meat). In addition, the blended space constructs a new meaning from juxtaposition of the two jobs (butchers kill and surgeons heal) thus arriving at the intended meaning of surgeons being incompetent. As such, this theory provides a more accurate account for novel meaning comprehension than the Conceptual Metaphor theory (Lakoff and Johnson 2011) as it focuses more on online meaning construction.

The Career of Metaphor model is one of the more current theories aiming at explaining the underlying mechanisms of metaphor comprehension. The theory was proposed by Bowdle and Gentner (2005) and is based on the structure mapping theory (Gentner 1983). Similarly to the Conceptual Metaphor theory (Lakoff and Johnson 2011), the structure mapping theory is founded on the assumption that when we interpret a metaphor, we map the structure of a familiar domain (source domain) onto a less familiar or abstract domain (target domain). This process involves identifying correspondences between the two domains, allowing us to understand the target domain through the lens of the source domain. For example, in the metaphor *the mind is a computer* (Bowdle and Gentner 2005), the familiar and well-known structure of the computer, such as, being able to process information, store data, and undergo complex operations, helps us understand the brain functions by drawing analogies between them. Importantly, the proposed theory is especially insightful with regard to novel metaphor processing as it explains the differences in processing novel versus conventional meanings. The Career of Metaphor model considers that metaphors either express similarities or category memberships. Bowdle and Gentner (2005) state that novel metaphors can be comprehended via an analogy, whereas conventional metaphors are processed through categorization. The theory states that novel metaphors are initially interpreted through a process of comparison involving alignment of the relational structure between the source domain and the target domain.

However, conventionalization changes the way we process metaphors and instead of a comparison we understand the conventional metaphors through categorization, where the source domain directly provides categorical information about the target domain, making the metaphor more immediately accessible. Thus, with conventionalization, the metaphorical meaning can be accessed more directly and effortlessly. This was tested by Bowdle and Gentner (2005) who elicited the switch from comparison to categorization in the lab by exposing the participants repeatedly to novel similes. In the experiment the participants first went through a training session, where they were presented with triads of novel similes constructed on the base term. Two of the phrases were provided, while for the third phrase participants were asked to fill in the target considering the meaning of the first two provided phrases (e.g., (1) *An acrobat is like a butterfly*, (2) *A figure skater is like a butterfly*, (3) (blank) *is like a butterfly*). In the second part of the experiment participants were presented with pairs of novel and conventional figurative phrases in form of similes (comparison) and metaphors (categorization), and were asked to decide which phrase is more sensible in each pair. However, some novel phrases were constructed on the base terms used in the first part of the experiment. The results showed that repeated exposure in the first part of the experiment increased participants' preference to metaphor (categorization) form. This shows that conventionalization facilitates shift from comparison to categorization.

In comparison, the Graded Salience Hypothesis (GSH) proposed by (Giora 1997) states that metaphor comprehension is governed by salience. According to the theory, when encountering figurative expressions, the more salient meaning is processed first. Salient meanings can be defined as more readily accessible, frequent, conventional, familiar and in agreement with the context (Giora 1997). Drawing on this theory, the salience of an expression (salient vs. less salient) influences how it is understood, with highly salient expressions being processed more quickly and automatically than less salient ones, which require more effortful processing. Hence, highly conventional metaphors or idioms, e.g. *spill the beans* will be process directly, arriving at the figurative meaning immediately as it is the more salient one (Gibbs 1980). Additionally, the GSH distinguishes between two types of processing governed by salience: sequential processing and parallel processing. Sequential processing happens when a less salient meaning is intended e.g., the literal meaning of a conventional idiom. In this case, the conventional understanding is the more salient and therefore is processed first before the intended less salient meaning is derived. Similarly, if the context supports a less salient meaning, the initially accessed more salient meaning is eventually overridden in favour of the contextually appropriate less salient meaning. Meanwhile, parallel processing will be activated if an expression has more than one salient meaning e.g., a conventional metaphor can have equally salient figurative and literal meaning, thus both meanings will be activated in the beginning (Blasko and Connine 1993). As illustrated, this theory stresses the key role that salience plays in figurative language processing.

Another aspect in metaphor comprehension that should be considered is aptness. Aptness refers to the extent to which the source domain provides an accurate representation of the important features of the target domain, with facilitated comprehension when those features are salient and relevant to the topic (Thibodeau and Durgin 2011; Jones and Estes 2006). For example, in the metaphor '*Memory is a warehouse',* the '*warehouse*' represents the category "vast but bounded spaces for storage" and equally '*memory*' can be considered a member of this category. Hence the metaphor should be easier to understand because of the shared connections between the vehicle and topic (Thibodeau and Durgin 2011: 1). This view also aligns with the Graded Salience Theory which underlines the importance of salience in metaphor comprehension (Giora 1997). More recently, Holyoak and Stamenković (2018) argued that processing of figurative utterances depends not only on conventionality but aptness of a given metaphor, which also can be "a better predictor of various performance measures than conventionality" (Holyoak and Stamenković 2018: 652). For example, Jones and Estes (2006) compared influence of conventionality and aptness in a study on metaphor comprehension. They reported that aptness and not conventionality predicted participants' preference for metaphors over similes (in a task where participants evaluated which of these is more sensible/natural), as well as ease of metaphor comprehension (in a task where participants read metaphorical utterances and pressed spacebar once they comprehended them). Jones and Estes (2006) also observed that aptness contributes to the processing speed of figurative utterances. The more apt and familiar the metaphor, the faster we are able to access its metaphorical meaning, even in the absence of prior context (Blasko and Connine 1993). Moreover, it has been noticed that the degree of aptness correlates with the novelty of a metaphor. In their study, Bowdle and Gentner (2005) observed a generally higher rating on aptness for metaphors considered conventional as compared to a lower rating for metaphors classified as novel.

The Relevance Theory emphasises that human cognition is designed to prioritize and maximize the relevance of meaning, aiming for the listener's optimal ratio of cognitive effort to cognitive effects in language interactions. This means that the listener aims at gaining the most cognitive benefit with the least cognitive effort in communication (Sperber and Wilson 1995 as cited in Gibbs and Colston 2012: 100). Therefore, this theory enforces the direct view of processing figurative language and according to the theory, metaphors are processed similarly to literal language. The idea underlining the importance of relevance is grounded in the assumption that human cognition is naturally oriented toward maximizing relevance (Gibbs and Colston 2012: 100). In this process, when our perception identifies potentially relevant stimuli, it automatically draws on our prior knowledge and assumptions, and as the result interprets the information in the most efficient and beneficial way (Wilson 2011). According to this theory, when encountering a metaphor, the addressee considers possible interpretations in order of accessibility. The search for the sufficient interpretation stops once the requirement of relevance has been fulfilled (Gibbs and Colston 2012: 101). Contrary to indirect access models, potentially increased cognitive effort in processing novel figurative language should not be attributed to the three-part process (accessing literal meaning first, then rejecting it and finding the figurative meaning) but rather to the strength of the implicature (implied meaning) of that metaphor (Gibbs and Colston 2012: 101). Based on the differentiation between strong and weak implicatures, we can interpret differences in conventional metaphor versus novel metaphors processing. The Relevance Theory indicates that conventional metaphors evoke at least one strong implicature and some weak implicatures as opposed to novel metaphors which elicit several weak implicatures instead (Gibbs and Colston 2012: 103). With novel metaphors, the addressee does not need to access and understand all weak implicatures condensed in the metaphor but rather is able to understand the meaning by only accessing enough interpretations to perceive the utterance as sufficiently relevant to the context (Gibbs and Colston 2012: 103). Therefore, the theory underlines the importance of context which can facilitate the metaphorical interpretation over the literal one, depending on how relevant the metaphor is in the given setting.

Another theory that emphasises the role of context in figurative language comprehension is the Constraint Satisfaction View (Katz and Ferretti 2001). This theory underlines the importance of all available linguistic and non-linguistic cues that people are exposed to when processing a metaphorical utterance to choose the most fitting interpretation of the utterance in the given situation (Gibbs and Colston 2012: 96). Katz and Ferretti (2001) supported the view with a study on proverb comprehension in a self-paced reading task with a yes/no comprehension question. In their study, the researchers tested familiar (e.g., *Lightning never strikes the same place twice*) and unfamiliar proverbs (e.g., *Straight trees have crooked roots*) embedded in contexts that were either facilitating their figurative or literal meanings. The results showed that familiar proverbs were read faster than unfamiliar proverbs enforcing the role of familiarity and conventionality in figurative language processing. Additionally, the unfamiliar proverbs were processed faster when embedded in contexts that biased their metaphorical interpretation. Yet, after inspection of the reading latency of the last two words, the figurative meaning of unfamiliar proverbs was not always fully processed even by the final word of the phrase, indicating more cognitive effort. Therefore, Gibbs and Colston (2012: 100) point out that contrary to the standard pragmatic view, graded salience hypothesis, and underspecification view, the Constraint Satisfaction View stresses the importance of many factors contributing to processing figurative language and thus does not support the idea that figurative language always requires a more complex and slower processing in comparison to literal language.

Moreover, The Underspecification View (Frisson and Pickering 2001) also supports the importance of context in metaphor processing. This theory proposes that upon encountering a metaphor, the mind retrieves a generalized and underspecified meaning of the word that can accommodate different interpretations depending on the context. The proposal can be explained with the following example: *Mrs. Graham is quite certain that they disarmed just about every critic who was opposed to spending more money on art* (Frisson and Pickering 2001 as cited in Gibbs and Colston 2012: 93). According to the theory, when the metaphorically used word *disarmed* is first encountered, it is processed with an underspecified meaning. Therefore, it is processed without interpreting the verb as either literal (e.g., *taking away physical weapons*) or metaphorical (e.g., *reducing opposition*). However, as the sentence unfolds further with the additional context, the meaning of the verb becomes specified. Additionally, the strength of the preceding context modulates the speed of processing. Hence, if the context is strong the meaning is accessed faster and with a neutral context the processing is slower (Gibbs and Colston 2012: 93).

The Deliberate Metaphor Theory proposed by Steen (2015) offers yet another perspective on metaphor use and comprehension in language. This theory proposes a distinction between deliberate and non-deliberate use of metaphors. The author argues that metaphors can be used deliberately on this example: "*Imagine your brain as a house filled with lights. Now imagine someone turning off the lights one by one. That's what*

Alzheimer's disease does. It turns off the lights so that the flow of ideas, emotions and memories from one room to the next slows and eventually ceases. And sadly -- as anyone who has ever watched a parent, a sibling, a spouse succumb to the spreading darkness knows -- there is no way to stop the lights from turning off, no way to switch them back on once they've grown dim. At least not yet." (Steen 2014: 182). According to the author, this example is a prove of deliberate use of metaphor where the reader is explicitly invited to form cross domain mappings by imagining the scenario. Steen (2015) argues that deliberate and non-deliberate metaphors are processed differently, with non-deliberate metaphors being processed automatically and deliberate consciously reflecting on the mappings. This theory has been criticized by Gibbs (2015) who questioned the distinction, pointing to the graded activation of metaphors, which according to psycholinguistic studies, depends on context, conventionality, or individual differences.

The discussed Direct Access Theories propose that figurative language can be understood without requiring an initial literal interpretation, challenging earlier models that assumed a sequential processing approach. The above discussion highlights the factors contributing to metaphor comprehension. Importantly, the aforementioned theories propose that meaning is accessed directly and provide a more accurate account for online processing of meaning. This aspect is relevant especially for processing of novel metaphors, which require increased effort in meaning construction. Additionally, the selected theories underline the importance of contextual cues, aptness, salience, categorization, comparison, or conventionality in construction of figurative meaning. The direct access approach significantly reshapes the understanding of how metaphorical expressions are processed and how meaning is derived. This discussion provides the theoretical foundation, which is of importance for the study reported in Chapter 5. In the study, novel metaphors were investigated with particular focus on the aspect of animacy violation within a non-constraining sentence context.

1.4. Special role of conventionality in metaphor processing and comprehension

In regards to selected models of figurative language comprehension, it is important to focus on the aspect of conventionality, which some models explicitly view as a determining factor in figurative language processing. In this section, I will focus on metaphors, addressing conventionality as a spectrum rather than a binary distinction between conventional and novel metaphors.

Handl (2011) describes the process of conventionalization of linguistic items and points to factors contributing to that process. She recognizes motivation as a factor influencing conventionalization. Cognitive linguistics, which approaches language in connection to other cognitive processes, such as perception or reasoning, rather than perceive language as an isolated entity, contributes to a thorough understanding of metaphors. Based in cognitive linguistics, human behaviour is generally motivated by both conscious and unconscious triggers. Thus, language as a subset of human behaviour is inherently motivated (Handl 2011: 53). Motivation refers to non-arbitrary connections between linguistic signs and their meanings (e.g., in an onomatopoetic word, content of the item motivates the form of an item). With such motivational links we retrieve meanings easier and repeated use of an item strengthens and creates more links to its meanings. Thus, frequency of use of a linguistic item has a direct influence on conventionalization, where higher frequency corresponds to greater conventionalization. Furthermore, Handl (2011) discusses two traditional notions that can lead to conventionalization of a lexical unit: institutionalization and lexicalization. Institutionalization refers to socio-pragmatic aspects and the process by which new linguistic items become widely accepted and integrated into the shared lexicon of a speech community (Handl 2011: 57). Lexicalization, in comparison, refers to a gradual historical process through which a linguistic unit becomes a fixed, conventional part of the lexicon, often losing some of its original morphological transparency and motivation, usually enhanced by frequency of use (Handl 2011: 57-58). Importantly, for novel or uncommon linguistic occurrences to be comprehended we already need to have a conceptual basis and connections that allow for their interpretation. Moreover, with time and frequency of use, the lexical units become more entrenched in the speaker's mind and thus are accessed more automatically (Handl 2011: 64). Entrenchment of a linguistic item is closely connected to the notion of salience, which refers to the prominence and ease of activation of a mental concept for the item. For novel or less familiar meanings, salience will be dependant on situational relevance or prior exposure to them. Having identified factors contributing to the process of conventionalization of linguistic items, the concept of conventionality of figurative language should be addressed. For the relevance to the study, I will discuss conventionality focussing on metaphors, acknowledging that other forms of figurative language (e.g., idioms, metonymy) can also be discussed with the aspect of conventionality.

Building on this discussion, conventionality is not a binary concept and should be perceived and addressed as a spectrum. Lakoff and Johnson (1980: 53) approached conventionality of metaphors by distinguishing three types of imaginative metaphors that can be contrasted with literal language. All examples are explained by the authors on the basis of the metaphor THEORIES ARE BUILDINGS. Based on the Conceptual Metaphor theory by Lakoff and Johnson (2011) which states that a conceptual metaphor is constructed on mappings between a concrete concept from which the metaphor is drawn (source domain e.g., *buildings*) and an abstract concept (target domain e.g., *theories*), which is understood in terms of the source domain. The first type of metaphor is based on "extensions of the used part of a metaphor e.g., *These facts are the brick and mortar of my theory*" (Lakoff and Johnson 1980: 53). In this example the metaphor is built on the conventional parts of the base metaphor by familiar mappings to the outer shell of the building. However, it extends the conventional aspect of the source domain (BUILDINGS), which is the outer shell, by using materials making up the building in that metaphor. Another type of metaphors are "instances of the unused part of the literal metaphor e.g., *His theory has thousands of little rooms and long, winding corridors*" (Lakoff and Johnson 1980: 53). This metaphor goes beyond the conventional scope of the parts of the source domain (BUILDINGS) that are commonly mapped into the target domain (THEORIES) and uses rooms and corridors to construct the metaphor. The last type mentioned by Lakoff and Johnson are "instances of novel metaphors, e.g. *Classical theories are patriarchs who father many children, most of whom fight incessantly*" (Lakoff and Johnson 1980: 53). Based on the example we can see that novel metaphors introduce entirely new mappings from unrelated domains (PATRIARCHS) to understand the target concept (THEORIES). These three types can be viewed as in Fig. 1.

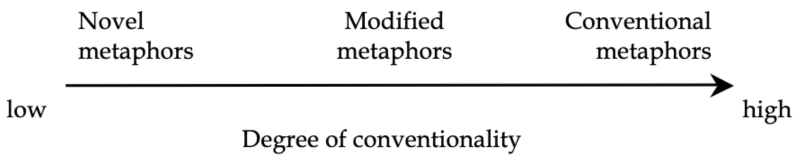

Fig. 1. The conventionality of metaphors as a continuum proposed by Handl (2011) based on Lakoff and Johnson (1980: 53)

Based on this visualization of the continuum, one would consider the first and the second type of metaphors discussed above proposed by Lakoff and Johnson (1980) as modified metaphors as they are based on conventional mappings but extend them introducing new aspects into the mappings. The third type can be considered novel metaphors as it introduces a new and creative conceptualization of the target domain. Adding to the discussion, Lakoff and Turner (1989) distinguish between two levels of conventionality: linguistic and conceptual. With such differentiation, novel metaphors can be defined both as unconventional on the linguistic and conceptual level. Modified metaphors and conventional metaphors, in comparison, are conventional on the conceptual level but modified metaphors can be less conventional on the linguistic level. However, the distinction is not straightforward and should be perceived in terms of degree. For instance, a metaphor constructed by extending existing correspondences is not conventional, but still is less novel than a metaphor created on entirely new correspondences (Handl 2011: 78).

Identifying criteria to determine the conventionality of metaphors can be demanding. There have been several proposals on how to distinguish metaphors according to their degree of conventionality. Traugott (1985: 22-23) proposed three parameters (reference, conceptualization, and distance) to distinguish conventional metaphors. Reference relates to the subjectively perceived figurativeness of a given utterance, which allows for distinguishing between more straightforward metaphorical instances and less obvious ones. Conceptualization evaluates the contrast between the source and target domain linked in a metaphor. Therefore, if the categories of the source domain and target domain are different (e.g., animate vs. inanimate) the metaphor can be more taxing as it involves more conceptual changes. This is especially true for novel figurative items. In comparison, conventionalized figurative expressions do not require increased processing effort to the same extent as novel expressions (e.g., Pynte et al. 1996; Lai et al. 2009; Rataj et al. 2018). At last, distance refers to whether the features selected from the source domain and mapped onto the target domain are central or typical within the source domain. Thus, metaphors are more conventional when they are based on prominent and central aspects of the source domain.

To investigate how the degree of conventionality affects cognitive processing, Sana et al. (2021) tested the interaction between conventionality and inhibitory control (ability to suppress irrelevant meanings). The study included two experiments. In

Experiment 1 participants were presented with metaphors (e.g., *a friend is an anchor*) and anomalous phrases (e.g., *fruits are furniture*) and performed a semantic decision task deciding if the presented phrases made sense. Importantly, the metaphors differed in levels of conventionality (conventional vs novel). The other task employed in this experiment was a flanker task, where the participants were presented with a row of arrows and had to respond to the middle arrow. In the congruent trials the middle arrow was pointing in the same direction as the others, whereas in the incongruent trial the arrow pointed in the other direction as the others. The flanker task was used to measure inhibitory control. The results showed that participants with poorer inhibitory abilities (scored higher reaction times, which mean that they were more susceptible to dominant and irrelevant responses) responded significantly faster to conventional metaphors in comparison to novel metaphors. In Experiment 2, the researchers exchanged the flanker task and manipulated the inhibitory load. In this experiment, half of the participants performed an odd-even task deciding if a presented number is even or odd (lower load), and half of the participants performed a n-back task, where they were presented with a sequence of numbers between each phrase and were asked to decide if the currently showed number was the same two trials before (higher load). The results showed that in the higher load condition conventional metaphors were judged more accurately. Overall, the results suggest that the degree of conventionality influences how much cognitive control is required.

In psycholinguistics the proposed theory and parameters were adopted into criteria of metaphoricity or semantic relatedness (Handl 2011: 81). To provide a thorough view on metaphors, in a study conducted by Katz et al. (1988), the researchers normed 464 metaphors on four defining aspects of metaphors: comprehensibility, metaphoricity, imagery variables, and sentence characteristics (e.g., familiarity, semantic relatedness). The selected dimensions allow for close inspection and control of metaphors, especially in a research design. When scored on a rating scale for metaphoricity, very conventional metaphors or frozen/dead metaphors, which are no longer discerned as strikingly figurative or nonliteral, will be evaluated as less metaphorical in comparison to novel metaphors. Therefore, adopting rating scales on the aforementioned parameters provides a reliable determination of the degree of conventionality of metaphors, which is particularly important within studies investigating figurative language.

1.5. Conclusions

This chapter examined the differences between literal and figurative language, highlighting and illustrating the intricate nature of figurative meaning. Furthermore, it provided a review of key theoretical accounts that have laid groundwork for studies on metaphors. Research on metaphors has evolved from early accounts of metaphors (Grice 1975) to the most acknowledged theories (e.g., Gibbs et al. 1993; Giora 1997; Lakoff and Johnson 2011). This theoretical review illustrated the development of approaches and perspectives on metaphors. For example, while the Conceptual Metaphor theory (Lakoff and Johnson 2011) provided powerful foundation, subsequent views like The Career of Metaphor model (Bowdle and Gentner 2005) or the Conceptual Blending theory (Fauconnier and Turner 2003) have broadened the understanding of novel and conventional metaphor processing and comprehension. The discussion of the theoretical background is particularly important for the present study reported in Chapters 4 and 5, which focuses on the role of animacy violation in novel metaphor processing and comprehension.

Chapter 2: Review of research methodologies and relevant studies

2.1. Introduction

The chapter is devoted to review of the research methods commonly used to investigate metaphor processing. This chapter discusses prominent research methods focusing on capturing both subjective judgements and objective measures of cognitive processes. The review focuses on questionnaires, reaction times (RTs), accuracy rates and electrophysiological measures. Discussion of this methodology is crucial for the study presented in Chapters 4 and 5, particularly because it employs methodological triangulation combining questionnaires, reaction times, and electrophysiological measures (EEG) to thoroughly investigate the role of animacy violations in novel metaphor processing and comprehension. Establishing this methodological background provides a solid foundation to the subsequent discussion of significant studies and theories, as well as the study design presented in Chapters 4 and 5.

2.2. Selected research methods used to investigate meaning processing and comprehension

An intriguing perspective on metaphors is provided by Gibbs (2017) who compares them to butterflies. He states that although butterflies can and are looked at when they are dead "neatly arranged, and pinned down, under glass" (Gibbs 2017: 317), they are also interesting to be studied in the wild interacting with the world. With this metaphor he stresses the point that metaphors can be analysed with various methodologies depending on the discipline. Literary scholars analyse metaphors in the context of a given literary genre, philosophers can analyse metaphors regarding conceptual aspects, and psycholinguists and neurolinguists adopt experimental methodology and neuroimaging techniques to investigate metaphor processing and comprehension. For this project, I will focus on selected methodologies relevant to the present study discussed in Chapters 4 and 5: questionnaires, reaction time (RT) method, and electroencephalography. To preface this section, it is important to note that a majority of experimental studies, including the one discussed in Chapters 4 and 5, integrate diverse measures including e.g., behavioural and brain imaging measures. Triangulation provides a greater perspective on observed phenomena by taking advantage of strengths of the measures. Additionally, multi-method methodology can enhance the validity of findings.

2.2.1. Questionnaires

Questionnaires are a common tool used for research. This method can be insightful as a stand-alone research tool but can prove especially beneficial in addition to other methodologies as a complementary source of information or a safeguard of stimuli selection before the main experiment. This research tool allows us to collect factual, behavioural and attitudinal data (Dörnyei and Dewaele 2022: 6-7). Factual data includes demographic information (e.g., age, gender), education background, socioeconomic status, and especially interesting in language research, language profile. As behavioural data we understand information about participant's habits, actions or behaviours (e.g., language learning strategies), and the attitudinal data includes participant's values, beliefs, interests, opinions, and attitudes.

The commonly used type of measurement in questionnaires are rating scales, especially the Likert scale. The scale was first introduced in the 1930s by Rensis Likert and now is a frequently used tool in research. The standard Likert scale focuses on the level of agreement with a given statement with possible responses ranging from *strongly disagree* to *strongly agree* where the participants choose the extent to which they agree or disagree with a given item. However, it is common to use other response labels (e.g., from *never* to *always*) that suit the research question better and can measure various attributes and aspects of language, communication, or psychological responses e.g., frequency of use, metaphoricity, valence, or familiarity. The Likert scale usually uses a 5-point or a 7-point rating scale, where each item on the scale has an assigned number for scoring purposes (e.g., *strongly disagree* = 1, *strongly agree* = 5) allowing for a middle option (e.g., *neither agree nor disagree* = 3). However, the design is up to the researchers to choose the most fitting scale for the research question and can range from a 2-point scale to a 9-point scale (Dörnyei and Dewaele 2022: 30-31).

The Likert scale is highly advantageous for norming studies in language research allowing to safeguard the stimuli selection for behavioural or physiological experiments. Studies have shown that in research on metaphor, conventionality, familiarity, meaningfulness, metaphoricity and cloze probability have significant influence on metaphor processing. Apart from cloze probability, the mentioned factors are measured using rating scales. Familiarity questionnaires check how frequently the participants encounter a given expression (by rating how often they encounter each sentence). Moreover, familiarity corelates with conventionality as typically novel metaphors are unfamiliar and conventional metaphors are familiar (Wang 2022). Meaningfulness questionnaires capture how interpretable are given utterances by asking the participants to evaluate how meaningful each of them is. In metaphor studies this measure is administered to ensure that metaphors and literal sentences are seen as meaningful (e.g., Jankowiak et al. 2017; Rataj et al. 2018). Moreover, metaphoricity is measured to show to what extend provided utterances are understood figuratively (e.g., Coulson and Van Petten 2002; Yang et al. 2013). At last cloze probability checks if the critical word is predictable due to the preceding context. Cloze probability can be measured by providing sentences and asking participants to fill in the missing critical word with a word that first comes to their mind creating meaningful utterances (e.g., Kutas and Federmeier 2011; Bambini et al. 2016; Jankowiak et al. 2017) Studies show that novel metaphors score lower on cloze probability in comparison to literal sentences (Wang 2022).

One of the benefits of using questionnaires is that this tool allows for large data collection since many questionnaire studies can be conducted remotely using online software (e.g., Qualtrics 2005). The usefulness of questionnaires as a data collection method was extremely visible during the COVID pandemic, when questionnaire studies were conducted with increased online participation (Dörnyei and Dewaele 2022: 2). This in turn enabled many scientists to continue their work despite the unfavourable circumstances for research. Another benefit is the time and effort needed to conduct questionnaire studies including data analysis. A well-structured questionnaire allows for efficient data collection and statistical analysis in comparison to other more time-consuming research methods like interviews or electroencephalography.

However, despite the advantages, questionnaires pose some limitations. The main criticism about questionnaire data concerns the validity and reliability of the findings. One of the challenges is the questionnaire construction. The overall design, as well as questions, should be concise and easily comprehensible to avoid participants' confusion and thus discrepancy in the collected data. Within the design, the researchers need to account for participant's possible lack of commitment by considering the length of the questionnaire to avoid the fatigue effect or adding *trap questions* (e.g., easy common-sense questions) that check if the participants are focused (Dörnyei and Dewaele 2022: 9). Particular attention should be devoted to providing clear and detailed instructions, explicitly informing participants of the task at hand (Dörnyei and Dewaele 2022: 22,152). On the same note, questions should be designed and formulated to accurately measure the intended phenomena. Additionally, items should be randomized to minimalize bias caused by the order (Dörnyei and Dewaele 2022: 140) and increase validity. Similarly, to avoid bias of response tendencies, there should be an equal number of items per experimental condition (i.e., each type of stimulus, such as sentence categories in language research) and filler items to balance the stimuli set. An unequal proportion of items per condition can influence how the participants perceive and evaluate the later items in the task (Birnbaum 1999).

Apart from the design, self-report research measures can provide unreliable results due to participants providing false responses intentionally or unintentionally. Some answers might not resemble the truth due to participants misunderstanding the questions, forgetting, or falsely remembering something. Sometimes, the issue may be the acquiescence bias (the participants agree with the sentence even when they are unsure and/or avoid strong responses on the ends of the spectrum) or the halo effect (the participants overgeneralize on a specific topic) (Dörnyei and Dewaele 2022: 13).Participants can also provide false information intentionally. This is mainly due to the social desirability or prestige bias, which means that they want to be perceived positively, hence they provide responses they see as more desirable or acceptable than what they actually feel or believe, especially in the case of sensitive questions (e.g., asking about illegal activities) (Dörnyei and Dewaele 2022: 10). However, it is worth noting that this bias can also affect factual questions (e.g., level of education, income, marital status, habits) (Oppenheim 1992) as the participants may desire to be perceived as more prestigious. Importantly, some factual or demographic questions (e.g., about marital status or mental health) can also be considered sensitive as they may evoke negative emotions in participants. To minimalize the bias, researchers ensure anonymity and/or avoid sensitive questions (Dörnyei and Dewaele 2022: 18).

These challenges can be mitigated through careful design, which can help reduce bias, improve response accuracy, and provide a more comprehensive understanding of the data collected, ultimately strengthening the overall validity of the research. However, the biggest limitation of questionnaires is that the method does not allow for in-depth inquires as their structured format limits the ability to explore complex, nuanced perspectives, follow up on participant responses or investigate cognitive processes. Therefore, the method is most insightful when combined with other methodologies.

2.2.2. Behavioural studies: the reaction times method

Reaction time (RT) methods, which focus on the time that passes from the onset of a stimulus to the response, usually measured in milliseconds (ms) (Grigore et al. 2015), are an important approach. This method was first employed to measure behavioural responses by Donders (1969). In his study he discovered and described three types of reactions varying on their time latencies. Since then, psychologists agree on the three types of reaction times being: simple reaction times, recognition reaction times, and choice reaction times (Luce 1986). Simple reaction times refer to the experimental task where there is one stimulus (e.g., light, sound, or word) and one response. Recognition reaction times are obtained in a go/no-go task, where participants are exposed to two types of stimuli: one that they should respond to, and one that should be ignored and serves as a distraction. At last, choice reaction times are recorded in a task where participants are asked to choose a key that corresponds to a given type of stimuli in a set that includes multiple stimuli (Harald Baayen and Milin 2010). Moreover, the types of experimental tasks can be combined to obtain other types of reaction times. For example, decision reaction times are a combination of simple and choice designs, where the participants are presented with one stimulus at a time (e.g., a word) and asked to choose a response corresponding to the type of presented stimuli. The number of possible responses is the same as the number of stimulus types (Harald Baayen and Milin 2010). The reaction times technique is widely used in many fields to research behavioural responses including investigation of various cognitive and linguistic processes e.g., semantic relatedness (Rataj et al. 2023), mood and bilingualism (Naranowicz et al. 2022b) or creativity (Rataj et al. 2018b). The reaction times technique is very often combined with neurophysiological methods to provide a broader insight. Importantly, in research on metaphor comprehension RT measures revealed differences in processing speed between literal and metaphoric expressions.

One of the strengths of the RT method is that it is relatively easy to administer with usually minimal equipment requirements (e.g., a computer with a keyboard, and software for stimuli presentation e.g., Presentation). Therefore, it is cost-efficient and quite easily accessible for researchers. Furthermore, compared to subjective reporting research methods like questionnaires, the reaction times are an objective measure. Hence, they provide more reliable and standardized data in comparison to self-reported measures. A major strength of the RT method is its versatility of application. It can be used in studies on attention, perception, decision-making, or motor control in various fields. This allows for a broad applicability in many research areas. Moreover, while the reaction times technique is primarily employed to measure the response speed, this method also allows to record the accuracy of the obtained responses. Collected accuracy rates can be analysed to further investigate certain behavioural processes. Noteworthily, accuracy rates can be an indicator of the comprehension of the instructions or task difficulty.

As a limitation of the RT method, we can identify the speed-accuracy trade off (SAT), a phenomenon in which increased response speed results in a decrease in accuracy of the responses or vice versa (Heitz 2014). This observation processes can be accounted for by modulating the research design to prioritize either speed or accuracy of responses. Katsimpokis et al. (2020) point to two commonly used experimental practices used to control for SAT depending on the aim of the study: experiment instructions and response deadline. In experiment instructions researchers can explicitly ask participants to focus on either response speed or accuracy, biasing SAT depending on the aim of the study. Response deadlines can have a form of text or sound after each trial to inform the participants that their response was not sufficiently fast. Such deadlines can quicken participants' decision making.

Another aspect that needs to be considered in RT studies is that depending on the aim of the study, the stimuli selection must be carefully controlled to avoid the impact of confounding variables, which could disrupt or bias the findings. Although the process can be challenging, control of the selected stimuli allows for deeper insight into the mechanisms underlying cognitive processes. In language processing studies we need to control the linguistic units (e.g., words) for several variables relevant to the studies, e.g. word frequency as high-frequency words are processed faster than less frequent words (McNamara 2005: 129), word length as shorter words have a faciliatory effect on language processing, whereas longer words are processed more slowly (Balota et al. 2002), or familiarity because more familiar lexical items are recognized faster than unfamiliar lexical items (Balota et al. 2006a). Therefore, stimulus selection can be time consuming as it needs to account for potential confounding variables that can influence the reaction times.

Another factor to consider when using the RT method in research is the individual differences that might influence the results such as age or fatigue among other factors. Studies have shown that reaction time tends to shorten from infancy through the late 20s and then progressively slow down with age, becoming more pronounced in complex tasks (Der and Deary 2006). Moreover, the slower reaction times may be due not only to physical factors but also to cognitive strategies like increased caution and attention allocation, affecting processing speed (Botwinick 1966; Redfern et al. 2002). Another considerable factor influencing the RTs is fatigue. Although all research methods are susceptible to the influence of fatigue, in the case of the RT method mental fatigue and sleepiness has a significant impact on reaction times especially in more complex tasks (Singleton 1953). At last, while the RT method can provide valuable insights into language processing, it provides a combined temporal measure of cognitive processes occurring during the task including language processing and decision-making. The RT method indicates the time taken for a response, on which we can base our interpretations. However, it does not allow to investigate the exact mental operations or strategies underlying the observed reaction times. Despite that, reaction time methods remain a valuable tool in psycholinguistics especially in combination with other research methods such as electroencephalography (EEG) or eye-tracking, as it allows for a complementary perspective on cognitive processes.

The versatility of the reaction time methods makes them valuable in language research, as they allow to capture the speed and effort of language processing. This is especially relevant for studies on metaphor processing and comprehension. Most commonly employed tasks in research on metaphor comprehension are self-paced reading task, lexical decision task and semantic decision task. The self-paced reading task measures the time that the participant spends on reading sentences, phrases or words, typically by pressing a button/key to move to the next segment. Moreover, the task can have different designs of stimuli presentation (e.g., whole sentence presentation or word-by-word presentations) (Kaiser 2013: 139-140). In the lexical decision task participants are presented (visually or verbally) with a string of letters and are instructed to decide if the presented lexical unit is an existing word usually by pressing corresponding button/key (Kaiser 2013: 138). Lastly, in the semantic decision task, which was used in the study discussed in Chapter 5, the participants make judgements of presented words or utterances. The judgments can be made about the meaning (e.g., deciding if a presented sentence is meaningful), semantic category (e.g., deciding if the word is an animal or object) or semantic features of a presented word or phrase (e.g., deciding if the given word refers to something that is animate (Jones and Estes 2012: 47-48). In language comprehension studies the semantic decision task usually takes form of a binary choice, in which the participants choose between two options (e.g., asking the question "Is this sentence meaningful?" providing two options: "yes" and "no")(Gomez et al. 2007).

2.2.3. Electroencephalography

In comparison to behavioural methodologies that give a glimpse into language processing and comprehension, one of the most insightful technologies available to psycholinguists has become electroencephalography (EEG). This non-invasive technique allows for an accurate tracing of cognitive mechanisms underlying language processing as well as language comprehension when meaning is established. Used for the first time on a human by Hans Berger, a German psychologist, in 1929 (Tudor et al. 2005), the method gave researchers opportunities to extensively study the human brain. Since the 1930s when the method was first developed, it proved to be beneficial for both scientific and clinical research. Although researchers were sceptical at first (Stone and Hughes 2013), the EEG technique is now not only frequently used in medicine and psychology but has also become popularized in many other research areas including linguistics. In linguistics, electroencephalography allows researchers to observe multiple language aspects, including emotions (Naranowicz et al. 2022a), creativity (Jończyk et al. 2020; Rataj et al. 2018a), semantic relatedness (Rataj et al. 2025) and novel metaphor comprehension (e.g., Rataj et al. 2018; Jankowiak et al. 2017; Jankowiak 2021; Li et al. 2022), which are central to the study reported in Chapters 4 and 5.

