Institut für Linguistik

The Role of Attribute Prominence in Conceptual Processing

Inaugural-Dissertation
zur Erlangung des Doktorgrades der Philosophie (Dr. phil.)
durch die Philosophische Fakultät der
Heinrich-Heine-Universität Düsseldorf

Vorgelegt von
Svetlana Mnogogreshnova
aus
Krasnojarsk, Russland

Erstgutachter: Prof. Dr. Dr. Peter Indefrey
Zweitgutachterin: Prof. Dr. Wiebke Petersen

Düsseldorf, Dezember 2023

D61

Abstract

In the current thesis, the influence of attribute prominence on the speed and accuracy of semantic access is investigated in a series of four experiments focusing on the following attributes and semantic categories: parts and colors of animals, colors and parts of food (fruit and vegetables), affordance and parts of tools. Chapter 1 provides a review of the theoretical background relevant for the topic of the present thesis, such as existing theories of conceptual representation, visual word recognition, priming, previous research on conceptual information processing, and the role of color in object recognition. Chapter 2 deals with the methodological aspects of the experiments described in the current thesis, namely feature and image databases used for the selection of stimuli, experimental paradigms used in the experiments, and the software used for conducting experiments and analyzing the data. Chapter 3 provides information about methods and results of all four experiments as well as a discussion of the findings. Experiment 1 utilized a property verification task with verbal stimuli from all three semantic categories, which showed that primary attributes were generally retrieved more quickly and accurately than secondary attributes. However, it was not yet clear whether this effect could have arisen on the lexical level. In Experiments 2 and 3, a two-alternative forced choice task was used with pictorial stimuli from two semantic categories: animals and food, both of which had color and part attributes that differed in their prominence for these categories. Experiment 2 included isolated attributes in the form of color patches and parts of objects, whereas in Experiment 3 images depicting whole objects were used. These two experiments revealed a color correction bias observed for food items for which color was the primary attribute, and allowed to choose a degree of color change that would be comparable with changes in part attributes in terms of speed and accuracy of perceptual discrimination and, therefore, could be used for the subsequent experiment. Experiment 4 was a property verification task with pictorial stimuli, in which original objects, objects with altered part attributes, or objects with altered color attributes were presented. The data showed that it takes longer to reject an incongruent part attribute, if it is a secondary attribute of the given semantic category, and that both color and part attributes are retrieved more accurately, if they are primary attributes of the semantic category in question. In Chapter 4, implications of these experimental findings for conceptual representation research are discussed and future research possibilities are outlined.

Zusammenfassung

In der vorliegenden Dissertation wird der Einfluss der Prominenz von Attributen auf die Geschwindigkeit und Präzision des semantischen Zugriffs anhand von vier Experimenten untersucht, dabei liegt der Fokus auf den folgenden Typen von Attributen und semantischen Kategorien: Teil- und Farbattribute von Tieren, Farb- und Teilattribute von Nahrungsmitteln (Obst und Gemüse), Zweck- und Teilattribute von Werkzeugen. Kapitel 1 bietet eine Übersicht über Theorien der Konzeptrepräsentation, der visuellen Worterkennung und des Primings, sowie eine Übersicht über den Stand der Forschung zur Verarbeitung von konzeptuellen Informationen und zur Rolle von Farben in der Objekterkennung. Kapitel 2 beschäftigt sich mit den methodologischen Aspekten der in der vorliegenden Dissertation beschriebenen Experimente, insbesondere mit den verfügbaren Datenbanken von konzeptuellen Eigenschaften und Bilddatenbanken, die bei der Vorbereitung der Stimuli berücksichtigt wurden, mit den experimentellen Paradigmen, die bei den Experimenten benutzt wurden, und mit der Software, die für die Durchführung der Experimente und für die Auswertung der Daten eingesetzt wurde. Kapitel 3 beschreibt die Methoden und Ergebnisse der vier Experimente und diskutiert die Forschungsergebnisse. In Experiment 1 wurde eine Eigenschaftsverifizierungsaufgabe mit verbalen Stimuli aus allen drei semantischen Kategorien benutzt. Die Ergebnisse zeigten, dass prominentere Attribute schneller und präziser abgerufen werden als Attribute, die weniger prominent für die entsprechenden Kategorien waren. Allerdings konnte nicht ausgeschlossen werden, dass dieser Effekt auf der lexikalischen Ebene entstanden ist. In den Experimenten 2 und 3 wurde eine Forced-Choice-Aufgabe mit Bildern benutzt, die Objekte aus den semantischen Kategorien „Tiere“ und Nahrungsmittel“ zeigten. Beide Kategorien haben Farb- und Teilattribute, die sich jedoch in ihrer Prominenz für diese Kategorien unterscheiden. In Experiment 2 stellten die Stimuli isolierte Attribute in Form von Farbfeldern oder Objektteilen dar, während im Experiment 3 Bilder, die Objekte als Ganzes darstellten, benutzt wurden. In diesen zwei Experimenten wurde eine Tendenz zur Farbkorrektur bei Objekten aus der Kategorie „Nahrungsmittel“, für die Farbattribute prominenter sind als für Tiere, gefunden. Darüber hinaus wurde der Farbveränderungsgrad bestimmt, der mit Teilveränderungen hinsichtlich der Reaktionszeiten und Fehlerrate vergleichbar war und somit im darauffolgenden Experiment benutzt werden konnte. Experiment 4 war eine Eigenschaftsverifizierungsaufgabe mit Bildstimuli, die unveränderte Objekte, Objekte mit geänderten Teilattributen oder Objekte mit geänderten Farbattributen darstellten. Die

Daten zeigten, dass man länger braucht, um ein inkongruentes Teilattribut abzulehnen, wenn es für die gegebene Kategorie weniger prominent ist, und dass man weniger Fehler sowohl bei Farb- als auch bei Teilattributen macht, wenn diese eine hohe Prominenz für die entsprechenden semantischen Kategorien haben. Im Kapitel 4 wird die Bedeutung dieser Ergebnisse für die Forschung im Bereich der konzeptuellen Repräsentation diskutiert und es werden zukünftige Forschungsmöglichkeiten beschrieben.

Content

List of Figures

List of Tables

INTRODUCTION

It is known that concepts are not the smallest units in the mental lexicon and can be further subdivided into smaller elements. There have been a number of theories trying to describe the structure of concepts with respect to various phenomena, such as facilitation, inhibition, semantic interference etc., which arise from the interaction between concepts and/or conceptual elements. According to the frame theory as one of the most influential theories on how the human mental lexicon is organized, concepts are represented in the form of frames which are defined as "recursive attribute-value structures" (Petersen, 2007: 151), and there are several kinds of constituents within a conceptual frame. One of them is represented by attributes, which describe certain properties of a given concept. For example, concrete concepts which denote tangible objects (e.g., *cat*, *book*) can have various types of attributes, such as size, shape, color, etc. These attributes are activated in a particular order, and "the relative timing of each feature is informative about the role the feature plays in the word representation of the object" (Lam, Dijkstra & Rueschemeyer, 2015: 6). In other words, attributes which play a more prominent role in the structure of a given concept are supposed to be accessed earlier than other, less prominent attributes of that concept. This prominence of a particular attribute might be assessed based on how frequent a particular attribute is mentioned in response to a concept label. Such data is usually published in the form of feature databases, and can be used for research purposes.

For any given semantic category, some attributes tend to be more important than others, e.g., body parts are an important feature of an animal, since they are often used to distinguish between the concepts in that particular category, whereas colors might vary for objects of the same kind (e.g., for different cat breeds), thus providing less distinctive information for identifying an object concept and, therefore, for animals, colors can be characterized as a less important attribute in comparison with body parts. However, for some members of the category the prominence of a particular attribute might be reversed, for instance, a zebra is one of the members of the semantic category "animals", but its color attribute is more prominent than its body parts, because it is more distinctive, whereas the same body parts can be found in another concept from the same category, such as a horse. One of the questions investigated in the current dissertation is whether attribute prominence for a particular item can override attribute prominence for the semantic category as a whole, when it comes to semantic access via concept labels and via pictorial stimuli.

The current series of studies aimed to investigate whether attribute prominence for the semantic category or attribute prominence for individual items have an influence on the order of processing of the respective attributes within the frame structures of concrete concepts that belong to certain semantic categories. The experiments focus on such semantic categories as animals, tools, and food, and their attributes: color and part attributes of animals and food, as well as affordance and part attributes of tools. The information on the prominence of these attributes for the semantic category as a whole and for individual items was determined based on the data about members of those semantic categories provided in the feature database compiled by McRae et al. (2005).

The present thesis comprises four chapters. It starts with Chapter 1 that provides a theoretical background on the topic of the dissertation and the phenomena that are relevant for the experiments described in it. Section 1.1 outlines the most influential theories of concept representation, such as the accounts assuming binary features, the feature list theory, non-decompositional accounts of concept representations, the prototype theory, and the frame theory. Section 1.2 is devoted to the topic of concept activation and priming, including types of priming, theories explaining the existence of priming effects and the differences between strategic and automatic priming. Then Section 1.3 gives an overview of previous research on time-course of conceptual information processing and encoding. Next, Section 1.4 provides some information on the process of visual recognition of words, as well as on the mental lexicon and lexical access, and subchapter 1.5 deals with color perception and the role of color in object recognition.

Chapter 2 deals with methodological aspects of the studies involved in the current thesis. Section 2.1 provides an overview of some existing feature databases and a justification of the choice of one of them in accordance with the purposes of the current studies. Then, Section 2.2 gives an overview of some of the existing image databases and the information about the sources of the images that were later used as the stimuli in some of the experiments described in the current thesis. Section 2.3 describes experimental paradigms that were utilized in the experiments described in the present thesis, such as the property verification task and the two-alternative forced-choice task. Section 2.4 describes the software that was used to conduct the experiments described in the present thesis. Finally, the software that was used for the statistical analysis of the collected data and the procedure of the analysis are described in Section 2.5.

In Chapter 3, the details about the four experiments are presented, including the information about participants, stimuli, procedure and results of statistical analysis, as well as the discussion of the results of each of the experiments involved.

Chapter 4 then combines the results of the experiments described in the current thesis by presenting a general discussion and implications of the findings and discusses possible opportunities for future research on the topic.

In conclusion, the findings of the described experiments are summarized and achievement of the aims of the current dissertation is assessed.

1. CONCEPTUAL REPRESENTATION, PRIMING, VISUAL WORD RECOGNITION, AND THE ROLE OF COLOR IN OBJECT RECOGNITION

1.1. Theories of Conceptual Representation

All knowledge about everything in the world around us is stored in our memory in the form of concepts. A concept is defined as a mental representation of a certain referent (i.e., a tangible entity or an abstract phenomenon existing in the real world) in the human mind. Based on their referents, concepts can be classified into concrete and abstract. Concrete concepts are representations of living and non-living things, i.e., they have tangible, measurable physical referents in the real world (e.g., *cat*, *stone*, etc.), whereas abstract concepts are notions, phenomena, and ideas, which do not have a unique tangible referent (e.g., *art*, *love*, etc.). A word or an expression which is used to refer to a particular concept is known as a concept label. Taken together, concepts constitute a conceptual system. Activation of a certain concept within a conceptual system is the basis of successful speech production and speech comprehension.

However, it is known that a concept is not the smallest unit of knowledge. Over the decades, multiple theories of conceptual representation have been proposed, which tried to describe the inner structure of concepts, the ways in which concepts relate to each other, and various phenomena that arise based on these relations. This chapter will give an overview of the most prominent theories of conceptual representation, including their strong and weak points, in order to justify the choice of one of those theories for the current study.

1.1.1. Binary Features

The binary features theory was developed by Roman Jakobson and was initially referred to as distinctive features and was used in phonology to describe characteristics of phonemes, using sets of binary features indicating place and mode of their articulation (Jakobson, Fant & Halle, 1952; Gussenhoven & Jacobs, 1998). For example, Chomsky and Halle considered a symbol such as [i] to be an abbreviation for the following combination of distinctive features (Chomsky & Halle, 1968: 64):

[+ segment] [+ vocalic] [– consonantal] [+ high] [– low] [– back] [– round] [– tense]

Binary features allowed to compare phonemes to each other and to describe certain constraints which they might be subject to. Due to its efficiency, this approach was later also implemented in other branches of linguistics. Particularly, in semantics binary features were used to describe meanings of words. According to this approach, a concept was characterized by a set of features which could assume a positive or a negative value (e.g., [+HUMAN] or [–HUMAN]), in which case a feature would be considered a predicate. This theory could be successfully used to distinguish between concepts and describing simple relations between them. For instance, hyponymy was reflected in the fact that one of two concepts has an additional feature in its structure in comparison with the other concept in that pair. Concepts were considered logically incompatible, if they differed in at least one feature. However, if the difference between them was restricted only to a single feature, they were considered complementary opposites, for example:

boy [+HUMAN] [–ADULT] [+MALE]

girl [+HUMAN] [–ADULT] [–MALE]

Binary features are useful for characterizing concepts from domains with clear relations between different terms, like concepts of family members (e.g., *mother*, *father*, *sister*, *brother*, etc.). In this case, whenever a particular property has to be verified, the corresponding feature has to be found in the structure of the given concept. Priming effects, in turn, would be based on two concepts having one or more features in common. However, this theory becomes less efficient, if the number of binary features necessary to describe a concept becomes too large, for example, when it comes to more complicated relationships between concepts, like the order of the days of the week or months of the year, instead of just binary relations. Apart from that, the meaning of some words, like relational nouns and most verbs, also cannot be described using binary features. In other words, binary features can only be applied to a limited number of concepts (Löbner, 2013).