Electroencephalography measures the electrical activity of the brain by recording how neurons of the brain communicate with each other via electrical impulses, which produce small voltage changes that can be detected at the scalp's surface by placing electrodes on the scalp (Luck 2014: 45). The set up in language research usually contains 64 electrodes (see Fig. 2) but depending on the purpose of the study the set can have fewer (32) or more (128) electrodes. The electrodes are fastened in a dedicated cap and when placed on the participant's head they gather the signal that is later amplified and recorded (see Fig. 3). The recorded signal is then pre-processed and analysed as discussed below.

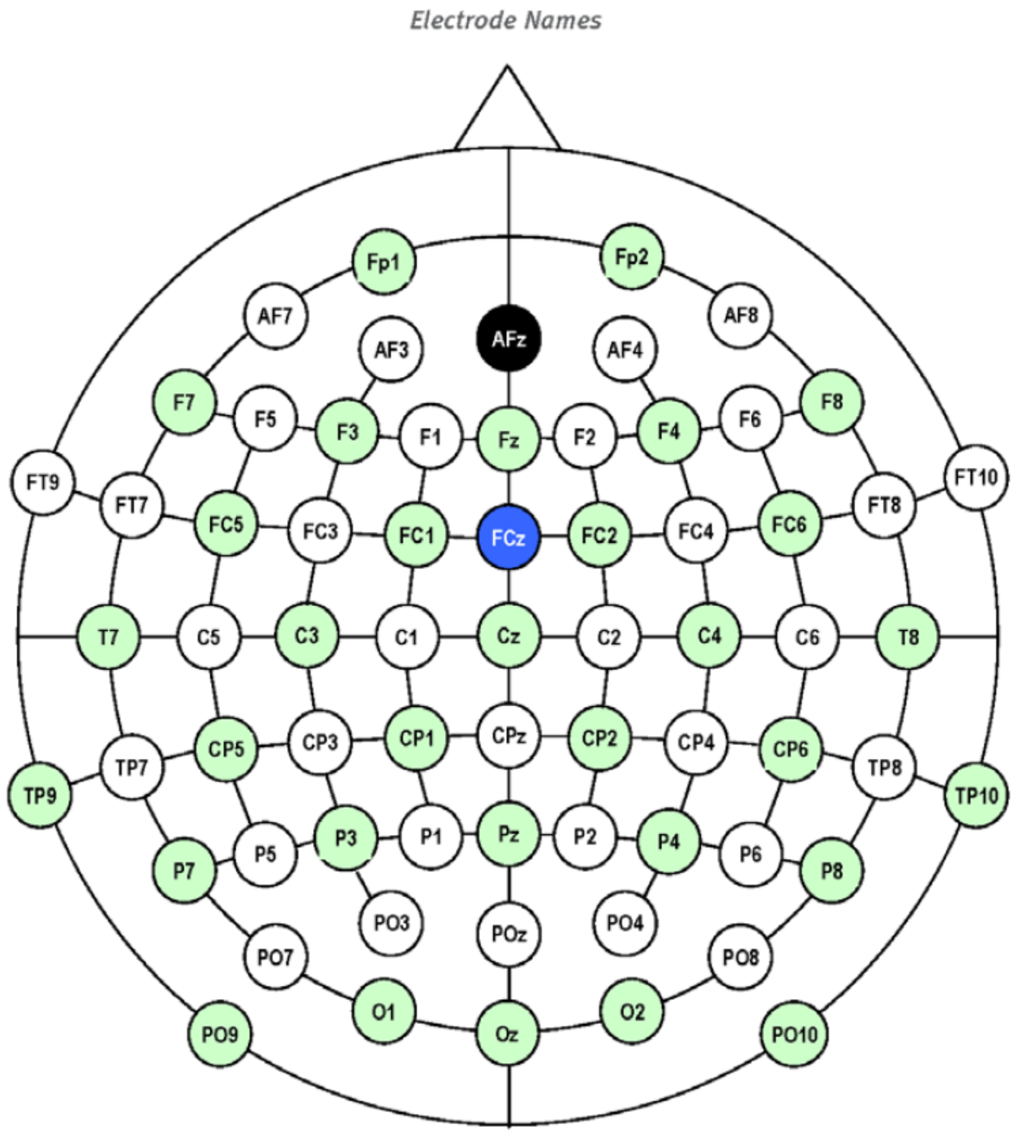

Fig. 2. Distribution of 64 electrode channels (after Rupasov et al. 2012: 6) The graph shows an example of an experimental setting in an ERP study (after Beres 2017).

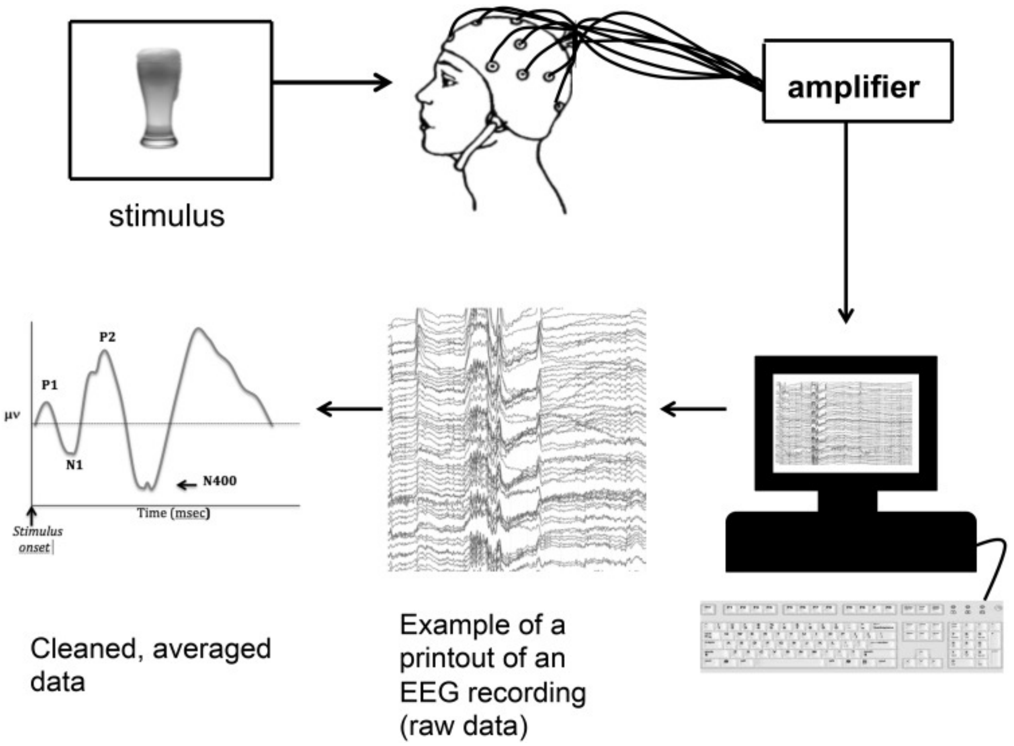

Fig. 3. The graph shows an example of an experimental setting in an ERP study (after Beres 2017).

2.2.4. Event-related potentials

The EEG technique, which is a non-invasive method of recording brain responses to stimuli, allows researchers to investigate complex language comprehension and cognitive processes. However, the continuous EEG recording is not very informative in such experiments, therefore the analysis is focused on the event-related potentials (ERPs), which are components recorded as responses to cognitive events (e.g., to a word, which appears on the computer monitor) (Kemmerer 2015). With such an approach we can observe brain responses to a time-locked stimulus (e.g., word, picture, sound) within an experimental task, capturing neural activity evoked by a controlled experimental event.

The ERP method is very insightful in language research as it allows to observe and analyse how the brain processes linguistic information in real time. However, to obtain valuable data for the analysis we need to extract the event-related potentials from the continuous EEG recording. The ERPs are small amplitudes indistinguishable in a raw

EEG recording, which is a mix of noise (random background activity) and signal (the ERPs). Therefore, to eliminate the noise and isolate the signal, the EEG data is time-locked to a stimulus event and then averaged across trials (Cohen 2015: 98). To analyse ERPs, the continuous recording has to be segmented into short segments, time-locked to stimulus presentation, called epochs. The short segments of interest, which are time-locked to the stimulus onset, usually include time span from 100-200 ms pre-stimulus baseline and last up to 800-1500 ms post-onset (Luck 2014). This time window allows for observation of the neural event elicited by a stimulus shortly before stimulus presentation, during and after. After extracting epochs, baseline correction is required. Some factors like skin potentials and hydration, or static charges in the electrode can cause the overall voltage of the signal to drift vertically, which in turn will result in error variance in the amplitude measurements (Luck 2014). Hence baseline correction is conducted to minimize the vertical offset. With this correction, the average voltage of the pre-stimulus period is calculated as a reliable representation of the voltage offset and with subtracting it from the entire epoch, we ensure that the brain response to the stimulus is measured from a consistent starting point of around zero voltage, making the results more reliable. Moreover, the EEG recording contains artefacts unrelated to brain activity, such as those caused by eye blinks and eye movements, muscle activity (e.g., facial expressions), heartrate and other noise such as electrical noise e.g., picked up from external electrical devices (Luck 2014). To establish a clear signal for the ERP analysis a specialized software is employed. The software allows for automatic artefact detection, manual artefact inspection, or independent component analysis (ICA) that helps establish the type of artefact interference and enables artefact removal. That process ensures that data selected for the analysis include brain activity and are not a result of noise. Event-related potentials, which are small waveforms, are difficult to detect in a single trial. Therefore, they need to be averaged together from many trials of the same stimulus presentation or event. Averaging enhances the noise-to-signal ratio, where the noise cancels itself out and specific brain responses remain consistent. This allows researchers to isolate the neural responses to an event.

As a result of the data processing, the obtained ERP components are suitable for statistical analysis. The most prominent and frequently studied event-related potentials in language research are: P100, N100, P200, P300, N400, P600 (as shown in Fig. 4). The letters correspond to the polarity of the peaks, with P standing for positive deflections, and N for negative deflections relative to the reference. The number corresponds to their approximate latency in milliseconds. The early components (P100, N100, P200) are known to be elicited by perceptual stimuli (e.g., a word) automatically as it is presented. In comparison, the later components (later than 250 ms) reflect cognitive processes during tasks. However, as Beres (2017) points out, the distinction between early and late components is illustrative and each peak is often task dependent. In next sections I will focus in detail on the event-related potentials relevant for the study on figurative language comprehension discussed in the thesis: the N400 (including the FN400), the Late Positive Complex (LPC), and the sustained negativity effect.

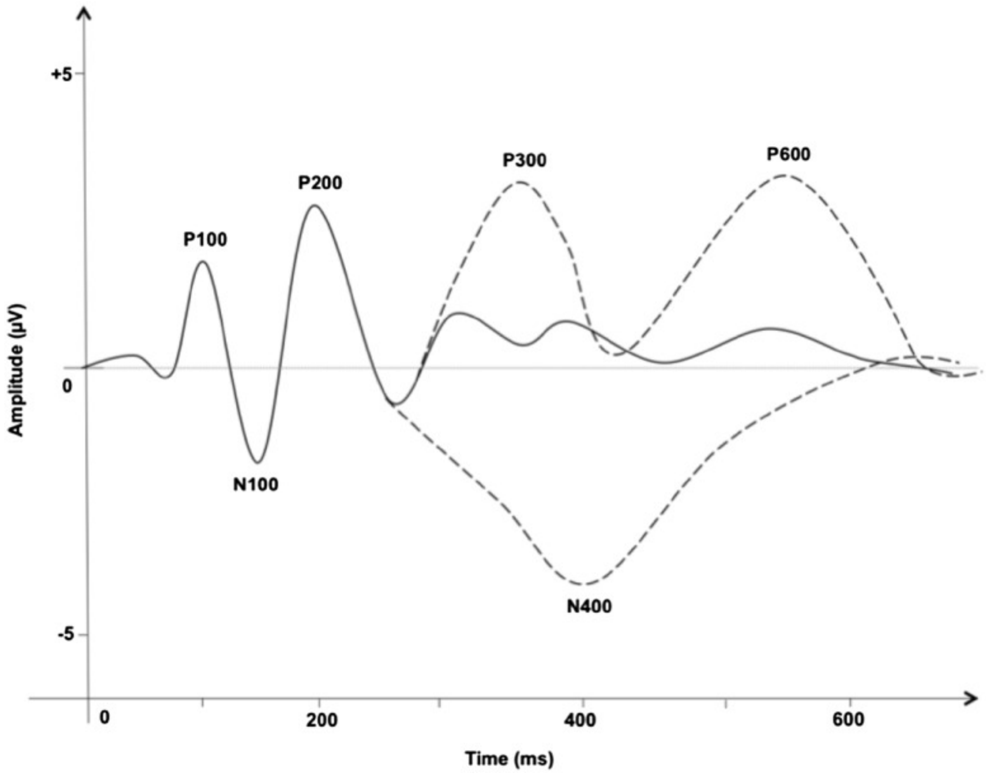

Fig. 4. Visualisation of most studied event-related components in language research (Luck 2014). The solid line represents ERPs related to basic perception, typically elicited in experiments using perceptual stimuli (e.g., a picture), while the dashed lines represent components elicited by particular cognitive or linguistic manipulations (after Beres 2017).

2.2.4.1. The N400

The N400, one of the most well-studied components in language research, is a negative-going deflection of a waveform that peaks around 400 ms after critical word onset. It is most prominent over the centro-parietal electrode sites and presents itself with a slightly larger amplitude over the right hemisphere compared to the left hemisphere (Kutas and Hillyard 1980; Luck 2014: 101). The N400 was first reported in the study by Kutas and Hillyard (1980), in which the participants read meaningless sentences (e.g., *I like my coffee with cream and dog*) and meaningful sentences (e.g., *I like my coffee with cream and sugar*) with a word-by-word presentation. The results showed larger N400 amplitudes in response to anomalous sentences as compared to meaningful sentences reflecting the brain response to the violation of the semantic expectancy. With such results, the study showed that the N400 is responsive to processes of meaning construction and semantic integration. Recently, the N400 component has been investigated in e.g., language comprehension, face and gestures processing, and semantic memory amongst others (Beres 2017). Yet the component is not merely a response to any type of violation or unexpected stimuli but is closely associated to meaning processing. One of the most well documented discoveries for the N400 is its sensitivity to incongruent stimuli in lexical and semantic priming tasks, where the N400 amplitudes were more enhanced when the target word was unrelated or weakly related e.g., semantically, categorically, or associatively to the preceding prime word (Kutas and Federmeier 2011).

The N400 has provided deeper insight and understanding of cognitive processes and semantic memory, and has been mostly discusses on three accounts (Delogu et al. 2019). The access/retrieval account proposes that the amplitude reflects cognitive effort of knowledge retrieval from long term memory with increased N400 suggesting more difficult access to semantic information. The effect can be facilitated or inhibited by the preceding context (Kutas and Federmeier 2011). In comparison, the integration account states that the N400 reflects the effort needed to integrate the elicited word meaning with the preceding context (e.g., Kutas and Federmeier 2011; Hagoort and Van Berkum 2007; Hagoort et al. 2009; Lotze et al. 2011). This approach combines the access/retrieval account and integration account concluding that the N400 displays both memory retrieval and context integration processes. Recently, Eddine et al. (2022) proposed yet another account of the N400: a predictive coding account. This account is based on the assumption that the brain makes constant expectations and the N400 is interpreted as an indicator of lexico-semantic prediction error at both word-level (e.g., lexical frequency, orthographic neighbourhood, semantic richness, repetition priming, semantic priming) and sentence-level (e.g., the effects of cloze probability, semantic overlap between expected and presented words). For example, to investigate semantic priming, the Predictive Coding Model (Eddine et al. 2022) on which the account is based, was presented with a sequence of a prime word (e.g., *lime*), two blank interactions, and with a target word (e.g., *corn*). In this example, the prime and target words share some semantic features (both can be considered plants). The results showed smaller lexico-semantic prediction errors for primed targets in comparison to unprimed targets. The model also showed smaller prediction errors (smaller N400 amplitudes) for high cloze probability words (e.g., *bread*) that low cloze probability words (e.g., *couch*) when preceded by a sentence context (e.g., *She spread the butter on her…*). Overall, these results suggest that the N400 reflects the amount of new lexico-semantic information encoded in the input that the brain has not already predicted.

In figurative language processing, the increased N400 has been interpreted to reflect the cognitive demands of mappings between domains and conceptual integration in figurative utterances in comparison to literal sentences (Kutas and Federmeier 2011; Cieślicka and Heredia 2019). Moreover, the N400 effect has been observed across different modalities such as speech production (Strijkers and Costa 2011) or sign language (Kutas et al. 1987). However, Kutas and Federmeier (2011) point out that the N400 effect is modality-dependent and manifests itself differently depending on the stimuli type (e.g., a word, sound, picture), especially in topography (e.g., novel metaphors evoke the N400 effect usually over cento-parietal region, pictures evoke a more frontal N400 effect) (Kutas and Federmeier 2011).

In studies on the N400 component, some researchers observed a more frontal negativity similar in morphology, timing and response pattern to the traditional N400 effect, maximal between 300 and 500 ms post-stimulus (henceforth, FN400) (Voss and Federmeier 2011). Throughout the debate on the function of the FN400 it was established that it is one of two components reflecting familiarity-based recognition, with the FN400 being sensitive to differences between previously encountered items (old) and previously unencountered items (new) in old/new recognition tasks (e.g., Woodruff et al. 2006; Yu and Rugg 2010; Bridger et al. 2012). The distinction between the N400 effect being associated with semantic processing and the FN400 elicited by old and new items has been challenged in a study conducted by Voss and Federmeier (2011). In their study, the researchers explored the connection between the FN400 and the N400 by incorporating a short-term semantic priming manipulation into a continuous recognition task. In this experimental design, each critical word was presented twice within the session and additionally preceded with a semantically related prime word either for their initial or second appearance. The participants evaluated each critical word performing two judgments: valence judgement (deciding if the word is positive or neutral/negative) and a recognition judgement (deciding if the word is old or new). The ERP responses for non-primed critical words were separated for analysis based on whether they were presented the first or second time to mimic the traditional old/new effect. When comparing the task conditions, the researchers found that the ERP differences between the semantic priming and old/new conditions during the 300–500 ms window were indistinguishable. Hence, the results lead them to conclude that there are no functional differences between the FN400 and N400. The study questioned the distinction between the FN400 as a marker of familiarity-based memory access and its differentiation from the N400 reflection of conceptual-semantic processing. However, the study has been criticised on a few accounts. Especially, on the assumption that the incorporation of semantic priming into a recognition task and then dividing the non-primed stimuli can accurately represent the familiarity-based recognition usually observed in standard old/new experimental designs. Even though the old/new responses were analysed separately, with short intervals between the repetition of the words, the electrophysiological response to the second presentation might have been influenced by semantic operations elicited by their first presentation (Bridger et al. 2012), thus leading to the incorrect conclusion. Since then, many studies rejected the hypothesis and differentiated the FN400 from the N400 across different tasks (e.g., Bridger et al. 2012; Stróżak et al. 2016; Leynes et al. 2024). Another interpretation of the differences between FN400 and N400 was introduced by The Unexpected Fluency Attribution (UFA) model proposed by Mecklinger and Bader (2020), which distinguished between the FN400 and the N400 in terms of familiarity. The UFA view distinguished between relative familiarity, which is based on recent encounters (e.g., elicited in a laboratory setting) and absolute familiarity, which incapsulates lifetime knowledge and experience. With such classification they argue that the FN400 reflects changes in relative familiarity and the N400 indicates recognition processes based on changes in absolute familiarity.

Research results have also shown a connection between frontal negativity and working memory. In a study by Kiss et al. (2007), the researchers investigated the differences between storage/maintenance (retaining and rehearsing information in mind for a short period) processes and control (actively managing and updating information, monitoring) processes in a running memory task. In the maintenance condition participants were presented with a series of single stimuli (digits 0-9 and letters) and the participants were asked to remember the first 2 items of the presented stimuli. After the series the participants were presented with a probe and decided if it matched the first two items or not. For the updating condition the procedure stayed the same, but the participants were instructed to remember the last two items of the series. Because the participants could not predict the length of the series, they had to update their memory each time a new item appeared. The results showed centro-parietal positivity in the time window of 450-650 ms post-stimulus onset for the maintenance condition at the second item in the series (they were asked to remember the first two items) and no similar activity for the subsequent items. The researchers argued that the positivity at the beginning of the series resembles increased effort in encoding to form a stable working memory representation with reduced demands for the rest of the series. As for the updating condition, the researchers observed increased frontal negativity. In the time window of 450-650 ms post-stimulus onset they observed an increased negativity over the right frontal region. In the time window of 650-900 in the left frontal region the maximum negativity was observed when updating of the full memory set was needed. The authors claimed that the frontal negativity effect pointed to increased processing load or task difficulty demands. Additionally, Kiss et al. (2007) suggest that the frontal negativity reflects initiation of working memory updating.

Frontal negativity in relation to working memory was also observed in an EEG study by Boudewyn et al. (2013). In their study, participants listened to short passages where the critical word was either lexically associated or associated with the prime word (e.g., *He was not prepared for the fame and fortune/praise*). Additionally, the preceding context was manipulated to either be congruent (e.g., context about winning a prize) or incongruent (e.g., context where the protagonist was arrested) with the last sentence. The aim was to investigate variability in working memory and cognitive control. The listening task was combined with a listening span test in which the participants were asked to listen to the sentences for comprehension and memorize the final word of each sentence in a set. After the set, the participants were asked to recall them. The listening span task was used to index individual differences in verbal working memory capacity (calculated as the number of correctly recalled items). The results showed a congruence effect with right frontal negativities in the 300-500 ms time window observed for participants with high listening span. For low listening span participants, the effect was observed in the left posterior site. The researchers state that such topographical difference should be interpreted as reflection of additional or distinct neural processes in working memory for higher span participants (who were able to recall more words). To conclude, the findings in the studies on working memory suggest that the frontal N400 effect can be modulated by working memory demands, particularly integration and updating in cognitively demanding tasks.

2.2.4.2. The late positive complex (LPC)

Another frequently observed component in language research is the late positive complex (LPC), which is relevant to the present study discussed in Chapter 5 as it has been linked to reanalysis processes observed in studies on animacy violation (e.g., Kuperberg et al. 2007; Szewczyk and Schriefers 2011) and metaphor comprehension (e.g., Coulson and Van Petten 2002; De Grauwe et al. 2010; Rutter et al. 2012). The LPC refers to a group of positive-going deflections of a waveform that is observed between 500-900 ms after the critical word onset. Originally, the late component was referred to as the P600 and was observed in response to syntactic errors e.g., *John hoped the man to leave* or syntactically unpreferred sentence endings i.e. garden path sentences, which are grammatically correct sentences involving an ambiguity of meaning that is only resolved at the sentence ending (e.g., *The broker persuaded… to sell the stock was sent to jail*, where *persuaded* is initially read as a main verb) (Osterhout and Holcomb 1992; Frisch et al. 2002; Gouvea et al. 2010). In addition to its sensitivity to syntactic irregularities, the LPC component also reflects sentence complexity (e.g., "who" questions such as *Emily wondered who the performer in the concert had imitated for the audience's amusement*) (Kaan et al. 2000: 164) with larger amplitudes for more complex sentences in comparison to simpler ones, despite both being syntactically correct (e.g., Osterhout and Holcomb 1992; Osterhout and Mobley 1995; Kaan et al. 2000). Moreover, in a study by Kaan et al. (2000), sentences with long-distance wh-dependencies (e.g., *Emily wonders who the performers in the concert imitate*) elicited the P600 effect relative to control sentences without such dependencies, suggesting that the late component reflects processes of syntactic integration.

The P600 was the first late positive component to be described, and although was initially associated with purely syntactic anomalies, since then it has been observed also in response to semantic aspects of sentences, first reported by Kuperberg et al. (2003). In their study, the researchers noticed the late positive effect as a response to thematic role animacy violations (e.g., *For breakfast the eggs would only eat toast and jam*), which they attributed to a monitoring process, checking process or an online process of repairing the sentence structurally to make sense. The semantic P600 was also found in studies by Burkhardt (2006, 2007) and Schumacher (2011) in response to sentences requiring additional processing at the discourse level. Furthermore, the late positive component was observed in studies on figurative language. For example, Regel et al. (2011) conducted a study where they observed that ironic statements embedded in stories, compared to literal ones, elicited the late component. Moreover, this effect was also observed in studies on metaphor comprehension (e.g., Coulson and Van Petten 2002; De Grauwe et al. 2010; Rutter et al. 2012; Yang et al. 2013), which will be discussed in detail later in this chapter.

Currently, the late components are collectively referred to as the late positive complex (LPC) (Yang et al. 2019). As discussed above, the LPC is no longer purely associated with language violations and incorporates the component's different functions, scalp distributions, latency and duration depending on the underlining mental processes (Delogu et al. 2019). Additionally, the LPC is thought to reflect "the word-by-word construction, reorganization, or updating of a mental representation of what is being communicated" (Brouwer et al. 2012: 138).

2.2.4.3. Sustained negativity

Noteworthily, within the late positive complex time window (500-900 ms) some researchers have observed a prolonged negativity without a well-defined peak and a more anterior topography in comparison to the N400 (Cui et al. 2018). The sustained negativity has been attributed to ongoing semantic integration processes (Tang et al. 2017) and continued effort to find the appropriate meaning (Goldstein et al. 2012). Importantly, the sustained negativity has been frequently observed in research on novel metaphor comprehension, which is the key interest in the study reported in Chapter 5.

Arzouan et al. (2007) observed the sustained negativity over the right hemisphere sites, elicited by novel metaphorical dyads (e.g., *conscience storm*), suggesting additional meaning integration for novel metaphors in comparison to conventional metaphors (e.g., *transparent intention*) and literal pairs (e.g., *problem resolution*). Moreover, the results can indicate access to the non-literal route for novel metaphor processing. Similarly, Rutter et al. (2012) found a graded effect for late negativity with most negative amplitudes for nonsensical phrases, followed by novel metaphoric phrases, and then literal phrases. The researchers interpreted this effect as continuation of the N400 component and thus a reflection of greater cognitive effort in semantic integration of two distinct concept. Furthermore, in the study on creative language processing by Rataj et al. (2018) a reduction in the LPC responses to novel metaphors were interpreted as reflecting an overlap with the sustained negativity effect for this type of sentences. Importantly, the observed amplitudes presented a similar topography to the N400 effect. Consistently with previous research, the results were attributed to increased demands in memory due to mapping processes of two distinct concepts or task complexity. Comparably, similar results were found in other studies on monolingual novel metaphor processing (e.g., Regel et al. 2011; Goldstein et al. 2012; Spotorno et al. 2013). Interestingly, similar results have been also observed in studies on bilingual novel metaphor comprehension by Jankowiak et al. (2017, 2021), which showed that regardless of the language, novel metaphors are cognitively taxing. Additionally, the authors interpret the results in favour of the Career of Metaphor model (Bowdle and Gentner 2005), which states that creative, unfamiliar and novel figurative expressions may rely on processes of identifying correspondences and comparisons between the source and target domains for a meaningful interpretation. Taken together, the mentioned studies illustrate the complexity of novel metaphor processing and the underlying mechanisms.

2.2.4.4. Strengths of electrophysiological research

The most prominent strength of electrophysiological methods is their excellent temporal resolution, which allows to precisely investigate neural processes millisecond by millisecond (Kappenman and Luck eds. 2011). Moreover, Cohen (2015: 24-26) states that EEG has high temporal resolution, precision and accuracy. Resolution refers to sampling rate that is to the number of data samples per given time unit (e.g., a few hundred or one thousand samples per second). Precision is defined as "the certainty of the measurementat each time point" (Cohen 2015: 24). Lastly, accuracy refers to the measure of how closely the timing of the EEG signal resembles the timing of brain electrical activity. Therefore, the high resolution, precision and accuracy of the ERP technique can be incredibly insightful in studies on language processing as it captures the rapid, complex, and multi-stage cognitive processes.

Moreover, the ERP technique is advantageous in comparison to behavioural measures as it allows for a continuous observation of brain activity in meaning processing. For example, in a reaction time study there is no data available before the presentation of the stimuli and between the presentation of the stimuli and the participant's response. In comparison, electroencephalography records continuous brain activity during the experiment, which provides information about the pre-stimulus state of the brain, neural responses to the stimulus, as well as the executive and cognitive processes before and after the response to the presented stimuli (Luck 2014: 25). Thus, the ERP technique is considered the most advantageous and informative technique to uncover and study human rapid mental processes such as perception, cognition, or emotions (Luck 2014: 27).

2.2.4.5. Limitations of electrophysiological research

Given the strengths discussed above, it is also important to acknowledge certain limitations associated with the use of the electrophysiological method. The main limitation of EEG is relatively low spatial resolution, precision and accuracy in comparison to other imaging methods with high spatial resolution such as fMRI (functional Magnetic Resonance Imaging) and MEG (Magnetoencephalography) (Cohen 2015: 26-27). Low spatial resolution, which refers to the ability to distinguish between different brain locations, can stem from the relatively low number of electrodes used in the typical EEG setup (most commonly 64 electrodes in language research, but the number can range up to 256). Moreover, each electrode records signal from multiple brain regions and not exclusively the neural activity directly beneath it. Hence, the obtained signal from an electrode cannot be confidently used to localize the source of activity. The setup also leads to lesser consistency in how the measurements represent a specific location (spatial precision) and reduced ability to pinpoint the actual anatomical source of neural activity (spatial accuracy). This limitation needs to be taken into consideration to avoid overinterpreting the results and making claims about precise locations of neural activity based just on the scalp EEG data.

Although the extensive time required to conduct studies is a common feature of experimental research on language processing and comprehension, it is still worth noting that compared to behavioural measures (e.g., reaction times) using EEG is more time consuming. Apart from the experiment, the preparation for the experiment, which includes head measuring, capping, placing electrodes and gelling needs to be considered. As well as cleaning and disinfecting the equipment after the experiment, which additionally extends the time devoted to one EEG session. Furthermore, in contrast to other behavioural measures the electrophysiological raw recording involves extensive data preprocessing to obtain data ready for statistical analysis. Adding to the aspect of time we need to conduct an EEG experiment, collecting a sufficient data sample can be somewhat time-consuming and demanding, taking into consideration the standard recruitment criteria for EEG language research (e.g., normal or corrected-to-normal vision and hearing, no neurological conditions, no implanted electronic/metal devices, not being diagnosed with psychiatric/mental disorders), which can narrow the number of participants qualified to participate in the study.

2.3. Electrophysiological research on metaphor processing and comprehension

Metaphors have long interested linguists as a means of expressing abstract ideas in the language. Using electrophysiological methods allows to uncover and examine the intricate cognitive mechanisms operating during figurative language comprehension. In this section I will focus on most prominent results in event-related potential studies on metaphor comprehension, significant for the study reported in Chapter 5. This section begins with addressing the influence of conventionality on metaphor processing. Then the impact of context is reviewed and at last the role of executive functions in metaphor comprehension will be addressed. All aforementioned aspects are of high significance in reference to the ERP study presented in Chapter 5.

2.3.1. The role of conventionality in metaphor processing and comprehension.

One of the most consistent results in figurative language processing was first found in a self-paced reading study conducted by Pynte et al. (1996), where the researchers compared comprehension of literal (e.g., *Those animals are lions*), familiar metaphoric sentences (e.g., *Those fighters are lions*) and unfamiliar metaphoric sentences (e.g., *Those apprentices are lions*) focusing on the changes in the N400 component. The results showed increased N400 amplitudes after the critical word (the last word of a sentence) in both types of metaphors in comparison to literal sentences. Since then, consistent results have been observed in many EEG studies on metaphors (Jankowiak et al. 2017; Rataj et al. 2018b; Lai et al. 2009; Coulson and Van Petten 2002). These studies have shown that one of the most important aspects that influences the N400 amplitudes is found to be the degree of conventionality. Furthermore, conventionality was investigated in a study conducted by Arzouan et al. (2007), which consisted of two experiments. In Experiment 1 participants performed a semantic decision task and were asked to decide whether presented word pairs conveyed meaning, and the task in Experiment 2 was a self-paced reading task, in which participants were instructed to read presented word pairs and think about their meaning. The study design included four types of word pairs: conventional metaphoric (e.g., *lucid mind*), novel metaphoric (e.g., *ripe dream*), literally related (e.g., *problem resolution*), and unrelated (e.g., *indirect blanket*) word pairs. The study aimed to compare cognitive processes between conventional and novel metaphorical phrases. The results showed a graded N400 effect with smallest amplitudes found for literal relations, intermediate for conventional metaphors followed by novel metaphors, and largest to unrelated word pairs in both experiments. Such an effect was attributed to retrieval of conceptual knowledge with the process being noticeably more demanding for novel metaphors in comparison to conventional metaphors and literal semantically related pairs. Moreover, in comparison to conventional metaphors, which revealed a left-hemisphere bias, novel metaphors showed a right-lateralized N400 effect. Regarding the LPC component, the amplitudes were attenuated due to sustained negativity for novel metaphoric phrases relative to semantically related and conventional metaphoric phrases. Additionally, exclusively novel metaphors showed a right-hemisphere late negativity.

Furthermore, Lai et al. (2009) conducted a study with a multiple-choice semantic decision task, which focused on the comparison of conventional (familiar and interpretable, e.g., *Every point in my argument was attacked*) and novel (less familiar but interpretable, e.g., *Every second of our time was attacked*) metaphors. The results showed more pronounced N400 responses to anomalous sentences (unfamiliar and least interpretable, e.g., *Every drop of rain was attacked*), novel metaphors, and conventional metaphors, compared to literal sentences (familiar and most interpretable, e.g., *Every soldier in the frontline was attacked*) in the early N400 time window (320-440 ms). In the late N400 time window (440-560 ms), however, the effect was more pronounced for anomalous sentences and novel metaphoric sentences relative to conventional metaphoric sentences and literal sentences. The effect had a centro-parietal distribution. Such findings suggest an ongoing processing effort for novel metaphors, similar to anomalous utterances. These results show that metaphor conventionality modulates metaphor comprehension.

Moreover, De Grauwe et al. (2010) investigated the processing of familiar nominal metaphors in a form of "A is B" by examining both the N400 and the LPC. In the study, researchers compared familiar nominal metaphors (e.g., *Unemployment is a plague*) with literal sentences (e.g., *Cholera is a plague*) and with semantically anomalous sentences (e.g., *Metal is a plague*) in a semantic decision task. The results showed that anomalous sentences and metaphors evoked larger N400 in comparison to literal sentences in the early N400 window (325-400 ms). In terms of the LPC, the familiar metaphorical sentences evoked a larger positivity effect than literal and anomalous sentences. The late positive responses were interpreted as ongoing attempt to understand the meaning of the metaphors and a process of resolving a conflict between implausible literal interpretation and plausible figurative interpretation.

A study by Goldstein et al. (2012) also focused on the role of conventionality in metaphor comprehension. The selected novel metaphoric, conventional metaphoric, literal and anomalous word dyads were adapted from the already discussed study by Arzouan et al. (2007). The ERP experiment by Goldstein et al. (2012) consisted of two parts. In the first part of the experiment participants performed a semantic decision task (50% of trials) and an explanation task (50% of trials). In the explanation task the participants were asked to explain the meaning of the given dyad or use it in a sentence. In the second part of the experiment, the participants were presented with the same stimuli as in the semantic decision task during EEG recording. The study revealed the effect of conventionalization on metaphors processing, reflected in attenuated N400 responses to previously explained (in the first part of the experiment) novel metaphors in comparison to unexplained novel metaphors. This finding indicates easier retrieval of semantic links between the target and source domain established during the explanation task. Interestingly, the results showed a reversed N400 effect for conventional metaphors with increased N400 amplitudes for previously explained relative to unexplained conventional metaphors. The researchers stated that this result probably reflects a shift from categorization to comparison elicited by the explanation task. That is, when the participants were explaining conventional metaphors in a similar manner to novel metaphors, they most likely reflected on the metaphors. This resulted in establishing new links between the domains and modulating the meaning integration process. Moreover, the analysis showed reduced amplitudes for novel metaphors in the LPC time window possibly caused by overlapping sustained negativity. The LPC amplitudes were increased for the unexplained novel metaphors in comparison to explained novel metaphors. This right-biased late negativity was attributed by the researchers to a continued process of reanalysis in attempt to derive meaning. The reduced LPC amplitudes for explained novel metaphors suggest reduced processing effort needed for meaning reanalysis.

Rutter et al. (2012) further investigated the influence of conventionality on metaphor comprehension and tested novel metaphor (e.g., *The clouds have danced over the city*) processing in comparison to literal (e.g., *The clouds have moved over the city*) and meaningless (e.g., *The clouds have read over the city*) stimuli. A graded effect was observed with the most negative amplitudes for anomalous sentences, then novel metaphors, followed by literal sentences in the early N400 time window (350-500 ms). The increased N400 amplitudes for novel metaphors and anomalous sentences were interpreted in line with the Conceptual Blending theory (Fauconnier and Turner 2003) as increased effort in establishing connections between two semantically unrelated and distant concepts. Moreover, the researchers observed sustained negativity in the late component time window resembling the N400 pattern, which according to them reflected continued effort in the reanalysis processes and integration of the concepts. The results can also be a marker of increased working memory load caused by the delayed response procedure employed in this study, which required participants to retain the meaning of a given utterance in their working memory until response (Rataj et al. 2018b).

A graded N400 effect was also observed by Rataj et al. (2018). Identically, the smallest N400 amplitudes were found for literal sentences (e.g., *Supposedly, there is a new theory*), the largest for anomalous sentences (e.g., *In the room, there stood an old, comfortable theory*), and intermediate amplitudes for novel metaphoric sentences (e.g., *He was amazed by another moldy theory*) in both the semantic decision task and the reading task with comprehension questions. The observed results suggest differences in processing effort between literal sentences, anomalous sentences, and novel metaphoric sentences. The observed N400 effect can reflect greater activation in semantic memory for novel metaphors in comparison to literal sentences. Regarding the LPC, the overlapping late negativity in the study by Rataj et al. (2018) reduced the amplitudes for novel figurative utterances in comparison to anomalous sentences illustrating additional effort needed to integrate novel metaphoric meanings.

Bambini et al. (2018) conducted a study investigating literary metaphors taken from poems and novels (e.g., *grass of velvet*) embedded in their original context, and literal expressions (e.g., *throne of velvet*). In the experiment participants read passages and completed a word-matching task after each trial. They were informed that some passages were literary but were not informed explicitly about the metaphors. The results showed that metaphors elicited increased both N400 and sustained negativity amplitudes in comparison to literal utterances. Moreover, the effects were stronger at frontal sites. The N400 effect has been interpreted as indication that metaphors result in increased difficulty in lexical access as they are novel and complex. The sustained negativity has been interpreted as reflecting increased working memory load evoked by novel metaphors and their context.

Studies described above show that regardless of the variety of research designs the results on metaphor comprehension tend to be consistent about the influence of conventionality in metaphor research. Moreover, studies on bilingual metaphor comprehension show a similar pattern of results. Jankowiak et al. (2017) conducted a semantic decision experiment on novel metaphoric, conventional metaphoric, literal, and anomalous dyads presented in Polish (L1) and English (L2). The graded effect was found in the late N400 time window (400-500 ms) in both the native and non-native context with the most attenuated amplitudes for literal sentences, followed by conventional metaphoric, novel metaphoric and anomalous word pairs. These results illustrated the increased cognitive effort in processing novel metaphors in comparison to conventional metaphors in both L1 and L2. Yet, modulation of results by language condition was observed in the late positivity time window. In L1, novel metaphors elicited smaller LPC amplitudes than literal utterances. However, in L2, both novel and conventional metaphors resulted in reduced LPC amplitudes relative to literal utterances. Such findings suggest a continued processing difficulty for novel metaphorical utterances in native language and for both novel and conventional metaphors in a non-native context. That assumption can also be supported by the N400 amplitudes being delayed about 20 ms for English stimuli in comparison to Polish stimuli. In another study, Jankowiak et al. (2021) compared novel nominal metaphors (e.g., *Love is a monastery*), novel similes (e.g., *Love is like a monastery*), literal sentences (e.g., *This monument is a monastery*), and anomalous sentences (e.g., *A carpet is a monastery*). Similarly to the previous study, the N400 effect was graded in L1 (Polish) with most robust amplitudes for anomalous sentences, then novel metaphors and novel similes, and smallest for literal sentences. However, unlike in the previous study, novel metaphors and similes converged with anomalous sentences in L2 (English) in the N400 time window. All three conditions resulted in increased amplitudes in comparison to literal sentences. The results imply that novel metaphoric utterances and similes are processed as nonsensical utterances in the initial stages in a bilingual context. Nonetheless, the results in the LPC time window did not reveal any between-language differences with an observed sustained negativity in L1 and L2. In line with studies on novel metaphor processing in the monolingual context, prolonged negativity for novel nominal metaphors was interpreted as continued meaning integration for novel or complex metaphorical meanings, or activation of a non-literal route during processing of semantically intricate novel metaphors (Jankowiak et al. 2021). Recently, Miller et al. (2025) tested novel metaphors (e.g., *The earthquake inhaled the city*), literal sentences (e.g., *The earthquake devastated the city*) and anomalous sentences (e.g., *The earthquake defrosted the city*) to investigate novel metaphors comprehension in second language conducted on L2 speakers of English in a self-paced reading task with comprehension questions. The results were analysed together with results from the study by Jończyk et al. (2020) conducted on English monolinguals using the same study design and stimuli. The results obtained from English monolinguals showed a graded N400 effect with most robust responses to anomalous sentences, followed by novel metaphors, and smallest for literal sentences. However, the results from L2 English speakers showed increased N400 amplitudes for anomalous sentences in comparison to literal sentences with novel metaphors failing to reach significant difference from either the anomalous or literal conditions. These results can be an indicator of lesser sensitivity to novel meanings in L2. The analysis of behavioural results showed lower accuracy exhibited by L2 speakers in comparison to monolinguals. The researchers observed correlation of proficiency and accuracy rates revealing that accuracy increased as proficiency did. Overall, the lack of a graded N400 effect for L2 speakers was interpreted by the researchers as a result of reduced sensitivity to anomalies exhibited by L2 speakers as well as lower accuracy rates for lower-proficiency participants.

The aforementioned studies illustrate the influence of the degree of conventionality on metaphor comprehension, indicating that novelty increases cognitive demands, requires increased effort to access and retrieve meaning from long-term memory, and results in prolonged difficulty in meaning integration.

2.3.2. The role of context in metaphor comprehension

In addition to the discussed body of research on metaphors, another important factor that plays a role in metaphor processing is context. In their study, Pynte et al. (1996) investigated the influence of context in metaphor comprehension by embedding novel and conventional metaphors in relevant and irrelevant contexts. The results showed increased N400 amplitudes and reduced LPC amplitudes for metaphors embedded in irrelevant context than relevant context. This finding showed a significant influence of context on metaphor comprehension. Further analysis of the N400 showed that novel metaphors elicited larger amplitudes when embedded in irrelevant context (e.g., *They are not idiotic: those fighters are squirrels*) as compared to reduced amplitudes for conventional metaphors in relevant context (e.g., *They are not cowardly: those fighters are lions*). Moreover, the N400 amplitudes were increased for novel metaphors in relevant context (e.g., *They are not cowardly: those apprentices are lions*) relative to conventional metaphors in irrelevant context (e.g., *They are not idiotic: those fighters are lions*). Such results underline the importance of both metaphor conventionality and context.

Moreover, Yang et al. (2013) tested the influence of context on metaphor comprehension. In this study, participants read novel metaphors (e.g., *The words are iron*) or literal sentences (e.g., *The metal is iron*) preceded by either a relevant or irrelevant prime word in a reading task. Through the manipulation of context five stimuli conditions were created: congruent target metaphor (e.g., *Powerful* - *The words are iron*), congruent source metaphor (e.g., *Hard – The words are iron*), congruent literal (e.g., *Hard – The* *metal is iron*), incongruent metaphor (e.g., *Silly – The words are iron*) and incongruent literal (e.g., *Silly – The metal is iron*) conditions. Contrary to the researchers' assumptions, no significant N400 effect was observed. However, the results revealed an effect of congruency in the LPC time window, with increased LPC responses to metaphors and literal sentences in incongruent conditions compared to congruent conditions. Within the incongruent condition metaphors elicited larger LPC responses than literal sentences. While in the congruent condition a graded effect was observed in the LPC time window with largest amplitudes observed for congruent target metaphors, followed by congruent literal sentences, and smallest amplitudes for congruent source metaphors. The results were interpreted by the researchers in line with the Conceptual Blending theory (Fauconnier and Turner 2003) by proposing that the increased late positivity points to greater effort in retrieval and integration of conceptual information from working memory. Regarding the graded effect, the researchers argued that the amplitudes for congruent source metaphor (e.g., *Spicy – The girl is cinnamon*) were smallest because the prime word *spicy* was related to the word *cinnamon* but did not support the metaphorical meaning (i.e., *The girl is sexy*), hence no reanalysis occurred. Whereas the congruent target metaphors (e.g., *Sexy – The girl is cinnamon*) showed increased reanalysis effort marked by the largest LPC amplitudes because the prime word allowed for metaphorical interpretation of the sentence.

Similarly, Jankowiak and Cieślicka (2025) investigated whether semantically related context facilitates novel metaphor processing. In their study, Spanish-English and English-Spanish bilinguals read novel metaphors (e.g., *Motivation is an engine*), novel similes (e.g., *Viruses are like travelers*), literal sentences (e.g., *These enclosures are cages*), and anomalous sentences (e.g., *An elbow is sugar*) in English, preceded by prime words either semantically related (e.g., STRENGTH - *Motivation is an engine*) or unrelated (e.g., BUN - *Motivation is an engine*) to the meaning of the presented sentences. Additionally, the language context was manipulated, thus creating four priming conditions: English related (English prime word was related to the meaning of the sentence), English unrelated (English prime word was unrelated to the meaning of the sentence), Spanish related (Spanish translation of the English related prime word), and Spanish unrelated (Spanish translation of the English unrelated prime word). The participants were instructed to decide if the presented sentence was meaningful or meaningless. The results showed a sentence type effect in both the N400 and LPC time windows. In the early time window, anomalous sentences elicited most pronounced N400 responses indicating increased lexico-semantic difficulty at the early processing stage. In contrast, in the late time window, increased LPC responses were observed for novel metaphors and similes. According to the authors, such result may reflect increased cognitive effort required for meaning integration for unfamiliar figurative language. Moreover, the results revealed an interaction between priming language and sentence type in the N400 time window in within-language priming, with anomalous sentences eliciting increased amplitudes relative to other conditions. Furthermore, an interaction between priming language and prime-sentence relatedness was observed in both the N400 and LPC time windows. The results showed increased N400 responses for the unrelated condition relative to the related condition, suggesting that related context can reduce cognitive effort in early semantic processing. In comparison, the unrelated condition elicited smaller LPC responses as compared to the related condition. The researchers suggest that the observed pattern reflects increased late semantic integration demands when the prime word is related to the sentence, as the related prime word needs to be considered and incorporated into final meaning interpretation. In contrast, unrelated primes can be disregarded as irrelevant early on, thus reducing later integration effort. Both effects were only present in the within-language condition. As for behavioural data, the researchers observed a sentence type effect with anomalous sentences eliciting shorter reaction times in comparison to other sentence types. However, no relatedness and language effects were observed. Acknowledging that the study was conducted with bilingual participants and primary focused on cross-language priming effects in bilingual novel metaphor processing, its findings nonetheless show the influence of context on language comprehension, especially for complex and novel expressions like novel metaphors.

Bambini et al. (2016) conducted a study on the influence of context cues on metaphor processing. In their study, in Experiment 1, the researchers embedded metaphorical (e.g., *shark-lawyer*) and literal (e.g., *shark-fish*) noun pairs into passages with minimal context, e.g. for literal *Do you know what that fish is? A shark* vs. for metaphorical *Do you know what that lawyer is? A shark*. In the study design, the researchers implemented an adjective matching task to conceal the presence of metaphors from the participants. As instructed, the participants selected one of two adjectives deciding which matches the preceding sentence better (e.g., for the utterances based on the noun *shark* the adjectives were *ferocious* vs. *geographical*). In contrast however, in Experiment 2, the noun pairs were embedded in supportive context (e.g., for literal noun pairs *That fish is really aggressive. It is a shark* vs. for metaphoric pairs *That lawyer is really aggressive. He is a* *shark*). Experiment 1 showed increased N400 and LPC amplitudes for metaphors relative to literal utterances. Whereas in Experiment 2, there was no significant effect in the N400 time window, while metaphors evoked enhanced LPC amplitudes in comparison to literal sentences. These findings suggest that the N400 is prone not only to metaphoricity, but also to context manipulations. With a supportive context, the lack of the N400 effect can be attributed to decreased processing costs as the concept has been already partially activated by the contextual cues. The observed LPC effect was interpreted by the researchers as a reflection of later pragmatic processes needed to interpret the meaning of the metaphorical utterance. The presence of the LPC effect in both experiments with minimal and supportive context shows that context does not impede the pragmatic mechanisms involved in metaphor interpretation.

Recently, Tang et al. (2025) conducted a study investigating the influence of context on scientific metaphors in Chinese. Scientific metaphors are figurative utterances based on concepts and specialized terminology taken from secondary school natural science subjects (e.g., mathematics, chemistry) (Tang et al. 2017). The stimuli consisted of four conditions: scientific metaphors based on similar shapes (e.g., *The conductor is a tunnel*) (SMS), scientific metaphors based on similar functions or concepts (e.g., *The antibodies are weapons*) (SMF), scientific literal expressions (e.g., *The conductor is an object*) (SL), and unrelated sentences (e.g., *The conductor is a cup*) (UR). The study consisted of two ERP experiments. In Experiment 1 the stimuli were presented without context in a semantic-decision task. In Experiment 2 the stimuli were preceded either by relevant or irrelevant visual context (e.g., a picture of a neuron) in a self-paced comprehension task, where the participants were instructed to go to the next trial only after they understood the picture. The effect of context was observed with reduced N400 amplitudes for all types of stimuli in relevant context condition as compared to no-context condition. The influence of context was also observed for the two types of metaphors. The results of the no-context conditions showed increased N400 responses for scientific metaphors based on similar functions (SMF) in comparison to scientific metaphors based on similar shapes (SMS) and scientific literal expressions (SL) suggesting increased processing difficulty for SMF. However, in the relevant context condition the difference disappeared indicating that supporting context can reduce processing difficulty. Similarly, in the no-context condition SMS elicited increased N400 responses in comparison to SL, and in relevant condition the difference disappeared. This observation shows that a supporting or relevant context may ease sentence comprehension. Furthermore, the analysis showed prolonged latencies in the LPC time window for all conditions in the presence of irrelevant context thus suggesting slower reanalysis process and meaning integration. Lastly, the behavioral results showed higher accuracy rates for metaphors preceded by relevant context in comparison to irrelevant context condition. As for RTs, no significant effect of context was observed. Such results can indicate that supporting context can facilitate access to metaphorical meaning, improving comprehension rather than processing speed.

Taken together the studies have shown that relevant and supportive context can facilitate metaphor processing and comprehension. These findings are relevant to the present study as they have shown that the N400 and the LPC time window amplitudes are prone to context manipulations. In the present study the sentence context was minimal.

2.4. Executive functions in metaphor comprehension and production

Recently, Menashe et al. (2024) investigated the role of attention in metaphor processing relative to the degree of conventionality. The study implemented a metaphor novelty evaluation task on nominal conventional metaphors (e.g., *Feeling happy is like visiting a field of flowers*) and novel metaphors (e.g., *Feeling happy is like visiting a field of cotton candy*) with addition of a short attention network test (ANT). The ANT was used to examine attentional functions focusing on three attentional networks: alerting, orienting, and executive control. In support of the Career of Metaphor model (Bowdle and Gentner 2005), the ERP results showed that the N400 peaked earlier for conventional metaphors, which require meaning retrieval from long-term memory in comparison to novel metaphors, for which the process of meaning retrieval and comparison of two concepts is more demanding. Moreover, N400 amplitudes for novel metaphors in the midline frontal regions of interest were associated with the executive control network. Such results suggest that executive control functions contribute to comprehension of novel metaphors. The lack of comparable results observed for conventional metaphors may suggest differences in how attention is recruited during the evaluation of conventional and novel metaphors.

Moreover, Silvia and Beaty (2012) investigated the relationship between intelligence and the ability to generate creative metaphors. In the study, the researchers measured fluid intelligence, which refers to the ability to consciously solve novel problems, through six fluid intelligence tasks. After the tasks, the participants were given brief descriptions of metaphors and were asked to create metaphors based on their personal experience (about the most boring school class and the most disgusting thing they ate or drank). The produced metaphors were then rated by three expert judges on creativity, novelty, and cleverness. The results showed that the participants who scored higher in fluid intelligence were able to produce more creative metaphors. Importantly, research shows a relationship between metaphor production and comprehension, and working memory.

Chiappe and Chiappe (2007) conducted two experiments. Experiment 1 included a listening span test, a Stroop task and a metaphor comprehension task. In the listening span test participants listened to sentences and were asked to fill in the last word with the most predictable word. Additionally, after each set, they were asked to recall all final words. To measure inhibitory control, in the Stroop task participants were presented with a string of letters and were asked to name the ink colour. For the congruent condition names of colours were written in that colour (e.g., *BLUE written in blue*) and for the incongruent condition names of colours were in a different colour (e.g., *BLUE written in red*). In the metaphor comprehension task, the participants read metaphors (e.g., *Lawyers are sharks; My brother is a garden*) and were asked to provide interpretations. The set of metaphors differed in level of topic constraint (high constraint e.g., lawyer vs. low constraint e.g., brother) and vehicle ambiguity (e.g., *shark* has few clear attributes, whereas e.g., *garden* allows for more associations). The results showed that participants with a high memory span provided better and faster metaphor interpretations and performed faster and more accurately in the Stroop task. Experiment 2 included a retrieval fluency task where participants were asked to name as many things as possible in a provided category. This task replaced the Stroop task and was used to measure inhibitory control. Instead of metaphor comprehension, Experiment 2 included a metaphor generation task. In this task, participants were provided with metaphors with a missing vehicle (e.g., *Some billboards* (blank) and an expected meaning (e.g., *billboards are something noticeable and unattractive*). The participants filled in the blanks with a vehicle which fitted the provided attributes (e.g., *warts*). The results revealed that participants who scored higher in the listening span and retrieval fluency tasks provided more apt metaphors in comparison to participants with lower scores. Taken together these results show that working memory capacity influences metaphor comprehension and production.

Furthermore, Mashal (2013) compared the role of working memory in familiar and unfamiliar metaphor comprehension. The study included three experiments. In Experiment 1 participants performed a digit span test where they heard a sequence of digits and were asked to recall the sequence in the same (requires short term retention) or reverse (requires executive control) order. Then they performed a metaphor comprehension task. The participants were provided with familiar (e.g., *defence line*) and unfamiliar (e.g., *transparent moment*) utterance and were asked to choose the best interpretation out of four provided alternatives. The results showed higher accuracy for familiar metaphors. Moreover, the accuracy in metaphor comprehension (regardless of familiarity) showed a positive corelation with backward digit span. These results were interpreted as an indicator of the importance of working memory in metaphor comprehension. In Experiment 2, the researchers used a metaphor recognition task in which participants first listened to a short two-word expressions and then were provided with the first word of the dyads and chose the second word from four options. The more difficult version of the task was also introduced with a longer list of dyads to memorize. In Experiment 3 the free recall task was employed. The participants listened to a short list of dyads and were asked to recall as many as possible. The results showed that unfamiliar metaphors resulted more often in semantic errors (participants chose semantic distractors i.e. related in meaning to the correct target) than in phonological errors (participants chose phonological distractors i.e. similar in phonology to the correct target). Experiment 3 also showed that unfamiliar metaphors were more difficult to recall. To conclude, the results were interpreted in support of the relationship between working memory capacity and metaphor comprehension, specifically increased executive functions needed to suppress irrelevant information.

To conclude, the studies have shown that executive functions play a significant role in novel and conventional metaphor processing and comprehension. These findings suggest that novel metaphors require increased engagement of working memory and that executive control can contribute to their comprehension.

2.5. Behavioural responses in electrophysiological studies on metaphor comprehension

Although behavioural methods, in particular reaction times and accuracy rates, do not provide such an in-depth view of metaphor processing as the event-related potential method, they add an important perspective to the picture of the underlying processes of understanding metaphors. In the study by Arzouan et al. (2007), the novel metaphoric phrases elicited longest RTs, followed by semantically unrelated dyads, conventional metaphors and semantically related dyads in both semantic decision and self-paced reading tasks. Such results were in line with electrophysiological results that indicated increased processing demands for novel metaphoric expressions. Additionally, there was no significant difference in RTs between conventional metaphors and semantically related pairs. Similar results were obtained in a study by Lai et al. (2009) in which the researchers observed a graded effect in reaction times with longest RTs for novel metaphors, then anomalous sentences, conventional metaphors and with shortest RTs for literal sentences in the multiple-choice semantic decision task. Comparably, Goldstein et al. (2012) noticed longest RTs for novel metaphors followed by anomalous sentences, then conventional metaphors and at last shortest for literal sentences in a semantic decision task. As previously discussed in the description of the study, novel metaphors were tested in two conditions: explained and unexplained. The provided explanation modulated reaction times resulting in shorter RTs for explained novel metaphors in comparison to unexplained metaphors. These results show the influence of conventionality on metaphor comprehension. Moreover, the graded effect was also observed in a delayed decision task in the study by Rutter et al. (2012). Such a graded effect in reaction times has been observed in multiple studies, demonstrating longer responses to novel metaphors in comparison to anomalous sentences, conventional metaphors and literal sentences (e.g., Rutter et al. 2012; Tang et al. 2017; Rataj et al. 2018; Jankowiak et al. 2017, 2021; Menashe et al. 2024). The longer RTs can be interpreted as a reflection of increased difficulty in processing novel metaphoric utterances. Moreover, such findings demonstrate how the degree of conventionality influences metaphor comprehension showing that meaning construction required in processing novel meanings is more demanding and time-consuming than meaning retrieval for conventional meanings (e.g., Arzouan et al. 2007; Goldstein et al. 2012; Jankowiak et al. 2017).

Another behavioural aspect which provides valuable insight into novel metaphor processing are accuracy rates. Studies show that novel metaphors are more taxing in comparison to conventional metaphors resulting in lower accuracy rates (e.g., Arzouan et al. 2007; Lai et al. 2009; Rataj et al. 2018; Jankowiak et al. 2017, 2021). Noteworthy, task accuracy should not be considered as an objective measure of correctness but rather as the degree of agreement with scores obtained in the normative tests. Moreover, similarly to patterns observed in event-related responses and reaction times, lower accuracy for novel meanings further indicates that the degree of conventionality impacts comprehension. Low accuracy scores for novel metaphors show that they are often classified as anomalous, which suggests increased effort in categorizing novel metaphorical expressions as either meaningful or meaningless, especially in a binary decision task (Rataj et al. 2018b) with a time limit to provide a response (Arzouan et al. 2007).

2.6. Conclusions

The methodological approaches reviewed in this chapter, including questionnaires, reaction times, accuracy measure, and electrophysiological measures, offer valuable insights into the cognitive mechanism underlying metaphor processing, especially when employed together. Questionnaires are very useful for norming stimuli before the experiments, the reaction time technique measures processing speed and efficiency, and event-related potentials allow researchers to investigate the time course of specific cognitive processes. This has been illustrated by the studies on novel metaphor processing and comprehension discussed in this chapter. Moreover, the reviewed research emphasized the role of conventionality, context, and executive functions in novel metaphor processing and comprehension. Taken together, this chapter has outlined research methods employed in the present study and summarized the most relevant findings in novel metaphor research that provide the foundation for the analyses and results discussed in Chapters 4 and 5.

Chapter 3: The role of animacy in cognition and language comprehension

3.1. Introduction

Animacy is a fundamental concept in both language and cognition and this chapter aims to explore the role of animacy across both domains. Traditionally, the concept of animacy refers to the distinction between living and non-living entities. However, animacy extends beyond pure binary biological distinction and proves to be rather a gradient and context-dependent concept, which will be acknowledged in this chapter. To illustrate the complexity of the concept of animacy, the chapter first provides an overview of definitions of animacy, exploring different approaches and attempts to operationalize it in research. Later, it addresses how the animate vs. inanimate distinction emerges early in cognitive development. Importantly, the role of animacy in cognition and language processing is discussed to provide foreground for the study discussed in Chapters 4 and 5. By examining both behavioural and neurophysiological evidence this chapter aims to illustrate the role of animacy in memory, attention, grammar and language processing. Lastly, this chapter draws attention to the potential influence of animacy on figurative language processing and underlines the importance of further research on the concept. It is crucial to address this aspect at it is extremely relevant to the study reported in Chapters 4 and 5, which investigated the role of animacy violations in novel metaphor processing.

3.2. Defining animacy

Across numerous studies, animacy has been shown to be a fundamental concept in human cognition, influencing attention, memory, grammar acquisition and language processing. To underline the importance of the concept of animacy in the understanding of the world, researchers have focused on the cognitive differences between the perception of animate vs. inanimate entities. There have been multiple attempts to fully define and operationalize the concept of animacy, employing different methodologies and manipulating various variables, and there is now a clear agreement in research that animacy is not purely a binary concept (living vs non-living). Early accounts of animacy focused on the differentiation between animate and inanimate entities. Gelman and Spelke (1981) highlighted fundamental differences between animate and inanimate entities, emphasizing the autonomous agency of animate entities, their capacity for growth and reproduction, and their cognitive abilities including perception and thinking. However, based on more recent studies a more graded approach seems more suitable. Yamamoto (1999) proposed the General Animacy Scale (Fig. 5.) illustrating the gradient nature of animacy. The theory is based on human's anthropocentric view of the world, in which the human being in recognized as most animate and therefore is situated in the centre of the figure. By the same token, the further from the centre the entity is on the scale, the less animate it is.

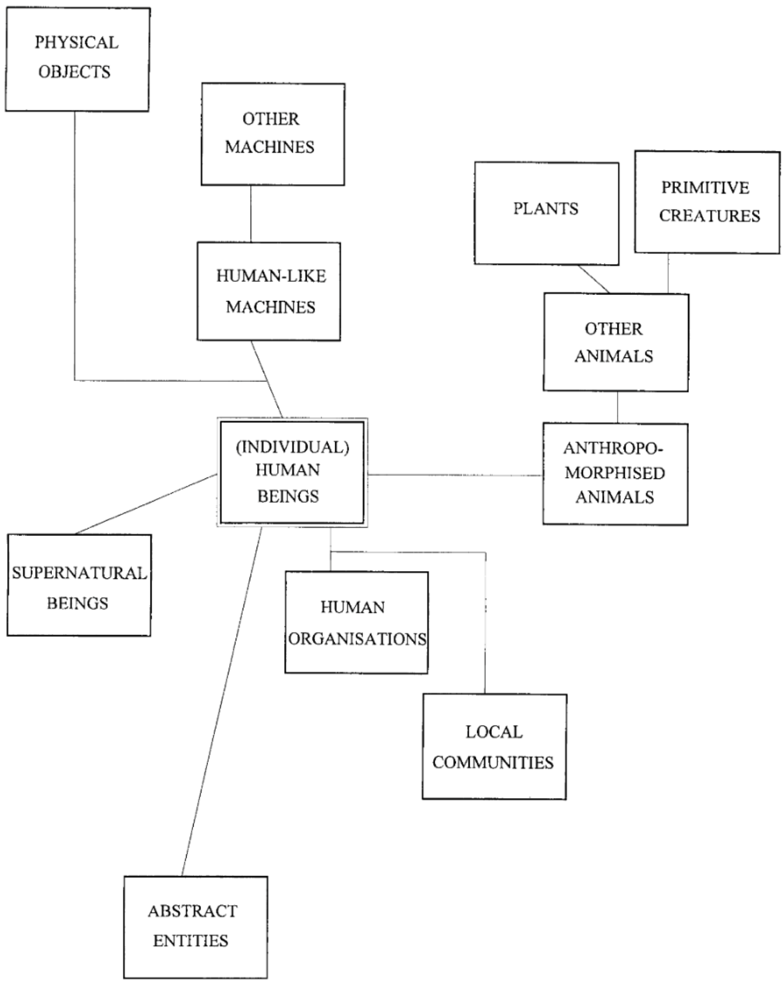

Fig. 5. The General Animacy Scale gradience (after Yamamoto 1999)

Moreover, animacy has been conceptualized as a hierarchy (e.g., Silverstein 1986; Dahl 2000; Aissen 2003). Within this approach, animacy is understood as a continuum with humans being perceived as most animate, followed by animals, plants, and inanimate objects, placing abstract concepts as least animate (Paczynski and Kuperberg 2011). Positioning humans as the prototypical animate entity is also consistent with Yamamoto's (1999) anthropocentric theory. Acknowledging the gradience of animacy has significant methodological implications, requiring researchers to accommodate the nuanced understanding of the concept in their study designs.

In attempt to operationalize animacy, VanArsdall and Blunt (2022) carried out a norming study to investigate different dimensions of animacy to provide a more detailed account. They selected 1200 relatively concrete nouns, which were rated across six scales reflecting key markers of animacy: propensity to movement, ability to reproduce, tendency to act in a goal-directed manner, ability to think, similarity to a person, and a living/non-living judgement. Such range of scales offered a broader and more nuanced understanding of the complexity and graded nature of the concept of animacy.

The multidimensional approach revealed participants' sensitivity to the intricacy of animacy. For instance, for movement, the ratings were higher not only for vehicles but also for weather phenomena in comparison to immobile entities, e.g. buildings. What is more, moving bodies of water (e.g., ocean) received higher scores than still bodies of water (e.g., puddle). Such ratings show that the participants focused more on the real-life movement rather than the ability of self-generated motion. Furthermore, to underline the participants' understanding of the complex nature of animacy, VanArsdall and Blunt (2022) illustrate this by ratings of ambiguous words (e.g., *pitcher*) or words for plants (e.g., *tulip*). For example, *tulip* was rated highly as living and able to reproduce, but also as not able to move, not being goal-oriented, not able to think and being unlike a person. Similarly, the word *robot* was rated as non-living and incapable of reproduction, somewhat capable to think and somewhat similar to a person, while receiving high ratings for movement and goal-directness. Such ratings prove that animacy is not perceived as a simple distinction of animate vs inanimate entities. Overall, the study draws attention to the graded nature of animacy, while emphasising the importance for researchers to take into consideration participants' grasp of physical, mental and categorical markers of animacy.

Further evidence that participants recognize animacy as a continuum was supported by an eye-tracking study conducted by Lowder and Gordon (2015). They tested simple sentences and sentences with relative clauses, manipulating the subjects of the sentences (instruments vs natural forces). The experiment compared subject-verb integration in sentences combining instruments (e.g., *hammer, crowbar, sword, pistol*) or natural forces (e.g., *hurricane, rain, earthquake, river*) as subjects with action verbs (e.g., *carried, wounded, damaged*). The results showed longer reading times for simple sentences than for sentences with relative clauses. Moreover, reading times were longer for sentences with instruments as subjects in comparison to natural forces. Such findings provided evidence for easier integration of natural forces with action verbs in comparison to instruments, where the results showed greater processing difficulty. Thus, natural forces are perceived as more likely to occupy the role of agent in the sentence structure, whereas instruments demand additional processing to integrate the subject's semantic features with the verb's thematic requirements. The effect was observed for simple sentences, where the action verb was the main verb, but not for relative clauses. These results resembled real-life perception of natural phenomena, often manifested through anthropomorphism, for example by describing them with animate adjectives (e.g., harsh wind, active volcano, calm sea). The results may reflect the focus on and vigilance towards agency, which is predicted to have evolved for survival (Nairne et al. 2013).

With the aim to operationalize the concept of animacy Trompenaars et al. (2021) enumerated three levels of animacy that are of interests in cognitive science and linguistics. The first level is the biological distinction between living and non-living entities based on purely biological criteria. This level has traditionally a binary differentiation of animacy (living vs non-living) and is considered universal. Thus, regardless of cultural or cognitive differences, the biological distinction between living and non-living entities remains an objective measure. Besides biological animacy, the authors list cognitive animacy and linguistic animacy, with the latter two showing greater diversity in terms of universality and granularity. In comparison to biological animacy, the other two levels are neither considered fully universal nor dichotomous. Cognitive animacy refers to the way people conceptualize and interpret an entity as animate through ascribed animate traits such as motion, intentionality or agency. Subsequently, the cognitive conceptualization can find its reflection in grammar and thus result in linguistic animacy, which is the third level mentioned by Trompenaars et al. (2021).

Importantly, the cognitive construct of animacy can vary in degree and depends on the context constraints. Hence, given the right context even inanimate objects can take on linguistic animacy, which can be reflected through various grammatical features such as word order, pronouns, or voice. To showcase how animacy can be manifested through grammar, Trompenaars et al. (2018) analysed two texts by the same author and compared linguistic expressions of two narrators: inanimate and human. The inanimate narrator in one of the analysed novels is a canvas that eventually becomes a painting. This narrator cannot move and communicates only with the readers to tell the story. To investigate how animacy is reflected in both texts, the researchers compared the use of grammatical functions, voice, thematic roles and verb types.

The results showed that, even though pronouns usually fulfilled the role of subject, in the story with the inanimate narrator the pronouns took the object role significantly more often than in the story with the human narrator. Regarding voice, the researchers predicted more passive constructions for the inanimate narrator. This hypothesis was partially confirmed as the story about the painting has indeed slightly more passive constructions. However, it was also observed that the passive voice was rarely used in both stories, hence the comparison is insufficient. As for thematic roles, the inanimate narrator was assigned the role of Agent less often and more often the role of Theme in comparison to the human narrator. Surprisingly, contrary to what Trompenaars et al. (2018) predicted, the inanimate narrator took the role of Experiencer more often than the human narrator. Additionally, the inanimate narrator acted more as an observer and not actor and used fewer action verbs in comparison to the animate narrator. In conclusion, there were no significant linguistic differences, nor clear-cut distinction between the animate and inanimate narrators. All these observations lead to the conclusion that linguistic animacy reflects conceptual animacy, rather than biological animacy, which can vary in the degree of animacy.

3.3. Animacy and cognitive functions

3.3.1. Attention

The ability to distinguish between animate and inanimate objects and agents is one of the fundamental cognitive processes available for humans. The ability to distinguish between human and non-human entities is not only crucial for survival but also helps us make sense of the surrounding world (New et al. 2007). Studies show that from early development humans can recognize animacy cues and differentiate between animate and inanimate objects. An early study by Poulin-Dubois et al. (1996) conducted on 9-month-old and 12-month-old infants tested their perception of animacy. In this study, a robot moved around a room on verbal commands. The study showed that the infants' level of attentiveness was increased when the robot moved in comparison to a human moving when given the same commands. Poulin-Dubois et al. (1996) suggest that infants can distinguish animate objects from inanimate objects based on the motion cues. Träuble et al. (2014) differentiate between static information such as facial features or appendages resembling legs and dynamic information such as movement patterns or self-propellednes that can be used to assess animacy. These perceptual cues can help distinguish between animate and non-animate entities, however the aspect of animacy is more complex.

A study by Abrams and Christ (2003) focused on the influence of animacy on attention. The researchers investigated motion in the experiment based on a visual search task in which the participants were presented with four items on the screen – one target and three distractors. Each of the presented four items underwent a different motion condition from the time it first appeared on the screen: continuous motion, static presentation, motion offset (appeared moving and then stopped), and motion onset (appeared static and began moving later in the trail). The participants were asked to identify the target among these items. The results showed that there was no advantage of motion itself in comparison to static presentation. However, the findings revealed that motion onset captured attention. The researchers argued that motion onset is a marker of animacy as it indicates that an object has its own internal energy source. They propose that the sensitivity to motion onset is rooted in biology and evolution. Rapid detection of a motion onset is crucial for identifying predators or prey and therefore is crucial for survival.

Supporting this claim, Pratt et al. (2010) conducted a follow-up study to extend the investigation of how animate motion captures attention. In their study, the researchers compared animate motion and inanimate motion, hypothesizing that animate motion should capture visual attention. In the task the participants were presented with four moving white objects on a black screen and were asked to identify targets involving objects in a reaction times experiment. The movement of the objects was manipulated to include animate motion (sudden motion that was not an effect of collision). The results showed that across six experiments, the animate motion (unpredictable movement, not through collision with another object) triggered quicker responses in comparison to inanimate motion (predictable movement obtained through collision).

Additionally, New et al. (2007) conducted a study inspecting visual attention mechanisms. The researchers put forward the animate monitoring hypothesis and posed that, as a result of evolutionary adaptation, humans present heightened attention to other humans and animals in comparison to categories of lower survival importance (e.g., plants, trees, mountains). In the task, the participants were presented with alternating pairs of images. In each trial a picture (showing complex natural scenes) was briefly presented, followed by an alternated version of the first picture. The two pictures differed by a single object and the participants were asked to identify the change as quickly and accurately as possible. The results supported the hypothesis and showed that participants were quicker to identify humans and animals in comparison to non-animate objects (e.g., plants, cars, mountains). This shows that the human attention system shows category-specific sensitivity. The mentioned studies emphasise how animacy is one of the key concepts embedded in cognition and evolutionarily important for survival.