1.1.2. Feature Lists

Another prominent theory which was used to describe conceptual structures is the feature list theory. A conceptual representation in the feature list theory is based upon a set of certain qualities which allow us to perceive an object in question as an instance of a particular category. According to this theory, each concept can be represented by a list of characteristics (features), which, in turn, can be divided into obligatory and varying ones (some researchers, e.g., Harley (2014), refer to these types of features as defining features and characteristic features, respectively). This distinction was based on the distinction between the essence and accidents of things suggested by Aristotle. Obligatory features represent the essence of a concept, describing requirements that an entity needs to meet, in order to be identified as that particular concept, or as a member of a particular category. Varying features represent accidents, i.e., characteristics which might be different across entities corresponding to the same concept, or across concepts within a given category. The features in question cannot be decomposed further into any smaller elements (Harley, 2014; Taylor, 2003).

Rumelhart, Hinton and McClelland (1986) also considered conceptual features to be one of the aspects of the parallel distributed processing (PDP) model, assigning them to processing units within that model. According to their approach, excitatory and inhibitory connections between different features within the same feature list were possible, depending on their co-occurrence (Barsalou, 1992; Rumelhart et al., 1986).

Thus, the feature list theory is more flexible than the binary features and, therefore, can better account for the processes that enable us to verify properties of concepts. Whenever a property of a concept is presented in a sentence like "A is B", a feature comparison has to be carried out, in order to estimate the degree of overlap between the concepts mentioned in that statement. If the overlap is very high and no conflicting information is found (e.g., "*A robin is a bird.*"), the property is confirmed; whereas if the degree of overlap is too low and/or conflicting features are detected (e.g., "*A robin is a pig.*"), the property is rejected. If there is a moderate degree of overlap between the concepts from the statement (e.g., "*A penguin is a bird.*", in which case the fact that a penguin cannot fly is in conflict with the general ability of birds to fly), a more detailed comparison needs to be conducted, which causes longer reaction times in a property verification task (Harley, 2014).

However, a serious drawback of this theory is that it has only one level, i.e., all its elements have equal status and, therefore, this theory cannot explain their inter-correlation, e.g., if a living being flies and builds nests, it probably lays eggs as well (Harley, 2014), or some of the complex relationships existing between features of a particular concept, like the fact that some features might exert a certain restriction on other features of the same concept, e.g., when we talk about the concept of travel, its price and speed exert certain restrictions on each other (Gamerschlag, Gerland & Osswald, 2014). Another problematic aspect of the feature list theory, according to Harley (2014), is that it is concentrating specifically on sentence verification tasks, and even a simple manipulation like reversing the order of concepts in such sentences produces results that are not accounted for by the feature comparison procedure.

1.1.3. Non-Decompositional Theories of Concept Representation

There are multiple theories that describe the structure of human memory and its functioning principles in an attempt to explain the existence of various phenomena, including priming effects.

One of the most prominent theories is the spreading activation theory, which serves as a basis for several models, the most influential of which are Anderson's spreading activation model, also known as ACT* (Anderson, 1983), and Collins and Loftus's model (Collins & Loftus, 1975). The common assumption of these two models is that human memory is represented in the form of a network that consists of numerous interconnected nodes. Whenever a participant receives a certain stimulus, e.g., visually or auditorily, a relevant node, i.e., the mental representation of the corresponding concept, is activated in the memory. This activation automatically spreads further to the nodes that share a connection with the source node. Apart from that, there is also a possibility for the activation to go in the opposite direction. Thus, access to related nodes is provided and, as a consequence, retrieval of the corresponding concepts and lexical items is facilitated, provided that the level of initial activation is high enough and that the connection between the nodes in question is strong enough. This mechanism is also influenced by the strength of nodes themselves. In other words, the more often a certain concept is encountered, the more activation it can convey to further nodes within the network. The links between the nodes can have different strength, depending on the frequency of co-occurrence of the concepts in question. It is also assumed that the amount of activation gradually decreases, as it spreads further from the source node (Anderson, 1983). Hence, the further a target node is situated from the source node, the longer it takes for it to become activated and the less activation it gets. This activation keeps spreading from the source node to the nodes that are connected with it for as long as the concept is actively processed. In this case, conceptual processing is supposed to be serial, meaning that only one concept at a time can be processed actively, and attentional control is used to switch between active processing of different concepts (Collins & Loftus, 1975; McNamara, 2005).

Collins and Loftus (1975) propose a slightly different description of human memory. According to their view, it is additionally divided into two levels: a conceptual network and a lexical network. In other words, the authors distinguish between concepts as such and their lexical representations (labels). Both levels are functioning as separate networks and are organized based on different principles: concepts are arranged according to their semantic similarity to each other, whereas lexical items are organized based upon their phonemic and orthographic similarity. Nevertheless, the boundary between these two networks is not impenetrable, since conceptual nodes are still connected to their lexical representations. This way of organization of human memory not only explains the existence of priming effects caused solely by semantic similarity or solely by associative relatedness of the stimuli, but also provides an explanation for an increase in priming effects caused by a combination of both semantic and associative types of prime-target relatedness.

Another difference between the two versions of the spreading activation theory is that in Collins and Loftus's model a prime and a target from one and the same trial are processed successively, i.e., one after the other, and the target can be pre-activated, even if the source of activation itself is not active anymore; while in ACT* it is assumed that, for facilitation to occur, both the prime and the target have to be sources of activation (McNamara, 2005: 16).

Another influential theory explaining priming effects is the compound cue theory (Ratcliff & McKoon, 1988; McKoon & Ratcliff, 1989). According to this theory, the reaction time of a response to a particular target item depends on the passive process of comparison between the target and the memory representation of the preceding prime. In this case, both the prime and the target are considered in combination with each other, and such a combination is referred to as a compound cue. However, if a task requires participants to ignore primes and react only to targets, as is the case in a lexical decision task, then each target has a priority over its prime within a compound cue (Ratcliff & McKoon, 1988; McKoon & Ratcliff, 1989). A compound cue which contains two words that are semantically or associatively related to each other is perceived as more familiar than a compound cue that consists of completely unrelated words, which explains shorter reaction times in related conditions and, consequently, the emergence of priming effects (McNamara, 2005).

One more theory explaining the structure of human memory is the so-called theory of distributed associative memory (TODAM). According to this theory, concepts are represented by “vectors of attributes” (Murdock, 1982: 610). Thus, when an item is being recognized, its vector is compared to the one that is stored in the memory of a participant. Every new item is added to the corresponding vector in the memory. Information about an associative relatedness between two concepts is represented as a convolution of their respective vectors and is added to the memory vector separately. The process of item recognition is based upon the vector of the stimulus item and its comparison to the memory vector, whereas random noise is also taken into consideration. Whenever a prime and its association to a target are given as input information, an output would be a vector that is similar to the vector of the target, which could explain facilitation of subsequent recognition of the corresponding concept (Murdock, 1982).

Some authors claim that the theory of distributed memory can explain only priming effects arising due to semantic relatedness of the stimuli (Thompson-Schill, Kurtz & Gabrieli, 1998), while the spreading activation theory was designed to account not only for semantic, but also for associative priming effects (Collins & Loftus, 1975). For this reason and because of its compatibility with the frame theory (see Section 1.1.5), the spreading activation theory is taken as a basis for the current experiments.

1.1.4. Prototype Theory

The prototype theory is based on categorization, which is one of the fundamental cognitive processes that enables human mind to classify concepts into categories based upon certain conditions (Taylor, 2003). According to this theory, concepts are viewed as members (exemplars) of a particular category. Membership within a category is determined by how close a concept is to a prototype, i.e., to a typical member of the category in question. Generally a member’s typicality within a particular category is determined not by a concept’s frequency, but by its family resemblance with respect to other members of the same category (Murphy, 2002), which can be explained in the following way: “members of a category come to be viewed as prototypical of the category as a whole in proportion to the extent to which they bear a family resemblance to (have attributes which overlap with those of) other members of the category” (Rosch & Mervis, 1975: 575). In other words, prototypes as the most typical representatives are used as reference points for their respective categories and, consequently, membership in a category can be graded, meaning that there can be a difference in the status of concepts within a particular category: for example, a robin is considered a better example of a bird than a penguin. This, in turn, leads to the fact that there are no firm boundaries between different categories.

Apart from the distinctions between category members, semantic categories themselves can have different levels. For example, the category “furniture” represents a superordinate level, “chairs” is a category of the basic level, and “armchairs” is a category of a subordinate level. Members of categories on the basic level are usually acquired first, information about them is accessible earlier than information about members of categories on higher or lower levels of categorization. Objects on higher levels of categorization are easily distinguishable from each other, whereas members of subordinate level categories are more similar to each other, and the basic level of categorization represents “the level that has the most distinctive attributes and provides the most economical arrangement of semantic memory” (Harley, 2014: 334).

One of the advantages of this approach is that the prototype theory is able to explain the phenomenon of color categorization, which the binary features or the feature list theory cannot account for (Löbner, 2013). The prototype theory is also able to explain the fact that verification of property statements is quicker for prototypical category members than for atypical exemplars of a category, as well as for the fact that labels for more typical members of a category are acquired earlier and retrieved faster than labels for less typical members (Harley, 2014).

On the other hand, the prototype theory is less efficient, when it comes to abstract notions, for which a prototype is more difficult to define (Harley, 2014). Still, some researchers successfully used this theory for investigating abstract concepts, for example, Coleman and Kay (1981) studied the English concept of a lie using 8 scenarios that represented various combinations of three distinctive features of a lie, namely the factual falsity of the proposition, the fact that the speaker believes the proposition to be false, and the intent of the speaker to deceive the addressee. The fact that the scenarios received different rating scores supported the claim that membership of the category “lie” was graded. Moreover, the authors also found out that the three features they used as criteria played different roles in defining whether something counted as a lie: the belief of the speaker that what they say is false was the most important conceptual element, followed by the intent to deceive the listener, and the factual falsity was the least important of the three criteria. Thus, for some abstract concepts there seems to exist a certain prototype that other concepts might be compared to, in order to establish their category membership.

However, a major weakness of the prototype theory, according to Harley (2014), is its circularity: category membership of objects is said to be based on their similarity, and their similarity is explained by the fact that they belong to the same semantic category. Apart from that, the coherence of some semantic categories is not explained by the prototype theory, especially when it comes to the categories for which there is no easily identifiable prototype (Harley, 2014), like the noun class including "women, fire, and dangerous things" from the Dyirbal language in Australia (Lakoff, 1987: 5).

1.1.5. Frame Theory

As an alternative approach, the notion of frames was introduced in the cognitive science in 1970s by Charles Fillmore. He defined it as "any system of concepts related in such a way that to understand any one of them you have to understand the whole structure in which it fits; when one of the things in such a structure is introduced into a text, or into a conversation, all of the others are automatically made available" (Fillmore, 1982: 111). However, at first Fillmore considered frames to be non-recursive (Gamerschlag et al., 2014), while proponents of the modern version of the frame theory describe frames as "recursive attribute-value structures" (Petersen, 2007: 151). It means that a frame can include another frame as one of its elements. In the frame structure, the central node (i.e., the label of the concept itself) is connected to other nodes representing its properties. The values of these properties can also be frames and can be connected to further elements. This way of organization enables a quick transition from one concept to another, which, among other things, makes priming effects possible. That is a reasonable advantage of the frame theory for investigation of phenomena that are based on connections between concepts or between concepts and their structural elements.

There are two main types of frames, depending on what exactly they represent: predicative frames and concept frames. Predicative frames aim to describe a particular event or an action, and all participants and details of the described process are parts of the frame structure in this case. In contrast to that, concept frames capture a certain entity, and its properties serve as frame elements (Gamerschlag et al., 2014: 4).

Nouns that represent labels for concept frames are classified into four types, according to their relationality and inherent uniqueness: sortal nouns, individual nouns, relational nouns and functional nouns (Löbner, 2015). Individual nouns are inherently unique, but not relational. Concepts of this type must have a unique reference, even though it can be different depending on time (e.g., *pope*, *weather*). Proper names also belong to individual nouns. Functional nouns have both inherent relationality and uniqueness. They usually denote a property or a part of something, therefore attributes in a frame can be regarded as functional concepts themselves, which proves the recursive nature of a frame structure. Sortal nouns are not relational and have no inherent uniqueness, they denote categories of concrete objects or abstract notions, e.g., *dog* or the word *noun* itself. Relational nouns "include terms for potentially multiple roles such as brother, neighbor, friend, colleague, contemporary, etc. or multiple parts of objects" (Löbner, 2015: 46-47). Due to existence of polysemy, one and the same word could belong to two types, depending on its meaning in a particular context or situation (Petersen, 2007; Gamerschlag et al., 2014).

Frames consist of three basic elements. These are attribute-value sets, constraints and structural invariants. Attributes and values are properties and their instances in a particular concept. Type signatures are applied to attributes to ensure that their values are subordinate concepts of those attributes, i.e., that the values correspond to appropriateness specifications of the attributes, such that an attribute like "taste" could only assume values like "sweet", "bitter", etc. (Petersen, 2007). Constraints are restrictions that are exerted on a frame's attributes or values. There are several types of constraints defined by Barsalou (1992), such as attribute constraints, value constraints, contextual constraints and optimizations. Attribute constraints are relations showing that variations in the values of one attribute can cause variations in the values of another attribute within the same concept frame. For example, if we consider the frame "trip", faster transportation is usually more expensive, which shows how the value of the attribute "speed" can affect the value of the attribute "price". Value constraints are more specific, e.g., in the frame "vacation", if the attribute "location" has the value "Rockies", the attribute "activity" is likely to have the value "snow skiing". Contextual constraints represent general constraints of a particular situation due to physical or cultural reasons, for instance, the dependence of speed and duration of transportation on one another or relation between income and tax in a particular country. In contrast, optimizations are based upon the agent's aim, e.g., optimization of the speed, price and comfort of transportation. In other words, constraints reflect variations within a frame. As far as structural invariants are concerned, they "represent relatively constant relations between a frame's attributes", including spatial, temporal, causal, intentional and other types of relations (Barsalou, 1992: 37).