Significantly, recently it has been observed by Loucks et al. (2023) that not all animates capture attention equally. The aim of their study was to test the animate monitoring hypothesis, which poses that all animates are expected to be more attentionally salient in comparison to inanimate objects. To test whether the hypothesis can be generalized to all animates, the researchers compared attention reactions to mammals, non-mammals (e.g., birds, reptiles, insects), and inanimate entities. The results showed that mammals were identified faster than non-mammals and inanimates. To further investigate the hypothesis, the authors dissected the non-mammal category into individual categories: insects, herpetofauna (reptiles and amphibians), and birds. The results demonstrated that generally non-mammals were less readily recognized as living entities. These findings expand on the animacy monitoring hypothesis and suggest investigating animacy as a more graded rather than binary concept in attention. Additionally, the authors emphasise the need for more research on animacy. To address this gap, the study discussed in Chapter 5 investigated selected aspects of animacy violations on language comprehension.

3.3.2. Memory

Another important aspect of animacy in cognition is its influence on memory processes. Acknowledging animacy as a core dimension that emerges early in human development, Nairne et al. (2013) explored its role in memory retention. The researchers based their hypothesis once again on the evolutionary importance of animacy, proposing that memory systems are adaptably shaped to prioritize animates in memory processing and retention. In their study, the participants were presented with equal number of animate (e.g., *duck, soldier*) and inanimate (e.g., *hat, violin*) words in three trials. After each trial the participants were presented with a short distractor task and then asked to recall as many words as they could from the trail. Importantly, the words were previously carefully matched across 10 dimensions (e.g., familiarity, frequency, imagery) to exclude influence of other viables. The results showed, as expected, a strong recall advantage for animates in comparison to inanimates.

In comparison to the study conducted by Nairne et al. (2013), which investigated recall from long-term memory, Daley et al. (2020) examined the animacy effect in working memory. Working memory developed from the earlier notion of short-term memory. However, unlike short-term memory, which primarily holds information temporarily, working memory additionally involves active manipulation of information during a cognitive task (Baddeley 2012). In their study, Daley et al. (2020) conducted two experiments. In Experiment 1 the participants performed a serial order recall task, in which they recalled the presented list of 3, 4, 5, or 6 words containing animate (e.g., *bee*, *duck*, *hamster*) or inanimate (e.g., *glove*, *pearl*, *sock*) nouns in the order of their presentation. After the experiment, most participants reported rehearsal as their strategy to memorize the words. Therefore, Experiment 2 involved articulatory suppression, in which participants were instructed to repeat an irrelevant word aloud continuously throughout the task, consequently preventing them from using the rehearsal strategy. The results from both experiments revealed an animacy effect for 4, 5, and 6 words sets but not 3 words set. Animate words were recalled more accurately as compared to inanimate words. Although the articulatory suppression significantly decreased recall accuracy, the animacy effect persisted. Such results stress the robustness and automaticity of the animacy effect. Moreover, the animacy effect in recall tasks was found in multiple studies with consistent results in different languages: English (e.g., Nairne et al. 2013; Popp and Serra 2018; DeYoung and Serra 2021) , French (e.g., Bonin et al. 2015), Chinese (e.g., Li et al. 2016) or Portuguese (e.g., Félix et al. 2024). The consistency observed in studies employing manipulations of different task conditions (e.g., arousal, contextual details, articulatory suppression) emphasizes the significant role animacy plays in working memory.

A study by Bonin et al. (2014) further explored the animacy effect expanding their experiment to include picture stimuli. Moreover, they checked both quantity and quality of recognition as well as investigated if animate items evoke a sensory and/or perceptual experience. In the study, Experiment 1 and 2 employed free recall tasks. Prior to the free recall task in Experiment 1, participants performed a categorization task, where they were presented with words and were asked to categorize each as animate or inanimate. The free recall task was unexpected for the participants to mimic a real-life scenario as usually people do not have time to find a strategy to memorize items. In comparison, in Experiment 2 participants were presented with pictures in the categorization task. The third experiment involved a recognition memory paradigm where the participants decided if they "remember" (R-response) or "know" (K-response) the previously presented stimuli. They were instructed to give the R-response if they could recollect the learning process (e.g., they remembered the word because it evoked some association or other personal experience). Whereas the K-response was given when the participants were sure about recognizing the word but were not able to recall any specific memory of a learning sequence. The results revealed that not only were animates recognized more often than inanimates, but they also received more R-responses than K-responses in comparison to inanimate words, where the difference was not observed. In Experiment 4, sensory experience ratings were used. In this task participants judged the degree to which each presented word evokes a sensory experience (e.g., taste, smell, sound). Noteworthy, in all four experiments the findings showed that both animate words and pictures were remembered more accurately than inanimate ones, regardless of the task. The researchers argued that these results indicate richer encoding of animate entities and better-preserved traces in memory. Such interpretation supports the view that human memory is evolutionarily tuned to solve survival issues (e.g., finding food or being protected from predators). Lastly, the results showed that animates do not evoke more sensory and/or perceptual responses thus ruling out the assumption that the animacy effect can be attributed to animates' richer sensory, perceptual or semantic features. Instead, the results similarly to the previous studies support the functionalist view that memory is tuned to prioritise fitness-relevant information (i.e., information crucial for survival and/or reproduction) (Nairne 2010).

The discussed research highlights the importance of animacy and how it influences human cognition. There is substantial evidence that animacy significantly affects attention, and studies show that infants as young as 9 month old present heightened attention in response to animate motion (Poulin-Dubois et al. 1996). This phenomenon has been tested in various visual tasks measuring attentional responses to animates vs inanimates (e.g., Abrams and Christ 2003; Pratt et al. 2010; New et al. 2007). Although, the results consistently show that animates capture attention more than inanimates, Loucks et al. (2023) points out that animates do not attract attention to the same degree as they found that mammals are identified faster that non-mammals or inanimates. Altogether, the findings are in line with the animate monitoring hypothesis proposed by New et al. (2007). This theory states that animate advantage in attention is the result of evolution and suggests that this advantage is rooted in survival. Heightened attention to animates was and still is crucial for survival i.e. identifying threat or food. Moreover, the aforementioned studies on the aspect of animacy in memory show that animates take priority in memory retention both in long-term (e.g., Nairne et al. 2013) and working memory (e.g., Daley et al. 2020). Taken together, these studies show that animacy is not only an important aspect of cognition but also a significant variable in psycholinguistic studies that needs to be taken into account. The substantial influence of animacy on fundamental cognitive functions emphasizes the importance of further research on this aspect. In the study reported in Chapter 5, animacy is the key variable as the study investigated novel metaphors with and without animacy violations by examining neural activity using ERPs.

3.4. Electrophysiological research on the aspect of animacy in language processing

As mentioned above, research has shown that the distinction between animate and inanimate entities emerges early on in life. Moreover, drawing on the discussed research in the previous section, animacy has direct influence on attention and memory, both of which are fundamental cognitive mechanisms involved in language processing. Therefore, this section in devoted to electrophysiological research which has investigated animacy in language processing and comprehension. Noteworthy, languages differ on grammaticalization of animacy. Moreover, languages differ in the degree to which they enforce animacy constraints in grammar. For instance, some languages such as Jakaltek and Halkomelem categorically ban inanimate subjects, whereas Swedish or English exhibit tendency to disfavour them without categorically banning them (Bayanati and Toivonen 2019). Despite cross-linguistic differences in animacy reflection in grammar, animacy remains, as discussed below, one of more influential concepts in language processing. This section concentrates on ERP research on the role of animacy in language processing, as the electrophysiological methodology was employed in the study described in Chapter 5 to investigate animacy in novel metaphor processing.

A study conducted by Kuperberg et al. (2003) investigated processing of two types of conceptual relationships between noun-phases (NPs) and verbs using electroencephalography. The stimuli included three types of unambiguous sentences, each containing a critical verb that assigned the thematic role of 'agent' to the preceding subject in the NP. In sentences without violations (e.g., *For breakfast the boys would only eat…*) the NP was a probable animate agent (*boys*) for the following verb (*eat*), considering the preceding context (*For breakfast*). The other two types of sentences included either non-thematic role pragmatic violations or thematic role animacy violations with improbable agents for the verbs based on the preceding context. In sentences created with pragmatic violations (e.g., *At breakfast the boys would only bury…*) the thematic role was filled by a likely agent (*boys*) for the matched verb (*bury*). However, it was pragmatically incoherent with the preceding context (*At breakfast*). In comparison, the sentences with animacy violations (e.g., *For breakfast the eggs would only eat…*) featured inanimate NPs (*eggs*), which are more suitable to occupy the theme role not the agent role (eggs are eaten but eggs cannot eat). In the experiment, the participants read the three types of sentences and decided if a presented sentence was plausible. For the analysis, all ERP components were time-locked to the critical verb in the sentences (e.g., *eat*). The results confirmed the predictions and showed significantly increased N400 amplitudes for the non-thematic role pragmatic violations in comparison to non-violated verbs. However, the thematic role animacy violations, even though the sentences were rated by the participants as less plausible, resulted in a nonsignificant N400 effect. Nonetheless, animacy violations elicited robust P600 responses. According to the authors the robust P600 effect elicited by animacy violations reflects the effort involved in resolving the mismatch between the inanimate agent and the thematic requirements of the critical verb.

To further investigate the results obtained from their previous study, Kuperberg et al. (2007) conducted another study extending the number of experimental conditions. They retained the non-violation condition (e.g., *For breakfast the boys would eat toast and jam*), pragmatic non-thematic role violation (e.g., *For breakfast the boys would plant flowers in the garden*), semantically related animacy thematic role violation (e.g., *For breakfast the eggs would eat toast and jam*), and added semantically unrelated animacy thematic role violation (e.g., *For breakfast the eggs would plant flowers in the garden*). The additional condition was introduced to investigate potential factors influencing the N400 and P600 waveforms, especially relatedness. Therefore, in the semantically unrelated animacy thematic role violation condition, the critical verb was not related to neither the agent nor any word in the preceding context. This study replicated the significant N400 effect for pragmatic non-thematic role violations. Additionally, both the related and unrelated animacy thematic role violations failed to elicit the N400 effect, however, they resulted in a robust P600 effect. The results showed that animacy plays a role in integration of thematic relationships between the NP agent and following verb during online processing. The absence of a significant N400 effect was attributed by the authors to the possibility that further semantic processing was abandoned once the thematic role violation was detected. The P600 effect can reflect the processing cost of the mismatch between expected and assigned thematic roles, especially in cases involving animacy violations. Thus, the late response can be a marker of animacy violations.

Based on results from yet another study, Paczynski and Kuperberg (2011) concluded that animacy expectations are not primarily driven by thematic roles. In their experiment they tested implausible sentences assigning animate nouns to the Patient role (usually associated with inanimate nouns) e.g., *At the homestead the farmer penalized the meadow for laziness*, and assigning inanimate nouns the Experiencer role (prototypically needs animate nouns) e.g., *At the homestead the farmer interested the meadow in some work*. Both violation conditions (animate noun in the Patient role, and inanimate nouns in the Experiencer role) showed an N400 and P600 effect relative to the non-violating conditions. However, the effects were not modulated by thematic roles. These results show that the influence of animacy on online processing emerges rather from direct mapping between linear word order and animacy hierarchy, where an animate noun precedes an inanimate noun (Aissen 2003).

In their study, Szewczyk and Schriefers (2011) investigated whether animacy holds a special status in language processing. They conducted an ERP study to determine if animacy is processed differently from other semantic features. The materials included short stories in which the final sentence was structured to create a strong expectation for either an animate or inanimate object noun. Thus, the participants read short stories in which the last sentence either fitted the preceding context, included semantic violation without animacy violation, or included animacy violation. No significant difference was observed between the two violations for the N400 time window. However, animacy violations elicited significantly larger P600 amplitudes in comparison to semantic violations. The researchers interpreted the larger P600 amplitude as an indication of increased processing cost, greater conflict or less well-formed representations for animacy violations.

Furthermore, to examine processing fluency of animate over inanimate entities in language, Czypionka et al. (2023) conducted an EEG study on simple German words and noun-noun compound words. The researchers investigated the N400 time window expecting to observe processing cost related to lexical access. The materials constructed for the lexical decision task included compound nouns (e.g., *Schlittenhung, "*sled dog*")*, in which animacy of both the modifier (e.g., *schlitten* 'sled') and the head (e.g., *hund* 'dog'), as well as animacy of the full form compounds were manipulated. The selected compounds were semantically transparent i.e. the overall meaning of the compound can be derived from the constituents (e.g., *Schlittenhung, "*sled dog*"*). Thus, the selected compounds fell into four categories: animate- animate (e.g., *Wolfshund* 'wolfhound') and inanimate-animate (e.g., *Schlittenhung, "*sled dog*"*) combinations that resulted in full-form animate compounds, as well as inanimate-inanimate (e.g., *Tishdecke* 'table cloth') and animate-inanimate (e.g*., Pferdedecke* 'horse rug') combinations which resulted in full form inanimate compounds. Noteworthy, full form animate compounds referred only to animals, and inanimate to concrete objects. The results revealed that full form inanimate compounds evoked the largest N400 amplitude, followed by a moderate effect for compounds, which consisted of one animate and one inanimate noun. The full form animate compounds evoked the least negative amplitude. For comparison, the study also included a lexical decision task experiment only on simple single animate (animals) and inanimate words (concrete objects). Here, the inanimate words elicited larger N400 amplitudes than animate words despite being matched on length and frequency. In comparison to simple single words, the compounds elicited smaller amplitudes. Taken together, the results showed that for animates lexical access is less costly, thus reducing processing cost compared to inanimates. Additionally, in the study the researchers manipulated constituent animacy in compounds. The experiment on compound nouns revealed that animacy facilitated lexical access in an additive manner, which they interpret as a sign that semantic properties of the constituents are accessed automatically. Moreover, the researchers highlight animacy as notable factor influencing lexical access.

The influence of animacy has been also recently observed in language production in a study by Wang et al. (2024). The authors conducted a picture-word interference ERP experiment in Mandarin Chinese. The participants were presented with a picture (target, e.g., a picture of a paper bag) with a distractor word (e.g., coaster, castle, baby, leopard) layered on the picture and were asked to name the target represented by the picture ignoring the superimposed word. The animacy congruency of target and distractor (congruent: e.g., a picture of a paper bag and a distractor word *coaster*; incongruent: e.g., a picture of a paper bag and a distractor word *baby*) was manipulated alongside the Mandarin Chinese classifiers. In Mandarin Chinese, the classifier selection is to some extent dictated by semantic features including animacy of a noun. The results showed an animacy congruency effect, where the animacy-incongruent condition evoked larger N400 amplitudes in comparison to the animacy-congruent condition. This finding suggests that animacy-congruent pairs were closely situated in the semantic network, allowing stronger spreading activation. Such activation facilitates semantic processing, as reflected by a decreased N400 amplitude. This interpretation can be supported with Levelt’s Model of Speech Production (Levelt 1989) according to which, words that belong to the same semantic category tend to be in close proximity in the semantic network. Moreover, these words share overlapping semantic features, e.g., animacy, which enhances activation between related concepts (Levelt et al. 1999; Levelt 1999). As for the late component, the results showed an increased P600 response to classifier-incongruent pairs relative to classifier-congruent pairs suggesting that classifiers are processed syntactically.

As discussed above, animacy has been perceived as one of the fundamental concepts in cognition and language. Furthermore, the findings show that animacy influences language processing regardless of the modality (e.g., reading, picture naming). However, Nieuwland and Van Berkum (2006) showed how context can modulate processing and comprehension of animacy violations. In their research, verb-object animacy violations were either presented in isolation or embedded in a supportive context. In the study participants listened to short, six-sentence-long cartoon-like stories containing animacy violations in two experiments. In Experiment 1 the main entity in the story was either a man in the animate condition or an object in the inanimate condition. In the inanimate condition all six sentences contained an animacy violation (e.g., *Once upon a time, a psychotherapist was consulted in her home office by a yacht/sailor with emotional problems. The yacht/sailor confided in her that everything in life had gone wrong and started crying. The psychotherapist consoled the yacht/sailor by stating that everybody experiences these kinds of trouble every now and then. But the yacht/sailor doubted whether to continue outlining his problems to her. The psychotherapist advised the yacht/sailor to be honest not only with her, but especially with himself. At that moment the yacht/sailor cried out that he was absolutely terrified of water*). However, the analysis was conducted only on the first, third and fifth sentence. The N400 was time-locked to the critical word (e.g., *yacht/sailor*) onset. The results revealed a strong N400 effect for the initial sentence of the stories (e.g., *Once upon a time, a psychotherapist was consulted in her home office by a yacht/sailor with emotional problems.*), reflecting surprise upon encountering animacy violation. Nevertheless, by the end of the story, the effect vanished. That shows that animacy violation has a stronger effect without any initial context. However, the effect can be neutralized with further supporting context. To exclude any potential confounding variables that could account for the disappearing N400 effect (e.g., stories being perceived as implausible and the participants becoming uninterested in them) the researchers conducted another experiment. In Experiment 2 the stories revolved around an inanimate entity (e.g., a peanut) in a cartoon like context (e.g., the peanut falling in love). By the end of the stories the inanimate object was either described based on animacy violation (e.g., *The peanut was in love*) that suited the context (story about a peanut singing a love song) or without animacy violation but inappropriate for the story context (e.g., *The* *peanut was salted*). The results showed a significant N400 effect elicited for the latter condition and lack thereof for context appropriate animacy violations. Taken together, the results showed that a supporting context can play a significant role in language comprehension, neutralizing or even overpowering animacy violations and real-world implausibility. Moreover, it supports the one-step model of language comprehension according to which global and local contexts can both immediately affect semantic processing (e.g., Berkum et al. 1999). Importantly, the study by Nieuwland and Van Berkum (2006) emphasizes the role of context in language processing. Taken together with the findings discussed in Chapter 2, it is clear that context has strong influence on language processing, especially novel metaphor comprehension. Hence it is one of the most important variables to control in experimental research. Based on the results, sentences in the present study were constructed with minimal context preceding animacy violations for clarity.

To conclude, the discussed studies showcase the effect of animacy and animacy violations in language processing. It has been observed that animacy violations and animacy incongruencies can elicit increased N400 amplitudes (e.g., Nieuwland and Van Berkum 2006; Paczynski and Kuperberg 2011; Czypionka et al. 2023; Wang et al. 2024). Decreased N400 amplitudes observed in animacy congruent stimuli (Wang et al. 2024), as well as for animate simple words and compounds (Czypionka et al. 2023) can be interpreted as less costly lexical access for animates in comparison to inanimates. Such interpretation can be supported by Levelt's Model of Speech Production (Levelt 1989), which states that words sharing semantic features (e.g., animacy) are closer in proximity in the semantic network, and the overlap of the shared features can boost activation reducing processing cost. Furthermore, the observed robust P600 responses were attributed to increased cost of processing and integration for animacy violations (Kuperberg et al. 2007) in attempt to resolve the mismatch created by animacy violations (Kuperberg et al. 2003b). Moreover, the P600 results from the study conducted by Szewczyk and Schriefers (2011) suggest that animacy violations are significantly more difficult to resolve than semantic violations, thus needing increased processing effort. The interpretation of increased N400 and P600 amplitudes observed by Paczynski and Kuperberg (2011) was based on the animacy hierarchy (Aissen 2003), which suggests linguistic preference for animates preceding inanimates in syntax. Lastly, Nieuwland and Van Berkum (2006) highlight the role of context in language processing showing decreased N400 amplitudes to animacy violations after exposure to preceding supportive context. All aforementioned studies are relevant to the ERP study discussed in Chapter 5.

3.5. The role of animacy in metaphor comprehension: evidence from ERP studies

Although the aspect of animacy has gained attention in research on language and cognition, there is still scarcity of studies on animacy in metaphor processing and comprehension. There have been many studies conducted on figurative language using the ERP method both in the monolingual and bilingual contexts. Importantly, the results have been mostly consistent. Studies show that metaphors elicit increased N400 amplitudes in comparison to literal utterances reflecting greater semantic integration (e.g., Lai et al. 2009; Jankowiak et al. 2017; Rataj et al. 2018). As for the LPC time window, the observed increased responses for metaphors in comparison to literal utterances suggest reanalysis and ongoing integration processes crucial to derive the intended meaning (e.g., Coulson and Van Petten 2002; Arzouan et al. 2007; De Grauwe et al. 2010). Furthermore, novel metaphors are more cognitively taxing than conventional metaphors without supporting context, thus resulting in increased N400 and LPC responses (e.g., Lai et al. 2009; Rataj et al. 2018). However, to the best of my knowledge, there is still a gap in research on how animacy might further modulate the effects observed in the aforementioned studies.

Ji et al. (2020) explored the role of animacy in metaphor comprehension in Mandarin Chinese in a lexical decision task. The stimuli sentences were selected from contemporary Chinese literature and, when necessary, adapted to conform to the Subject-Verb-Object structures. In the study two experimental variables were manipulated: animacy of the sentence-initial noun (animate vs. inanimate actor) and sentence type (metaphoric vs. literal sentences). Thus, creating four conditions: animate metaphors (e.g., *The teachers light up the future*), inanimate metaphors (e.g., *The sufferings light up the future*), animate literal sentences (e.g., *The man lit the firewood*), and inanimate literal (e.g., *The match lit the firewood*) (Ji et al. 2020). Moreover, the conditions differed in noun concreteness with metaphors being constructed with abstract nouns (e.g., *future*) and literal sentences with concrete nouns (e.g., *firewood*). Behaviourally (accuracy rates and reaction times) there were no differences between the four conditions. The ERP results showed that metaphors with inanimate actors elicited significantly larger N400 amplitudes at the verb position in comparison to literal and metaphorical sentences with animate actors. However, animacy did not modulate the N400 at the final processing stage measured at the noun. According to the authors, this finding shows increased difficulty in integration of inanimate nouns with verbs. Contrary to the expectations, the analysis of the LPC revealed reduced P600 responses at the final stage in inanimate metaphors relative to animate metaphors. The observed reduced LPC amplitudes differ from previously conducted studies on animacy, which showed an increased processing effort for animacy violations (e.g., Kuperberg et al. 2003, 2007; Paczynski and Kuperberg 2011; Szewczyk and Schriefers 2011). With current results, Ji et al. (2020) argue that animacy violations may facilitate the integration process in metaphor reanalysis. In addition to these findings, the researchers observed a noun concreteness effect in the early processing stage (300-420ms). This means that concreate nouns were processed more easily than abstract nouns. Although it did not influence the integration process (460-640ms), the researchers emphasize that this effect should be controlled in further research.

A study by Li et al. (2022) investigated the processing of Chinese verbal metaphors. The study included four types of stimuli: Subject-verb metaphor (SVM) (inanimate subject followed by action verb, thus creating animacy violation e.g., *The company grasped the opportunity*), Verb-object metaphor (VOM) (animate subject followed by action verb and an abstract object e.g., *The boss grasped the opportunity*), Literal-abstract (LA) (inanimate subject with abstract verb with similar meaning to the action verbs, creating animacy violation e.g., *The company sought the opportunity*), and Literal-concrete (LC) (animate agent and concrete object e.g., *Xiaozhe grasped the rope*).The study employed a reading task in which participants read sentences and then answered semantic decision comprehension questions. The results of ERPs time-locked to the verb onset showed increased N400 amplitudes for the verb-object metaphors and literal concrete sentences compared to subject-verb metaphors and literal-abstract sentences. However, in the LPC time window, the results showed more positive amplitudes for subject-verb metaphors and literal-abstract sentences, both of which included animacy violations (inanimate subjects paired with animate verbs), relative to the other two conditions. The increased LPC amplitudes for animacy violations were interpreted as reflecting reanalysis of the meaning, tuning for abstract interpretation even before the object appeared. The results of ERPs time-locked to the object onset revealed significantly more negative N400 amplitudes for literal-concrete condition than the other three conditions. Moreover, the LPC amplitudes were increased for subject-verb metaphors and the amplitudes for verb-object metaphors as compared to both literal conditions. The LPC results indicate that upon arriving on the object, a semantic conflict arose between the verb and the object. This conflict triggered sentence-level reanalysis to derive meaning, which caused increased cognitive demands.

Recently, another study investigated animacy in metaphor processing in the context of bilingualism. Li et al. (2025) conducted an ERP study exploring the role of noun animacy in processing action metaphors in Chinese (L1) and English (L2). The materials were constructed to follow the SVO structure, and the object nouns varied in animacy levels. Thus, there were two conditions of action metaphors employed in the semantic judgement task: high animacy condition (e.g., *Music awakens the teenager*) and low animacy condition (e.g., *Music awakens the morning*) in both languages. The manipulation created animacy violations as typically in the SVO sentence structure animate nouns occupy the role of the subject acting as agents, whereas the object position is often occupied by inanimate nouns. In the semantic decision task, the participants were presented with sentences and decided if a given sentence is metaphorical or literal. The behavioural results showed longer reaction times for high-animacy condition than for low-animacy condition. Moreover, the reaction times were shorter for L1 than for L2. The longer RTs for high-animacy condition were interpreted by the researchers as an indicator of more challenging semantic comprehension, resulting in decreased processing efficiency. The neurophysiological results showed larger N400 amplitudes for L1 than L2. The animacy effect was observed only in L1 with high animacy objects evoking increased N400 responses relative to low animacy objects. The authors interpreted this result as a possible reflection of increased cognitive demands in semantic processing for animacy violations (SVO sentences with high animate objects) in L1. Lack of such results in L2 can be an indication of insufficient syntactic proficiency in the second language and thus reduced sensitivity to animacy cues. Additionally, the ERP analysis showed that the N170 amplitudes were smaller for L2 high animacy objects in comparison to low animacy objects, possibly indicating greater visual familiarity for low animacy nouns. However, no such difference was observed for L1, suggesting high proficiency of L1 and equal familiarity of low and high animacy nouns. The analysis of P200 showed a language effect with the L1 objects evoking larger amplitudes than L2 objects. Such observation suggests that L1 information is likely more salient and therefore draws attention more in the early stages of language processing. Moreover, no animacy effect was observed for P200. Two potential limitations were noted by the researchers. One limitation is the lack of control of the types of predicate verbs. Additionally, given the diversity of metaphors, it remains unclear whether the findings obtained for action metaphors in this study can be generalized to other types of metaphors.

3.6. Conclusions

In conclusion, there is still a gap in research on the aspect of animacy violation in figurative language comprehension. Previous research has shown that animacy violations elicit distinct ERP responses in comparison to semantically congruent utterances. It has been observed that animacy is a fundamental concept in our cognition affecting attention (e.g., Abrams and Christ 2003; New et al. 2007; Pratt et al. 2010) and memory retention (e.g., Nairne et al. 2013; Daley et al. 2020). The findings, which show that animates capture attention significantly more and are memorized better and more accurately than inanimates result from human evolutionary priority of survival (e.g., fast and accurate detection of food or threat). Moreover, the role of animacy is reflected in numerous ERP studies (e.g., Nieuwland and Van Berkum 2006; Paczynski and Kuperberg 2011; Czypionka et al. 2023; Wang et al. 2024). These studies have shown that animacy violations can evoke more negative N400 amplitudes, and also tend to elicit robust LPC responses marking ongoing effort in resolving the semantic mismatch. However, a vast majority of the effects reported in this chapter have been observed in the context of literal language. There is still a significant gap in literature on the role of animacy and animacy violations in novel metaphor processing. Additionally, most previous studies on animacy and metaphors have been conducted in Mandarin Chinese, a language that notably differs from Polish in grammar, morphology, and word order. Hence, the discussed results may not be directly applicable to Polish. Therefore, the study reported in the following chapters makes a significant contribution as it aims to fill the gap in existing research.

Chapter 4: Evaluation of novel metaphors with and without animacy violation: normative studies

4.1. Introduction

As discussed in Chapter 1, many factors influence processing and comprehension of metaphoric language. Studies have shown that factors such as lexical frequency and familiarity contribute to the degree of conventionality of metaphors, which is known to significantly influence metaphor processing (e.g., Pynte et al. 1996; Arzouan et al. 2007; Goldstein et al. 2012; Rutter et al. 2012; Rataj et al. 2018; Jankowiak et al. 2021). Similarly, context substantially impacts metaphorical meaning construction and needs to be taken into consideration in research on metaphors (e.g., Pynte et al. 1996; Yang et al. 2013; Bambini et al. 2016; Tang et al. 2025). Moreover, it has been observed that factors like word length, semantic richness, grammatical number and case additionally impact processing of linguistic stimuli (Balota et al. 2006b, 2004). Taken together with the sensitivity of electrophysiological responses to language variables discussed in Chapter 2, it is essential to carefully control the experimental stimuli. Therefore, several rating scale studies were conducted in the process of material selection for the ERP study reported in Chapter 5. These normative studies included ratings of meaningfulness, metaphoricity, familiarity and cloze probability of created and selected novel metaphors with animacy violations, novel metaphors without animacy violations, literal sentences and anomalous sentences.

4.2. Aims and hypotheses

The aim of the normative studies is to ensure that selected experimental stimuli are well controlled and accurately represent experimental conditions, which is important for valid interpretation of ERP results. Based on existing research and literature on metaphor processing and comprehension, the selected stimuli were tested on proposed above dimensions and the following hypotheses were formulated.

Hypothesis 1. ***A graded effect is expected with literal sentences being judged as most meaningful, followed by both types of novel metaphors, and anomalous sentences as least meaningful.*** A similar pattern has been observed in studies on metaphor processing and comprehension (e.g., Rataj et al. 2018; Jankowiak et al. 2017). Novel metaphors were expected to be perceived as less meaningful than literal sentences and more meaningful than anomalous sentences because this type of stimuli can be ambiguous and more taxing in judgements. This can be supported by the Career of Metaphor model (Bowdle and Gentner 2005), which proposes that novel metaphors are comprehended through comparisons between the target and source domain. Therefore, the judgements were expected to be less consistent relative to literal and anomalous sentences.

Hypothesis 2. ***Novel metaphors are expected to be rated as more metaphorical than literal sentences.*** This hypothesis is also based on existing findings in studies on metaphor comprehension (e.g., Jankowiak et al. 2017; Rataj et al. 2018; Ji et al. 2020). The stimuli chosen for the present study were created to convey novel and creative meanings, which should be pronounced in the metaphoricity ratings in comparison to literal sentences, which were expected to be perceived as not metaphorical.

Hypothesis 3. ***Novel metaphors are expected to be evaluated as unfamiliar, whereas literal sentences are predicted to be judged as familiar but only at a moderate level.*** The studies on novel metaphors have shown a similar pattern of literal sentences being rated as more familiar than metaphors, although still falling around the middle point of the rating scale (e.g., Lai et al. 2009; Jankowiak et al. 2017; Rataj et al. 2018; Ji et al. 2020). All stimuli employed in the present study were created specifically for the purpose of the ERP study. Novel metaphors were expected to be perceived as unfamiliar due to their novelty and creativity. In comparison, literal sentences are more likely to be encountered in real life, thus can be perceived as more familiar.

Hypothesis 4. ***Low cloze probability scores are expected across all four stimuli conditions.*** Cloze probability tests are conducted to assure that the critical words will not be anticipated from the preceding sentence context in the ERP experiment. This is important as the N400 component is sensitive to probability and prediction (e.g., Kutas and Hillyard 1984). This proposed hypothesis is in line with existing research, which showed low cloze probability scores across conditions, with significantly higher cloze probability for literal sentences in comparison to other conditions (e.g., Bambini et al. 2016; Rataj et al. 2018; Ji et al. 2020).

4.3. Participants

The experimental stimuli were rated by native speakers of Polish. The participants were recruited online at the Faculty of English at Adam Mickiewicz University. Raters whose mean rating in any experimental condition exceeded ≥ 3 *SD* were excluded from further analysis. Raters who participated in the normative studies were excluded from taking part in the main ERP experiment. The number of participants and their demographic information is provided in Table 1.

Table 1. Demographic information for participants in the four normative studies, including number of participants, their gender and mean age.

Survey type	Number of raters included in the analyses	Mean age
Meaningfulness	N = 112 (81women)	M_{age} = 21.12, SD = 2.29
Metaphoricity	N = 112 (91 women)	M_{age} = 20.43, SD = 2.21
Familiarity	N = 112 (88 women)	M_{age} = 20.05, SD = 1.75
Cloze probability	N = 112 (79 women)	M_{age} = 21.28, SD = 2.45

4.4. Material selection

All the materials were carefully created for the purpose of the ERP study to accurately represent the four experimental conditions: novel metaphors with animacy violation (e.g., *To jest kolejna zmęczona książka 'This is another tired book'*), novel metaphors without animacy violation (e.g., *To jest kolejna zatruta książka 'This is another contaminated book'*), literal sentences (e.g., *To jest kolejna podarta książka 'This is another torn book'*), and anomalous sentences (e.g., *To jest kolejna pszenna książka 'This is another wheat book'*). To minimize the impact of context (e.g., Pynte et al. 1996; Yang et al. 2013; Bambini et al. 2016), sentences in all conditions differed only by one word, i.e., the adjective preceding the target word. Targets and adjectives for the project were selected from the Polish corpus SUBTLEX-PL (Mandera et al. 2015). The critical words were controlled for length (between 3 to 13 letters), frequency of use (3-5 Zipf scale), number, grammatical case, concreteness and liveliness. The critical words chosen for the study were singular, concrete, inanimate nouns in Nominative case. All compound nouns, abstract nouns, proper names, Polish-English cognate words and Polish-English interlingual homographs were removed from the list. Polysemous words were also excluded from the pool of target words (e.g., *pociąg* as a vehicle (*train*) vs attraction). All these criteria are of particular relevance for studies employing the electroencephalography technique and are known to modulate the results. Research shows that word length influences lexical access, with longer words (8-13 letters) causing inhibition, whereas shorter words (3-5 letters) resulting in facilitation in language processing (Balota et al. 2006b). Similarly, high-frequency words result in faster recognition to low-frequency words, which are processed slower (Balota et al. 2002). This aspect is relevant for the N400, which has shown reduced amplitudes for high-frequency items and increased for low-frequency items (e.g., Kutas and Hillyard 1984) . Additionally, the critical words were controlled for concreteness as concrete words (e.g., *car*) elicit faster processing due to imagery-based facilitation, whereas abstract words (e.g., *truth*) are hard to imagine resulting in slower processing (Balota et al. 2004). This pattern is also reflected in reduced N400 responses for concrete and increased for abstract words (e.g., Kounios and Holcomb 1994). Furthermore, the selected critical words needed to be inanimate. As expected, this step raised many questions about the level of animacy as discussed in Chapter 3. All sentences were five words long, with the final two positions occupied by an adjective followed by a critical word, which together determined the four experimental conditions. The initial set of stimuli included 116 novel metaphors with animacy violations, 116 novel metaphors without animacy violations, 116 literal sentences and 116 anomalous sentences.

4.5. Procedure

This study was approved beforehand by the Ethics Committee for Research Involving Human Participants at Adam Mickiewicz University (nr akt KE/1/2022). To ensure the validity of the created stimuli sentences, several questionnaires were conducted online using the online survey tool Qualtrics (Qualtrics 2005). The sentences were evaluated in four normative studies: a cloze probability test, and ratings of: meaningfulness, familiarity, and metaphoricity. To avoid the fatigue bias, the stimuli set was divided into 8 blocks for each measure in the normative study. No critical word was repeated within a block, and each appeared in only one condition per block. Sentences were randomized within blocks, and filler sentences were added to balance the conditions. Block assignment was randomized across participants. Each block was filled out by 14 participants. The participants received links to the questionnaires and filled out the questionnaire in one sitting. The participants were compensated for filling out the questionnaires by course credits or by receiving a gift card. The instructions used for all aforementioned questionnaires are provided in Appendix B. Each sentence was evaluated for meaningfulness, familiarity and metaphoricity using a 7-point Likert scale. Novel metaphors with animacy violations, novel metaphors without animacy violations, and literal sentences were tested on all three dimensions, whereas anomalous sentences were assessed only for meaningfulness.

4.6. Results

In the normative pretests, ratings of sentence meaningfulness, familiarity, and metaphoricity were analysed using nonparametric tests. Differences across sentence types were examined using the Friedman test. Pairwise comparisons were conducted using pairwise Wilcoxon signed-rank test with Bonferroni adjustment for multiple comparisons.

Participants evaluated meaningfulness of the created sentenced on a scale from 1 (very meaningless) to 7 (very meaningful). The Friedman test showed a significant effect of sentence type, $\chi^2(3) = 236.46$, $p < .001$. The pairwise comparisons with the Bonferroni correction revealed a graded effect with literal sentences ($M = 6.55$, $SE = 0.05$) being judged as most meaningful, followed by novel metaphors with animacy violations ($M = 3.91$, $SE = 0.08$), $p < .001$, and novel metaphors without animacy violations ($M = 3.73$, $SE = 0.08$), $p < .001$, with anomalous sentences being rated as least meaningful ($M = 2.74$, $SE = 0.06$), $p < .001$. Both types of novel metaphors were rated as significantly more meaningful than anomalous sentences, $p < .001$ (Fig.6). The comparison between novel metaphors with animacy violations and without animacy violations showed no significant difference in meaningfulness ratings, $p = .77$ (Fig. 6).

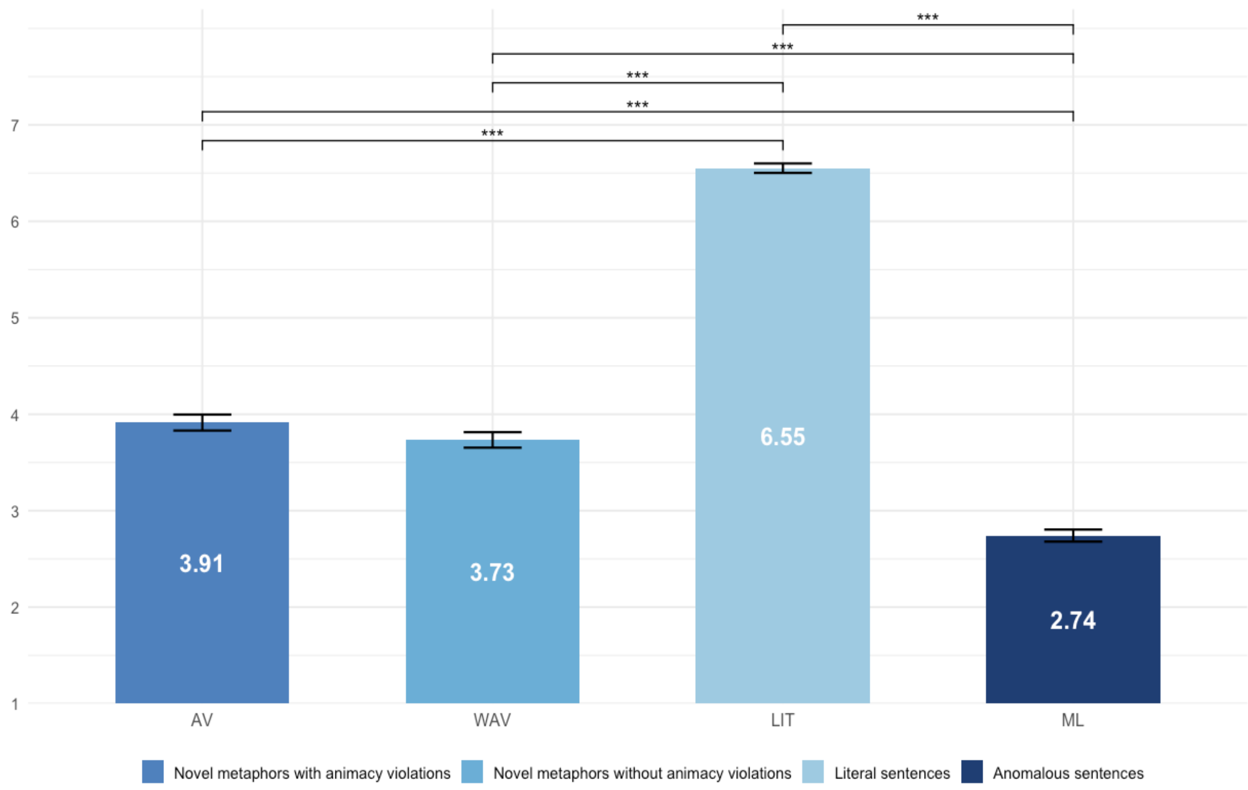

Fig. 6. Meaningfulness ratings for novel metaphors with animacy violations (AV), novel metaphors without animacy violations (WAV), literal sentences (LIT), and anomalous sentences (ML) on a scale from 1 (very meaningless) to 7 (very meaningful)

To assess metaphoricity of the constructed stimuli, participants rated how metaphorical given sentences were on a scale from 1 (very literal) to 7 (very metaphorical). The analysis revealed a significant effect of sentence type, $\chi^2(2) = 163.81$, $p < .001$. Pairwise comparisons with the Bonferroni correction showed that novel metaphors with animacy violations ($M = 5.72$, $SE = 0.05$) were rated as more metaphorical than novel metaphors without animacy violations ($M = 5.34$, $SE = 0.06$), $p < .001$. Additionally, both types of novel metaphors were judged as more metaphorical than literal sentences ($M = 1.73$, $SE = 0.05$), $ps < .001$ (Fig. 7).

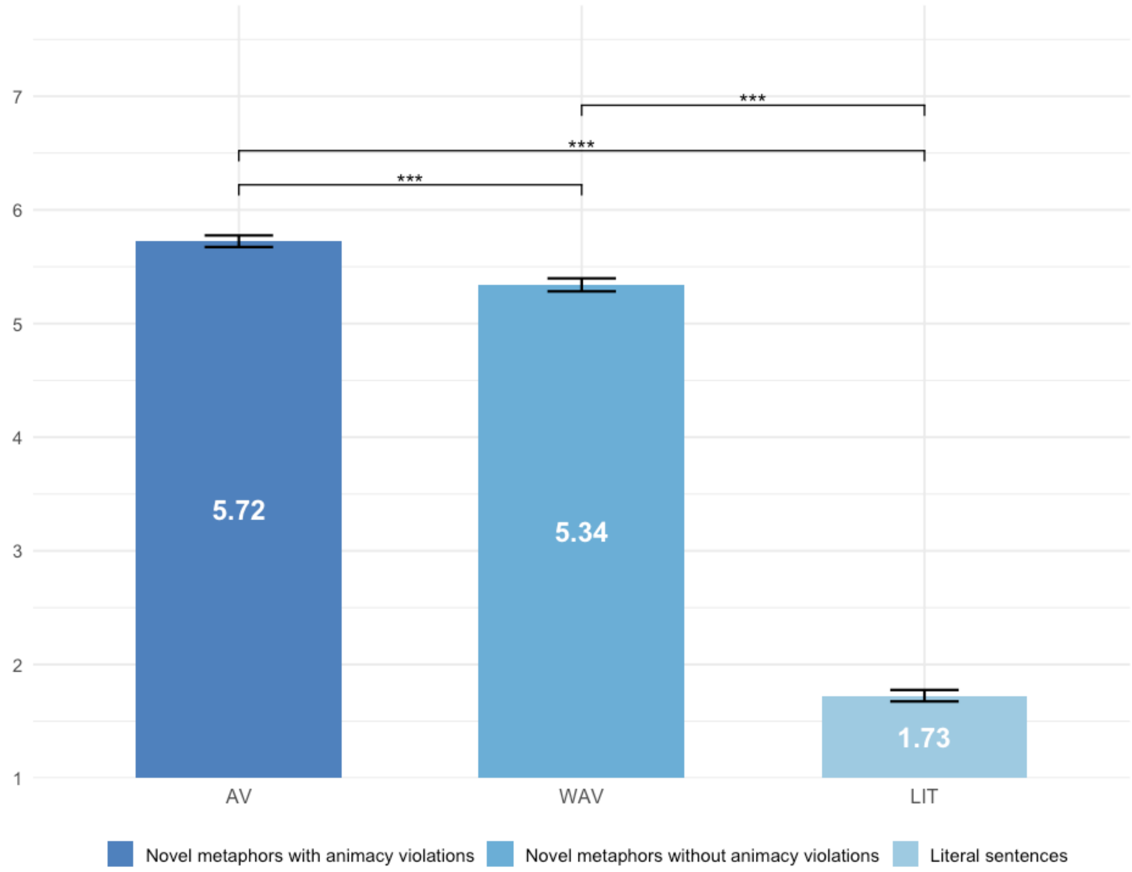

Fig. 7. Metaphoricity ratings for novel metaphors with animacy violations (AV), novel metaphors without animacy violations (WAV), and literal sentences (LIT) on a scale from 1 (very literal) to 7 (very metaphorical)

Participants assessed familiarity of the sentences by deciding how often they had encountered the presented sentences on a scale from 1 (very rarely) to 7 (very frequently). The analysis showed a significant effect of sentence type, $\chi^2(2) = 159.6$, $p < .001$. Pairwise comparison with the Bonferroni correction revealed that literal sentences ($M = 5.32$, $SE = 0.08$) were regarded as significantly more familiar than both novel metaphors with animacy violations ($M = 1.62$, $SE = 0.04$), $p < .001$, and novel metaphors without animacy violations ($M = 1.63$, $SE = 0.04$), $p < .001$ (Fig. 8). No difference was observed between novel metaphors with and without animacy violations.

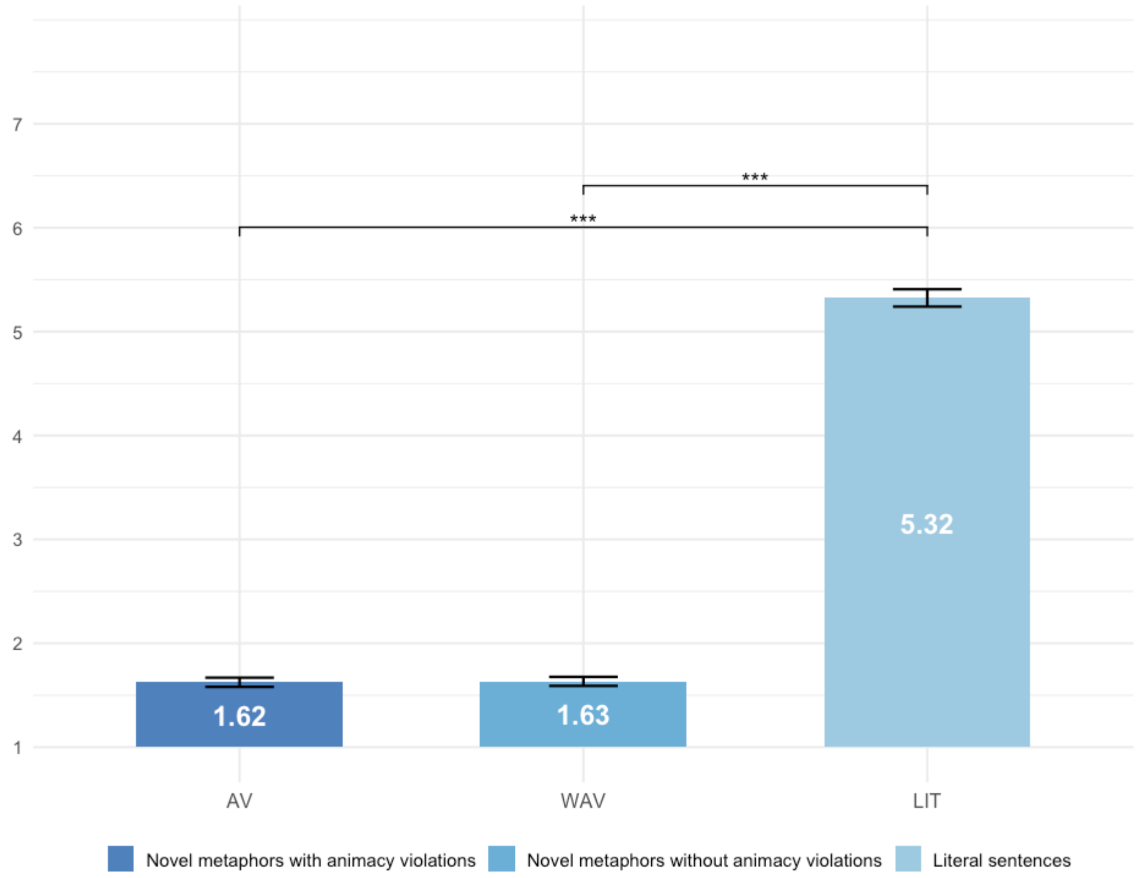

Fig. 8. Familiarity ratings for novel metaphors with animacy violations (AV), novel metaphors with-out animacy violations (WAV), and literal sentences (LIT) on a scale from 1 (very rarely) to 7 (very frequently)

To ensure that the critical words were not predictable based on the preceding sentence context, a cloze probability task was employed. In the task, the participants were presented with the sentences without target words (e.g., *To jest kolejna zmęczona* ____, *'This is another tired* _____') and filled in the critical word with a noun that first came to their minds and fitted the context creating a meaningful sentence. The results confirmed that target words were generally unpredictable across all four conditions. Novel metaphors with animacy violations ranged between 0%-7% (*M* = .21%, *SD* = 0.1) and novel metaphors without animacy violations showed a mean cloze probability of 0% (*M*= 0.0, *SD* = 0.0). Moreover, anomalous sentences ranged between 0%-14% (*M* = .20%, *SD* = 1.53). The cloze probability scores for literal sentences were slightly higher, ranging between 0%-36% (*M* = 2.6%, *SD* = 5.7). The analysis showed the main effect of sentence type, $F(3, 309) = 23.32$, $p < .001$, $\eta^2_p = .19$. Post hoc comparisons showed that literal sentences had significantly higher cloze probability than novel metaphors with animacy violations, novel metaphors without animacy violations, and anomalous sentences, *ps* <.001. No significant differences were observed between novel metaphors with animacy violations, novel metaphors without animacy violations, and anomalous sentences, *ps* > .49

4.7. Discussion

The results of the normative study were similar to existing research on novel metaphor processing and comprehension. Similar to other studies on metaphors (e.g., Rataj et al. 2018; Jankowiak et al. 2017), ratings of meaningfulness showed a graded effect with literal sentences being perceived as most meaningful, followed by novel metaphors, and anomalous sentences judged as least meaningful. The results support Hypothesis 1. Importantly for the design of the ERP study, there were no significant differences between both types of novel metaphors. Moreover, as expected in Hypothesis 2, novel metaphors were evaluated as very metaphorical and literal sentences as literal (e.g., Jankowiak et al. 2017; Rataj et al. 2018; Ji et al. 2020). In the study conducted by Ji et al. (2020), no differences were observed between metaphors with and without animacy violations. Interestingly, results in the present study revealed a significant difference in metaphoricity between the two types of metaphors, with novel metaphors with animacy violations being judged as more metaphorical than novel metaphors without animacy violations. Furthermore, as proposed in Hypothesis 3, the familiarity ratings showed the pattern previously observed in other studies on metaphoric language (e.g., Lai et al. 2009; Jankowiak et al. 2017; Rataj et al. 2018; Ji et al. 2020). The results showed a significant difference between novel metaphors and literal sentences. Comparably to the pervious findings, literal sentences received ratings near the midpoint on the 7-point Likert scale, which means that they were perceived as neither very familiar nor very unfamiliar. Lastly, as predicted in Hypothesis 4, the cloze probability test showed low scores across all four conditions with literal sentences showing slightly higher scores than novel metaphors and anomalous sentences. Similar results were observed in other studies (e.g., Bambini et al. 2016; Rataj et al. 2018; Ji et al. 2020). Taken together, the findings of the normative studies validated the stimuli, demonstrating that the selected sentences accurately represented the intended categories.

4.8. Conclusions

The normative studies described in this chapter established ratings of meaningfulness, metaphoricity, and familiarity for novel metaphors with animacy violations, novel metaphos without animacy violations, literal and anomalous sentences. Moreover, the cloze probability test was conducted to establish predictability of the selected stimuli. As discussed in Section 4.7, the obtained results were mostly in line with standard applied in most prominent studies on novel metaphor processing and comprehension. Taken together, the findings of the normative studies validated the stimuli, demonstrating that the selected sentences accurately represented the intended categories.

Chapter 5: The influence of animacy violation on novel metaphor comprehension: an ERP study

5.1. Introduction

In recent years, interest in figurative language processing, especially in novel metaphor comprehension (e.g., *This is another contaminated book*), has increased. Importantly, the existing body of research shows consistent results on metaphoric language comprehension. At the same time, the aspect of animacy is acknowledged as one of the fundamental dimensions, which influences our cognition, but not widely studied in the context of metaphor comprehension. The main aim of this study was therefore to examine the role of animacy violation (i.e., a combination of an inanimate noun with an adjective suggesting animacy e.g., *a happy carpet*) in novel metaphor processing in native speakers of Polish, using the event-related potentials technique. Additionally, behavioural data was collected to trace the reaction times and accuracy in the semantic judgment task.

5.2. Aims and hypotheses

Metaphors play a significant role in communication. Through years, research has explored metaphoric language comprehension using various methodologies, and the obtained results have shown that novel metaphors require increased processing demands in comparison to literal sentences (e.g., Jankowiak et al. 2017; Rataj 2014; Rataj et al. 2018). Additionally, animacy has been recognized as the key concept in human cognition (e.g., Nairne et al. 2013; Bonin et al. 2014; Daley et al. 2020). Studies have shown that emerging from the evolutionary need for survival, humans are tuned to animate cues, which is evident in domains such as attention (e.g., Abrams and Christ 2003; Pratt et al. 2010; Wang et al. 2024) and language processing (e.g., Kuperberg et al. 2007; Paczynski and Kuperberg 2011; Czypionka et al. 2023). However, there is a gap in research on the role of animacy and animacy violations in metaphoric language processing, with only a few studies published to date (Ji et al. 2020; Li et al. 2022, 2025). Hence, building on the existing research on metaphoric language comprehension and the aspect of animacy, the present study aimed to investigate the role of animacy violations in novel metaphor comprehension. This ERP study compared brain responses to novel metaphors with animacy violation (e.g., *To jest kolejna zmęczona książka 'This is another tired book'*), novel metaphors without animacy violation (e.g., *To jest kolejna zatruta książka 'This is another contaminated book'*), literal sentences (e.g., *To jest kolejna podarta książka 'This is another torn book'*), and anomalous sentences (e.g., *To jest kolejna pszenna książka 'This is another wheat book'*).

Hypothesis 1. ***Novel metaphors are expected to evoke increased N400 amplitudes relative to literal sentences with a centro-parietal distribution. Moreover, a graded effect is expected with most negative amplitudes elicited by anomalous sentences, moderate by novel metaphors with and without animacy violations, and least negative by literal sentences***. A graded N400 effect has been frequently observed in studies on metaphor comprehension (e.g., Arzouan et al. 2007; Coulson and Van Petten 2002; Lai et al. 2009; Rataj et al. 2018). Based on the existing research, it was assumed that novel metaphors would elicit increased N400 amplitudes in comparison to literal sentences, as the processing and comprehension of novel metaphors requires meaning construction (e.g., Coulson and Van Petten 2002; Lai et al. 2009; Rataj et al. 2018). This interpretation can be explained within The Career of Metaphor model, which states that novel metaphors are processed through comparison, establishing correspondences between the source and target domains (Bowdle and Gentner 2005). Moreover, the centro-parietal distribution has been frequently observed in studies on metaphor processing and comprehension (e.g., Coulson and Van Petten 2002; Arzouan et al. 2007; Rataj et al. 2018; Jankowiak et al. 2021). Based on these results a similar scalp distribution of the N400 effect was expected.

Hypothesis 2. ***Novel metaphors are predicted to elicit increased sustained negativity (600-800 ms) amplitudes in comparison to literal sentences.*** In comparison to the findings for metaphor comprehension in the N400 time window, the results for the late time window are less consistent. Some researchers observed the LPC effect with metaphors eliciting more pronounced positivity in comparison to literal sentences (Coulson and Van Petten 2002; De Grauwe et al. 2010; Bambini et al. 2016), other studies have observed sustained negativity in the later time window with novel metaphors evoking more negative amplitudes as compared to literal sentences (e.g., Arzouan et al. 2007; Lai et al. 2009; Rutter et al. 2012; Jankowiak et al. 2021). Moreover, the reduction of the LPC amplitudes has been interpreted as an overlap between sustained negativity and the LPC, with the sustained negativity being attributed to continued mapping processes between two conceptual domains (e.g., Arzouan et al. 2007; Rataj et al. 2018). The effect has been attributed to continued semantic integration processes needed to arrive at the novel meaning (Arzouan et al. 2007) or ongoing analysis of the sentence meaning attempting to resolve a conflict (e.g., implausible sentences) and derive intended meaning (De Grauwe et al. 2010). Furthermore, the sustained negativity mimicking the N400 pattern can indicate increased memory load due to recollection of the full sentence needed for meaning integration, which can additionally be reinforced by a delayed response procedure, or be related to task complexity (Rataj et al. 2018b). Taken together, the findings are in line with the Career of Metaphor model (Bowdle and Gentner 2005), which states that novel metaphors are processed via analogy, thus sustained negativity can reflect the demanding mapping process of two distant concepts. Considering these findings, sustained negativity was expected in the present study.

Hypothesis 3. ***Animacy violation is predicted to modulate the N400 responses to novel metaphors.*** Due to scarcity of research on the influence of animacy violations on novel metaphor comprehension (Ji et al. 2020; Li et al. 2022, 2025), this hypothesis was largely based on electrophysiological research on animacy violations. In this research, several studies have shown a more pronounced N400 amplitudes for stimuli containing animacy violations, interpreted as a reflection of increased lexical processing demands (Paczynski and Kuperberg 2011; Czypionka et al. 2023; Wang et al. 2024). As for metaphorical language, both Ji et al. (2020) and Li et al. (2025) found increased N400 amplitudes for metaphors with animacy violations. In the study by Ji et al. (2020), metaphors with animacy violations were compared to metaphors without animacy violations, and literal sentences. Li et al. (2025) compared low-animacy and high-animacy metaphors and observed larger N400 amplitudes for metaphoric sentences with animacy violations in high-animacy condition (e.g., *Debt crushes the worker*) than low-animacy condition (e.g., *Debt crushes the company*).

Hypothesis 4. ***Animacy violation is expected to influence the sustained negativity.*** In studies conducted by Szewczyk and Schriefers (2011) and Kuperberg et al. (2003) sentences with animacy violation elicited robust LPC amplitudes in comparison to sentences without animacy violation. In contrast, more recently, in the study conducted by Ji et al. (2020) the results showed reduced LPC amplitudes in response to metaphors with animacy violation in comparison to metaphors without animacy violation. However, the limitation of this study was the use of abstract nouns as target words in metaphors, which are processed differently than concrete words, which were used as targets in literal sentences. In their study, the researchers noticed the concreteness effect, which suggests that concrete words were processed faster and easier than abstract nouns (Ji et al. 2020) (see detailed description in Chapter 3). In comparison, Li et al. (2022) observed more positive LPC responses for metaphors containing animacy violations. Noteworthy, both studies were conducted on Mandarin Chinese, which has distinct grammar, morphology, and word order. The results of the present ERP study will, therefore, contribute to the existing research on figurative language processing and lay groundwork for future research on the role of animacy violation in novel metaphor processing.

Hypothesis 5. ***Novel metaphors will elicit slower reaction times and lower accuracy rates in comparison to literal sentences. Animacy violation is predicted to elicit slower reaction times and reduced accuracy***. Novel metaphors have been observed to elicit increased reaction times and lower accuracy scores as compared to literal utterances (e.g., Goldstein et al. 2012; Tang et al. 2017; Jankowiak et al. 2021; Rataj et al. 2018). This effect can be the result of novel metaphors being more cognitively taxing than literal sentences. This can be especially visible in a binary semantic decision task, where literal sentences are easily qualified as meaningful, whereas the decision may not be as easy for novel metaphors (Rataj et al. 2018b). Moreover, in line with the Career of Metaphor (Bowdle and Gentner 2005), meaning construction for novel metaphors is more time consuming due to the process of establishing links between two distant domains (Jankowiak et al. 2021). Thus, lower accuracy rates and longer RTs were expected for novel metaphors in comparison to literal sentences. Moreover, animacy violation has been observed to be more cognitively taxing (e.g., Kuperberg et al. 2003b; Szewczyk and Schriefers 2011), for which reason it was predicted that novel metaphors with animacy violations would result in slower reaction times and lower accuracy in comparison to novel metaphors without animacy violations. It also should be acknowledged that due to the delayed response paradigm used in the semantic decision task in the present study, the assumptions regarding behavioural data should be regarded as tentative, given the limited number of studies employing this design.

5.3. The event-related potential study

5.3.1. Participants

The initial sample included 34 participants. However, two participants were excluded from the data set because of low quality of the recorded EEG signal. Additionally, one participant was excluded from the analyses due to low accuracy score (below 50%) on anomalous trials. The final sample included 31 native speakers of Polish (25 women, M_{age} = 21.9 *SD* = 1.42) recruited from the Faculty of English and the Faculty of Polish and Classical Philology at Adam Mickiewicz University in Poznań. The participants received course credits or gift cards. Moreover, all participants were right-handed ($M_{laterality\ index}$ = 85.8, *SD* = 14.6). The score was measured using an online Handedness Questionnaire (Cohen 2008, adapted from Oldfield 1971). All participants reported normal or corrected-to-normal vision and no history of language and neurological disorders.

5.3.2. Material selection

Based on the normative studies described in Chapter 4, 12 critical words were excluded from the final set. Since each critical word was used across all four conditions, their rejection applied to all four conditions. Thus, the final set of stimuli selected for the ERP experiment included 104 novel metaphors with animacy violation, 104 novel metaphors without animacy violations, 104 literal sentences, and 104 anomalous sentences. The targets in the final set were similar in frequency of use (3-5 Zipf scale; *M* = 3.8, *SD* = 0.44) and length (between 5 to 13 letters; *M* = 6.58, *SD* = 1.36). The final experimental materials used in this study are provided in Appendix A. The selected stimuli for the EEG study were counterbalanced across four lists to avoid repetition of critical words within a set. One participant saw only one list. Each list consisted of 26 novel metaphors with animacy violation, 26 novel metaphors without animacy violation, 26 literal sentences, and 26 anomalous sentences. Additionally, to minimalize response bias, each list included 52 anomalous filler sentences adapted from Rataj et al. (2018) to balance the set, ensuring an equal number of meaningful and meaningless sentences. The filler items were not used in the analyses. Prior to the main experiment, participants performed a practice block including 18 sentences that were not included in the main experiment. The order of presentation of the sentences within each list was randomized for each participant. Each participant completed one list including in total 156 sentences.

5.3.3. Procedure

The EEG study was conducted applying guidelines suggested for collecting data with human participants and was approved beforehand by the Ethics Committee for Research Involving Human Participants at Adam Mickiewicz University. The participants were provided with the information about the study and procedures alongside with an informed consent. Moreover, the participants in the EEG study filled out a self-report questionnaire (e.g., about neurological or language difficulties). Additionally, the participants were asked to fill the Edinburgh Handedness Questionnaire (EHQ; Oldfield 1971) to ensure that all participants were right-handed.

The EEG experiment was conducted at the Neuroscience of Language Laboratory (NeuroLang Lab) at the Faculty of English at Adam Mickiewicz University. The experiment was conducted in a dim and quiet cabin in the laboratory. The participants were seated 70 cm away from the monitor. The stimuli were presented on the computer screen in front of the participants using Presentation software (Neurobehavioral Systems). The sentences were presented in a random order using rapid serial visual presentation (RSVP) with one word presented on a computer screen at a time. The stimuli were presented in the centre of the screen in white Consolas font on a grey background. The time sequence for stimuli presentation was as follows: fixation cross for 200-500ms (random) followed by a screen for 400-1000 ms (random). Then all words in the sentence were presented for 300 ms and separated by a blank screen (200 ms). After the last word in the sentence a fixation cross appeared for 700-900 ms (random) followed by a question mark, which indicated that the participants should respond. The delayed response design was used to ensure that motor artifacts did not influence the EEG signal during the 1000 ms time window after the target onset. The question mark remained on the screen until response or for 2000 ms. The intertrial interval lasted for 1000-1500 ms. In the semantic decision task participants decided if a presented sentence is meaningful or meaningless by pressing a corresponding key on the keyboard. Participants used the M and Z keys. Key assignment was counterbalanced across participants.

5.3.4. EEG signal recording and processing

EEG signal was recorded from 64 active actiCAP slim electrodes (BrainVision Analyzer Gilching, Germany): FP1, FP2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, C2, C4, T8, TP9, CP5, CP1, CP2, CP6, TP10, P7, P3, Pz, P4, P8, PO9, O1, Oz, O2, PO10, AF7, AF3, AF4, AF8, F5, F1, F2, F6, FPz, FT7, FC3, FC4, FT8, FCz, C5, C1, Cz, C6, TP7, CP3, CPz, CP4, TP8, P5, P1, P2, P6, PO7, PO3, POz, PO4, PO8 at the standard extended 10/20 positions, with the reference electrode located at the FCz channel. The signal was later re-referenced offline to common average reference (e.g., Rataj et al. 2018a; Jankowiak et al. 2021; Li et al. 2025). In addition, to monitor horizontal eye movements bipolar electrodes were positioned horizontally next to the outer canthus of the eyes (hEOG). To monitor vertical eye movements bipolar electrodes were placed above and below the right eye (vEOG). The EEG signal was amplified by the BrainVision actiCHamp amplifier (Brain Products), at a 1000 Hz sampling rate. Impedances were kept below or at 10 kΩ for each electrode and ERPs were time-locked to the onset of the final word of each sentence.

The signal was pre-processed and analysed using BrainVision Analyzer 2.2. software (Brain Products). First, the signal was filtered offline with a high-pass filter set at 0.1 and a low-pass filter set at 30 Hz. Then the Independent Component Analysis (ICA) was performed to correct for ocular artefacts, and the signal was re-referenced. Data was segmented from 2500 ms before stimulus onset to 4000 ms afterward. Next, the baseline correction was applied, with the baseline defined as -100 ms to 0 ms before stimulus onset. Then the data was inspected for muscular artefacts or excessive noise contaminating the signal. Trials that contained such artefacts were excluded to ensure that only brain activity was analysed. Trials that were above maximal allowed voltage step of 50 μV/ms, or with voltage differences higher than 300 μV (minimal allowed amplitude: -150 μV, maximal allowed amplitude: 150 μV) were rejected from further analysis. At last, the data was segmented per experimental condition: novel metaphors with animacy violation, novel metaphors without animacy violation, literal sentences, and anomalous sentences.

5.3.5. Summary of statistical procedure

Accuracy rates, reaction times and ERP results were analysed using repeated measures ANOVA with sentence-type as the within-subject factor. Twenty-five electrodes were used for the analysis of the EEG signals. For the anterior-posterior axis, the following electrodes were selected: F3, F1, Fz, F2, F4 (frontal); FC3, FC1, FCz, FC2, FC4 (fronto-central); C3, C1, Cz, C2, C4 (central); CP3, CP1, CPz, CP2, CP4 (centro-parietal); P3, P1, Pz, P2, P4 (parietal). For the laterality axis, the following electrodes were chosen: F3, FC3, C3, CP3, P3 (left); F1, FC1, C1, CP1, P1 (left medial); Fz, FCz, Cz, CPz, Pz (mid-line); F4, FC4, C4, CP4, P4 (right); FC1, FC2, C2, CP2, P2 (right medial). For the ERP analysis all responses were selected. Including all trials ensures that neural comparisons reflect experimental manipulation and minimalizes the risk of response-related neural activity distorting effects (VanRullen 2011). Based on the collapsed grand average waveform of all four conditions and all electrodes, the relevant time windows were selected (Luck and Gaspelin 2017). As a result, the analyses were conducted for the following time windows: between 270–500 ms (the N400 time window) and between 600–800 ms (the sustained negativity time window). Accuracy was calculated as the percentage of correct responses in the semantic decision task. Responses were considered correct if participants judged novel metaphors and literal sentences as meaningful and anomalous sentences as meaningless. For statistical analyses, the accuracy rates were subjected to an arcsine transformation (Laurencelle and Cousineau 2023). Accuracy rates and reaction times were analysed using repeated measures ANOVA with sentence-type as the within-subject factor. Significance values for pairwise comparisons were corrected for multiple comparisons using Bonferroni adjustment. Since Mauchly's test indicated a violation of the sphericity assumption, the Greenhouse-Geisser correction was applied. Corrected p-values are reported using the original degrees of freedom.

5.3.6. Behavioural results

5.3.6.1. Accuracy rates

The results of the repeated measures ANOVA showed a significant main effect of sentence type, $F(3, 90) = 105.24$, $p < .001$, $\varepsilon = .600$, $\eta^2_p = .78$. Bonferroni-adjusted pairwise comparisons with the Bonferroni correction showed that literal sentences ($M = 97.77\%$, $SE = 0.47$) were rated significantly more accurately than novel metaphors with animacy violations ($M = 57.07\%$, $SE = 3.16$), novel metaphors without animacy violations ($M = 52.11\%$, $SE = 2.65$), and anomalous sentences ($M = 78.54\%$, $SE = 2.40$), *ps* $< .001$. Moreover, anomalous sentences were also rated significantly more accurately than novel metaphors with animacy violations ($p = .001$) and without animacy violations ($p < .001$). The two types of novel metaphors did not differ significantly ($p = .55$). Mean accuracy rates for each sentence type are provided in Fig. 9.

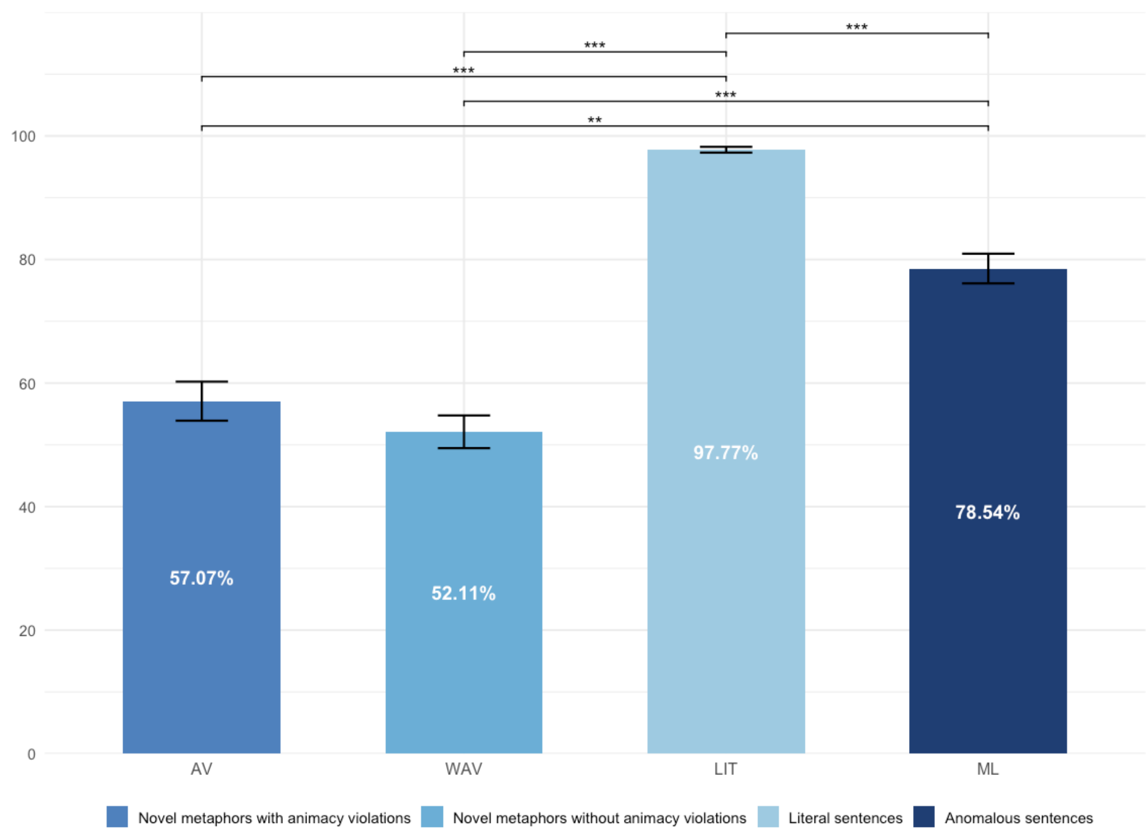

Fig. 9. Accuracy rates (%) for novel metaphors with animacy violations (AV), novel metaphors with-out animacy violations (WAV), literal sentences (LIT), and anomalous sentences (ML).