According to Barsalou (1982), there are two kinds of components within frames, namely context-independent and context-dependent elements. The information that constitutes context-independent properties of a particular concept is activated every time this concept is encountered, while context-dependent features are only activated, if they are important in a particular context, in which the concept is presented. If this assumption is right, it would mean that context-independent information about a particular concept is accessed, even if the corresponding word is presented for a very short period of time, such as when conscious processing, i.e., conscious extraction of any additional information, is avoided.

All in all, frames are more elaborate and flexible structures in comparison with feature lists. A feature list is a one-level set of independent specific characteristics of a given concept, while a frame represents a complex hierarchical system, including attributes and their values along with structural invariants and constraints. Empirical evidence reported by Barsalou (1992) confirms that people and animals can perceive the difference between attributes and values and correlation between some properties, which justifies the existence of the attribute level in conceptual structures and makes frame theory more suitable for investigation of concepts and their properties. Moreover, frames can represent not only simple concepts, but the fact that they are recursive also gives them an opportunity to represent categories and their members, prototypes, conceptual combinations and even sequences of events (Barsalou, 1992). This versatility and flexibility of frame structures was the reason why this theory was chosen as the basis for the current study.

1.2. Concept Activation and Priming

Irrespective of the theories describing the way conceptual structures are organized, researchers generally agree that concepts are not isolated from each other, but rather co-exist in an intricate network that enables a transition from one concept to another in the course of speech production or comprehension.

One phenomenon that is based on this kind of transition and is relevant for the current study is the phenomenon of priming, since there is a semantic connection between attribute values and the respective concept labels. Priming arises when a certain stimulus (prime) influences the perception of a subsequently presented item (target), in which case facilitation effects are observed in trials with related stimuli as compared to trials with unrelated stimuli. For example, in Experiment 1 verbal stimuli were used and attributes were presented before congruent or incongruent concept labels, which might lead to the emergence of priming effects in the congruent condition in comparison with the incongruent one, meaning that in a congruent condition an attribute facilitates the perception of the subsequently presented concept label, leading to shorter reaction times and lower error rates.

In this chapter, several theories describing the way human memory is organized are outlined, and the relevant aspects of priming are discussed, namely types of relatedness between primes and targets, and types of priming.

1.2.1. Types of Prime-Target Relatedness

It is known that priming effects can be caused by different types of relatedness between primes and targets, such as associative relatedness (Meyer & Schvaneveldt, 1971; Den Heyer, Briand & Dannenbring, 1983; Hirshman & Durante, 1992; Brown & Hagoort, 1993; Neely, VerWys & Kahan, 1998; Kahan, Neely & Forsythe, 1999; Brown, Hagoort & Chwilla, 2000; Ortells, Abad, Noguera & Lupiáñez, 2001; Abad, Noguera & Ortells, 2003; Noguera, Ortells, Abad, Carmona & Daza, 2007; Balota et al., 2008), semantic relatedness (Fischler, 1977; Valdés, Catena & Marí-Beffa, 2005) or a combination of both semantic and associative relatedness (Perea & Gotor, 1997; Thompson-Schill, Kurtz & Gabrieli, 1998; Brown & Besner, 2002; Perea & Rosa, 2002; Bouaffre & Frédérique, 2007; Sánchez-Casas et al., 2012). The difference between semantic and associative relatedness was first described by Ira Fischler. In his study (Fischler, 1977), he used a lexical decision task with words that were not associated with each other, but shared certain conceptual components (e.g., *nurse – wife*) and compared them to associatively related words (obtained via a free association test) and pairs of unrelated items (generated by re-assigning target words within the two previously described types of stimuli). Primes and targets were presented on the screen simultaneously and remained visible until a response was given, whether both letter strings presented in a trial were real words. As a result, a positive correlation between the strength of semantic relatedness and facilitation effects was observed, whereas associative strength correlated negatively with the amplitude of facilitation effects. In most of the other studies, semantically related words are represented by synonyms, antonyms, category coordinates etc., and semantic priming effects are generally smaller than associative priming effects; it is also often the case that semantically related stimuli are strongly associated with each other. In such cases, associative relatedness of primes and targets can boost already existing semantic priming effects (Lucas, 2000).

However, the studies mentioned above investigated the phenomenon of direct priming, i.e., situations in which a prime and a target are directly related to one another. There is also a different type of priming, namely mediated priming, in which a prime and a target are connected not directly, but through a common associate that is not presented in a trial, e.g., *milk – (cow) – bull* or *lion – (tiger) – stripes* (De Groot, 1983; Hill, Strube, Roesch-Ely & Weisbrod, 2002; Hill, Ott & Weisbrod, 2005). Unlike these examples, the order of stimuli presentation in Experiment 1 of the present thesis was reversed, i.e., instead of forward association (from a concept to an attribute), backward association was observed (from an attribute to a concept). And still, this might mean that there is a possibility of observing some residual priming effects in the incongruent condition as well, given that incongruent attributes are taken from other concepts from the same semantic category.

As far as visual attributes, like color or form, are concerned, some researchers investigated those, albeit mostly by using pictorial stimuli in categorization, naming or matching tasks. The majority of such studies claimed that a congruent color and/or form facilitates object recognition (Price & Humphreys, 1989; Wurm, Legge, Isenberg & Luebker, 1993; Hayward, 1998; Naor-Raz & Tarr, 2003), whereas an incongruent color may cause inhibition (Therriault, Yaxley & Zwaan, 2009), and that such results are more likely to be observed in experiments involving high color diagnostic objects (Tanaka & Presnell, 1999; Redmann, FitzPatrick, Hellwig & Indefrey 2014).

Moreover, Tyler and Moss (1997) investigated functional and perceptual attributes using an auditory lexical decision task performed by healthy subjects and patients having a brain lesion. For the purpose of this dissertation, we will consider their experiments involving healthy participants. Tyler and Moss (1997) used 52 words denoting living beings and 68 words denoting inanimate things in a property generation pre-test, aiming to obtain the most salient functional and perceptual attributes for a subsequent auditory lexical decision task, for which living beings and inanimate things were selected as primes and the generated properties were used as targets. Functional properties were represented by a place where an animal lived, e.g., *crocodile – river*, or by what an inanimate thing was used for, e.g., *cherry – eat*; whereas perceptual attributes included color, shape, size, or texture, e.g., *crocodile – green*. These types of attributes were compared to superordinate and co-ordinate prime-target pairs, such as *crocodile – animal* and *crocodile – elephant*, respectively. Association strength between primes and targets was kept to a minimum, since the authors were looking for semantic, and not associative, priming effects. For each of the four conditions (functional attribute, perceptual attribute, superordinate, co-ordinate), related and unrelated prime-target pairs were compiled and an equal number of word-nonword pairs was added. The participants heard word pairs through the headphones with an inter-stimulus interval of 200 ms and had to perform a lexical decision task on target words. The results of this experiment revealed that priming effects were comparable between the four prime-target relatedness types as well as the types of primes (animate vs. inanimate), and neither perceptual, nor functional attributes showed any differential priming effects in favor of animate or inanimate prime words. The authors also assumed that, "although the prime word activates all of the semantic features with which it is associated, the activation rise-time of each feature may be a function of its saliency within the representation" (Tyler & Moss, 1997: 521). In their second experiment, Tyler and Moss (1997) used 100 words denoting man-made objects as primes and their functional properties (a typical way to use an object, a location where an object is used, adjectival description of an object's function) and perceptual properties (visible parts, colors, shapes) collected in a separate pre-test as targets. An example for a prime-target pair in the functional condition would be *blouse – wear*, and in the perceptual condition *blouse – button*. The authors used an additional pre-test to determine the so-called "isolation point" of the prime (the point at which people could first identify the upcoming word) and presented target words visually for 54 ms in the center of the screen either at the isolation point of a prime or at its offset. The analysis revealed that there were significant priming effects both if primes were cut off at the isolation point and if they were used as a whole. In the perceptual condition, though, priming effects were only observed, if the target was shown at the offset of the prime. Tyler and Moss (1997) concluded that perceptual properties are automatically activated upon hearing concepts of living or non-living things, but their activation is slower than that of functional properties.

1.2.2. Strategic and Automatic Priming

According to Brown and Hagoort (1993), there are three main factors which contribute to priming effects. The first of them is the activation spreading from a prime to a target, which occurs automatically as soon as the stimuli are presented. The second factor involves the expectations that a participant forms with respect to target words depending on their contexts (e.g., using information from the corresponding primes), in which case conscious processing is involved. The third factor is semantic matching, which is also attributed to controlled processing and consists of matching each target to the preceding prime. When it comes to automatic priming, when there is not enough time to form expectations or perform semantic matching, the latter two factors cannot contribute to facilitation of target perception and, therefore, priming effects are significantly smaller (Brown & Hagoort, 1993; Brown & Besner, 2002; Balota et al., 2008). And even though the results of strategic processing can be used to investigate functioning principles of human memory and mental lexicon, these processes are highly dependent on experimental tasks and instructions, which is why "it is generally thought that only automatic processes reliably reflect long-lasting organizational structure of semantic networks" (Lucas, 2000: 619).

As a consequence, priming effects investigated in a particular experiment can result from automatic or strategic processes, depending on the timing of stimuli presentation, which determines the time available for their processing. Automatic priming does not involve awareness and occurs at short stimulus-onset asynchronies (SOAs) between primes and targets, while strategic priming requires controlled processing for facilitation to occur and takes place at somewhat longer SOAs (Brown & Hagoort, 1993; Lucas, 2000). As far as the exact timing is concerned, there is no unanimous opinion among researchers on that point. In some studies, primes are presented for a very short period of time, which varies in different experiments: 57 ms (Sánchez-Casas et al., 2012), from 10 to 40 ms (Brown & Hagoort, 1993) or less than 50 ms (Hirshman & Durante, 1992). Some researchers consider all prime durations within a range from 50 to 250 ms to be suitable for automatic priming (Garman, 1991: 294; Lucas, 2000: 622). McNamara (2005) considers inhibition to be an indicator of conscious processes and, since it is not observed at SOA of 200 ms or shorter, he recommends to use these particular timing conditions for investigation of automatic processing. However, other factors, such as characteristics of lists of stimuli, like high or low relatedness proportions, can also play a role in distinguishing between automatic and strategic processing. In particular, the influence of such factors is minimal under the conditions of automatic priming (McNamara, 2005: 72), whereas in studies with longer SOAs a clear positive correlation between the number of related stimuli and the amplitude of priming effects is observed (den Heyer et al., 1983; Brown et al., 2000)

In some experiments designed for automatic priming investigation, masking is implemented, which means that a certain pattern is presented immediately before (forward masking) or after (backward masking) a stimulus to prevent or impede its identification and thus to eliminate possible effects caused by controlled processing of primes. A mask itself can consist of a string of certain symbols, for example, many researchers used hash marks (Balota et al., 2008; Brown & Hagoort, 1993; Sánchez-Casas et al., 2012; Perea & Gotor, 1997; Perea & Rosa, 2002), asterisks (Neely et al., 1998); ampersands (Brown & Besner, 2002; Noguera et al., 2007), random letter strings (De Groot, 1983; Valdés et al., 2005; Ortells, Kiefer, Castillo, Megías & Morillas, 2016), plus signs or Xs (Hirshman & Durante, 1992; Neely et al., 1998), or an array of random symbols (Balota et al., 2008) etc. Under masking conditions, a prime itself is usually presented for a duration of less than 100 ms. To make sure that participants are not aware of the primes, a separate verification task is sometimes designed to check it. Some researchers believe that human perception can adapt to a certain mode of stimulus presentation, which means that data received from a verification task presented before an experimental task might not be reliable enough. Therefore, it is recommended to perform a verification task either during the experiment (e.g., with experimental and verification trials alternating during the course of a session) or after the whole experiment is over (Hirshman & Durante, 1992). The duration of presentation of stimuli in a verification task should be identical to the duration of presentation of stimuli in the experimental task. There are several ways to design a verification task: it can be a forced choice task, in which participants are presented with a prime and then have to choose a stimulus matching it from a list of several alternatives (Brown & Hagoort, 1993), a lexical decision task, in which participants have to guess whether the prime that has just been shown is a word or a nonword (Cheesman & Merikle, 1986), a free recall task, in which participants have to report the whole prime or a part of it on their own (Hirshman & Durante, 1992), or a semantic categorization task, in which participants have to identify the semantic category of the prime that has just been presented (Ortells et al., 2016). In case of successful masking, participants typically do not perform above chance level in these kinds of prime verification tasks (Brown & Hagoort, 1993; McNamara, 2005). Still, even though participants do not perceive primes consciously, there can be relatively small, but statistically significant priming effects observed under these conditions (Garman, 1991; Lucas, 2000).

Since Experiment 1 is a property verification task, which means that participants would have to react not only to concept labels (targets), but also to take attributes (primes) into consideration, it would be necessary to ensure that they have enough time to read both primes and targets and to be able to consciously match them. This, in turn, means that the priming effects we would encounter in the congruent condition as a result of this procedure would be strategic in nature.