5.3.6.2. Reaction times

The analysis of reaction times was conducted only on correct responses. For this analysis, reaction times below 150 ms were excluded from the data set. Since Mauchly's test indicated that the sphericity assumption was violated, the Greenhouse-Geisser correction was applied, and corrected p-values are reported using the original degrees of freedom. The repeated-measures ANOVA showed a significant main effect of sentence type, $F(3, 90) = 17.56$, $p < .001$, $\eta^2_p = .36$. Pairwise comparisons using the Bonferroni correction showed that literal sentences evoked the shortest reaction times and differed significantly from novel metaphors with animacy violations, novel metaphors without animacy violations, and anomalous sentences, all *ps* < .001. No significant differences were observed between novel metaphors with animacy violations, novel metaphors without animacy violations, and anomalous sentences, all *ps* >.05. Mean reaction times for each sentence type are provided in Fig.11.

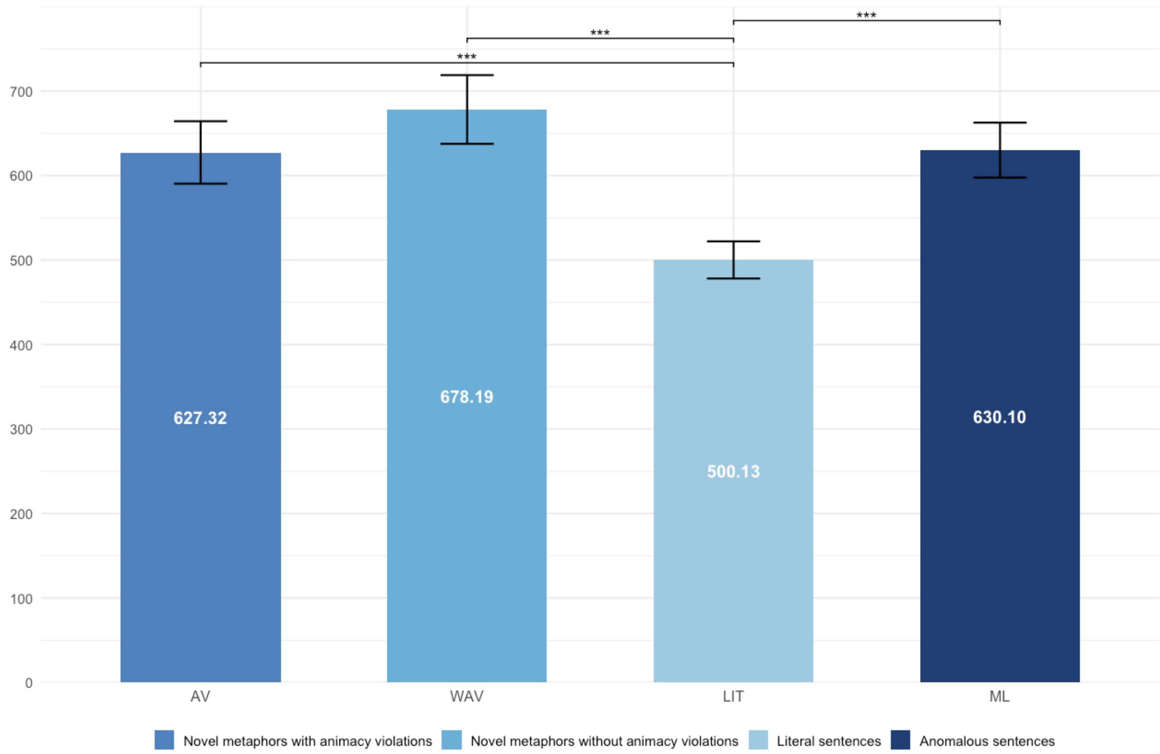

Fig. 10. Reaction times (ms) for novel metaphors with animacy violations, novel metaphors without animacy violations, literal sentences, and anomalous sentences

5.3.7. The N400 time window (270-500ms)

The results showed that the three-way anterior-posterior axis by laterality by sentence type interaction was nonsignificant, $F(48, 1440) = 0.95$, $p = .565$, $\eta^2_p = .001$. Similarly, the interaction between sentence type and laterality was nonsignificant, $F(12, 360) = 0.81$, $p = .544$, $\eta^2_p = .001$. A significant interaction was observed between sentence type and anterior-posterior axis, $F(12, 360) = 9.87$, $p < .001$, $\eta^2_p = .037$. To deconstruct the observed interaction, separate analyses were performed on each level of the anterior-posterior axis. The results showed a significant sentence type effect over frontal, $F(3, 90) = 10.00$, $p < .001$ $\eta^2_p = .068$, fronto-central, $F(3, 90) = 5.59$, $p = .003$, $\eta^2_p = .032$, and parietal sites, $F(3, 90) = 7.66$, $p < .001$, $\eta^2_p = .053$ sites. It was also marginally significant over the centro-parietal sites, $F(3, 90) = 2.74$, $p = .054$, $\eta^2_p = .019$. The effect was nonsignificant over the central sites, $F(3, 90) = 0.62$, $p = .602$, $\eta^2_p = .004$. As the effect of sentence type was strongest over frontal sites and also present over the fronto-central sites (see Table 2, Fig. 11, 12), the follow-up analysis was conducted on the average of electrodes from these two groups and included the following electrodes: F3, F1, Fz, F2, F4, FC3, FC1, FCz, FC2, FC4. The analysis showed a main effect of sentence type, $F(3, 90) = 8.42$, $p < .001$, $\eta^2_p = .052$. Pairwise comparisons with the Bonferroni correction showed that novel metaphors with animacy violations elicited significantly more negative amplitudes than novel metaphors without animacy violations ($p = .016$) and literal sentences ($p < .001$). Moreover, amplitudes for anomalous sentences were significantly more negative than those elicited by novel metaphors without animacy violations ($p = .02$) and literal sentences ($p = .02$). No significant difference was observed between novel metaphors with animacy violations and anomalous sentences ($p = 1$). Estimated marginal means showed that metaphors with animacy violations elicited the largest N400 amplitudes, followed by anomalous sentences, novel metaphors without animacy violations, and least negative for literal sentences (Table 2, Fig. 11).[1]

---

[1] A supplementary analysis was conducted on correct responses in the N400 time window. The pattern of results was consisted with the reported findings.

Table 2. Mean amplitudes for novel metaphors with animacy violations (AV), novel metaphors without animacy violations (NAV), literal sentences (LIT), and anomalous sentences (ML) over fontal and fronto-central electrode positions within the 270-500ms time window.

Electrode position	AV	NAV	LIT	ML	Pairwise comparisons
Frontal	*M* = -1.36 *SE* = 0.31	*M* = -0.45 *SE* = 0.28	*M* = -0.23 *SE* = 0.32	*M* = -1.33 *SE* = 0.33	AV vs LIT ($p < .001$) AV vs NAV ($p = .01$) LIT vs ML ($p = .004$) ML vs NAV ($p =.006$)
Fronto-central	*M* = -1.09 *SE* = 0.27	*M* = -0.45 *SE* = 0.25	*M* = -0.31 *SE* = 0.30	*M* = -0.89 *SE* = 0.30	AV vs LIT ($p = .001$) AV vs NAV ($p =.03$)
Frontal, Fronto-central	*M* = –1.23 *SE* = 0.28	*M* = -0.45 *SE* = 0.25	*M* = -0.27 *SE* = 0.30	*M* = -1.11 *SE* = 0.31	AV vs LIT ($p < .001$) AV vs NAV ($p =.016$) LIT vs ML ($p = .02$) ML vs NAV ($p = .02$)

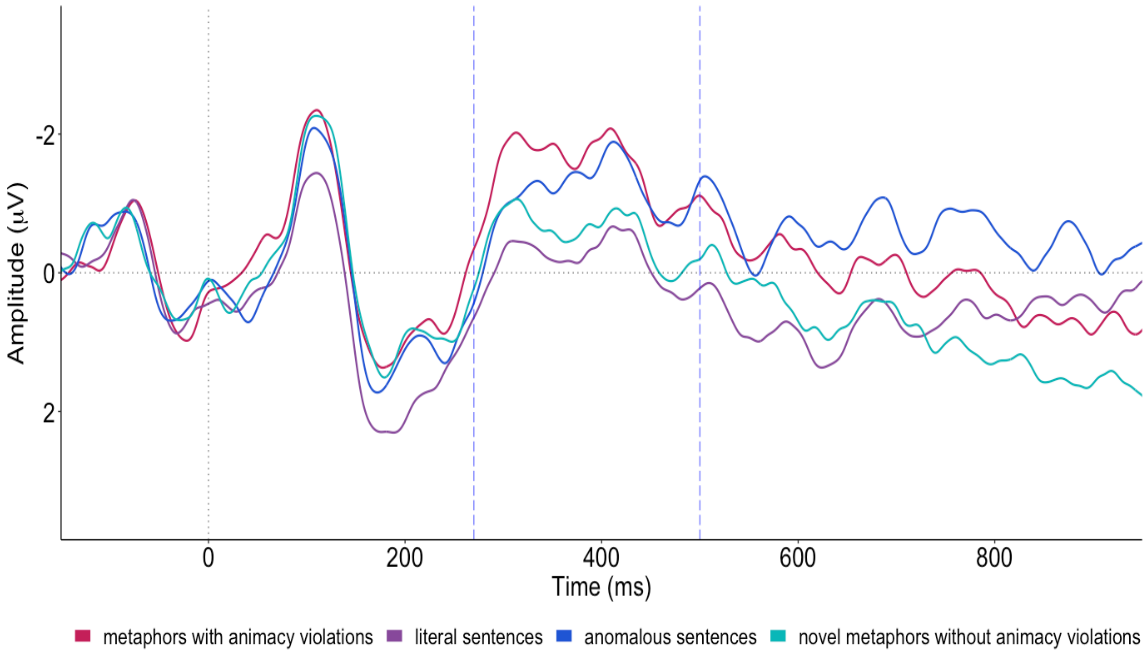

Fig. 11. Grand averages for novel metaphors with animacy violations (red line), novel metaphors without animacy violations (green line), literal sentences (purple line), and anomalous sentences (blue line) over frontal and fronto-central electrodes with marked time window 270-500ms.[2]

[2] A supplementary analysis was conducted to inspect the baseline. No significant difference between conditions was observed.

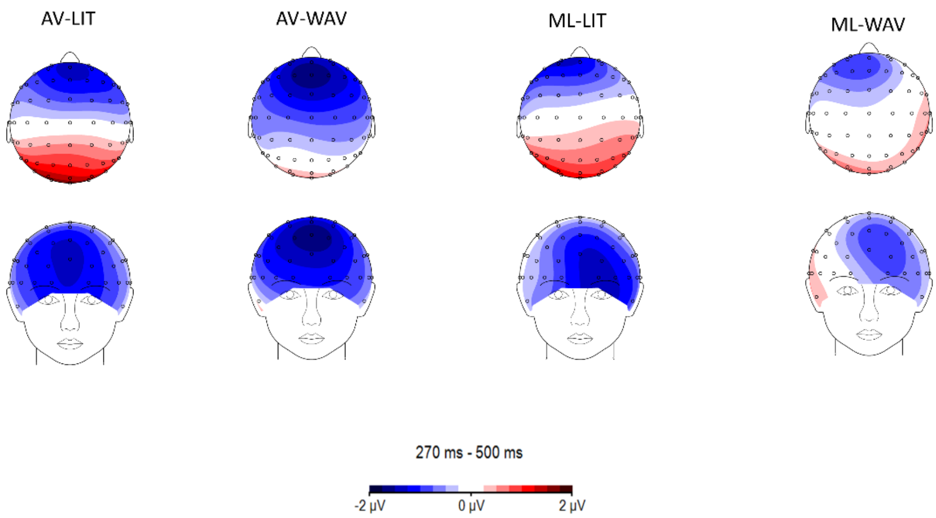

Fig. 12. Difference topographical maps of the N400. The maps show differences between novel metaphors with animacy violations (AV), novel metaphors without animacy violations (WAV), literal sentences (LIT), and anomalous sentences (ML).

5.3.8. The sustained negativity time window (600-800ms)

The results showed that the three-way anterior-posterior axis by laterality by sentence type interaction was nonsignificant, $F(48, 1440) = 1.28$, $p = .095$, $\eta^2_p = .002$. Moreover, the two-way interaction of sentence type by laterality also showed an nonsignificant effect, $F(12, 360) = 1.25$, $p = .289$, $\eta^2_p = .002$. The results revealed a significant main effect of sentence type, $F(3, 90) = 4.88$, $p = .007$, $\eta^2_p = .010$ and an interaction between sentence type and the anterior-posterior axis, $F(12, 360) = 6.33$, $p < .001$, $\eta^2_p = .015$. To deconstruct the interaction additional analyses were conducted. The main effect of sentence type was significant over the frontal, $F(3, 90) = 11.49$, $p < .001$, $\eta^2_p = .053$, fronto-central, $F(3, 90) = 7.59$, $p < .001$, $\eta^2_p = .029$, and central, $F(3, 90) = 3.47$, $p = .028$, $\eta^2_p = .014$ sites. The main effect was nonsignificant over the centro-parietal, $F(3, 90) = 0.58$, $p = .609$, $\eta^2_p = .003$, and parietal sites, $F(3, 90) = 1.11$, $p = .347$, $\eta^2_p = .006$ (Table 3, Fig. 13, 14). Based on observed effects on frontal and fronto-central sites, where the effect was strongest, a region of interest analysis was conducted by averaging across the electrodes encompassing these areas (F3, F1, Fz, F2, F4, FC3, FC1, FCz, FC2, FC4).[3] This analysis showed a main effect of sentence type, $F(3, 90) = 10.21$, $p < .001$, $\eta^2_p = .044$. The pairwise comparisons with the Bonferroni correction revealed that anomalous sentences evoked significantly more negative amplitudes than novel metaphors without animacy violation, $p < .001$ and literal sentences, $p = .002$. No significant difference was observed between novel metaphors with animacy violations and novel metaphors without animacy violations ($p = .1$), as well as anomalous sentences ($p = .1$). However, novel metaphors with animacy violation elicited more negative amplitudes than literal sentences ($p = .04$). Additionally, a linear effect was observed, $F(1, 1236) = 45.05$, $p < .001$, $\eta^2_p = .040$.[4]

Table 3. Mean amplitudes for novel metaphors with animacy violations (AV), novel metaphors without animacy violations (WAV), literal sentences (LIT), and anomalous sentences (ML) over frontal and fronto-central electrode positions within the 600-800ms time window.

Electrode position	**AV**	**NAV**	**LIT**	**ML**	**Pairwise comparisons**
Frontal	*M* = -0.33 *SE* = 0.41	*M* = 0.36 *SE* = 0.38	*M* = 0.38 *SE* = 0.43	*M* = -1.10 *SE* = 0.43	AV vs ML (*p* = .04) LIT vs ML (*p* < .001) ML vs NAV (*p* < .001)
Fronto-central	*M* = 0.05 *SE* = 0.38	*M* = 0.63 *SE* = 0.35	*M* = 0.77 *SE* = 0.42	*M* = -0.25 *SE* = 0.38	AV vs LIT (*p* = .02) LIT vs ML (*p* = .01) ML vs NAV (*p* = .004)
Frontal, fronto-central	*M* = -0.14, *SE* = 0.39	*M* = 0.50 *SE* = 0.35	*M* = 0.58 *SE* = 0.41	*M* = -0.67 *SE* = 0.40	AV vs LIT (*p* = .04) LIT vs ML (*p* = .002) ML vs NAV (*p* < .001)

[3] A supplementary analysis of the frontal, fronto-central, and central sites combined showed the same pattern of results.
[4] A supplementary analysis was conducted on correct responses in the sustained negativity time window. The pattern of results was consistent with the reported findings.

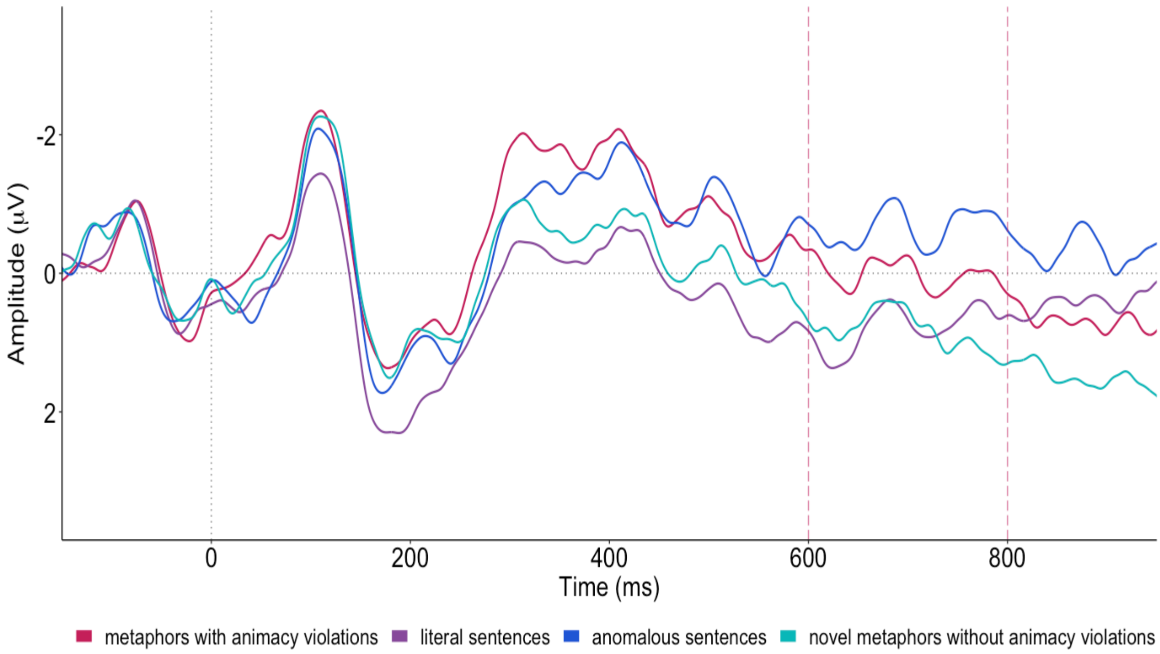

Fig. 13. Grand averages for novel metaphors with animacy violations (red line), novel metaphors without animacy violations (green line), literal sentences (purple line), and anomalous sentences (blue line) over frontal and fronto-central electrodes with marked time window 600-800ms.

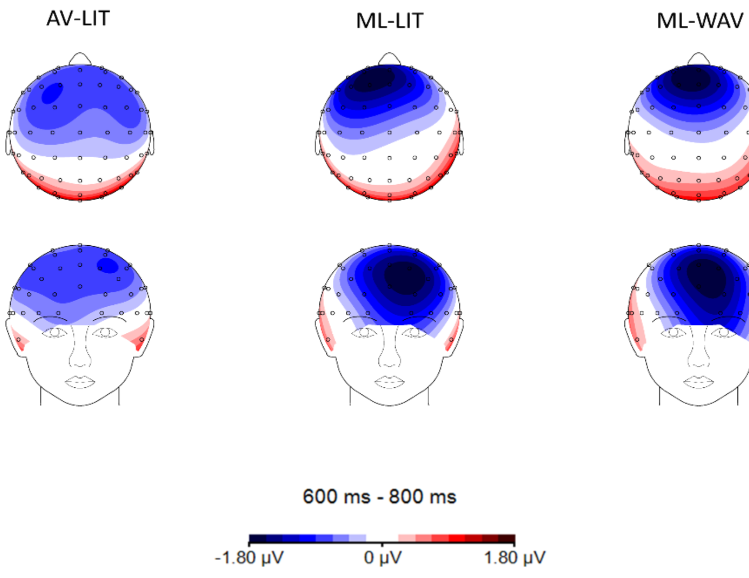

Fig. 14. Difference topographical maps of the sustained negativity. The maps show differences between novel metaphors with animacy violations (AV), novel metaphors without animacy violations (WAV), literal sentences (LIT), and anomalous sentences (ML).

5.4. Discussion and conclusions

The experiment aimed to investigate the role of animacy violations in novel metaphor processing and comprehension. The ERP measure was used to capture real-time neural responses to novel metaphors with animacy violations and without animacy violations. Additionally, reaction times and accuracy rates were collected to provide a comprehensive account of the findings. The ERP study employed a two-choice semantic decision task with a delayed response procedure including four types of stimuli: novel metaphors with animacy violations, novel metaphors without animacy violations, literal sentences and anomalous sentences. As discussed in Section 5.2., it was hypothesized that differences between conditions would modulate the N400 and sustained negativity amplitudes. For the N400 time window, a graded effect was predicted with largest negativity for anomalous sentences, moderate for novel metaphors and smallest for literal sentences. Novel metaphors with animacy violations were predicted to elicit larger N400 amplitudes than novel metaphors without animacy violations, and literal sentences. Moreover, in the sustained negativity time window, novel metaphors were expected to evoke more pronounced responses than literal sentences. Additionally, animacy violation was predicted to modulate sustained negativity time window with increased amplitudes for novel metaphors with animacy violations as compared to novel metaphors without animacy violations and literal sentences.

In the N400 time window (270-500ms), novel metaphors with animacy violations elicited most negative amplitudes, followed by anomalous sentences, novel metaphors without animacy violations, and literal sentences. There was no significant difference between novel metaphors with animacy violations and anomalous sentences. Interestingly, novel metaphors without animacy violations did not differ from literal sentences. A linear effect was observed in the N400 time window, but differed from the graded effects consistently reported across numerous studies (e.g., Arzouan et al. 2007; Rutter et al. 2012; Jankowiak et al. 2017; Rataj et al. 2018; Jończyk et al. 2020; Jankowiak et al. 2021). The frequently observed graded effect, which was predicted in Hypothesis 1, shows most pronounced N400 response to anomalous utterances, followed by novel metaphors, and least pronounced for literal utterances. In previous research, such an effect has been interpreted in line with The Career of Metaphor model (Bowdle and Gentner 2005), which states that novel metaphors are comprehended through comparison and require establishing links between the source and target domain. Moreover, the findings can support the Conceptual Blending theory, which proposes that the new metaphorical meaning is derived from combining meaning from different mental spaces (e.g., Coulson and Van Petten 2002, Rutter et al. 2012). The increased N400 amplitudes for novel metaphors as compared to literal sentences have been interpreted as reflecting an increased difficulty of forming cross-domain mappings. Additionally, the difference in the N400 amplitudes between novel metaphors and literal sentences has been linked to an increased engagement of semantic memory in the processing of novel metaphors (Rataj et al. 2018b).

In the present study, the most pronounced negativity in the N400 time window were evoked by novel metaphors with animacy violation, and not as in previous studies by anomalous sentences. However, the difference between the two conditions was not significant. Similar results for novel metaphors were observed by Lai et al. (2009), in which responses to novel metaphors did not differ from responses to anomalous sentences in the 440-560 ms time window. Importantly, animacy violation was not controlled in that study. In the present study, a similar effect was observed only for novel metaphors with animacy violation, which points to increased processing demands for both types of utterances due to lack of readily available mappings between domains (Bowdle and Gentner 2005), additionally enhanced by animacy violations in the case of novel metaphors.

As predicted in Hypothesis 3, animacy was expected to modulate the effect in the N400 time window. This prediction was based on results from several studies that have reported increased N400 amplitudes in response to animacy violations in language processing. Paczynski and Kuperberg (2011) observed more pronounced negativity in the N400 time window for animacy violations created by assigning animate nouns in the Patient role (usually associated with inanimate nouns), and inanimate nouns in the Experiencer role (typically needs animate nouns). The observed N400 modulation was interpreted by the researchers as response to violation of the animacy hierarchy (Aissen 2003) present also in language. As argued by the authors, grammatically many languages prefer placing animate nouns before inanimate nouns. Hence, when the animacy-based expectation is violated, it disrupts lexico-semantic processing and results in larger N400 responses. Wang et al. (2024) conducted a study investigating picture-word interference manipulating animacy congruence. The results showed that animacy incongruent condition (e.g., a picture of a paper bag and a distractor word *baby*) evoked increased N400 amplitudes in comparison to the animacy congruent condition (e.g., a picture of a paper bag and a distractor word *coaster*). The results were interpreted in line with the Levelt's Model of Speech Production (Levelt 1989), which proposes that items belonging to the same semantic category demand less cognitive resources in processing as they are located in closer proximity in the semantic network. Thus, the authors claim that items sharing semantic features (e.g., animacy) require reduced processing demands, and evoke attenuated N400 responses. By analogy, animacy incongruencies (i.e., animacy violations) demand enhanced semantic integration. Moreover, Czypionka et al. (2023) reported the facilitatory influence of animacy on lexical access. The researchers observed that animate single words and compounds were more easily processed than inanimate, showing reduced N400 responses. The results suggest that animacy is a significant semantic feature that is automatically accessed and facilitates processing.

The difference between novel metaphors with animacy violations and novel metaphors without animacy violations observed in the present study is similar to the study conducted by Ji et al. (2020). In the study by Ji et al. (2020), animacy violations elicited increased N400 amplitudes at the verb position in the sentences, indicating higher processing effort and increased difficulty in the integration of the inanimate nouns with verbs. However, no such N400 modulations were observed measured at the noun in the last part of the sentences. Noteworthily, while the materials used for the study by Ji et al. (2020) were constructed combining inanimate nouns with verbs, in the present study animacy violation was established by combining animate adjectives with inanimate nouns. Therefore, the difference in materials indicates that the present study extends prior findings by demonstrating that the animacy effect emerges not only in verb-based violations but also in adjective-noun metaphors. Li et al. (2025) also reported larger N400 amplitudes in response to high-animacy (unexpected in canonical SVO structures) novel metaphors relative to low-animacy novel metaphors and interpreted these results as a reflection of larger cognitive demands in novel metaphor processing caused by animacy violations.

Interestingly, the results of the current study did not show a significant difference between novel metaphors without animacy violations and literal sentences. This has not been supported by normative studies and behavioural data, which showed that novel metaphors differed significantly from literal sentences in meaningfulness, metaphoricity, familiarity ratings and the cloze probability test (Section 4.6). In the study conducted by Li et al. (2022), larger N400 amplitudes (time-locked to the verb onset) were observed for the verb-object metaphors (no animacy violation) and literal concrete sentences (no animacy violation) compared to subject-verb metaphors (animacy violation) and literal-abstract sentences (animacy violation). No significant differences were observed between literal sentences without animacy violation and metaphors without animacy violation. In comparison, Ji et al. (2020) observed increased N400 amplitudes for both animate and inanimate metaphors in comparison to literal sentences. In the study by Li et al. (2025) literal fillers were not included in the analyses. Studies on novel metaphor comprehension in which animacy violation was not controlled reported significant differences between novel metaphors and literal utterances (e.g., Pynte et al. 1996; Rutter et al. 2012; Rataj et al. 2018; Jankowiak et al. 2021)

Overall, despite scarcity of research on the aspect of animacy violation in novel metaphoric language processing, there is substantial evidence (as discussed in Chapter 3) that animacy is one of the fundamental cognitive dimensions, is evolutionarily significant, and organizes our semantic representations. The results obtained in the present study support these observations. The significant difference between novel metaphors with animacy violations and novel metaphors without animacy violations can be interpreted as indexing variation in cognitive load needed to construct meaning. Such observation can stem from different categories of the source (animate) and target (inanimate) domains for novel metaphors with animacy violations, making these metaphors more cognitively taxing than novel metaphors without animacy violations (Traugott 1985: 22-23), thus resulting in more negative N400 responses relative to novel metaphors without animacy violations.

Analyses of the late time window (600-800ms) revealed prolonged negativity with most negative amplitudes for anomalous sentences and novel metaphors with animacy violations, followed by metaphors without animacy violations and literal sentences. Similarly to the N400, no significant differences were observed between anomalous sentences and novel metaphors with animacy violations (Abraham et al. 2021). Also, no difference was found between novel metaphors without animacy violations and literal sentences. However, a linear effect was observed. The sustained negativity has been frequently observed in studies on novel metaphors (e.g., Arzouan et al. 2007; Regel et al. 2011; Goldstein et al. 2012; Rutter et al. 2012; Rataj et al. 2018; Jankowiak et al. 2021). Rataj et al.

(2018) interpreted the attenuated LPC responses to novel metaphors as an overlap with the sustained negativity effect elicited by such stimuli. In line with the results observed in the study by Rataj et al. (2018), the present study observed similar topography for both the N400 and sustained negativity (frontal and fronto-central). Suggested by the authors, sustained negativity can reflect increased working memory load stemming from effortful attempts in forming mappings between two distant domains for novel metaphors. This interpretation is consistent with the Career of Metaphor model (Bowdle and Gentner 2005), according to which novel and creative utterances are processed and comprehended through comparisons between the source domain and target domain. For anomalous sentences, the sustained negativity may indicate memory retrieval of the sentence context and the anomaly (Rataj et al. 2018b). In the present study, the significant difference between novel metaphors with animacy violations and without animacy violations observed in the N400 time window had disappeared in the later time window. Interestingly, although not observed in the analysis on the region of interest (frontal and fronto-central sites combined), anomalous sentences differed significantly from novel metaphors with animacy violations over the frontal sites, where effects were strongest, while no such effect was observed for the N400 time window. This could potentially indicate a trend towards successful meaning construction for novel metaphors with animacy violations and seems to show that while the metaphoricity effect persisted, the animacy effect did not.

Overall, amplitudes for novel metaphors without animacy violations showing no significant difference when compared to literal sentences in both time windows could point to smaller demands in meaning integration when juxtaposed with novel metaphors with animacy violations. The observed effect of animacy violation, especially in later stages of novel metaphor processing, is relatively novel in light of existing research. Contrary to previously noted increased processing effort in meaning integration for animacy violations, which resulted in more robust LPC responses (e.g., Kuperberg et al. 2003, 2007; Paczynski and Kuperberg 2011; Szewczyk and Schriefers 2011), Ji et al. (2020) observed attenuated P600 responses to metaphors with animacy violations, which the authors interpreted as indication that animacy violations in metaphors potentially facilitate late meaning integration. In comparison, Li et al. (2022) observed increased LPC amplitudes for metaphors with animacy violations at the verb processing stage interpreted as reanalysis oriented towards abstract interpretation. Although animacy violation was predicted to influence the sustained negativity in Hypothesis 4, none of these studies on animacy violations in metaphoric language reported sustained negativity observed in the present study.

Interestingly, the observed N400 and sustained negativity effects were strongest over the anterior sites. Although more frontal distribution for the late components has been observed in several previous studies (e.g., Coulson and Van Petten 2002; Bambini et al. 2016; Rataj et al. 2018), the topography for the N400 differs from the most frequently observed centro-parietal distribution in studies on novel metaphor comprehension. However, frontal negativity in the N400 time window has been observed in a study by Bambini et al. (2018), in which, the researchers investigated literary metaphors (taken from poems and novels e.g., *grass of velvet*) in comparison to literal utterances in a semantic decision task. The results showed that metaphors elicited increased negativity in the N400 time window followed by sustained negativity over frontal sites. The authors suggested that the fronto-central distribution of the N400 effect was likely the result of an overlap with subsequent frontal sustained negativity. Furthermore, Menashe et al. (2024) investigated conventional and novel metaphor processing and comprehension. The researchers observed N400 amplitudes with frontal distribution for both conventional and novel metaphors in a study examining attentional functions (alerting, orienting, and executive control). Moreover, the authors reported a significant positive correlation between the frontal N400 effect for novel metaphors and executive control network. The results were interpreted in line with the Career of Metaphor model (Bowdle and Gentner 2005), which proposes that novel metaphors require retrieval and comparison of semantic features. The authors also stated that both types of metaphors were semantically demanding because they were more complex and longer (sentences) in comparison to other studies.

In a larger context of previous research on the N400 component, there has been a discussion on the frontally distributed negativity which led to a differentiation between the standard N400 component and the frontal N400 (FN400)(Bridger et al. 2012; Stróżak et al. 2016; Leynes et al. 2024). The FN400 has been observed in old/new recognition tasks (e.g., Bridger et al. 2012) and interpreted as reflection of relative familiarity, which is a context-based familiarity resulting from experimental repetition (Mecklinger and Bader 2020). However, the frontally distributed N400 has also been linked to task demands and the process of updating in working memory (Kiss et al. 2007). In their study, Kiss et al. (2007) explored maintenance and control (updating) processes. In the maintenance condition, the task required the participants to retain and refresh information in mind for a certain period of time, whereas in the updating condition, the participants had to actively update and monitor information in the working memory. The results revealed frontally distributed negativity in the N400 time window for the updating condition. This finding has been interpreted as indication of heightened processing demands required for information updating in the working memory. Noteworthily, the Rapid Serial Visual Presentation (RSVP) technique employed in the present study, required participants to actively manage and update information, as they needed to integrate the final word with preceding sentence context to make the meaningful/meaningless judgements. Moreover, Boudewyn et al. (2013) noted that the frontal N400 can be modulated by working memory load and semantic integration processes, so combining the presented stimuli with preceding context. In their study, the participants listened to short stories with either congruent or incongruent context preceding the last sentence. The authors also investigated individual differences in a recall task. The results showed frontal negativity in the N400 time window for participants with better working memory capacity (based on the number of recalled items). Once again, the finding is relevant to the present study, as the word-by-word presentation requires semantic integration of the final word with the preceding sentence context.

All these results show that the frontally distributed N400 can be expected in experiments employing cognitively demanding tasks and stimuli, which increase the working memory load. This is particularly relevant to the present study as both mentioned factors could have contributed to the frontal topography of the N400 effect. In the present ERP experiment, the two-choice semantic task was used with a delayed response paradigm. This design reduced the possibility of decision related motor artifacts influencing ERP effects. At the same time, the delayed response task is cognitively demanding because the participants have to both keep the sentence in memory and withhold the responses, while they actively maintain and integrate the meaning of the presented stimuli. The retention of meaning until the response is taxing and increases working memory load, which was observed to influence the frontally distributed N400 (Kiss et al. 2007; Boudewyn et al. 2019). Moreover, as discussed, the Rapid Serial Visual Presentation (RSVP) technique, which has been used in studies on metaphors (e.g., Jończyk et al. 2020; Miller et al. 2025) and animacy (e.g., Nieuwland and Van Berkum 2006) contributes to increased memory load. In this technique, stimuli are presented word-by-word, each only for a brief duration (e.g., 200 ms). Such a stimulus presentation method requires participants to retain the sentence context in memory in order to successfully integrate it with the final word.

As for stimuli, animacy has been repeatedly shown to influence memory processes. The results across studies on long-term memory and working memory (e.g., Nairne et al. 2013; Daley et al. 2020; Bonin et al. 2014) showed that animates take priority over inanimates in memory retention, possibly due to richer encoding. The advantage of animates can be also explained in line with the animate monitoring hypothesis (New et al. 2007), which states that the priority of animates is evolutionary and stems from survival. Taking into consideration the observed results, it could be argued that upon arriving on animacy violations in novel metaphors in this study, the default preference of animates (e.g., animate noun) needs to be overridden in order to derive meaning. Such occurrence would increase working memory demands needed for this reanalysis and meaning integration. Taken together, the frontally distributed negativity observed in the N400 time window in the present study can be attributed to the role of animacy violation in metaphor processing as well as task demands.

Lastly, as predicted by Hypothesis 5, the behavioural results showed significant differences between novel metaphors and literal sentences. As expected, both types of novel metaphors were judged less accurately than literal sentences. Moreover, there was a linear effect with highest accuracy scores for literal sentences, followed by anomalous sentences, and lowest accuracy scores for novel metaphors. Animacy violation did not modulate the accuracy scores as there was no significant difference observed between novel metaphors with animacy violations and novel metaphors without animacy violations. The obtained accuracy results are in line with previous research on metaphor comprehension (e.g., Goldstein et al. 2012; Tang et al. 2017; Jankowiak et al. 2021; Rataj et al. 2018). Similarly, the analysis of reaction times showed that literal sentences elicited significantly shorter reaction times than novel metaphors with animacy violations, novel metaphors without animacy violations, and anomalous sentences. Longer RTs for novel metaphors, compared to literal utterances likely reflect greater cognitive demands required to establish links between distant concepts for novel metaphors in comparison to literal sentences, thus being more time consuming (Bowdle and Gentner 2005). This observation also has been frequently made in studies on metaphor processing (e.g., Goldstein et al. 2012; Tang et al. 2017; Jankowiak et al. 2021; Rataj et al. 2018).

Especially in a binary semantic decision tasks, novel metaphors can be more difficult to qualify as meaningful or meaningless in comparison to literal sentences, which are easily qualified as meaningful (Rataj et al. 2018b).

An alternative explanation could be the influence of familiarity and probability on stimuli processing. Although literal sentences were rated in the norming study as only moderately familiar, there was a significant difference observed in the familiarity ratings between both types of novel metaphors and literal sentences. Similarly, literal sentences got higher cloze probability score than the other three conditions. However, the ERP results do not support the familiarity explanation, as novel metaphors without animacy violations despite being rated as unfamiliar were processed similarly to literal sentences, both resulting in attenuated N400 amplitudes.