1.3. Previous Research on Conceptual Information Processing

There is a number of studies dealing with time course of encoding or processing of information on different levels. For example, psychologists investigated the sequence of gender and attractiveness processing in face perception using a combination of a dual-choice task and a go/nogo paradigm (Carbon, Faerber, Augustin, Mitterer & Hutzler, 2018), whereas linguists used similar methods to find out the time course of semantic and phonological encoding in the mother tongue (van Turennout, Hagoort & Brown, 1997; Schmitt, Münte & Kutas, 2000; Abdel Rahman & Sommer, 2003; Rodriguez-Fornells, Schmitt, Kutas & Münte, 2002; Abdel Rahman, van Turennout & Levelt, 2003; Chiu, 2012) or in the second language (Guo & Peng, 2007), some studies also investigated the time course of semantic and syntactic encoding (Schmitt, Schiltz, Zaake, Kutas & Münte, 2001), etc. As far as the methodology is concerned, electroencephalography (EEG) was used in most of such studies due to its high temporal resolution that allows to determine, which type of information becomes available earlier. Mainly two event-related potentials were used for that purpose, namely the lateral readiness potential (LRP) and the nogo-N200. An LRP is contralateral to the response hand and emerges as soon as the information that is used to determine which hand must be used for a response becomes available. A nogo-N200 is an event-related potential which serves as an indicator of response inhibition in nogo trials, i.e., when a response needs to be withheld. This ERP component is usually observed at frontocentral sites of the scalp and its peak latency (in some studies – its onset latency) marks the time point at which the conditions for withholding a response are met, i.e., when the corresponding information is available (Schmitt et al., 2000). It was also discovered that a nogo-N200 seems to be elicited by visual, but not by auditory stimuli (Falkenstein, Hoormann & Hohnsbein, 1999). Another ERP component that is mentioned in some studies and observed in nogo trials is P300 which is considered to be related to inhibition of an overt response as opposed to "inhibition of a premature response plan" indicated by N200 (Gajewski & Falkenstein, 2013: 278). Some researchers used only LRP (van Turennout et al., 1997), whereas more researchers used a combination of LRP and N200 effect for their studies (Schmitt et al., 2000; Schmitt et al., 2001; Carbon et al., 2018). Nevertheless, it is believed that an N200 effect alone can be sufficient to draw conclusions about the time course of certain processing stages (Guo & Peng, 2007). Its amplitude can be influenced by several factors. For instance, Benikos, Johnstone and Roodenrys (2013) investigated the influence of task difficulty on N200 by manipulating the time available for executing a response. Another factor which influences the amplitude of the above-mentioned event-related potential is the proportion of go and nogo trials in a block. Some researchers state that a nogo-N200 is larger, if nogo trials constitute less than 50% from the overall number of trials in a block (Enriquez-Geppert, Konrad, Pantev & Huster, 2010; Benikos et al., 2013). Moreover, participants' age could also play a role in terms of N200 latencies, since, according to Neumann, Obler, Shafer and Gomes (2008), older participants have later N200 latencies in comparison with younger people, which is why it is important to have a more or less homogeneous group of participants, in order to avoid excessive variance in the data.

Most of these studies aimed at investigating the sequence of information processing on different levels and not within one and the same level, but Abdel Rahman and Sommer (2003) manipulated the difficulty of semantic feature retrieval by using the go/nogo paradigm with two distinct types of semantic information and compared it to the speed of retrieval of phonological information. In their experiments, Abdel Rahman and Sommer used images depicting different views of eight kinds of animals which could be categorized by the size of the animal (small vs. large), its diet (herbivore vs. carnivore), or the initial sound of its name (consonant vs. vowel). In the first experiment, their participants had to categorize images by their semantic parameters (size or diet), while the initial sound was used to determine the go/nogo decision, i.e., whether a response should be made at all. The LRP signal determined by size was elicited earlier than the LRP signal determined by diet, indicating that information about the size of animals was available earlier than information on their diet. Also, there was a nogo-LRP signal detected for the easy semantic condition (i.e., size classification), whereas in the difficult semantic condition (i.e., diet classification), no nogo-LRP was observed, meaning that information about whether an animal was a herbivore or a carnivore was available after the information about the initial sound. In the second experiment, the go/nogo decision was determined by semantic information, while the response itself depended on phonological information. Reaction time in the easy semantic condition was 29 ms shorter than in the difficult semantic condition, LRP onset was similar in both conditions, but N200 latency was shorter in the easy semantic condition than in the difficult semantic condition, confirming that information about size of the animals was available earlier than information about their diet. No difference in error rates was observed between the two semantic conditions in either of the experiments. That means that it is possible to investigate the difference in time of activation of semantic attributes and design an experiment to determine whether prominent properties of a certain semantic category are available earlier than less prominent ones.

Previous studies on semantic memory retrieval show that behavioral methods, too, have been successfully used to analyze processing of conceptual information and semantic memory retrieval. For example, Ashcraft (1976) used a verification task with high-dominance and low-dominance property statements. In that study, property dominance was defined as “the position that some properties or attributes of concepts dominate the characterization of those concepts, in that they are more important, frequent, necessary or otherwise salient with respect to those concepts” (Ashcraft, 1976: 490). Ashcraft compared property dominance to the notions of relatedness and semantic similarity, while investigating the interaction between attribute dominance and priming between sentences. He collected normative data on conceptual properties from 35 people and used the frequency of occurrence of each property in response to a given concept as a measure for determining dominance of every property within a particular conceptual structure: properties mentioned by at least 50% of the participants were considered high-dominant (e.g., “*sparrow has wings*”), while properties mentioned by 25% of the participants or less were classified as low-dominant (e.g., “*sparrow has feet*”). In

Ashcraft's (1976) validation study, telegraphic sentences consisting of 3 words (a noun at the top, a connective verb in the middle, and a property at the bottom) were presented in lowercase letters on transparency paper using 35-millimeter slide frames and rear-projection triggered by participants pressing a button with their non-dominant hand. The participants then had to indicate whether the sentences they were presented with were true or false. In the subsequent experiment Ashcraft (1976) described a priming experiment, in which participants were presented with two sentences in each trial and had to verify them, and property dominance was manipulated both in prime sentences (between-subjects factor) and in target sentences (within-subjects factor), and priming effects were expected to differ depending not only on the dominance of the property expressed in the target sentence, but also depending on the dominance of the property expressed in the prime sentence. An inter-trial interval of 5 seconds was implemented. The results showed that low-dominance properties were verified more slowly than high-dominance properties, but low-dominance target sentences were not prone to priming effects irrespective of the dominance of the property in the first sentence, whereas high-dominance target sentences were equally primed by both high- and low-dominance properties in the first sentence. Having investigated priming effects in the absence of stimulus repetition, Ashcraft (1976) designed an experiment in which repetition priming was compared to semantic priming and was investigated with one or four unrelated trials intervening between the prime and the target sentences. The number of intervening trials was a between-subjects factor utilized, in order to find out how quickly priming effects would decay depending on the prominence of the properties. In the semantic priming condition, the first (prime) sentence described a low-dominance property, whereas the second one could be about a low- or a high-dominance property. Sentences with high-dominance properties were again found to have the largest priming effects which decayed slower and were still significant even after 4 intervening trials. Low-dominant sentences could only be primed by repetition. The results of the validation study as well as two subsequent experiments revealed that high-dominance properties were verified more quickly than low-dominance properties. It was taken as evidence for "a valid frequency-to-semantic dominance relationship which characterizes information accessibility in semantic memory" (Ashcraft, 1976: 493). However, in that study, the difference in the frequency of occurrence was relatively large and the difference in attribute types was not taken into consideration. The author also did not consider the time it took participants to reject incongruent properties, concentrating instead on the positive trials.

Katz (1981) also conducted a property verification task using such factors as property dominance and typicality. Property dominance was defined by how often a property was elicited to describe a particular object, whereas typicality depended on whether a given concept is a good example of an object having a certain property, e.g., “a dime is a good example of a round object whereas an apple is a much poorer (or atypical) example of something round” (Katz, 1981: 40). While these two parameters are independent of one another and can be isolated from each other in an experimental setting, they still show a relatively high correlation (r=0.46), with high-dominance properties usually being more typical. In the first experiment described by Katz (1981), participants had to verify simple sentences presented for 5 seconds. In those sentences, different combinations of dominance and typicality levels were represented:

a) high dominance and high typicality: *A globe is round.*

b) high dominance and low typicality: *A barrel is round.*

c) low dominance and high typicality: *A dime is round.*

d) low dominance and low typicality: *An apple is round.*

The results revealed that true sentences were verified more quickly than false statements (1585 ms vs. 1704 ms) and highly dominant properties were verified 81 ms more quickly and more accurately than low-dominant properties, while highly typical properties were verified 289 ms more quickly and more accurately than atypical properties. The interaction between these two factors revealed that dominance effects are mostly observed for atypical properties.

In the second experiment, participants heard the name of an object followed by a visual presentation of a property (or vice versa) and had to verify the property. To compile the stimuli, 5 properties were paired with 4 nouns. In 10 of the resulting word pairs, the dominance was kept at the medium level, 5 of the pairs were of high typicality and 5 pairs were of low typicality. In 10 other word pairs, typicality was kept at the medium level, whereas dominance was high in 5 pairs and low in 5 other pairs. The results revealed that the difference in reaction times between high and low typicality items was larger when a prime was a property than when a prime was an object. The difference in reaction times between high and low dominance items was larger when the stimuli were presented in reversed order (i.e., when the prime was an object and the target was a property).

Katz (1981) argued that the influence of these factors was based on the asymmetric accessibility of objects and properties: property dominance reflected object-to-property accessibility, while property typicality reflected property-to-object accessibility. This was proven by the second experiment described in the same article, in which participants performed a property verification task in which the order of presentation (properties first or objects first) was manipulated and either typicality or dominance of the properties was held constant to determine the order of presentation of the stimuli for which typicality or dominance have an effect on verification latencies and accuracy. In contrast to the current thesis, however, those experiments concentrated mainly on the positive responses, i.e., on trials in which the participants had to confirm a congruent attribute rather than reject an incongruent one. Moreover, neither attribute types nor semantic categories of the objects were considered in that study.

A potential problem with the studies by Ashcraft (1976) and Katz (1981) could be that they mainly concentrated on trials with positive responses, meaning that the results of such analyses are bound to be confounded by priming effects. To avoid a similar issue, negative trials are going to be of particular interest in the experiments described in the current thesis, since incongruent attribute values do not have a direct connection to the concept node in question, but participants would still need to access conceptual information, in order to reject an incongruent attribute.

1.4. Visual Recognition of Words, Mental Lexicon and Lexical Access

Experiment 1 utilizes verbal stimuli, which means that visual word recognition and lexical access are going to be relevant for the current thesis. Whenever participants are presented with visual verbal stimuli, the meaning of the presented stimuli is automatically retrieved. In order to access the meaning of a given word, first of all, this word has to be recognized. According to Taylor (1990), a familiar word is typically perceived as a whole, while unfamiliar words are recognized on a letter-by-letter basis. Depending on the circumstances, there are two types of processes involved in word recognition: bottom-up processing describes recognition of words based upon their perceptual features (e.g., vertical or horizontal lines, circles, etc.), while top-down processing involves context and world knowledge that help to form a set of possible word candidates, which are then checked against perceptual information about the word in question (Taylor, 1990).

The speed of word recognition can be influenced by various factors. Firstly, it depends on the type of stimuli which are being used. Reaction times are shorter for words in comparison with pseudowords and nonwords, while nonwords, i.e., letter strings containing letter combinations which are illegal in a given language, are rejected more quickly than pseudowords (Coltheart, Davelaar, Jonasson & Besner, 1977).

Another important factor is word frequency. It is taken into consideration by the most prominent models of word recognition, such as the serial search model, according to which words in the mental lexicon are ordered by their frequency (Forster, 1994), and the logogen model, in which each logogen, i.e., each mental device representing a certain word, has an individual threshold that depends on the word's frequency (Morton, 1969), and the interactive activation model, in which resting level of activation of a node representing a certain word is determined by the word's frequency (McClelland & Rumelhart, 1981). These assumptions attempt to explain faster reaction times for high frequency words. Word frequency can also interact with some other factors. For example, younger participants are said to have larger word frequency effects than older participants (Balota, Cortese, Sergent-Marshall & Spieler, 2004). Moreover, some researchers claim that word frequency causes particularly large effects in a lexical decision task (Balota & Chumbley, 1984; New, Ferrand, Pallier & Brysbaert, 2006). Taking these findings into consideration, it would be necessary to match the sets of stimuli to each other based upon their frequency, in order to eliminate unnecessary effects caused by this factor.

One more factor that can exert an influence on word recognition is the length of words. It can be measured in different ways, e.g., in letters or in syllables. Its influence as an independent factor is inconsistent, since some studies found effects that were explained by the length of words alone (Balota & Chumbley, 1984; New et al., 2006; Whaley, 1978), while others failed to reveal pure length effects (Hauk & Pulvermüller, 2004). New et al. (2006) found that number of letters and number of syllables affect reaction times in different ways. Particularly, the more syllables there are in a word, the longer it takes to recognize it. However, if word length is defined by the number of letters, recognition of short and long words is slower than recognition of words of a middle length, and this finding was independent of the number of syllables and grammatical category of words.

Some researchers also found a significant interaction between the length of words and their frequency, such that longer words were recognized slower, if their frequency was low (Balota et al., 2004). On the other hand, Hauk and Pulvermüller (2004) found no such interaction between word frequency and word length in behavioral data of their experiment.

Moreover, the speed of word recognition is also influenced by case effects. For example, Mayall and Humphreys (1996) reported that lexical decision task was subject to these effects to a greater extent than semantic classification or word naming, and reaction times for words which were written in mixed-case letters were longer than for words written completely in upper-case or completely in lower-case letters. As a way to control for this variable, all attribute values and concept labels in Experiment 1 of the current thesis were written only in upper-case letters.

The age of acquisition of a particular word can also affect the speed of perception of this word, affecting reaction times in a lexical decision task more than in a naming task (Cortese & Khanna, 2007) and causing larger effects for low-frequency words than for highly frequent stimuli (Bonin, Chalard, Méot & Fayol, 2001; Gerhand & Barry, 1999). As far as controlling for this factor in the current study is concerned, the experiments were conducted in German and all participants were native speakers of this language, which means that the age of acquisition of each word used as stimuli was as similar across participants as possible.

Other variables that are said to influence the speed of word recognition include orthographic neighborhood size and neighborhood frequency (Perea & Rosa, 2000). New et al. (2006) emphasize the importance of neighborhood size along with word frequency and word length. Balota et al. (2004) found facilitatory effects of a large neighborhood size for low-frequency words and inhibition for high-frequency words; this variable was also predicted by some word recognition models, for example, the serial search model by Forster (1994) and the interactive activation model by McClelland and Rumelhart (1981). However, Coltheart et al. (1977) found effects caused by neighborhood size only in their nonword data, such that participants had shorter reaction times when responding to nonwords with a large number of orthographic neighbors, while lexical decision latencies in case with real words were not affected by this factor.

For example, in one of the studies by Balota et al. (2004) the following set of variables was used: "phonological onsets, length in letters, orthographic density, objective frequency, subjective frequency, feedforward onset consistency, feedforward rime consistency, feedback onset consistency, feedback rime consistency, imageability, meaningfulness, number of associates, and estimates of semantic connectivity", aiming to compare their influence in naming and lexical decision tasks with participants from different age groups (Balota et al., 2004: 285). According to their study, the variables mentioned above were less predictive for older participants than for younger ones. Balota et al. (2004) also came to the conclusion that word frequency had a better predictive power when it came to the lexical decision task as compared to the naming task. It was also found that the degree to which the speed of word recognition was influenced by these factors depended on the type of task that was implemented in the corresponding studies. For example, according to Balota et al. (2004), semantic relatedness happens to have a larger effect on reaction times in the lexical decision task than on naming latencies in the speeded naming task.

There are several views on how the process of word recognition is organized. According to the view proposed by Chumbley and Balota (1984), word recognition is subdivided into two stages: firstly, a lexical entry in the mental lexicon is found and, secondly, the information about the corresponding word is retrieved from the accessed lexical entry. A similar view was supported by Seidenberg (1985), who divided the process of word recognition into three phases, namely a prelexical phase, in which an input is identified as a word and a lexical entry becomes activated; a lexical phase involves retrieval of information from the available lexical entry (e.g., the meaning of the word, its phonology, orthography, etc.), and an additional postlexical phase, at this level the information is integrated, in order to understand the whole context, in which the word occurs. According to another view that is called distributed semantic information, a familiarity index is assigned to each word in the mental lexicon of a person. Activation of any property of a word leads to an increase of this index and, when it reaches a certain threshold, the word itself becomes activated and, in turn, activates other words which have some properties in common with it (Taylor, 1990).

As soon as a word is recognized, its entry in a person's mental lexicon is accessed. An entry in a mental lexicon represents a collection of information that is relevant for the use of a particular word. According to Levelt (1991), such an entry consists of the meaning of the word in question accompanied by its syntactic, morphological and phonological properties (i.e., category of speech, syntactic arguments or roles it can take in a sentence, information about its grammatical form and phonological structure, etc.). Semantic and syntactic parts of a lexical entry comprise a lemma. The elements of a lexical entry are interconnected and may also include additional information, such as connotation of the corresponding word. Representations of various words in the mental lexicon are connected by two types of relations, namely intrinsic, which are based upon characteristics shared by words, and associative, based on how often certain words are used together in context. Intrinsic connections could be found between the words that are linked based upon their meaning, morphology (e.g., words sharing the same stem), or phonology (e.g., when certain sounds at the beginning or at the end of words are the same). These connections between word representations constitute the basis for priming effects, which implies that the speed of word recognition can also be influenced by preceding stimuli (Levelt, 1991).

The Lexical Quality Hypothesis (Perfetti, 2007) states that the speed and accuracy of word recognition is based on the quality of a number of criteria that include the word's orthography, phonology, grammar (morpho-syntax), meaning, and constituent binding, i.e., how tightly these elements are interconnected within a word's representation. For a given word, each of these criteria can be of low or high quality. The more specific and constant the information about a certain aspect of the word is (e.g., its spelling, pronunciation, etc.), the higher the quality of the corresponding criterion, which in turn leads to a more coherent and stable identity of the word in question and, as a result, more reliable word recognition results. Reading experience contributes to the reliability of this process by increasing relative frequency of word candidates and, therefore, ensuring faster and more effective word recognition and comprehension (Perfetti, 2007).

The mechanism of how a word is identified and how errors might be made in this process was described in the multiple read-out model (Grainger & Jacobs, 1996). According to this model described through the lens of a lexical decision task, there are three criteria playing a role in word recognition, namely the activation of a given word unit M (which is supposed to be a fixed parameter, since lexical access is automatic and cannot be slowed down or sped up at will), the total amount of activation in the lexicon Σ (its value is determined by the distribution of activation evoked by all word and nonword stimuli throughout the experiment), and time limit T (which depends on the task requirements concerning speed and accuracy). If M or Σ is reached before temporal deadline T, a positive response is given; otherwise, a negative response is given. Meanwhile, the speed of word recognition is determined by the earliest moment in time when either M or Σ criterion is reached. An error in response to a word (a false negative) is made, if T is set too low (i.e., if the time available for making a response is too short) or if M or Σ are set too high (i.e., if the amount of activation necessary for making a decision is not reached before the time limit). An error in response to a nonword (a false positive) is made under the opposite circumstances: if T is too high or if M or Σ are too low (Grainger & Jacobs, 1996).

1.5. Color Perception and the Role of Color in Object Recognition

One of the attributes investigated in this dissertation is color. Color perception plays a significant role in object recognition, allowing us to identify multiple objects in a single scene as well as to determine functionally important changes in the state of an object depending on its color variations, such as bleaching, aging, ripening, or rotting. The importance of such changes is based on the fact that they could potentially provide some information not only about the identity of an object, but also on its condition and its suitability for a particular purpose, e.g., whether a particular fruit is edible or not (Witzel & Gegenfurtner, 2018).

Color perception is essentially a process of interpretation, because, even though color is considered to be a property of an object, it is not just an external phenomenon based on physical parameters of an object's surface and illumination, rather it is constructed by the brain, which is illustrated by multiple visual illusions, such as the lilac chaser illusion (Hinton, 2005), in which a circle of blurred lilac discs is presented around a black fixation point on a grey background, and the lilac discs disappear one at a time in a clockwise direction, such that a viewer would see a gap in the circle of lilac discs running clockwise, and after some time one sees a green blurred disc instead of a gap, since green and lilac are complementary colors. This becomes possible due to a "misinterpretation" of color by the brain as a result of the so-called chromatic adaptation (Hurlbert, 2021).

Color is often considered to be distinct from other types of attributes of concrete objects. In particular, there are two phenomena which account for this fact, namely color diagnosticity and color constancy (Redmann, FitzPatrick, Hellwig & Indefrey, 2014; Witzel & Gegenfurtner, 2018; Hurlbert, 2021).

Color diagnosticity refers to the role colors play in the conceptual structures of various objects. Having encountered or handled an object multiple times, we memorize its typical color, which is also referred to as the memory color of that object. The notion of memory color was introduced by Ewald Hering in 1878 along with the idea about its influence on the perception of an object's color (Witzel & Gegenfurtner, 2020). Some researchers use the notion of memory color to define color diagnosticity, such that "objects with a memory color are called color diagnostic; objects that do not have a memory color are color neutral" (Witzel & Gegenfurtner, 2018: 479). Other authors also might refer to color neutral objects as low color-diagnostic objects. In this case a slightly different definition is used, according to which concrete objects can be classified into two categories: high color-diagnostic objects and low color-diagnostic objects. For high color-diagnostic objects, there is a typical color which is strongly associated with them (e.g., a typical color for a banana is yellow), while low color-diagnostic objects, such as cars, could have a wide range of colors and, therefore, none of these colors would be considered typical for a given object (Redmann, FitzPatrick, Hellwig & Indefrey, 2014).

As has been mentioned above, another important phenomenon connected to colors is color constancy. It is "a perceptual phenomenon, by which people perceive object colors as staying stable under changes in illumination" (Hurlbert, 2021: 69). For example, if a human observer has to match the color of a banana under illumination of different colors, there is a bias towards the memory color of the object, in this case the banana would seem more yellow than it actually is. This is one of the examples of the phenomenon of perceptual constancy, which is characteristic not only of color, but also of some other perceptual properties, e.g., size constancy allows us to estimate the real size of objects as they are moving away instead of thinking that they actually get smaller as they do so. All in all, perceptual constancy as such and color constancy in particular help humans transform two-dimensional perception into a three-dimensional stable representation of the world around us (Hurlbert, 2021). Nevertheless, it is important to remember that color constancy does not prevent people from detecting subtle color changes as such, since we are still aware of the changes in an object's color caused by different types of illumination, instead it only introduces a certain degree of bias towards the memory colors of the objects in question (Hurlbert, 2007).

Another study that investigated the connection between color information and object knowledge was the study by Mitterer and de Ruiter (2008). In their article, the authors describe two experiments aiming to investigate the influence of object knowledge on color perception, using line drawings of prototypically yellow objects (e.g., a banana), prototypically orange objects (e.g., a carrot), or neutral objects (e.g., a sock). The objects were presented in either of 7 hues: yellow, orange, or one of 5 ambiguous hues between those two values. In the first experiment, the participants were first divided into two groups, one of which was presented with prototypically yellow objects in yellow and prototypically orange objects in an ambiguous color and, therefore, was supposed to develop a bias towards categorizing the ambiguous color as orange. The second group saw prototypically orange objects in orange and prototypically yellow objects in an ambiguous color, and thus was supposed to develop a bias towards categorizing an ambiguous color as yellow. They subsequently had to categorize an ambiguous color presented on a color-neutral object. The expected biases were confirmed by the results of the experiment. In the second experiment, Mitterer & de Ruiter (2008) replaced color-diagnostic objects in the exposure phase with color-neutral objects, in order to test for the adaptation effect: if the results of the previous experiment were not caused by top-down influence of object knowledge, but rather were due to the fact that an ambiguous color was perceived as different from the unambiguous yellow or orange (depending on the group), then the results should be similar to the results of the first experiment. However, instead they found no such biases, which means that the findings of their first experiment must have been caused by the top-down influence of object knowledge. The authors, therefore, came to the conclusion that "observers use knowledge of an object's color to recalibrate their color categories" (Mitterer & de Ruiter, 2008: 633).

As far as the current thesis is concerned, this might have some implications for the experiments in which pictorial stimuli are utilized. The items used in those experiments are relatively high color-diagnostic objects, i.e., most of them have one typical color, for some only a very limited number of colors are possible and the most frequently mentioned one is represented in the picture (further information about creation of the stimuli is provided in the corresponding sections of the descriptions of the experiments). Taking the phenomenon of color constancy into consideration, it could be difficult for the participants to detect minor color changes not just because of the subtlety of the color change itself, but also because the bias towards the memory color of the object might make them think that the image in the condition with the color altered by 33% looks the same as the original image. As a way to control for that, two separate experiments were designed: in one of them, objects were depicted as a whole, whereas in the other experiment isolated attributes were presented, e.g., all colors were presented in the form of color squares. In this case, subtle color changes are expected to cause slightly higher error rates both in trials with color patches and with whole objects' depictions; however, when an object as a whole is depicted in an image, subtle color changes should be even harder to notice due to the effects of color constancy, leading to even higher error rates in the corresponding condition.

2. METHODOLOGY

In this chapter, methodological aspects of the current thesis are explained. Firstly, the materials, namely feature databases and image databases that were used to create stimuli for the experiments are described, including a short overview of the existing databases and reasons for choosing particular databases as sources of verbal or pictorial experimental stimuli. Secondly, the experimental paradigms, namely the property verification task and the two-alternative forced choice task, utilized in the experiments described in the current dissertation are going to be described. Thirdly, the information about the software used for programming and conducting the experiments is provided. The software in question includes Neurobehavioral Systems Presentation® Software, PsyhoPy and Pavlovia. Finally, some information on statistical analysis is provided, including the software that was used to perform reaction time analyses and error rate analyses as well as the procedure of data filtration and the analyses themselves.

2.1. Feature Databases

A feature database represents a collection of features for a wide range of concepts gathered from a large number of participants that is generally used for psycho- and neurolinguistic studies, e.g., to create stimuli for experiments. An advantage of using such feature databases is the fact that it allows to control for certain parameters of the stimuli, such as production frequencies of the features, their type (e.g., whether it is a visual, behavioral, or another type of feature), distinctiveness for a given concept etc., depending on which parameters the authors provide in their database. There are numerous databases in which various concepts and their semantic features are represented. These feature databases can differ from each other in several parameters which are described in this chapter.

First of all, they can be created in different languages, e.g., English (McRae, Cree, Seidenberg & McNorgan, 2005; Buchanan, Holmes, Teasley & Hutchison, 2013; Buchanan, Valentine & Maxwell, 2019; Devereux, Tyler, Geertzen & Randall, 2014; Vinson & Vigliocco, 2008), Czech (Konečná & Vaňkátová, 2017), Dutch (Ruts, De Deyne, Ameel, Vanpaemel, Verbeemen & Storms, 2004), Italian (Montefinese, Ambrosini, Fairfield & Mammarella, 2013), Spanish (Vivas, Vivas, Comesaña, García

Coni & Vorano, 2017), etc. Since semantic features represent conceptual knowledge that depends not on the language, but on the environment that people live in, this allows us to use a foreign semantic feature database after replacing concept labels and their attributes with the corresponding translation equivalents from the target language.

Another distinguishing characteristic of the databases is the number and age of respondents, from whom semantic features were collected, which allows to investigate conceptual knowledge of a particular age group. For example, in the database by Konečná and Vaňkátová (2017) all participants were children from 8 to 10 years old. However, in the majority of studies, semantic features are collected from adult native speakers (McRae et al., 2005; Buchanan et al., 2013; Buchanan et al., 2019; Devereux et al., 2014; Vinson & Vigliocco, 2008; Ruts et al., 2004; Montefinese et al., 2013; Vivas et al., 2017).