As for the results for anomalous sentences, usually in studies on metaphor processing a graded RTs effect is observed (with longest RTs for novel metaphors, followed by anomalous utterances and shortest for literal utterances) (e.g., Rutter et al. 2012; Goldstein et al. 2012; Lai et al. 2009). In this study no significant difference was observed between novel metaphors and anomalous sentences in reaction times. The potential reason for this observation could be that with two types of cognitively taxing novel metaphors, participants were more cautious in judgements of sentences including any type of anomalies, resulting in equally long reaction times. The accuracy pattern suggests that participants were actively searching for meaning, which could have led to slightly lower accuracy rates for anomalous sentences and no difference in RTs between novel metaphors and anomalous sentences. Moreover, the delayed response design could have contributed to the more attentive judgements as it required participants to withhold and retain the meaning.

Noteworthily, reaction times were obtained in a delayed response paradigm, in which participants responded 1 second after the critical word onset. Although most studies have employed immediate judgements in the semantic decision task (e.g., Lai et al. 2009; Jankowiak et al. 2017, 2021), some have used the delayed response paradigm and observed significant results (e.g., De Grauwe et al. 2010; Rataj et al. 2018, 2025). While the reaction times obtained in the delayed response design are not fully indicative of online processes, the results in the present study showed significant effects, thus are treated as meaningful and capturing processes related to semantic processing, integration as well as decision making.

5.5. Final remarks and suggestions

Taken together, the study investigated the role of animacy violation in novel metaphor processing and comprehension. It was hypothesised that novel metaphors would increase the N400 and sustained negativity amplitudes in comparison to literal sentences. Furthermore, animacy violation was predicted to modulate the N400 and sustained negativity amplitudes assuming that novel metaphors with animacy violations would be more cognitively taxing.

It has been observed that indeed animacy violation seems to influence how novel metaphoric language is understood. In the N400 time window, a linear effect of sentence type was observed. Although, pairwise comparisons failed to show significant differences between novel metaphors with animacy violations and anomalous sentences. In the same manner, no differences were observed between novel metaphors without animacy violations and literal sentences. Moreover, sustained negativity was observed in the LPC time window with a similar linear effect observed in the N400 time window, except for anomalous sentences evoking larger responses than novel metaphors with animacy violations. The findings were interpreted in line with the Career of Metaphor model (Bowdle and Gentner 2005), which suggests that novel and creative metaphors are processed by analogy between the source and target domain. The obtained results suggest increased working memory load in attempt to form mappings between the two domains.

Both the N400 amplitudes and sustained negativity showed frontal topography. Studies on novel metaphors usually report the N400 effect over the centro-parietal site. However, the frontally distributed N400 has been observed in studies on novel metaphors as well (e.g., Bambini et al. 2018; Menashe et al. 2024). Moreover, frontal N400 has been associated with increased working memory load caused by information updating and semantic integration processes (e.g., Kiss et al. 2007; Boudewyn et al. 2019). In the present study, the frontal distribution of both components points to increased working memory engagement and has been attributed to task complexity due to delayed response design, and more difficult stimuli manipulated by animacy violations.

Taken together, the present study obtained results similar to existing studies on animacy violations in novel metaphoric language i.e. observed increased N400 responses to novel metaphors with animacy violations. However, the observed sustained negativity and frontal distribution of both effects diverge from the results reported in the few studies on the aspect of animacy violations in novel metaphor processing and comprehension. Still, the literature on this topic remains limited and more research is needed.

While the present findings provide valuable insights, they also raise potential questions, which point toward directions for future research. This study employed two types of novel metaphors manipulating the aspect of animacy violations. It has been observed that novel metaphors with animacy violations were perceived similarly to anomalous sentences and novel metaphors without animacy violations similarly to literal sentences. This observation should be investigated further. Moreover, this study only focused on novel metaphoric language. Future research could benefit from including comparison to conventional figurative language with the aspect of animacy violations. This comparison would broaden the perspective on the role of animacy violations. Furthermore, the present study design affected the working memory load, which in turn could have influenced the distribution of the effect. Prospective studies could take into consideration a less straining task design (i.e., without the delayed response procedure) and compare findings. Lastly, most of the studies published on animacy violations in novel metaphor comprehension have been conducted in Mandarin Chinese, a language that differs substantially from, e.g. Polish in aspects of grammar, morphology, and word order. Therefore, examining this phenomenon in other languages would provide a more comprehensive view on the role of animacy violations in novel metaphor processing and comprehension.

Conclusion

The purpose of this dissertation was to investigate the role of animacy violation in processing and comprehension in novel metaphors. The existing body of electrophysiological research on metaphoric language comprehension has shown consistent results (e.g., Jankowiak et al. 2017; Rataj et al. 2018b; Lai et al. 2009; Coulson and Van Petten 2002). The obtained results have been often interpreted in line with the Career of Metaphor model (Bowdle and Gentner 2005), which proposes that novel and creative metaphoric expressions are comprehended by a comparison of two distinct concepts, and establishing correspondences between the source and target domains, and the Conceptual Blending theory (Fauconnier and Turner 2003), which states that new meaning is constructed by blending different mental spaces. The process of meaning construction for novel metaphors has been observed to result in increased N400 amplitudes in comparison to literal utterances, reflecting increased difficulty in forming cross-domain mappings (e.g., Coulson and Van Petten 2002; Lai et al. 2009; Rataj et al. 2018). Moreover, novel metaphors have been observed to influence the post-N400 effects. Some studies have observed more pronounced positivity for metaphors in comparison to literal sentences (Coulson and Van Petten 2002; De Grauwe et al. 2010; Bambini et al. 2016). While other studies have reported sustained negativity with more negative amplitudes for novel metaphors relative to literal utterances (e.g., Arzouan et al. 2007; Lai et al. 2009; Rutter et al. 2012; Jankowiak et al. 2021). This effect has been interpreted as continued mapping process between two domains and meaning integration (e.g., Arzouan et al. 2007; Rataj et al. 2018).

The aspect of animacy investigated in the present study has been observed to be one of the fundamental features of cognition. The ability to distinguish between animate and inanimate entities emerged from an evolutionary need for survival (New et al. 2007) and has been observed as early as infancy (Poulin-Dubois et al. 1996). Moreover, studies have shown that animacy is an important factor influencing attention (Loucks et al. 2023) and memory (Bonin et al. 2014). Animacy violation has been investigated in language research and has been observed to be cognitively taxing (e.g., Kuperberg et al. 2003b; Szewczyk and Schriefers 2011). The increased lexical processing demands linked to animacy violation have been reflected in larger N400 amplitudes (Paczynski and Kuperberg

2011; Czypionka et al. 2023; Wang et al. 2024) and the increased meaning integration difficulty for animacy violations has been associated with pronounced LPC responses (e.g., Kuperberg et al. 2003, 2007; Paczynski and Kuperberg 2011; Szewczyk and Schriefers 2011). However, only few studies have explored this aspect in novel metaphor processing and comprehension (Ji et al. 2020; Li et al. 2022, 2025) with a variability in study designs and obtained findings. By combining novel metaphors and the aspect of aniamcy violation, the present study aimed to fill this gap in existing research.

The ERP study discussed in the dissertation was preceded by a large-scale normative study, whose aim was to develop reliable materials. Rating scales included metaphoricity, meaningfulness, and familiarity ratings of the selected materials. Moreover, a cloze probability test was conducted to investigate critical word predictability. The normative studies were conducted to account for factors observed to influence patterns of results in event-related potential studies on novel metaphor. The ERP study employed a two-choice semantic decision task and recorded neural responses to novel metaphors with animacy violations, novel metaphors without animacy violations, literal sentences, and anomalous sentences. As predicted, the analysis of the N400 time window showed increased negativity for metaphors with animacy violation. The results showed larger negativity for novel metaphors with animacy violations and anomalous sentences than for novel metaphors without animacy violation and literal sentences. The results showed that novel metaphors with animacy violations demand increased semantic processing in comparison to literal sentences in line with previous studies (e.g., Jankowiak et al. 2017; Rataj et al. 2018; Jankowiak et al. 2021). Increased N400 amplitudes in response to novel metaphors with animacy violations relative to novel metaphors without animacy violations have been observed in the studies by Ji et al. (2020) and Li et al. (2025). In line with these observations, the results of the present study have shown that animacy significantly impacts novel metaphor processing. These findings point to enhanced difficulty in semantic integration for novel metaphors with animacy violations. Consistent with previous studies, these results can be interpreted as providing evidence for the Career of Metaphor model (Bowdle and Gentner 2005) as greater difficulty in establishing mappings between two distant concepts belonging to different categories of the source (animate) and target (inanimate) domains (Traugott 1985: 22-23). The obtained results also support the Conceptual Blending theory, which proposes that metaphorical meaning is derived through the integration of different mental spaces, resulting in construction of the new meaning.

This process can be cognitively demanding as it requires establishing mappings and meaning integration.

Moreover, analyses of the post-N400 time window showed prolonged negativity with a linear effect with most negative amplitudes for anomalous sentences followed by novel metaphors with animacy violations, metaphors without animacy violations and literal sentences. The sustained negativity has been previously observed in studies on novel metaphors (e.g., Arzouan et al. 2007; Regel et al. 2011; Goldstein et al. 2012; Rutter et al. 2012; Rataj et al. 2018; Jankowiak et al. 2021) and interpreted as reflecting increased demands in memory resulting from the mapping processes of two distinct concepts or task difficulty. In studies on animacy violation more robust LPC responses were observed for the animacy violation condition (e.g., Kuperberg et al. 2003, 2007; Paczynski and Kuperberg 2011; Szewczyk and Schriefers 2011), but studies on novel metaphors with animacy violations show less consistent results. Ji et al. (2020) observed attenuated LPC responses to metaphors with animacy violations, which was interpreted as an indication that animacy violation might facilitate late novel metaphor meaning integration. Whereas Li et al. (2022) reported increased LPC amplitudes for metaphors with animacy violations (at the verb processing stage) pointing to meaning reanalysis in the direction of abstract interpretation. The sustained negativity observed in the present study has not yet been reported in research on animacy violations in novel metaphor comprehension. The disappearance of significant differences between novel metaphors with and without animacy violations may suggest effective meaning construction for novel metaphors with animacy violations.

Interestingly, both the N400 and sustained negativity effects were observed over the anterior sites (frontal and fronto-central). This topography has been previously noted for sustained negativity in studies on novel metaphors (e.g., Coulson and Van Petten 2002; Bambini et al. 2016; Rataj et al. 2018). However, the N400 effect has most frequently been reported over the centro-parietal sites (Lai et al. 2009; Arzouan et al. 2007; Rataj et al. 2018; Jankowiak et al. 2021). The frontal topography for the N400 effect was unexpected, but has been previously found in studies on novel metaphors by Bambini et al. (2018) and Menashe et al. (2024). Additionally, frontally distributed N400 has been linked to increased working memory demands (e.g., Kiss et al. 2007; Boudewyn et al. 2019). In the present study, the frontally distributed negativity in the N400 time window has also been interpreted as a reflection of an increased memory load caused by information updating and greater semantic integration difficulty demanded in the delayed response semantic decision task with the Rapid Serial Visual Presentation (RSVP) technique. Moreover, animacy violation appears to affect the memory processes. As proposed by animate monitoring hypothesis (New et al. 2007), animates are more readily expected and processed than inanimates. In novel metaphors with animacy violations this expectancy needs to be over-ridden to derive meaning, thus increasing memory processing demands.

This study makes valuable contributions to the body of research on novel metaphoric language. It provides evidence for the important role animacy violation plays in metaphor processing and comprehension. Further research is needed to test whether similar findings can be replicated across different levels of conventionality and procedures.

Abstract

Metaphors are not merely ornaments of language. They are important communication tools that allow us to convey and understand abstract meanings through mappings with concrete conceptual domains (Lakoff and Johnson 2011). Early theoretical accounts (Grice 1975) proposed that metaphor comprehension requires activation and rejection of the literal meaning. However, with the growing body of research on metaphor processing and comprehension (e.g., Coulson and Van Petten 2002; Arzouan et al. 2007; Lai et al. 2009; De Grauwe et al. 2010; Rutter et al. 2012; Bambini et al. 2016; Rataj et al. 2018; Jończyk et al. 2020; Jankowiak et al. 2021), approaches to metaphoric language have changed, emphasising the importance of factors that can significantly influence how metaphors are processed and understood.

One of the most important factors in metaphor processing is conventionality. Studies have shown that more creative and novel metaphors require increased cognitive demands in meaning processing and integration than conventional metaphors and literal sentences (e.g., Lai et al. 2009; Rataj et al. 2018; Jankowiak et al. 2021). Novel metaphors have been investigated extensively in recent years using electrophysiological measures (EEG) and have reported consistent results. However, relatively little attention has been devoted to the role of animacy violations (i.e., combination of an inanimate noun with a verb or adjective suggesting animacy e.g., *healthy attitude*) in novel metaphor processing and comprehension. Animacy is considered to be one of the fundamental features in cognition (Nairne et al. 2013), influencing both attention (Abrams and Christ 2003) and memory (Daley et al. 2020). While animacy violation has been investigated in literal language research (e.g., Kuperberg et al. 2003b; Szewczyk and Schriefers 2011), there still remains a gap in research on metaphors. To the best of my knowledge, only a few studies have examined the influence of animacy violation on processing and comprehension of novel metaphors (Ji et al. 2020; Li et al. 2022, 2025). Noteworthily, these studies were conducted in Mandarin Chinese and reported some inconsistent results.

To address this gap, this PhD dissertation aimed to investigate the role of animacy violations in novel metaphor processing and comprehension using electrophysiological measures (EEG). Data was collected in normative studies and an ERP (event-related potential) experiment. The aim of the normative studies was to develop materials for the four experimental conditions. The ERP study measured behavioural and neural responses

to novel metaphors with animacy violations, novel metaphors without animacy violations, literal sentences, and anomalous sentences in a binary semantic decision task. The obtained results revealed a significant effect of animacy on novel metaphor processing. Both behavioural and ERP data point to greater cognitive demands required to derive meaning for novel metaphors with animacy violations in comparison to literal sentences. Furthermore, novel metaphors with animacy violations have been observed to require greater working memory engagement in meaning integration. Taken together, the obtained results showed that animacy violation impacts semantic processing and cognitive demands in novel metaphor processing and comprehension.

Streszczenie

Metafory nie są tylko ozdobnikami języka. Stanowią ważne narzędzia komunikacji, które pozwalają nam przekazywać i rozumieć abstrakcyjne znaczenia poprzez odwzorowania w konkretnych domenach pojęciowych (Lakoff i Johnson, 2011). Wczesne ujęcia teoretyczne (Grice, 1975) zakładały, że zrozumienie metafory wymaga aktywacji, a następnie odrzucenia znaczenia dosłownego. Jednak wraz z rozwojem badań nad przetwarzaniem i rozumieniem metafor (np. Coulson i Van Petten 2002; Arzouan i in. 2007; Lai i in. 2009; De Grauwe i in. 2010; Rutter i in. 2012; Bambini i in. 2016; Rataj i in. 2018; Jończyk i in. 2020; Jankowiak i in. 2021), podejścia do języka metaforycznego uległy zmianie, podkreślając znaczenie czynników, które mogą w istotny sposób wpływać na sposób przetwarzania i rozumienia metafor.

Jednym z najważniejszych czynników wpływających na przetwarzanie metafor jest konwencjonalność. Badania wykazały, że bardziej kreatywne i nowe metafory mogą wymagać większych zasobów poznawczych w zakresie przetwarzania i integracji znaczenia niż metafory konwencjonalne czy zdania dosłowne (np. Lai i in., 2009; Rataj i in., 2018; Jankowiak i in., 2021). Nowe metafory były w ostatnich latach szeroko badane z wykorzystaniem metod elektrofizjologicznych (EEG) i raportowano spójne wyniki. Jednak stosunkowo niewiele uwagi zostało poświęconej roli naruszeń żywotności (tj. połączenia rzeczownika nieożywionego z czasownikiem lub przymiotnikiem sugerującym żywotność, np. *zdrowe nastawienie*) w przetwarzaniu i rozumieniu nowych metafor.

Żywotność uznawana jest za jedną z fundamentalnych cech poznawczych (Nairne i in., 2013), wpływającą zarówno na uwagę (Abrams i Christ, 2003), jak i pamięć (Daley i in., 2020). Choć naruszenia żywotności były badane w lingwistyce (np. Kuperberg i in., 2003b; Szewczyk i Schriefers, 2011), wciąż istnieje luka badawcza dotycząca ich roli w nowych metaforach. Wedle mojego stanu wiedzy, jedynie kilka badań analizowało wpływ naruszeń żywotności na przetwarzanie i rozumienie nowych metafor (Ji i in., 2020; Li i in., 2022, 2025). Co istotne, badania te przeprowadzono w języku mandaryńskim i odnotowano w nich pewne niespójne wyniki.

Aby wypełnić tę lukę badawczą, niniejsza rozprawa doktorska miała na celu zbadanie roli naruszeń żywotności w przetwarzaniu i rozumieniu nowych metafor z wykorzystaniem metod elektrofizjologicznych (EEG). Dane zostały zebrane w badaniach normatywnych oraz w eksperymencie ERP (potencjałów wywołanych). Celem badań

normatywnych było opracowanie materiałów dla czterech warunków eksperymentalnych. W badaniu ERP mierzono reakcje behawioralne i neuronalne na cztery typy bodźców: nowe metafory z naruszeniami żywotności, nowe metafory bez naruszeń żywotności, zdania dosłowne oraz zdania bezsensowne, podczas gdy uczestnicy badania podejmowali decyzję semantyczną.

Uzyskane wyniki ujawniły istotny wpływ żywotności na przetwarzanie nowych metafor. Zarówno dane behawioralne, jak i potencjały wywołane wydarzeniem wskazują, że nowe metafory z naruszeniami żywotności wymagają większych zasobów poznawczych niż zdania dosłowne. Ponadto zaobserwowano, że nowe metafory z naruszeniem żywotności angażują w większym stopniu pamięć roboczą podczas integracji znaczenia. Podsumowując, uzyskane wyniki pokazują, że naruszenia żywotności wpływają na przetwarzanie semantyczne oraz zwiększają obciążenia poznawcze w rozumieniu nowych metafor.

References

Abrams, Richard A. and Shawn E. Christ. 2003. “Motion Onset Captures Attention”, *Psychological Science* 14, 5: 427-432.

Abraham, Anna, Barbara Rutter and Christiane Hermann. 2021. “Conceptual expansion via novel metaphor processing: An ERP replication and extension study examining individual differences in creativity”, *Brain and Language* 221: 105007.

Aissen, Judith. 2003. “Differential Object Marking: Iconicity vs. Economy”, *Natural Language & Linguistic Theory* 21, 3: 435-483.

Ariel, Mira. 2002. “The demise of a unique concept of literal meaning”, *Journal of Pragmatics* 34, 4: 361-402.

Aristotle. 1932. *Aristotle in 23 Volumes*. London: Harvard University Press. (http://data.perseus.org/citations/urn:cts:greekLit:tlg0086.tlg034.perseus-eng1:1459a) (date of access: 1 Aug. 2025).

Arzouan, Yossi, Abraham Goldstein and Miriam Faust. 2007. “Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension”, *Brain Research* 1160: 69-81.

Baddeley, Alan. 2012. “Working Memory: Theories, Models, and Controversies”, *Annual Review of Psychology* 63, Volume 63, 2012: 1-29.

Balota, David A, Gregory C Burgess, Michael J Cortese and David R Adams. 2002. “The Word-Frequency Mirror Effect in Young, Old, and Early-Stage Alzheimer’s Disease: Evidence for Two Processes in Episodic Recognition Performance”, *Journal of Memory and Language* 46, 1: 199-226.

Balota, David A., Michael J. Cortese, Susan D. Sergent-Marshall, Daniel H. Spieler and MelvinJ Yap. 2004. “Visual word recognition of single-syllable words”, *Journal of Experimental Psychology. General* 133, 2: 283-316.

Balota, David A., Janet M. Duchek, Susan D. Sergent-Marshall and Henry L. Roediger. 2006a. “Does expanded retrieval produce benefits over equal-interval spacing? Explorations of spacing effects in healthy aging and early stage Alzheimer’s disease.”, *Psychology and Aging* 21, 1: 19-31.

Balota, David A., Melvin J. Yap and Michael J. Cortese. 2006b. “Visual Word Recognition”, in: *Handbook of Psycholinguistics*. Elsevier, 285-375. (https://linkinghub.elsevier.com/retrieve/pii/B9780123693747500109) (date of access: 17 Aug. 2025).

Bambini, Valentina, Chiara Bertini, Walter Schaeken, Alessandra Stella and Francesco Di Russo. 2016. “Disentangling Metaphor from Context: An ERP Study”, *Frontiers in Psychology* 7. (http://journal.frontiersin.org/Article/10.3389/fpsyg.2016.00559/abstract) (date of access: 27 Dec. 2024).

Bambini, Valentina, Paolo Canal, Donatella Resta and Mirko Grimaldi. 2018. "Time Course and Neurophysiological Underpinnings of Metaphor in Literary Context", *Discourse Processes* 56, 1: 77-97.

Barlow, J.M., J.R. Kerlin and H.R. Pollio. 1971. *Training Manual for Identifying Figurative Language*. University of Tennessee, Department of Psychology. (https://books.google.pl/books?id=o6fDmgEACAAJ).

Bayanati, Shiva and Ida Toivonen. 2019. "Humans, Animals, Things and Animacy", *Open Linguistics* 5, 1: 156-170.

Beres, Anna M. 2017. "Time is of the Essence: A Review of Electroencephalography (EEG) and Event-Related Brain Potentials (ERPs) in Language Research", *Applied Psychophysiology and Biofeedback* 42, 4: 247-255.

Berkum, Jos J. A. Van, Peter Hagoort and Colin M. Brown. 1999. "Semantic Integration in Sentences and Discourse: Evidence from the N400", *Journal of Cognitive Neuroscience* 11, 6: 657-671.

Birnbaum, Michael H. 1999. "How to show that 9 > 221: Collect judgments in a between-subjects design", *Psychological Methods* 4, 3: 243-249.

Blasko, Dawn G. and Cynthia M. Connine. 1993. "Effects of familiarity and aptness on metaphor processing.", *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19, 2: 295-308.

Bonin, Patrick, Margaux Gelin and Aurélia Bugaiska. 2014. "Animates are better remembered than inanimates: further evidence from word and picture stimuli", *Memory & Cognition* 42, 3: 370-382.

Bonin, Patrick, Margaux Gelin, Betty Laroche, Alain Méot and Aurélia Bugaiska. 2015. "The 'How' of Animacy Effects in Episodic Memory", *Experimental Psychology* 62, 6: 371-384.

Botwinick, J. 1966. "Cautiousness in Advanced Age", *Journal of Gerontology* 21, 3: 347-353.

Boudewyn, Megan A., Adam R. Blalock, Debra L. Long and Tamara Y. Swaab. 2019. "Adaptation to Animacy Violations during Listening Comprehension", *Cognitive, affective & behavioral neuroscience* 19, 5: 1247-1258.

Boudewyn, Megan Ann, Debra L. Long and Tamara Y. Swaab. 2013. "Effects of Working Memory Span on Processing of Lexical Associations and Congruence in Spoken Discourse", *Frontiers in Psychology* 4. (https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2013.00060/full) (date of access: 13 Sept. 2025).

Bowdle, Brian F. and Dedre Gentner. 2005. "The Career of Metaphor.", *Psychological Review* 112, 1: 193-216.

BrainVision Analyzer. Gilching, Germany.

Bridger, Emma K., Regine Bader, Olga Kriukova, Kerstin Unger and Axel Mecklinger. 2012. “The FN400 is functionally distinct from the N400”, *NeuroImage* 63, 3: 1334-1342.

Brouwer, Harm, Hartmut Fitz and John Hoeks. 2012. “Getting real about Semantic Illusions: Rethinking the functional role of the P600 in language comprehension”, *Brain Research* 1446: 127-143.

Burkhardt, Petra. 2006. “Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials”, *Brain and Language* 98, 2: 159-168.

Burkhardt, Petra. 2007. “The P600 reflects cost of new information in discourse memory”, *NeuroReport* 18, 17: 1851-1854.

Chiappe, Dan L. and Penny Chiappe. 2007. “The role of working memory in metaphor production and comprehension”, *Journal of Memory and Language* 56, 2: 172-188.

Cieślicka, Anna and Roberto Heredia. 2019. “ERP Correlates of Figurative Language Processing”, in: *Approaches to the Study of Sound Structure and Speech*. Routledge, 244-257.

Cohen, Mak S. 2008. “Handedness Questionnaire” (https://www.brainmapping.org/shared/Edinburgh.php).

Cohen, Mike X. 2015. *Analyzing neural time series data: theory and practice*. Cambridge, Mass. [u.a]: MIT Press.

Coulson, Seana and Cyma Van Petten. 2002. “Conceptual integration and metaphor: An event-related potential study”, *Memory & Cognition* 30, 6: 958-968.

Cui, Ruifang, Yang Liu and Changquan Long. 2018. “FN400 and sustained negativity reveal a premise monotonicity effect during semantic category-based induction”, *International Journal of Psychophysiology* 134: 108-119.

Czypionka, Anna, Mariya Kharaman and Carsten Eulitz. 2023. “Wolf-hound vs. sled-dog: neurolinguistic evidence for semantic decomposition in the recognition of German noun-noun compounds”, *Frontiers in Psychology* 14: 1173352.

Dahl, Östen. 2000. “Animacy and the notion of semantic gender”, in: Barbara Unterbeck, Matti Rissanen, Terttu Nevalainen and Mirja Saari (eds.), *Gender in Grammar and Cognition*. De Gruyter Mouton, 99-116. (https://www.degruyterbrill.com/document/doi/10.1515/9783110802603.99/html) (date of access: 6 July. 2025).

Daley, Matthew J., Glenda Andrews and Karen Murphy. 2020. “Animacy effects extend to working memory: results from serial order recall tasks”, *Memory* 28, 2: 157-171.

De Grauwe, Sophie, Abigail Swain, Phillip J. Holcomb, Tali Ditman and Gina R. Kuperberg. 2010. "Electrophysiological insights into the processing of nominal metaphors", *Neuropsychologia* 48, 7: 1965-1984.

Delogu, Francesca, Harm Brouwer and Matthew W. Crocker. 2019. "Event-related potentials index lexical retrieval (N400) and integration (P600) during language comprehension", *Brain and Cognition* 135: 103569.

Der, Geoff and Ian J. Deary. 2006. "Age and sex differences in reaction time in adulthood: Results from the United Kingdom Health and Lifestyle Survey.", *Psychology and Aging* 21, 1: 62-73.

DeYoung, Carlee M. and Michael J. Serra. 2021. "Judgments of learning reflect the Animacy advantage for memory, but not beliefs about the effect", *Metacognition and Learning* 16, 3: 711-747.

Donders, F.C. 1969. "On the speed of mental processes", *Acta Psychologica* 30: 412-431.

Dörnyei, Zoltán and Jean-Marc Dewaele. 2022. *Questionnaires in Second Language Research: Construction, Administration, and Processing*. (3rd edition.) New York: Routledge. (https://www.taylorfrancis.com/books/9781003331926) (date of access: 27 Dec. 2024).

Eddine, Samer Nour, Trevor Brothers and Gina R. Kuperberg. 2022. "The N400 in silico: A review of computational models", in: *Psychology of Learning and Motivation*. Elsevier, 123-206. (https://linkinghub.elsevier.com/retrieve/pii/S0079742122000056) (date of access: 9 Sept. 2025).

Fauconnier, Gilles and Mark Turner. 2003. "Conceptual Blending, Form and Meaning", *Recherches en Communication* 19. (https://ojs.uclouvain.be/index.php/rec/article/view/48413) (date of access: 14 Sept. 2025).

Félix, Sara B., Marie Poirier, James S. Nairne and Josefa N. S. Pandeirada. 2024. "The breadth of animacy in memory: New evidence from prospective memory", *Psychonomic Bulletin & Review* 31, 3: 1323-1334.

Frisch, Stefan, Matthias Schlesewsky, Douglas Saddy and Annegret Alpermann. 2002. "The P600 as an indicator of syntactic ambiguity", *Cognition* 85, 3: B83-B92.

Frisson, Steven and Martin Pickering. 2001. "Obtaining a Figurative Interpretation of a Word: Support for Underspecification", *Metaphor and Symbol* 16, 3: 149-171.

Gelman, Rochel and Elizabeth Spelke. 1981. "The development of thoughts about animate and inanimate objects: implications for research on social", in: *Social Cognitive Development: Frontiers and possible futures*. Cambridge University Press, 43-66.

Genovesi, Chris. 2023. "Grice's Café: coffee, cream, and metaphor comprehension", *Frontiers in Communication* 8: 1175587.

Gentner, D. 1983. “Structure-mapping: A theoretical framework for analogy”, *Cognitive Science* 7, 2: 155-170.

Gibbs Jr., Raymond W. 2002. “A new look at literal meaning in understanding what is said and implicated”, *Journal of Pragmatics* 34, 4: 457-486.

Gibbs, Jr, Raymond W. 2017. *Metaphor Wars: Conceptual Metaphors in Human Life*. (1st edition.) Cambridge University Press. (https://www.cambridge.org/core/product/identifier/9781107762350/type/book) (date of access: 4 Feb. 2025).

Gibbs, Raymond W. 1980. “Spilling the beans on understanding and memory for idioms in conversation”, *Memory & Cognition* 8, 2: 149-156.

Gibbs, Raymond W. 2015. “Does deliberate metaphor theory have a future?”, *Journal of Pragmatics* 90: 73-76.

Gibbs, Raymond W., Darin L. Buchalter, Jessica F. Moise and William T. Farrar. 1993. “Literal meaning and figurative language”, *Discourse Processes* 16, 4: 387-403.

Gibbs, Raymond W. and Herbert L. Colston. 2012. *Interpreting figurative meaning*. Cambridge ; New York: Cambridge University Press.

Giora, Rachel. 1997. “Understanding figurative and literal language: The graded salience hypothesis”, *Cognitive Linguistics* 8, 3: 183-206.

Goldstein, Abraham, Yossi Arzouan and Miriam Faust. 2012. “Killing a novel metaphor and reviving a dead one: ERP correlates of metaphor conventionalization”, *Brain and Language* 123, 2: 137-142.

Gouvea, Ana C., Colin Phillips, Nina Kazanina and David Poeppel. 2010. “The linguistic processes underlying the P600”, *Language and Cognitive Processes* 25, 2: 149-188.

Grice, H. P. 1975. “Logic and Conversation”, in: Peter Cole and Jerry L. Morgan (eds.), *Speech Acts*. BRILL, 41-58. (https://brill.com/view/book/edcoll/9789004368811/BP000003.xml) (date of access: 4 Feb. 2025).

Grigore, Vasilica, Georgeta Mitrache, Mihaela Paunescu and Radu Predoiu. 2015. “The Decision Time, the Simple and the Discrimination Reaction Time in Elite Romanian Junior Tennis Players”, *Procedia - Social and Behavioral Sciences* 190: 539-544.

Hagoort, Peter, Giosuè Baggio and Roel M. Willems. 2009. “Semantic Unification”, in: Michael S. Gazzaniga (ed.), *The Cognitive Neurosciences*. The MIT Press. (https://direct.mit.edu/books/book/5453/chapter/3965003/Semantic-Unification) (date of access: 4 Feb. 2025).

Hagoort, Peter and Jos Van Berkum. 2007. “Beyond the sentence given”, *Philosophical Transactions of the Royal Society B: Biological Sciences* 362, 1481: 801-811.

Handl, Sandra. 2011. *The conventionality of figurative language: a usage-based study*. Tübingen: Narr.

Harald Baayen, R. and Petar Milin. 2010. “Analyzing reaction times”, *International Journal of Psychological Research* 3, 2: 12-28.

Heitz, Richard P. 2014. “The speed-accuracy tradeoff: history, physiology, methodology, and behavior”, *Frontiers in Neuroscience* 8. (http://journal.frontiersin.org/article/10.3389/fnins.2014.00150/abstract) (date of access: 27 Dec. 2024).

Holyoak, Keith J. and Dušan Stamenković. 2018. “Metaphor comprehension: A critical review of theories and evidence.”, *Psychological Bulletin* 144, 6: 641-671.

Jankowiak, Katarzyna. 2021. “Current trends in electrophysiological research on bilingual language processing”, *Language and Linguistics Compass* 15, 8: e12436.

Jankowiak, Katarzyna and Anna B. Cieślicka. 2025. “Cross-language priming effects in bilingual novel metaphor processing”, *Scientific Reports* 15, 1: 30971.

Jankowiak, Katarzyna, Marcin Naranowicz and Karolina Rataj. 2021. “Metaphors are like lenses: Electrophysiological correlates of novel meaning processing in bilingualism”, *International Journal of Bilingualism* 25, 3: 668-686.

Jankowiak, Katarzyna, Karolina Rataj and Ryszard Naskręcki. 2017. “To electrify bilingualism: Electrophysiological insights into bilingual metaphor comprehension”, *PLOS ONE* 12, 4: e0175578.

Ji, Haojie, Senqing Qi, Shiyang Xu, Jinxin Chen, David Yun Dai, Yadan Li and Weiping Hu. 2020. “The role of animacy in metaphor processing of Mandarin Chinese: An Event-Related Potential (ERP) study”, *Journal of Neurolinguistics* 56: 100915.

Jończyk, Rafal, Gül E. Kremer, Zahed Siddique and Janet G. van Hell. 2020. “Engineering creativity: Prior experience modulates electrophysiological responses to novel metaphors”, *Psychophysiology* 57, 10: e13630.

Jones, Lara L. and Zachary Estes. 2006. “Roosters, robins, and alarm clocks: Aptness and conventionality in metaphor comprehension”, *Journal of Memory and Language* 55, 1: 18-32.

Jones, Lara L. and Zachary Estes. 2012. “Lexical priming: Associative, semantic, and thematic influences on word recognition”, in: James Adelman (ed.), *Visual word recognition: Meaning and context, individuals and development, Vol. 2*. Current issues in the psychology of language, New York, NY, US: Psychology Press, 44-72.

Kaan, Edith, Anthony Harris, Edward Gibson and Phillip Holcomb. 2000. “The P600 as an index of syntactic integration difficulty”, *Language and Cognitive Processes* 15, 2: 159-201.

Kaiser, Elsi. 2013. “Experimental paradigms in psycholinguistics”, in: *Research Methods in Linguistics*. Cambridge University Press, 135-168.

Kappenman, Emily S. and Steven J. Luck (eds.). 2011. *The Oxford Handbook of Event-Related Potential Components*. (1st edition.) Oxford University Press. (https://academic.oup.com/edited-volume/34558) (date of access: 27 Dec. 2024).

Katsimpokis, Dimitris, Guy E. Hawkins and Leendert van Maanen. 2020. “Not all Speed-Accuracy Trade-Off Manipulations Have the Same Psychological Effect”, *Computational Brain & Behavior* 3, 3: 252-268.

Katz, Albert N. and Todd Ferretti. 2001. “Moment-By-Moment Reading of Proverbs in Literal and Nonliteral Contexts”, *Metaphor and Symbol* 16, 3 & 4: 193-221.

Katz, Albert N., Allan Paivio, Marc Marschark and James M. Clark. 1988. “Norms for 204 Literary and 260 Nonliterary Metaphors on 10 Psychological Dimensions”, *Metaphor and Symbolic Activity* 3, 4: 191-214.

Kemmerer, David. 2015. *Cognitive neuroscience of language.* New York, NY, US: Psychology Press.

Kiss, Ivan, Scott Watter, Jennifer J. Heisz and Judith M. Shedden. 2007. “Control processes in verbal working memory: An event-related potential study”, *Brain Research* 1172: 67-81.

Kounios, John and Phillip J. Holcomb. 1994. “Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory”, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20, 4: 804-823.

Kövecses, Zoltán. 2017. “Conceptual metaphor theory”, in: *Routledge Handbook of Metaphor*. Routledge. (https://www.researchgate.net/publication/311403391_Conceptual_metaphor_theory) (date of access: 3 Sept. 2025).

Kreuz, R. J., R. M. Roberts, B. K. Johnson and E. L. Bertus. 1996. “Figurative language occurrence and co-occurrence in contemporary literature”, in: R. J. Kreuz and M. S. MacNealy (eds.), *Empirical approaches to literature and aesthetics*. Ablex, 83-97.