Moreover, the number of concepts in a database can also vary, e.g., from 123 in Devereux et al. (2014) to 4436 in Buchanan et al. (2019). Types of those concepts can also be different. Most databases concentrate on nouns, but the database by Vinson and Vigliocco (2008), for example, includes nouns as well as verbs denoting body motion, light emission, contact, exchange, communication, sounds, sensation, etc., while Buchanan et al. (2013) also have adjectives in their database.

For our purposes, databases were considered that were collected from adults, since this matches the demographic parameters of the participants who would be recruited to take part in the experiments described in the current thesis. Moreover, the feature database also has to offer a sufficient number of stimuli from the semantic categories in question, namely “animals”, “tools”, “fruit and vegetables” (for convenience, the latter subset of items will be hereinafter referred to as “food”), accompanied by the data on production frequency of the features determining the prominence of the attributes. Since the focus of the current thesis lies not on lexical, but rather on conceptual and perceptual aspects of semantic processing, a feature database using a language other than German could be used. As a result, the feature database by McRae et al. (2005) was chosen as a source of the stimuli, which were then translated from English into German. There, defining features for 541 animate and inanimate basic-level concepts were collected from 725 participants. In order to resolve potential discrepancies in terminology, it should be stated that, while feature is a term mainly used in theories like the binary features or the feature list theory, within the frame theory this term corresponds to the notion of an attribute, and the actual features provided in the database, such as “*has legs*” or “*is green*”, would be referred to as attribute values.

2.2. Image Databases

Pictorial stimuli are widely used in studies dealing with attention or cognitive processes, as well as language, particularly when it comes to lexical access, processing of conceptual, syntactic, or phonological information, and the time-course of availability of those types of information (e.g., Schmitt, Münte & Kutas, 2000; Schmitt, Schiltz, Zaake, Kutas & Münte, 2001; Abdel Rahman & Sommer, 2003). To ensure that images used as stimuli in a particular experiment are homogeneous and comparable with each other, image databases are sometimes created, in which certain parameters of the constituent images are listed to make it possible for researchers to control the effects those parameters might cause to arise.

Images can also be used to investigate the effects produced by semantic or syntactic violations, in which case the original images have to be modified to depict the corresponding violations. Usually, such images represent not a single object, but a relationship between an object and a scene, which makes such databases suitable, e.g., for experiments that make use of the visual search paradigm. An example of such a database would be SCEGRAM – a collection of images with various degrees and combinations of semantic and syntactic violations developed specifically for research purposes by the Scene Grammar Lab of the Goethe University in Frankfurt (Öhlschläger & Võ, 2017). There, a semantic violation is represented in an image in such a way that a scene contains an object that is unlikely to be found in that particular environment (e.g., a cup on the toilet paper holder in the bathroom), whereas a syntactic violation is represented in an image, in which an object is found either in an unusual position within a congruent environment (e.g., a cup standing on the door of an opened dishwashing machine) or in a physically impossible position in an otherwise congruent environment (e.g., a cup hovering mid-air in the kitchen).

However, if the aim of a study is to investigate the influence of changes in certain attributes on conceptual processing of concrete objects, introducing a whole scene around that object would increase the visual complexity of the stimuli and create a lot of potential confounding factors. For the purposes of the current study, the stimuli needed to contain objects without any background, so that it would be easier for the participants to concentrate on recognizing a single object in each trial. There is a number of image databases which meet that requirement. One of them is the set of pictures published by Snodgrass and Vanderwart (1980) that includes 260 items "standardized on four variables of central relevance to memory and cognitive processing: name agreement, image agreement, familiarity, and visual complexity" (Snodgrass & Vanderwart, 1980: 174). Nevertheless, this database was deemed unsuitable for the purposes of the current study, since it is comprised of black-and-white line drawings, whereas color attributes are an important part of the research question in this thesis.

Thus, a database of color images would be preferred, in which objects are depicted in isolation, i.e., without any surroundings or other objects in their vicinity. Two databases were found that met those requirements. One of them was the Hatfield Image Test (Adlington et al., 2009) which includes a set of 147 images depicting objects from various semantic categories suitable for experimental and clinical research. The authors also published mean ratings for each item based on the following variables: name agreement, word frequency, age of acquisition, familiarity, visual complexity, and color diagnosticity. Unfortunately, very few items from this database depicted the stimuli chosen from the feature database by McRae et al. (2005) that was used to create stimuli for Experiment 1 and should also serve as a basis for selecting the stimuli for the rest of the experiments. Apart from that, all images in the Hatfield Image Test database had the size of 283×283 pixels, which made it difficult to modify images while maintaining a high quality of the stimuli. Therefore, the Hatfield Image Test database had to be rejected.

The second image database that met the requirements of the current study was the Bank of Standardized Stimuli (BOSS) created specifically for cognitive research purposes (Brodeur, Dionne-Dostie, Montreuil & Lepage, 2010). This set of pictures originally included 480 photographic images of objects and was later extended by adding 930 more pictures (Brodeur, Guérard & Bouras, 2014) along with their norms for name, object agreement, category, manipulability, viewpoint agreement, visual complexity, and familiarity. Apart from that, BOSS also provides line-drawn, grayscale, scrambled and blurred versions of the images. According to the authors, as of 2014, BOSS was "the largest existing photo bank providing norms for more than 15 dimensions" (Brodeur, Guérard & Bouras, 2014: 1). The fact that all images had the size of 2000×2000 pixels would allow to modify the selected items without losing their quality and level of details. Due to a wide variety of the stimuli represented in that databank, many images corresponding to the verbal stimuli of Experiment 1 could be found there. However, some concepts from the feature database by McRae et al. (2005) could not be found in the BOSS, and some of the other concepts had part attributes which were not shown in the images included in that database, e.g., the part attribute of an apple was "had seeds", but a picture of an apple from the BOSS stimuli depicted a whole apple, so the attribute was not visible and impossible to replace. Therefore, some of the images had to be taken from a different source, namely from the database "Freepng" (URL: https://www.freepng.ru), where images with transparent background in the PNG format were made available for non-commercial use. Unfortunately, no feature norms were provided for those images. More detailed information about the images and stimulus list composition in each of the experiments involved is provided in the corresponding sections of Chapter 3 of the current thesis.

2.3. Experimental Paradigms

There are two experimental paradigms which are relevant for the series of four experiments described in Chapter 3 of the current thesis: the property verification task and the two-alternative forced choice task. In particular, Experiment 1 utilized a property verification task with verbal stimuli, Experiments 2 and 3 used a two-alternative forced choice task with pictorial stimuli, and Experiment 4 used a property verification task with pictorial stimuli. In the current subchapter, information about these two types of tasks is provided, such as an explanation of the essence of each of these tasks and some theoretical aspects that needed to be considered while designing the experiments.

2.3.1. Property Verification Task

A property verification task is a type of task in which a concept label and a certain property are presented to participants who then have to decide whether the property belongs to the concept in question.

In this experimental paradigm, properties of objects refer to certain conceptual attributes, e.g., parts, shapes, colors, sounds, and so on. These properties can be classified according to their sensory modality, such as auditory, visual, tactile, etc. For example, a sound typically produced by an animal (e.g., barking of a dog) would be an auditory property, whereas the color of that animal would be a visual property. Modalities of various properties involved in a property verification task can influence participants' performance in terms of reaction times or error rates. According to Collins, Pecher, Zeelenberg and Coulson (2011), the effects of switching from one modality to another in subsequent trials of a verification task differ depending on the modality in question. Even though the analysis of the behavioral data obtained in their experiment failed to reveal any significant effects apart from lower accuracy in verification of auditory properties as compared to visual ones (the authors assumed that it was due to a lack of statistical power caused by a small number of participants), the pattern of reaction times corresponded to the expected results when it came to verification of visual properties (somewhat longer reaction times in the switch condition as compared to the non-switch condition), whereas the reversed pattern was observed for verification of auditory properties; the error rates also tended to be slightly higher in the switch condition. Talking about the switching effects, the authors were mainly referring to the electrophysiological data obtained in the same experiment, in which case switching elicited an N400 effect in the visual modality and a late positivity in the auditory modality, indicating different neural mechanisms engaged in property verification in different modalities.

There are at least three known theories that aim to describe and explain processes occurring during the performance of a property verification task. One of these theories is the feature matching model (Smith, Shoben & Rips, 1974). This model was originally developed to explain the process of verification of sentences stating category membership (e.g., "*A robin is a bird.*"), but the authors stated that this approach could be applied to feature verification as well. According to this model, the process of verification can be divided into two stages: in the first stage, features of a concept and a superordinate category are retrieved and compared to each other and a measure of similarity between them is determined, which is then compared to two reference parameters representing a high and a low level of similarity. If the measure of similarity exceeds the high level, a positive response is given; if the measure of similarity is less than the low level, a negative response is made. If, however, the similarity measure is situated between the high and the low levels, a second stage is required, in which feature weights are used to distinguish between defining and characteristic features, and the defining features of the concept are compared to the defining features of the superordinate category. A positive response is given, if the defining features of the concept include the defining features of the category or "are within the range of allowable values for the category" (Smith, Shoben & Rips, 1974: 223).

Another model is referred to as the semantic network account of property verification (Collins & Quillian, 1969), and it states that semantic information is stored at different levels in a network of nodes: at the level of the concept itself (e.g., the fact that a canary can sing), at a superordinate level just above the conceptual node (e.g., a canary's ability to fly is inferred from the fact that a canary is a bird and birds are able to fly), or at an even higher level (e.g., the fact that a canary has skin is inferred from the fact that a canary is a bird and a bird is a type of animals). Upon encountering a property statement, participants search for the property in the node of the concept as well as in the node of one or more superordinate levels. Depending on whether retrieving this information requires going through properties stored at one, two, or three levels, different reaction times should be observed. This theory was first proposed for storage of information in a computer memory, but Collins and Quillian (1969) tested it for the human memory in a series of three experiments (all of them included one-level sentences representing property statements or superset statements; the first experiment also included two-level and three-level sentences, the second one included a different set of two-level sentences, and the third experiment included the three-level sentences from the first experiment). The findings of these experiments largely corresponded to the expected results, namely that, if the required information was stored at a higher node, it took longer for the participants to confirm such a statement, but, when it came to rejecting false statements, the participants needed more time to give a response, if the information required to make a decision was stored close to the conceptual node in question. The authors assume that this search at different levels of nodes happens in parallel, since it is unlikely that all properties are retrieved from the first level before moving to the superordinate level. If the search is successful, the property is verified, i.e., a positive response is given; if the property is not found in the nodes or if the information stored in semantic memory contradicts the property statement in question, a negative response is made.

A more recent third account states that sensorimotor simulations are used to perform property verification tasks successfully (Solomon & Barsalou, 2004; Pecher, Zeelenberg & Barsalou, 2004). Thus, participants would construct a sensorimotor simulation of a concept and of a particular property, and then they would try to match these two simulations: successful match is followed by a positive response, whereas if the simulations do not match, a negative response is given. The evidence in favor of this approach was the fact that participants were faster and more accurate in their responses in a particular trial of a property verification task, if a feature from the same modality (e.g., visual, auditory, tactile, or taste-related) had been presented in the previous trial.

Neither the feature matching model, nor the semantic network account can explain this effect, since those approaches do not distinguish between various modalities of the properties involved in the task, which is why they are also referred to as amodal views.

To perform a property verification task, participants might use different strategies depending on the instructions received prior to the experiment as well as the type of stimuli involved in it. Some researchers argue that, if participants are presented with verbal stimuli in a property verification task, they might base their decision purely on association between the stimuli without actually accessing the conceptual information the words are referring to, unless those participants have been given particular instructions as to how they are supposed to perform the task (Glaser, 1992; Solomon & Barsalou, 2004). Nevertheless, even if verbal stimuli are used, there is a way to force participants to abandon the word association strategy by pairing the stimuli in such a way that there would still be a certain degree of association in word pairs even in incongruent trials. In this case, participants would not be able to rely solely on association between the stimuli and would have to access conceptual information, in order to give a correct response. A similar method was used by Solomon and Barsalou (2004), when they investigated whether participants used perceptual simulation while accessing conceptual information in a property verification task with part attributes. In their experiment, Solomon and Barsalou presented their subjects with two subsequent words in each trial: in trials in which positive responses were expected, a concept label followed by a word that represented a property of the concept with a high or a low degree of association with its concept (e.g., arrow – shaft), whereas in trials in which false responses were expected, the words denoted a pair of thematically or taxonomically related words (e.g., banana – monkey, table – furniture) or unrelated objects (e.g., asparagus – furniture). The subjects were assigned either a list of stimuli with associated false trials or with unassociated false trials, each of these groups was divided further into two subgroups, one of which was instructed to use mental imagery (perceptual simulation) to verify properties, whereas the other subgroup received neutral instructions with no explicit indication of a required strategy. Solomon and Barsalou reported that the presence of associated stimuli in the false trials slowed down reaction times and increased error rates in the false as well as in the true trials in both subgroups of participants (the one instructed to use mental imagery and the one that received neutral instructions). They also found an interaction between association and instruction, such that there was no significant difference between reaction times of subgroups with different instruction types in the unassociated condition, but when associated words were presented in the false trials, participants instructed to use mental imagery were significantly slower than those who received neutral instructions. The authors stated that it was caused by the fact that people who were instructed to use mental imagery constructed more detailed perceptual simulations when the word association strategy was no longer available. The results indicated that, when the participants were presented with unassociated stimuli in false trials, they tended to use the word association strategy to give a correct response, whereas when the stimuli in false trials were associated with each other, this strategy was no longer considered optimal and mental imagery was used instead (Solomon & Barsalou, 2004).

In Experiment 1 of the current thesis, verbal stimuli were used in a verification task, however, the fact that the same set of properties was re-assigned to different items within the same semantic category to form incongruent experimental trials allows us to conclude that the association between the stimuli might still be preserved to some extent and that, consequently, participants would have to access each concept upon seeing its label, in order to confirm or reject a particular property.

Pictorial stimuli with certain modifications have sometimes been used to investigate perception of semantic inconsistencies. Often, however, such stimuli are depictions of scenes, in which there is a target object that is either congruent or incongruent with a particular environment depicted in that scene. If an object is likely to be found in a given scene, but its position there is considered to be impossible, it is considered to be a syntactic violation. If an object is not expected to be seen in a particular scene altogether, such a picture represents a semantic violation (Öhlschläger & Võ, 2017). In Experiment 4, pictorial stimuli were used, in which objects from semantic categories "animals" and "food" were depicted on a transparent background with their original attributes or with a modified color or part attribute. The participants were asked to classify those images based on whether they represented truthful depictions of the objects in question. Considering that an altered attribute represents a violation of the conceptual structure of a given object, that experiment was essentially a property verification task, in which pictorial stimuli were utilized instead of verbal stimuli. In this case, therefore, a semantic violation is understood as a substitution of an element or an aspect of a picture representing an attribute of the depicted concept.

2.3.2. Two-Alternative Forced-Choice Task

One of the tasks that are widely used in studies involving a perceptual comparison of certain kinds of stimuli is a forced-choice task. In this type of task, participants are expected to make a choice between several stimuli based on a given criterion. The stimuli in a forced-choice task may differ in their modality, e.g., they could be auditory stimuli (Dumay & Gaskell, 2007), visual stimuli, like words (Rothe, Schulte-Körne & Ise, 2014) or pictures (Gere, Héberger & Kovács, 2021), or even tactile stimuli (Chancel & Ehrsson, 2020). The number of stimuli presented in a single trial also varies across experiments: researchers can implement this kind of tasks with different numbers of alternatives, e.g., with 2, 4 or more options to choose from (Gere, Héberger & Kovács, 2021).

There are also some potential problems associated with this type of task in the visual modality. For example, some researchers mention a decision bias as a confounding factor in such studies. Generally, it means that when participants are presented with visually complex alternatives to choose from, the differences between the stimuli might include not only the factor(s) which is (are) manipulated in the experiment, but also other parameters outside the researchers' control, and participants could use these unforeseen factors (or perhaps their individual preferences) as the basis for their decision in a given trial, despite being instructed to use a different criterion. Jogan & Stocker (2014) claim that, at least in a two-alternative forced-choice task, introducing a test item to the two reference items would help to avoid this kind of decision bias. In other words, showing a target picture and two alternatives at the same time, while asking participants to choose an alternative that matches the target picture would help to control for it, since having a target image on the screen at the same time would help participants focus on the parameter in question and make a more objective decision.

Therefore, in Experiments 2 and 3 of the current thesis a two-alternative forced choice task was implemented in such a way that all three images (the target picture and the two alternatives that the participants were expected to choose from) were presented on the screen simultaneously and remained visible until participants gave a response or until a time limit for a trial was reached.

2.4. Software Used to Conduct the Experiments

2.4.1. Neurobehavioral Systems Presentation® Software

Experiment 1 described in the current dissertation was programmed and conducted using Presentation® software (Version 18.0, Neurobehavioral Systems, Inc., Berkeley, CA, URL: http://www.neurobs.com) in the reaction time laboratory of Heinrich Heine University in Düsseldorf. The software used for this purpose allows to use visual or auditory stimuli, while controlling their position on the screen and timing of their presentation and recording information about participants' responses to the stimuli in question.

Scenarios for Presentation® software are written in the programming languages called Scenario Description Language (SDL) and Presentation Control Language (PCL). SDL is used for writing template files which contain descriptions of stimuli and their properties, while PCL is used to implement the scenarios themselves and control the mode of presentation of the stimuli in question. Template files and scenarios can be written in the form of one or several files and they essentially describe all the variables and settings that are relevant for a given experiment, determine the way input and output files are handled, describe all subroutines used within the experiment, state which response devices and buttons are active at which point in that experiment, define placeholders for certain types of stimuli and use the information from the input lists to put the corresponding items in those positions.

In particular, the program for Experiment 1 used input lists in the "txt" format containing verbal stimuli that were presented during the experiment and some information about the stimuli and experimental conditions, which would be relevant at the stage of statistical analysis. At the end of an experimental run log files containing all input information as well as the data provided immediately prior to starting an experimental run (e.g., the number of the list of stimuli used for this particular run, participant's identification number, etc.) and the data collected during the experiment (responses, reaction times) are saved in the "txt" format and are available for further processing and analysis.

2.4.2. PsychoPy and Pavlovia

Due to COVID restrictions, Experiments 2, 3 and 4 described in the current dissertation had to be conducted online, for this purpose PsychoPy and Pavlovia were used (Peirce et al., 2019). PsychoPy is a program which allows to create experiments involving various types of stimuli, including pictorial, verbal and auditory stimuli, while controlling spatial and temporal settings of their presentation. The program also allows to register responses given by participants using a range of devices, such as a microphone, a keyboard, or a mouse. There are three main interface modules in PsychoPy: Builder, Coder, and Runner. PsychoPy Builder is used to create experiments using the aforementioned settings, but without using any code explicitly. However, it is also possible to correct or add some code by using PsychoPy Coder, in which Python is used as a programming language. An experiment created this way can be tested or run locally on a computer using PsychoPy Runner.

In case an experiment is going to be conducted online, Pavlovia is used to run it (URL: https://pavlovia.org). An experiment can be uploaded directly from PsychoPy Builder into a Pavlovia account, where it can be manually assigned a status, depending on the stage of data collection: "pilot" (for a test phase), "active" (as long as data collection is in progress), or "inactive" (as soon as data collection is completed). In order to be able to collect data online, Pavlovia credits have to be purchased according to the expected number of participants. In Pavlovia, it is also possible for an experimenter to choose to save or to discard incomplete datasets collected in the course of an experiment. If an experimenter chooses not to save incomplete results, the data from a participant who failed to complete the experiment is not saved, in which case the corresponding Pavlovia credit will not be consumed for their experimental run, but rather reserved and released again for later use.

As a part of the uploading process, the program translates the Python code into Java Script, therefore, it is important to consider some differences in the syntax of these programming languages, which might lead to potential problems while running the experiment online. Although PsychoPy writes and translates the code of the experiment without the need for any user's intervention, certain symbols or combinations thereof used in lists of stimuli might be perceived by the program as a part of the code, e.g., the differences between symbols "\" and "/" in paths to certain files could play a role in the program's ability to find the files in question, while combinations like "\n" might be perceived as a new line character and disrupt the running program.

For the online experiments described in the current dissertation, PsychoPy Builder was implemented to program three routines for each experiment. Routine 1 was used for presentation of an instruction screen prior to the start of the actual experiment. At this stage, participants were informed about their task and could read a text presented on the screen at their own pace, after which they could indicate that they were ready to start the experiment by pressing a certain button. Routine 2 incorporated all settings for a trial within a particular experiment, including inter-trial intervals, timing of the stimuli and their position on the screen, and a time window for recording information about a keyboard response. A loop was constructed for Routine 2, in order to ensure the same mode and settings of stimuli presentation and response recording across all trials from a given list of stimuli, while keyboard responses and reaction times were registered by the program for each trial. All lists of stimuli were saved in "xls" format and indicated in the settings, such that their selection would be based on a certain parameter typed in prior to the instruction screen, e.g., based on a participant's number or their month of birth. The identity of a particular item was determined by a column in the list of stimuli, in which a path to that item was given. Finally, Routine 3 was implemented, in order to let a participant know that the experiment has come to an end and what they need to do, in order to close the program without disrupting the process of saving the data collected during their experimental run. The datasets from those participants, who completed the experiment, were saved in the "csv" format.

2.5. Statistical Analysis: Software and Procedure

The data of the experiments presented in the current thesis was analyzed with the help of R Studio, a program which utilizes R as a programming language and allows to conduct statistical analyses of large datasets containing multiple variables by using various types of tests and to create various types of graphs to visually present the data in question (R Core Team, 2022).

Before the actual statistical analysis takes place, the dataset needs to be prepared accordingly. The necessary manipulations, such as adjoining separate datasets with each participant's data into a common dataset for the whole experiment, data filtration, extracting, re-naming or adding variables, can be performed using the package "dplyr" (Wickham, François, Henry & Müller, 2022), which provides corresponding commands for these purposes. First, error rates per participant and per item are computed, in order to find out whether there are participants who failed to give at least a critical percentage of correct responses (e.g., 75%) or if there are items which caused particularly high error rates (e.g., >25%) due to not being familiar to the participants. Participants and items with unacceptably high error rates would have to be removed from the dataset, at which point the dataset can be considered ready for the error rate analysis.

Since the variable indicating whether a given response was correct or incorrect can only have one of these two values, a binary regression was used for the analysis of the error rates. It was done by using the "glmer" function and specifying the type of regression by adding "family=binomial" in that command.

To prepare the dataset for the reaction time analysis, further steps of data filtration are necessary. After the participants and items with particularly high error rates have been removed, all erroneous trials are filtered out and mean reaction times and standard deviations are calculated for each participant, each item and for the dataset as a whole. Each participant's and each item's mean reaction time should be within a certain range from the overall mean reaction time (e.g., 3 standard deviations). Participants or items which do not meet this requirement are filtered out. Also, each trial's reaction time should be within the same range from the corresponding participant's mean reaction time. The trials which fail to fulfil this criterion are filtered out, after which the dataset is considered ready for the reaction time analysis.

Reaction time as the predicted variable was checked for normality of distribution using the command "describe" from the package "psych" (Revelle, 2022) that provides statistical information about the variable, including its skewness, which in the case of normally distributed data should be close to zero. If normality of distribution is confirmed, the reaction time variable can be used in the analysis in its original form; however, if the reaction times are not normally distributed, a logarithmic function is used to transform the reaction times and the resulting new variable is used in the models.

Reaction times were analyzed using mixed linear models with the help of the "lmer" function from the package "lme4" (Bates, Maechler, Bolker & Walker, 2015). Apart from the fixed factors which are of interest in each particular experiment, a mixed linear model also includes random factors, such as participant number, item labels or numbers of lists of stimuli, which are included into the command in the form of random intercepts and slopes. Several candidate models with different combinations of random effects, including random intercepts and random slopes, are computed and compared with each other. The optimal random effects structure is selected based on the Akaike's Information Criterion, or AIC, which should be lowest for the best fitting model among the candidates (Bevans, 2022). For this purpose, the "aictab" command from the package "AICcmodavg" is used, which ranks the candidates, provides the number of parameters and an AIC score for each model, as well as the Delta_AIC, which is the difference between the AIC scores of the model in question and the next best model, the AIC weight that reflects the model's predicting power of each model separately, the cumulative AIC weight, and the log-likelihood (Mazerolle, 2020). Alternatively, the "AIC" command from the "stats" package can be used (R Core Team, 2022), however, it only gives an AIC score for each of the candidate models without ranking them or providing additional information. If the difference between AIC values of candidate models amounts to 2 or less, Bayesian Information Criterion (BIC) can be used to check, which model best fits the data. Similarly, "BIC" and "bictab" commands from the package "AICcmodavg" can be used to calculate it. Likewise, a model with the lowest BIC value is considered to be the best fit.

Apart from conducting the actual analysis, R Studio allows to create graphs representing the effects found in the course of the analysis using the package "ggplot2" (Wickham, 2016). In order to select or compute the data that is supposed to be represented in a graph, one can use the commands "summarySE" from the "base" package or "aggregate" from the "stats" package (R Core Team, 2022) and then use the resulting tables to build a graph. The graph itself is built layer by layer using the "ggplot" command that includes various elements determining the data that has to be represented, the type of the graph, its color or a combination of colors based on the levels of a certain variable, faceting, error bars, legends, labels, etc. If necessary, several graphs can be combined into a single figure using "ggarrange" command from the "ggpubr" package (Kassambara, 2020), which allows to choose settings for the alignment of the graphs, their common and separate labels, legends, etc.

The Open Science Framework, or OSF (Foster & Deardorff, 2017), was used to upload supplementary materials, including the scripts containing all commands used for statistical analysis of the data described in the present thesis (Mnogogreshnova, 2023, URL: https://osf.io/3fbn2/?view_only=ee43eb8a793e408ab8d6b504bc4daaff). These supplementary materials are also accessible via a QR code (see Appendix).

3. EXPERIMENTS

In this chapter, the information about a series of four experiments is provided, which were conducted to investigate the role of attribute prominence in conceptual processing. Each experiment is described in a separate subchapter that includes several sections with the information on participants in the experiment, on creation of stimuli and composition of lists of those stimuli, and on the procedure of the experiment. Then the results of the analysis are described, followed by a discussion of the findings.

3.1. Experiment 1

In Experiment 1, a property verification task with verbal stimuli was used, in which participants saw an attribute followed by a concept label and were asked to respond via a button press, whether the attribute they had just seen belonged to the subsequently presented concept. The concepts involved in this experiment belonged to one of three semantic categories: animals, food, or tools. For each semantic category, two types of attributes were selected that were classified as primary or secondary for a given semantic category based on the respective attribute's production frequency (i.e., number of times the attribute was mentioned in response to its concept label) provided in the feature database by McRae et al. (2005), which served as a source of the stimuli. Half of the trials contained congruent attribute values, and the other half contained incongruent attribute values taken from other members of the same semantic category. The hypothesis was that attribute prominence would affect the results in such a way that reaction times would be shorter in trials with primary attributes than in trials with secondary attributes both in the congruent and the incongruent condition, reflecting that primary attributes are assumed to be accessed prior to secondary attributes. The fact that there is a direct semantic connection between a congruent attribute value and its concept would mean that an effect of attribute congruence should be expected for primary as well as for secondary attributes, since this semantic connection would give rise to priming effects, given that congruent attribute values serve as primes and concept labels serve as target words. This congruence effect is probably going to be larger for primary attributes, given that, according to the frame theory, such attributes are more strongly connected to the concept nodes. However, incongruent attribute values are not a part of the conceptual structure of the subsequently presented concept, which would mean that reaction times in the incongruent condition would be more representative of the order in which the corresponding attributes are accessed, therefore, it would still take participants less time to reject an incorrect primary attribute than an incorrect secondary attribute. Therefore, is would be necessary to analyze the data from congruent and incongruent trials separately, to obtain more detailed results in both conditions, and pay particular attention to the incongruent condition when discussing the role of attribute prominence in retrieval of information from semantic memory. In addition, participants were also expected to make less errors in trials with primary attributes than in trials with secondary attributes both in the congruent and in the incongruent condition due to their prominence within the conceptual structure, since a stronger connection between a primary attribute and a concept node would be likely to make primary attribute confirmations or rejections less prone to errors.