Kuperberg, Gina R., Donna A. Kreher, Tatiana Sitnikova, David N. Caplan and Phillip J. Holcomb. 2007. “The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials”, *Brain and Language* 100, 3: 223-237.

Kuperberg, Gina R, Tatiana Sitnikova, David Caplan and Phillip J Holcomb. 2003a. “Electrophysiological distinctions in processing conceptual relationships within simple sentences”, *Cognitive Brain Research* 17, 1: 117-129.

Kuperberg, Gina R, Tatiana Sitnikova, David Caplan and Phillip J Holcomb. 2003b. “Electrophysiological distinctions in processing conceptual relationships within simple sentences”, *Cognitive Brain Research* 17, 1: 117-129.

Kutas, M., H. J. Neville and P. J. Holcomb. 1987. “A preliminary comparison of the N400 response to semantic anomalies during reading, listening and signing”, 39: 325-330.

Kutas, Marta and Kara D. Federmeier. 2011. “Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP)”, *Annual Review of Psychology* 62, 1: 621-647.

Kutas, Marta and Steven A. Hillyard. 1980. “Reading senseless sentences: Brain potentials reflect semantic incongruity.”, *Science* 207, 4427: 203-205.

Kutas, Marta and Steven A. Hillyard. 1984. “Brain potentials during reading reflect word expectancy and semantic association”, *Nature* 307, 5947: 161-163.

Lai, Vicky Tzuyin, Tim Curran and Lise Menn. 2009. “Comprehending conventional and novel metaphors: An ERP study”, *Brain Research* 1284: 145-155.

Lakoff, George and Mark Johnson. 1980. “The Metaphorical Structure of the Human Conceptual System”, *Cognitive Science* 4, 2: 195-208.

Lakoff, George and Mark Johnson. 2011. *Metaphors we live by: with a new afterword*. (6. print.) Chicago, Ill.: Univ. of Chicago Press.

Lakoff, George and Mark Turner. 1989. *More than cool reason: a field guide to poetic metaphor*. Chicago: University of Chicago Press.

Laurencelle, Louis and Denis Cousineau. 2023. “Analysis of proportions using arcsine transform with any experimental design”, *Frontiers in Psychology* 13. (https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.1045436/full) (date of access: 25 Sept. 2025).

Levelt, W. J., A. Roelofs and A. S. Meyer. 1999. “A theory of lexical access in speech production”, *The Behavioral and Brain Sciences* 22, 1: 1-38; discussion 38-75.

Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA, US: The MIT Press.

Levelt, Willem J.M. 1999. “Models of word production”, *Trends in Cognitive Sciences* 3, 6: 223-232.

Leynes, P. Andrew, Yashvi Verma and Alexandra Santos. 2024. “Separating the FN400 and N400 event-related potential components in masked word priming”, *Brain and Cognition* 182: 106226.

Li, Haoze, Nan Cui, Meng Jiang and Zuxuan Qin. 2025. “The impact of noun animacy on the processing of L1 and L2 action metaphors among Chinese English learners: An event-related potential study”, *Medicine* 104, 26: e43036.

Li, Ping, Xiaoyu Jia, Xinyu Li and Weijian Li. 2016. “The effect of animacy on metamemory”, *Memory & Cognition* 44, 5: 696-705.

Li, Ying, Xiaoxiao Lu, Yizhen Wang, Hanlin Wang and Yue Wang. 2022. "Is the Processing of Chinese Verbal Metaphors Simulated or Abstracted? Evidence From an ERP Study", *Frontiers in Psychology* 13: 877997.

Lotze, Heike K., Marta Coll, Anna M. Magera, Christine Ward-Paige and Laura Airoldi. 2011. "Recovery of marine animal populations and ecosystems", *Trends in Ecology & Evolution* 26, 11: 595-605.

Loucks, Jeff, Berit Reise, Rosselle Gahite and Shaun Fleming. 2023. "Animate monitoring is not uniform: implications for the animate monitoring hypothesis", *Frontiers in Psychology* 14: 1146248.

Lowder, Matthew W. and Peter C. Gordon. 2015. "Natural forces as agents: Reconceptualizing the animate–inanimate distinction", *Cognition* 136: 85-90.

Luce, R. D. 1986. "Response Times", *New York: University Press*.

Luck, Steven J. 2014. *An introduction to the event-related potential technique*. (2nd ed.) Cambridge, Mass.: MIT Press.

Luck, Steven J. and Nicholas Gaspelin. 2017. "How to get statistically significant effects in any ERP experiment (and why you shouldn't)", *Psychophysiology* 54, 1: 146-157.

Mandera, Paweł, Emmanuel Keuleers, Zofia Wodniecka and Marc Brysbaert. 2015. "Subtlex-pl: subtitle-based word frequency estimates for Polish", *Behavior Research Methods* 47, 2: 471-483.

Mashal, Nira. 2013. "The role of working memory in the comprehension of unfamiliar and familiar metaphors", *Language and Cognition* 5, 4: 409-436.

Mcelree, Brian and Johanna Nordlie. 1999. "Literal and figurative interpretations are computed in equal time", *Psychonomic Bulletin & Review* 6, 3: 486-494.

McNamara, Timothy P. 2005. *Semantic Priming*. (0 edition.) Psychology Press. (https://www.taylorfrancis.com/books/9781135432553) (date of access: 4 Feb. 2025).

Mecklinger, Axel and Regine Bader. 2020. "From fluency to recognition decisions: A broader view of familiarity-based remembering", *Neuropsychologia* 146: 107527.

Menashe, Shay, Nira Mashal and David Anaki. 2024. "N400 modulations in metaphor evaluation and its associations with attentional systems: A behavioral and ERP study", *Journal of Neurolinguistics* 72: 101203.

Miller, Adriana R., Rafał Jończyk, Holly A. Zaharchuk and Janet G. van Hell. 2025. "Unlocking Second Language Novel Metaphor Processing: Behavioral and ERP Insights From First and Second-Language English Users", *Psychophysiology* 62, 5: e70066.

Nairne, James S. 2010. “Adaptive Memory”, in: *Psychology of Learning and Motivation*. Elsevier, 1-32. (https://linkinghub.elsevier.com/retrieve/pii/S0079742110530019) (date of access: 4 June. 2025).

Nairne, James S., Joshua E. VanArsdall, Josefa N. S. Pandeirada, Mindi Cogdill and James M. LeBreton. 2013. “Adaptive Memory: The Mnemonic Value of Animacy”, *Psychological Science* 24, 10: 2099-2105.

Naranowicz, Marcin, Katarzyna Jankowiak and Maciej Behnke. 2022a. “Native and non-native language contexts differently modulate mood-driven electrodermal activity”, *Scientific Reports* 12, 1: 22361.

Naranowicz, Marcin, Katarzyna Jankowiak, Patrycja Kakuba, Katarzyna Bromberek-Dyzman and Guillaume Thierry. 2022b. “In a Bilingual Mood: Mood Affects Lexico-Semantic Processing Differently in Native and Non-Native Languages”, *Brain Sciences* 12, 3: 316.

New, Joshua, Leda Cosmides and John Tooby. 2007. “Category-specific attention for animals reflects ancestral priorities, not expertise”, *Proceedings of the National Academy of Sciences* 104, 42: 16598-16603.

Nieuwland, Mante S. and Jos J. A. Van Berkum. 2006. “When Peanuts Fall in Love: N400 Evidence for the Power of Discourse”, *Journal of Cognitive Neuroscience* 18, 7: 1098-1111.

Oldfield, R. C. 1971. “The assessment and analysis of handedness: the Edinburgh inventory”, *Neuropsychologia* 9, 1: 97-113.

Oppenheim, Abraham N. 1992. *Questionnaire design, interviewing and attitude measurement*. (New ed., reprinted.) London: Continuum.

Ortony, Andrew, Diane L. Schallert, Ralph E. Reynolds and Stephen J. Antos. 1978. “Interpreting metaphors and idioms: Some effects of context on comprehension”, *Journal of Verbal Learning and Verbal Behavior* 17, 4: 465-477.

Osterhout, Lee and Phillip J Holcomb. 1992. “Event-related brain potentials elicited by syntactic anomaly”, *Journal of Memory and Language* 31, 6: 785-806.

Osterhout, Lee and Linda A. Mobley. 1995. “Event-Related Brain Potentials Elicited by Failure to Agree”, *Journal of Memory and Language* 34, 6: 739-773.

Paczynski, Martin and Gina R. Kuperberg. 2011. “Electrophysiological evidence for use of the animacy hierarchy, but not thematic role assignment, during verb-argument processing”, *Language and Cognitive Processes* 26, 9: 1402-1456.

Petrun, Craig J and Susan M Belmore. 1981. “The On-Line Processing of Figurative and Literal”

Pollio, Howard R., Michael S. Fabrizi, Abigail Sills and Michael K. Smith. 1984. “Need metaphoric comprehension take longer than literal comprehension?”, *Journal of Psycholinguistic Research* 13, 3: 195-214.

Popp, Earl Y. and Michael J. Serra. 2018. "The animacy advantage for free-recall performance is not attributable to greater mental arousal", *Memory* 26, 1: 89-95.

Poulin-Dubois, Diane, Anouk Lepage and Doreen Ferland. 1996. "Infants' concept of animacy", *Cognitive Development* 11, 1: 19-36.

Pratt, Jay, Petre V. Radulescu, Ruo Mu Guo and Richard A. Abrams. 2010. "It's Alive!: Animate Motion Captures Visual Attention", *Psychological Science* 21, 11: 1724-1730.

Pynte, Joel, Mireille Besson, Fabrice-Henri Robichon and Jézabel Poli. 1996. "The Time-Course of Metaphor Comprehension: An Event-Related Potential Study", *Brain and Language* 55, 3: 293-316.

Qualtrics. 2005.

Rataj, Karolina. 2014. "Surfing the brainwaves of metaphor comprehension", *Poznan Studies in Contemporary Linguistics* 50, 1: 55-73.

Rataj, Karolina, Patrycja Kakuba and Walter J. B. van Heuven. 2025. "Overcoming expectations: Electrophysiological effects of inhibitory control demands in a non-linguistic task on subsequent semantic relatedness judgements", *Language, Cognition and Neuroscience* 40, 3: 377-395.

Rataj, Karolina, Patrycja Kakuba, Paweł Mandera and Walter J. B. Van Heuven. 2023. "Establishing semantic relatedness through ratings, reaction times, and semantic vectors: A database in Polish", *PLOS ONE* 18, 4: e0284801.

Rataj, Karolina, Deniece S. Nazareth and Frank van der Velde. 2018a. "Use a Spoon as a Spade?: Changes in the Upper and Lower Alpha Bands in Evaluating Alternate Object Use", *Frontiers in Psychology* 9. (https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2018.01941/full) (date of access: 22 Sept. 2025).

Rataj, Karolina, Anna Przekoracka-Krawczyk and Rob H.J. Van Der Lubbe. 2018b. "On understanding creative language: The late positive complex and novel metaphor comprehension", *Brain Research* 1678: 231-244.

Redfern, M. S., M. L.T.M. Muller, J. R. Jennings and J. M. Furman. 2002. "Attentional Dynamics in Postural Control During Perturbations in Young and Older Adults", *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 57, 8: B298-B303.

Regel, Stefanie, Thomas C. Gunter and Angela D. Friederici. 2011. "Isn't It Ironic? An Electrophysiological Exploration of Figurative Language Processing", *Journal of Cognitive Neuroscience* 23, 2: 277-293.

Rutter, Barbara, Sören Kröger, Holger Hill, Sabine Windmann, Christiane Hermann and Anna Abraham. 2012. "Can clouds dance? Part 2: An ERP investigation of passive conceptual expansion", *Brain and Cognition* 80, 3: 301-310.

Sana, Faria, Juana Park, Christina L. Gagné and Thomas L. Spalding. 2021. "The interplay between inhibitory control and metaphor conventionality", *Memory & Cognition* 49, 6: 1267-1284.

Schumacher, Petra B. 2011. "The hepatitis called …: Electrophysiological evidence for enriched composition", in: Jörg Meibauer and Markus Steinbach (eds.), *Linguistik Aktuell/Linguistics Today*. Amsterdam: John Benjamins Publishing Company, 199-219. (https://benjamins.com/catalog/la.175.10sch) (date of access: 4 Feb. 2025).

Silverstein, Michael. 1986. "7. Hierarchy of Features and Ergativity", in: *Features and Projections*. De Gruyter, 163-232. (https://www.degruyterbrill.com/document/doi/10.1515/9783110871661-008/html) (date of access: 6 July. 2025).

Silvia, Paul J. and Roger E. Beaty. 2012. "Making creative metaphors: The importance of fluid intelligence for creative thought", *Intelligence* 40, 4: 343-351.

Singleton, W. T. 1953. "Deterioration of performance on a short-term perceptual-motor task.", in: *Symposium on fatigue.* Oxford, England: H. K. Lewis & Co., 163-172.

Sperber, Dan and Deirdre Wilson. 1995. *Relevance: Communication and cognition, 2nd ed.* Malden: Blackwell Publishing.

Spotorno, Nicola, Anne Cheylus, Jean-Baptiste Van Der Henst and Ira A. Noveck. 2013. "What's behind a P600? Integration Operations during Irony Processing", *PLoS ONE* 8, 6: e66839.

Steen, Gerard. 2014. "Deliberate Metaphor Affords Conscious Metaphorical Cognition", *Cognitive Semiotics* 5, 1-2: 179-197.

Steen, Gerard. 2015. "Developing, testing and interpreting Deliberate Metaphor Theory", *Journal of Pragmatics* 90: 67-72.

Steen, Gerard J., Aletta G. Dorst, J. Berenike Herrmann, Anna A. Kaal and Tina Krennmayr. 2010. "Metaphor in usage", *cogl* 21, 4: 765-796.

Stone, James L. and John R. Hughes. 2013. "Early History of Electroencephalography and Establishment of the American Clinical Neurophysiology Society", *Journal of Clinical Neurophysiology* 30, 1: 28-44.

Strijkers, Kristof and Albert Costa. 2011. "Riding the Lexical Speedway: A Critical Review on the Time Course of Lexical Selection in Speech Production", *Frontiers in Psychology* 2. (http://journal.frontiersin.org/article/10.3389/fpsyg.2011.00356/abstract) (date of access: 4 Feb. 2025).

Stróżak, Paweł, Christopher W. Bird, Krystin Corby, Gwen Frishkoff and Tim Curran. 2016. "FN400 and LPC memory effects for concrete and abstract words", *Psychophysiology* 53, 11: 1669-1678.

Szewczyk, Jakub M. and Herbert Schriefers. 2011. "Is animacy special?", *Brain Research* 1368: 208-221.

Tang, Xuemei, Man Fu, Xiao Wang, Yixin Yao and Lexian Shen. 2025. “Context-modulating effect on processing scientific metaphors: Evidence from ERPs”, *Brain and Language* 266: 105582.

Tang, Xuemei, Senqing Qi, Xiaojuan Jia, Botao Wang and Wei Ren. 2017. “Comprehension of scientific metaphors: Complementary processes revealed by ERP”, *Journal of Neurolinguistics* 42: 12-22.

Thibodeau, Paul H. and Frank H. Durgin. 2011. “Metaphor Aptness and Conventionality: A Processing Fluency Account”, *Metaphor and Symbol* 26, 3: 206-226.

Träuble, Birgit, Sabina Pauen and Diane Poulin-Dubois. 2014. “Speed and direction changes induce the perception of animacy in 7-month-old infants”, *Frontiers in Psychology* 5. (https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2014.01141/full) (date of access: 15 May. 2025).

Traugott, Elizabeth Closs. 1985. “‘Conventional’ and ‘dead’ metaphors revisited”, in: Wolf Paprotté and René Dirven (eds.), *Current Issues in Linguistic Theory*. Amsterdam: John Benjamins Publishing Company, 17-56. (https://benjamins.com/catalog/cilt.29.04clo) (date of access: 4 Feb. 2025).

Trompenaars, Thijs, Lotte Hogeweg, Wessel Stoop and Helen de Hoop. 2018. “The language of an inanimate narrator”, *Open Linguistics* 4, 1: 707-721.

Trompenaars, Thijs, Teresa Angelina Kaluge, Rezvan Sarabi and Peter De Swart. 2021. “Cognitive animacy and its relation to linguistic animacy: evidence from Japanese and Persian”, *Language Sciences* 86: 101399.

Tudor, Mario, Lorainne Tudor and Katarina Ivana Tudor. 2005. “[Hans Berger (1873-1941)--the history of electroencephalography]”, *Acta Medica Croatica: Casopis Hravatske Akademije Medicinskih Znanosti* 59, 4: 307-313.

VanArsdall, Joshua E. and Janell R. Blunt. 2022. “Analyzing the structure of animacy: Exploring relationships among six new animacy and 15 existing normative dimensions for 1,200 concrete nouns”, *Memory & Cognition* 50, 5: 997-1012.

Vanrullen, Rufin. 2011. “Four Common Conceptual Fallacies in Mapping the Time Course of Recognition”, *Frontiers in Psychology* 2. (https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2011.00365/full) (date of access: 21 Sept. 2025).

Voss, Joel L. and Kara D. Federmeier. 2011. “FN400 potentials are functionally identical to N400 potentials and reflect semantic processing during recognition testing”, *Psychophysiology* 48, 4: 532-546.

Wang, Lin, Edward Wlotko, Edward Alexander, Lotte Schoot, Minjae Kim, Lena Warnke and Gina R. Kuperberg. 2020. “Neural Evidence for the Prediction of Animacy Features during Language Comprehension: Evidence from MEG and EEG Representational Similarity Analysis”, *The Journal of Neuroscience* 40, 16: 3278-3291.

Wang, Xin. 2022. “Normed dataset for novel metaphors, novel similes, literal and anomalous sentences in Chinese”, *Frontiers in Psychology* 13. (https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.922722/full) (date of access: 7 Sept. 2025).

Wang, Yufang, Jurriaan Witteman and Niels O Schiller. 2024. “The role of animacy in language production: evidence from bare noun naming”, *Quarterly Journal of Experimental Psychology* 17470218241281868.

Wilson, Deirdre. 2011. “PARALLELS AND DIFFERENCES IN THE TREATMENT OF METAPHOR 1IN RELEVANCE THEORY AND COGNITIVE LINGUISTICS”, *Studia Linguistica Universitatis Iagellonicae Cracoviensis* 128.

Woodruff, C. Chad, Hiroki R. Hayama and Michael D. Rugg. 2006. “Electrophysiological dissociation of the neural correlates of recollection and familiarity”, *Brain Research* 1100, 1: 125-135.

Yamamoto, Mutsumi. 1999. *Animacy and Reference: A cognitive approach to corpus linguistics*. Amsterdam: John Benjamins Publishing Company. (http://www.jbe-platform.com/content/books/9789027298768) (date of access: 6 July. 2025).

Yang, Fan-Pei Gloria, Kailyn Bradley, Madiha Huq, Dai-Lin Wu and Daniel C. Krawczyk. 2013. “Contextual effects on conceptual blending in metaphors: An event-related potential study”, *Journal of Neurolinguistics* 26, 2: 312-326.

Yang, Haopei, Geoffrey Laforge, Bobby Stojanoski, Emily S. Nichols, Ken McRae and Stefan Köhler. 2019. “Late positive complex in event-related potentials tracks memory signals when they are decision relevant”, *Scientific Reports* 9, 1: 9469.

Yu, Sarah S. and Michael D. Rugg. 2010. “Dissociation of the electrophysiological correlates of familiarity strength and item repetition”, *Brain Research* 1320: 74-84.

Appendix A: Experimental stimuli (in Polish)

No.	Context	Novel metaphor with animacy violation	No-animacy violation metaphor	Literal sentence	Anomalous sentence	Critical word
1	Zdenerwowało ją to	przekorne	burzliwe	kwaśne	giętkie	mleko.
2	Często niebezpieczna bywa	łakoma	zamszowa	skalna	uroczysta	szczelina.
3	Ona odwiedziła dzisiaj	przytomne	ciężkie	nudne	pszenne	miasto.
4	Mama lubiła to	milczące	kwaśne	rządowe	rozmieszane	zdjęcie.
5	Naukowców bardzo zaciekawiła	syta	papierowa	marna	radiowa	ziemia.
6	To jest kolejna	zmęczona	zatruta	podarta	parowa	książka.
7	Nagrodzone dziś zostało	łaskawe	pustynne	łatwe	kąpielowe	ciasto.
8	Zniszczyła się jej	nieśmiała	wieczna	zimowa	nadawcza	sukienka.
9	Miała przy sobie	anemiczny	mętny	nudny	złamany	scenariusz.
10	On spojrzał na	pijany	gorzki	antyczny	rozrzedzony	zegar.
11	Martwiła ją ta	uparta	bujna	zniszczona	potrzaskana	spódnica.
12	Na fotelu leżał	spragniony	ciasny	ważny	grzbietowy	rachunek.

No.	Context	Novel metaphor with animacy violation	No-animacy violation metaphor	Literal sentence	Anomalous sentence	Critical word
13	Szybko polubił to	mruczące	tytanowe	słodkie	poskładane	piwo.
14	Miał przy sobie	wierne	rześkie	papierowe	stanowe	pudełko.
15	Zmęczyły go wieczorem	zawzięte	wyrzeźbione	ochronne	informacyjne	spodnie.
16	Nagle pojawiła się	rozwydrzona	wełniana	biała	szklana	piana.
17	Rodzeństwo patrzyło na	dzielny	muzyczny	suchy	siłowy	trawnik.
18	Był to rzeczywiście	zatłoczony	pochmurny	miedziany	duchowy	garnek.
19	W dali słychać	nerwowy	religijny	stalowy	chrzestny	silnik.
20	W mieszkaniu zauważył	dumny	wytrawny	kolorowy	żywieniowy	kapelusz.
21	Niecierpliwił się przez	podstępny	jędrny	jasny	charytatywny	księżyc.
22	Przed nimi leżała	obrażona	mulista	zimna	piętrowa	brzoskwinia.
23	To była faktycznie	odważna	złota	droga	zaparkowana	kolacja.
24	On dokładnie opisał	wypoczęty	skuteczny	kupiony	okienny	portfel.
25	Marzyła jej się	odprężona	szumiąca	modna	gazowa	fryzura.
26	Zaskoczył go ten	wstydliwy	wiotki	brudny	pudrowy	zaułek.
27	Została już tylko	smutna	zaśniedziała	sucha	odbiorcza	kanapka.
28	Czekała na niego	mądra	wędkarska	pognieciona	nizinna	gotówka.
29	Na środku stał	bierny	skostniały	drewniany	językowy	statek.
30	Dał mi wczoraj	łyse	zardzewiałe	kupione	pieniężne	jedzenie.
31	Wszyscy patrzyliśmy na	śpiewający	musujący	mały	wypoczynkowy	pożar.
32	Zdziwił nas dzisiaj	łaskawy	prawny	długi	dymny	gwóźdź.
33	Obudziło go rano	marudne	kwaśne	intensywne	rabunkowe	słońce.

No.	Context	Novel metaphor with animacy violation	No-animacy violation metaphor	Literal sentence	Anomalous sentence	Critical word
34	W dali widzę	tańczący	rozlany	czysty	przeczytany	namiot.
35	Został tam tylko	egoistyczny	posiekany	stary	jajeczny	budynek.
36	W pokoju jest	śmiały	kryształowy	świeży	maszynowy	pomidor.
37	Chyba czeka nas	opanowana	sztuczna	ciepła	boczna	wiosna.
38	Zdziwiła ją ta	posłuszna	kudłata	brunatna	obrotowa	plama.
39	Trafiła nam się	uprzejma	krucha	zimna	wyprana	fala.
40	W domu leży	szczodry	piekący	zamknięty	nożny	dziennik.
41	Zdziwił nas jego	głodny	dostępny	mocny	marchewkowy	uścisk.
42	Słychać tam było	potulne	polne	zabytkowe	narciarskie	skrzypce.
43	Obok budynku jest	chętna	nastrojona	brązowa	napędowa	gleba.
44	W sklepie stała	przebiegła	skrzypiąca	nowa	napisana	świeca.
45	Jej uwagę zwróciła	waleczna	wymięta	ładna	terenowa	szuflada.
46	Rzeczywiście była to	zachłanna	słona	trudna	zamszowa	rozprawa.
47	Naszą uwagę przykuł	rozmowny	szorstki	zadbany	kruszony	ogród.
48	Blisko stąd jest	władcze	podarte	piękne	wyciskane	wybrzeże.
49	Spojrzała ukradkiem na	kulawy	cierpki	ogromny	fonetyczny	majątek.
50	Na zdjęciu widać	uzbrojony	czerstwy	letni	wyścigowy	kwiat.
51	Miał na sobie	roztargniony	ciekły	szary	płytki	szalik.
52	Wieczorem pokazała się	życzliwa	stęchła	biała	numeryczna	chmura.
53	W korytarzu był	wstydliwy	spróchniały	duży	operacyjny	owoc.
54	To jest kolejna	doświadczona	rozpruta	niebieska	biegunowa	filiżanka.
55	W apartamencie była	grzeczna	włochata	pakowana	cegłowa	żywność.

No.	Context	Novel metaphor with animacy violation	No-animacy violation metaphor	Literal sentence	Anomalous sentence	Critical word
56	Dostałam od niej	cierpliwe	zamglone	tanie	zbożowe	jabłko.
57	Wtedy spojrzała na	troskliwe	miejskie	nowe	dodatnie	łóżko.
58	Przed nim stała	hojna	nadmuchana	uszkodzona	gorsetowa	lodówka.
59	Oni wybrali sobie	dumny	rozpruty	kupiony	przelotny	ołówek.
60	W pokoju zastała	arogancki	wyprasowany	dziecięcy	przenośny	bałagan.
61	Dobrze znał ten	kapryśny	skisły	betonowy	smażony	chodnik.
62	W kącie leżała	pokorna	pikantna	szara	witaminowa	koszula.
63	Podziwialiśmy przez chwilę	zamożny	kwitnący	wypożyczony	widokowy	słownik.
64	To był bardzo	chciwy	bujny	drobny	dzianinowy	piasek.
65	Chłopca zdziwił jej	chudy	deszczowy	brzydki	bateryjny	uśmiech.
66	Na blacie leży	bogate	krzaczaste	czerwone	mieszkalne	ciastko.
67	Trafiła jej się	bosa	ognista	twarda	wschodząca	bułka.
68	Blisko znajduje się	naga	pomięta	ukryta	ręczna	jaskinia.
69	Wczoraj zaskoczyła nas	chytra	drewniana	narodowa	karciana	potrawa.
70	Przypomniał jej się	drażliwy	soczysty	zepsuty	herbaciany	głośnik.
71	Mężczyzna patrzył na	wrażliwe	alkoholowe	mokre	zagotowane	schody.
72	Czasem trafia się	zuchwały	gwiezdny	różowy	krzepliwy	latawiec.
73	Kupiłam w piątek	gibkie	betonowe	miękkie	naczelne	masło.
74	Tam leży ten	kulejący	grząski	pokaźny	skalisty	napiwek.
75	Musisz zobaczyć to	natrętne	soczyste	gęste	kieszonkowe	błoto.
76	Ciocię trochę denerwowała	gadatliwa	rozgotowana	czarna	dietetyczna	podłoga.

No.	Context	Novel metaphor with animacy violation	No-animacy violation metaphor	Literal sentence	Anomalous sentence	Critical word
77	Pod nami jest	łakoma	obfita	drewniana	gromka	skrytka.
78	Tutaj zawsze jest	śpiąca	żelazna	ładna	lustrzana	trawa.
79	Przypadkiem znalazła wczoraj	uległy	słoneczny	dopasowany	wybojowy	pierścionek.
80	Wzięła ze sobą	wyspany	spróchniały	ważny	gulaszowy	plecak.
81	Wszystkich zdziwiła ta	porywcza	prawnicza	zabawkowa	przestrzenna	strzała.
82	Na dole jest	obolała	rozpruta	metalowa	burzowa	puszka.
83	Była to zaskakująco	cwana	literacka	biała	pieczona	latarnia.
84	Przyszła sprawdzić ten	krzyczący	grząski	ceglowy	księgowy	komin.
85	Tata podał mi	wybredny	burzliwy	ciężki	piesza	czajnik.
86	W kącie leży	uszczypliwa	rozplątana	smaczna	sztormowa	cebula.
87	To może być	przebiegły	debiutancki	złożony	ujemny	żagiel.
88	Warta zobaczenia jest	skupiona	giętka	wapienna	gotowana	dolina.
89	Obok kanapy leżał	kapryśny	polny	zakurzony	rozbity	dywan.
90	Ona spoglądała na	nucący	rozdrapany	ekologiczny	obuwniczy	wiatrak.
91	W pudełku był	posępny	przekwitły	krótki	pitny	pasek.
92	Przed domem znalazł	życzliwe	potargane	wilgotne	płynne	drewno.
93	Miała w domu	zbuntowane	potargane	krzywe	sojowe	widły.
94	Potrzebne nam będzie	żwawe	płomienne	złamane	piesze	wiertło.
95	Na trawie leży	zgrabna	mozaikowa	mokra	rowerowa	sałata.
96	Obok drzewa stała	znużona	obwisła	działająca	mrożona	wyrzutnia.
97	Stąd widoczna jest	muskularna	wysłużona	sosnowa	kajakarska	kora.

No.	Context	Novel metaphor with animacy violation	No-animacy violation metaphor	Literal sentence	Anomalous sentence	Critical word
98	Przydałby się jej	czujny	bankowy	plastikowy	octowy	zderzak.
99	Po chwili zniknęła	ospała	wystygła	wyraźny	sypka	tęcza.
100	Przyśniła mi się	rozbrykana	puszysta	jaskrawa	szpinakowa	iskra.
101	Dziecko patrzyło na	speszony	rozlany	mocny	grochowy	płomień.
102	W mieszkaniu stoi	kulawa	panierowana	wspaniała	pożywna	choinka.
103	Na kanapie leżał	dumny	rześki	pomysłowy	napędowy	podarunek.
104	Na dywanie leżała	wycieńczona	parująca	złota	zachodząca	moneta.

Appendix B: Instructions in normative studies

1. Familiarity rating:

INSTRUKCJA
Na siedmiopunktowej skali należy ocenić jak często lub rzadko może Pan_Pani spotkać się z podanym wyrażeniem (które stanowią dwa ostatnie wyrazy zdania). Prosimy kierować się pierwszym wrażeniem po przeczytaniu zdania.

SKALA
1 – Bardzo rzadko
2 – Rzadko
3 – Raczej rzadko
4 – Ani rzadko, ani często
5 – Raczej często
6 – Często
7 – Bardzo często

PRZYKŁAD
Na stole leżał depilowany wisiorek. - bardzo rzadko (1)
Na stole leżał pozłacany wisiorek. - bardzo często (7)

2. Metaphoricity rating:

INSTRUKCJA
Na siedmiopunktowej skali należy ocenić jak dosłowne lub metaforyczne jest każde podane zdanie. Prosimy kierować się pierwszym wrażeniem po przeczytaniu zdania.

SKALA
1 – Bardzo dosłowne
2 – Dosłowne
3 – Trochę dosłowne
4 – Ani dosłowne, ani metaforyczne

5 – Trochę metaforyczne
6 – Metaforyczne
7 – Bardzo metaforyczne

PRZYKŁAD
Na stole leżał pozłacany wisiorek.- bardzo dosłowne (1)
Na stole leżał nieposłuszny wisiorek. - bardzo metaforyczne (7)

3. Meaningfulness rating:

INSTRUKCJA
Na siedmiopunktowej skali należy ocenić jak sensowne lub bezsensowne jest każde podane zdanie. Prosimy kierować się pierwszym wrażeniem po przeczytaniu zdania.

SKALA
1 – Bardzo bezsensowne
2 – Bezsensowne
3 – Raczej bezsensowne
4 – Ani bezsensowne ani sensowne
5 – Raczej sensowne
6 – Sensowne
7 – Bardzo sensowne

PRZYKŁAD
Na stole leżał depilowany wisiorek. - bardzo bezsensowne (1)
Na stole leżał pozłacany wisiorek. - bardzo sensowne (7)

<u>Niektóre zdania mogą wydawać się mniej sensowne niż inne, ale jednak niezupełnie bezsensowne. Na przykład:</u>

Na stole leżał skostniały wisiorek. - raczej sensowne (5)/ sensowne (6) - w tym zdaniu skostniały wisiorek może oznaczać zardzewiały, stary wisiorek.

Na stole leżał nieposłuszny wisiorek. - raczej sensowne (5)/ sensowne (6) - w tym zdaniu nieposłuszny wisiorek może oznaczać poplątany lub trudny do zapięcia wisiorek.

W tych przypadkach bardzo ważna jest Pana_i indywidualna opinia. Należy pamiętać, że w tym zadaniu nie ma poprawnych / niepoprawnych odpowiedzi.

4. Cloze probability test:

INSTRUKCJA
Prosimy o przeczytanie podanych zdań oraz o uzupełnienie luki wyrazem, który przyjdzie Pani_Panu na myśl. Prosimy kierować się pierwszym wrażeniem po przeczytaniu zdania.

PRZYKŁAD
Na stole leżał pozłacany ______.

odp. łańcuszek.

Appendix C: Informed consent form (in Polish)

ZGODA NA UDZIAŁ W BADANIU NAUKOWYM

Tytuł projektu: Przetwarzanie zdań w języku polskim: badanie ERP

Osoby przeprowadzające badanie:

mgr Patrycja Kakuba (patkak@amu.edu.pl), doktorantka, Wydział Anglistyki, UAM

dr Karolina Rataj (krataj@amu.edu.pl), Zakład Studiów Psycholingwistycznych, Wydział Anglistyki, UAM

Oświadczam, że zaznajomiłem/łam się z formularzem informującym o badaniu i zrozumiałem/łam przedstawioną w nim treść. Na wszystkie zadane pytania otrzymałem/am wyczerpujące odpowiedzi. Wyrażam dobrowolną i świadomą zgodę na udział w badaniu. Jestem również świadomy/a faktu, iż w każdej chwili mogę odstąpić od udziału w badaniu.

Wyrażam zgodę na przetwarzanie moich danych uzyskanych w trakcie eksperymentu, zgodnie z ustawą z dnia 10 maja 2018 r. o ochronie danych osobowych (Dz.U. z 2018 r. poz. 1000), wyłącznie w celach naukowych.

Niniejszy dokument, potwierdzający zgodę na udział w badaniach będzie przechowywany zgodnie z zasadami przechowywania dokumentacji poufnej.

Imię i nazwisko uczestnika badania: ………………………………………………

Data i podpis uczestnika badania: ………………………………………………

Oświadczam, że uczestnik badania zapoznał się z informacją o badaniu, a dane uzyskane podczas eksperymentu będą przechowywane zgodnie z zasadami przechowywania dokumentacji poufnej oraz przetwarzane zgodnie z ustawą z dnia 10 maja 2018 r. o ochronie danych osobowych (Dz.U. z 2018 r. poz. 1000).

………………………………………….

Data i podpis prowadzącego badanie

Appendix D: Handedness Questionnaire (in Polish) adapted from Oldfield (1971)

DATA__________________ ID:

KWESTIONARIUSZ PREFERENCJI RĘKI

Instrukcje

Zaznacz:

Której ręki wolisz używać kiedy wykonujesz daną czynność?

Czy kiedykolwiek używasz drugiej ręki kiedy wykonujesz daną czynność?

Której ręki wolisz używać, kiedy:	LEWEJ	Brak preferencji	PRAWEJ	Czy kiedykolwiek używasz drugiej ręki?
piszesz	☐	☐	☐	☐ Tak
rysujesz	☐	☐	☐	☐ Tak
rzucasz	☐	☐	☐	☐ Tak
używasz nożyczek	☐	☐	☐	☐ Tak
używasz szczoteczki do zębów	☐	☐	☐	☐ Tak
używasz noża (bez widelca)	☐	☐	☐	☐ Tak
używasz łyżki	☐	☐	☐	☐ Tak
używasz miotły (chodzi o rękę, która jest wyżej)	☐	☐	☐	☐ Tak
zapalasz zapałkę	☐	☐	☐	☐ Tak
otwierasz pudełko (chodzi o rękę, którą trzymasz wieczko)	☐	☐	☐	☐ Tak
trzymasz mysz komputerową	☐	☐	☐	☐ Tak
otwierasz drzwi kluczem	☐	☐	☐	☐ Tak
trzymasz młotek	☐	☐	☐	☐ Tak
trzymasz szczotkę do włosów lub grzebień	☐	☐	☐	☐ Tak
trzymasz kubek podczas picia	☐	☐	☐	☐ Tak