3.1.1. Methods

3.1.1.1. Participants

Forty students (8 lefthanders; 24 females; age: from 19 to 35 years, mean age: 23.75 years) from the University of Düsseldorf took part in Experiment 1 and were paid for their participation. All participants were native speakers of German and reported to have normal or corrected-to-normal vision. The study was approved by the local ethics committee of Heinrich Heine University Düsseldorf (study number 2019-436-Drittmittel).

3.1.1.2. Stimuli

The stimuli for the experiment were taken from the feature database compiled by McRae et al. (2005). Words denoting concrete concepts from three semantic categories were chosen from this database: animals (93 items), food (48 items) and tools (85 items). In each semantic category, two attributes that are present in all or in the majority of concepts were investigated: the chosen attributes were part and color attributes for animals and food items, and affordance and part attributes for tools. For each concept, one value of each attribute was selected based upon its naming frequency, i.e., in each case, the most frequent value of an attribute was chosen except for one item ("donkey"). For that particular item, the value "has legs" had the highest production frequency, namely 19, however, when the attribute values had to be re-assigned to different concepts within the same semantic category to create stimuli for the incongruent condition, that value proved to be very common among concepts from the semantic category "animals". As a result, the second most frequent value "has 4 legs" with the production frequency of 18 was selected for the concept "donkey" to ensure successful re-assignment of attribute values for the incongruent condition. The database often provided several values for a single attribute, e.g., the same animal could have legs, ears, and a tail, which would all be values of its part attribute. The attribute with the highest sum of naming frequencies of its values represented the primary attribute, whereas the attribute with a smaller sum of production frequencies represented the secondary attribute. For the semantic category "animals", the part attribute was primary and the color attribute was secondary; for food items, color was primary and the part attribute was secondary; for tools, affordance was the primary attribute, while the part attribute was secondary. Table 1 provides information on the total sum of production frequencies and the proportion of production frequencies of the primary and secondary attributes for each of the semantic categories involved.

Table 1. Production frequencies for primary and secondary attributes in semantic categories "animals", "food", and "tools".

Semantic category	Attribute	Sum and proportion of production frequencies
animals	all	12284
	part	3137 (25.5%)
	color	1257 (10.2%)
food	all	6639
	color	1105 (16.6%)
	part	920 (13.9%)
tools	all	8253
	affordance	2514 (30.5%)
	part	1502 (18.2%)

However, in each semantic category there were some items that deviated from the general pattern of their semantic categories, meaning that their most prominent feature was not the most prominent feature for most of the items from their semantic category.

For example, for most items from the semantic category “animals” body parts constitute a more prominent feature than color, but the item “*zebra*” shows a deviating pattern: total production frequency of its color attribute values (“*is black*”, “*is white*”, “*has stripes*”) is 59, which is higher than the total production frequency of part attribute values (features “*has 4 legs*”, “*has a mane*”, “*has a tail*” and “*has hooves*” were mentioned in total 49 times). All in all, there were 13 deviating items in the semantic category “animals”, 12 deviating items in “food” and 13 deviating items in “tools”. Thus, there were two factors reflecting attribute prominence: attribute prominence for the semantic category and attribute prominence for individual items, which allowed to compare them and determine which attribute prominence is a better predictor when it comes to reaction times and error rates.

Since the original database was compiled in English and the current experiment was conducted in German, the selected stimuli had to be translated. In the process of translation, several issues arose. Several items had to be deleted from the list due to the fact that their translations were identical: from each pair of similar words only one word was selected (e.g., “*dove*” and “*pigeon*” – “*Taube*”, “*elk*” and “*moose*” – “*Elch*”, “*turtle*” and “*tortoise*” – “*Schildkröte*”). It should also be noted that attribute values and their production frequencies for both concepts in each pair were similar. Apart from that, due to differences between the languages, some changes had to be made to features as well. Thus, an inconsistency was found in the original database, namely that the features “*has skin*” and “*has a skin*” were used to denote the same property of several concepts. This problem was solved by using a German equivalent “*hat eine Haut*” in all corresponding cases. In addition, differences between the languages also led to overlaps between features in some cases. For example, the feature “*has a shell*” has different translation equivalents depending on the concept in question: “*hat eine Schale*” for the concept “*nut*” and “*hat einen Panzer*” for the concept “*turtle*”. Moreover, the feature “*hat eine Schale*” is also used as a feature of the concept “*banana*” (the original feature was “*has peel*”).

The final set of stimuli included all three semantic categories with the following number of items: animals – 76 items, tools – 60 items, food – 40 items. As it has already been mentioned, two attributes were selected for items from each semantic category. For each concept, one most frequently given value of each attribute was chosen to create stimuli for congruent trials in which participants would give affirmative responses indicating that these values belong to their concepts. In order to avoid any response bias, an equal number of trials with incongruent attribute values was constructed, to which negative responses were expected, by randomly re-assigning attribute values within each semantic category, such that every concept would be paired with an incongruent value of every attribute.[1]

The stimuli were divided into two lists. Every concept was mentioned in each list twice (with a congruent value for one of its attributes and with an incongruent value of the other attribute), e.g., one group of participants would see an animal paired with a congruent color and an incongruent part, while another group would see the same concept with an incongruent color and a congruent part. As a result, in 50% of trials a negative response was expected. The program "Mix" (van Casteren & Davis, 2006) was used to pseudorandomize the stimuli within each list. For this purpose, a script file was written containing the following constraints for randomization: identical concepts or identical features must be at least 10 items away from each other, trials with the same kind of expected response or with items from the same semantic category must not be repeated more than 2 times in a row. Then those two lists of stimuli were reversed to obtain two additional lists which would help to avoid any effects that might be caused by a particular order of stimuli or by the fact that participants' reaction time could become longer due to their fatigue toward the end of the experimental run.[2]

3.1.1.3. Procedure

In Experiment 1, the stimuli were presented in capital letters in the middle of a computer screen in front of participants. Firstly, they saw a certain feature that was presented to them for 1500 ms (e.g., IST ROT) followed by a concept label denoting either an object for which the preceding feature is congruent (e.g., TOMATE) or an object for which this feature is incongruent (e.g., BANANE) presented for 500 ms. Upon presentation of the concept label, participants had to decide whether the property

[1] The only item that might be considered an exception from this principle was "Maulwurf" ("mole") that was assigned the attribute value "hat Augen" ("has eyes") which originally was a congruent attribute value for the item "Eule" ("owl"). Strictly speaking, a mole does have eyes, but, given that a mole's eyes are virtually non-functioning, and that all other animals from the stimuli had full-functioning eyes, this attribute was kept as an incongruent attribute value for the item "Maulwurf".

[2] All lists of stimuli are available via the Open Science Framework, or OSF (Mnogogreshnova, 2023, URL: https://osf.io/3fbn2/?view_only=ee43eb8a793e408ab8d6b504bc4daaff). In addition, these materials are also accessible using a QR code (see Appendix).

belonged to the subsequently presented concept. The decision was made by pressing one of three buttons on the response pad in front of them: if the property was considered suitable for the presented concept, they pressed a green button indicating a positive response; if this was not the case, they pressed a red button giving a negative response; if for some reason they were not sure whether the property belonged to the concept or were not familiar with a particular concept or property, they could use a yellow button to indicate their uncertainty. The yellow button was used to identify items that were unfamiliar to the participants and therefore should be excluded from further analysis and use in subsequent experiments.

3.1.2. Results

The original dataset contained 14080 trials. In order to prepare the dataset for the analysis, several steps of filtration were taken. Firstly, the percentage of errors for each participant was checked. None of the participants had error rates exceeding 25%, therefore, no participants were removed from the analysis. Secondly, all items causing error rates of 25% or higher were filtered out, since high error rates could indicate that the participants were not familiar enough with the concepts in question, therefore, such concepts should not be included into further analysis. This led to the elimination of 29 items (13 items from the semantic category “animals”, 8 items from the semantic category “food” and 8 items from the category “tools”). At this stage of filtration, 11760 trials remained in the dataset and an error rate analysis was performed.

For the analysis of reaction times, all trials in which participants gave incorrect responses were eliminated from the dataset and the remaining trials were filtered in such a way as to make sure that each participant’s mean reaction time and reaction time to each item were no more than 3 standard deviations shorter or longer than the overall mean reaction time. In addition, reaction time in each trial had to be within 3 standard deviations from the respective participant’s mean reaction time. Trials that did not meet this criterion were eliminated from the dataset. All in all, 25.9% of the total number of trials were filtered out. The total number of remaining datapoints amounted to 10438.

Error Rates

The data on error rates in trials with congruent and incongruent attributes for each level of attribute prominence for the semantic category for items from each of the semantic categories involved in Experiment 1 is represented in Figure 1 below.

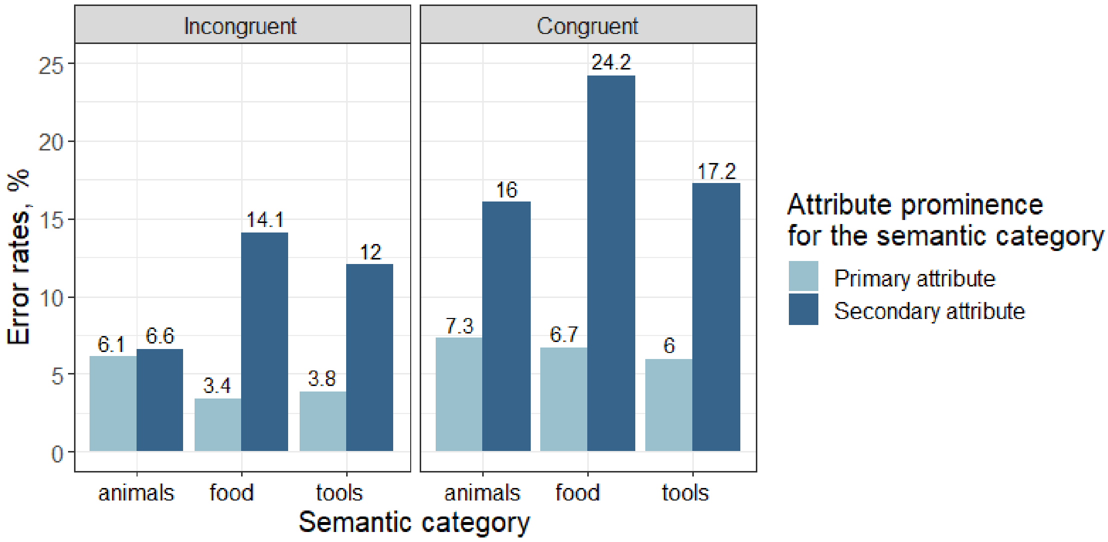


Figure 1. Error rates in Experiment 1 for each level of attribute prominence for the semantic category.

For the error rate analysis, a binary logistic regression model was computed, which included the following fixed factors: semantic category (animals, food, tools), attribute prominence for the semantic category (primary attribute, secondary attribute) and attribute congruence (congruent attribute, incongruent attribute), as well as participants, concepts, and attribute values as random factors with by-participant, by-concept, and by-value slopes for attribute congruence. An additional model with attribute prominence for the semantic category replaced by attribute prominence for individual items was then computed. These two models were compared to one another using an analysis of variance and the AIC values of those models. Removing attribute prominence for the semantic category from the model containing it resulted in a larger effect ($Chisq(1)=32.912$, $p<0.001$) than removing attribute prominence for individual items from the model containing it ($Chisq(1)=11.766$, $p<0.001$). Next, an AIC value was obtained for each model, which revealed that attribute prominence for the semantic category as a whole was a better predictor of error rates than attribute prominence for individual items. AIC values for both models are presented in Table 2.

Table 2. Comparison of binary logistic regression models containing attribute prominence for the semantic category vs. attribute prominence for individual items in Experiment 1.

	K	AICc	Delta AICc	AICcWt	Cum.Wt	LL
exp1.er_prominence.for.category	14	6432.94	0.00	1	1	-3202.45
exp1.er_prominence.for.item	14	6454.08	21.15	0	1	-3213.02

Therefore, the model that included attribute prominence for the semantic category was used in the subsequent analysis of error rates. However, even though the deviating items were few in number, a model including attribute prominence for individual items was computed and used in a separate analysis, so that it would be possible to determine, whether it would yield a different pattern of results as compared to attribute prominence for the semantic category as a whole.

The analysis of error rates revealed a significant main effect of attribute prominence for the semantic category ($Chisq(1)=32.912$, $p<0.001$) with higher error rates in trials with secondary attributes in comparison with primary attributes.

There was also a significant main effect of attribute congruence ($Chisq(7)=444.17$, $p<0.001$), showing higher error rates in the congruent condition in comparison with the incongruent one.

The main effect of semantic category did not reach the threshold of statistical significance ($p=0.075$).

As for the interactions, a three-way interaction between semantic category, attribute prominence for the semantic category and attribute congruence was not significant ($p>0.1$). Neither the interaction between semantic category and attribute congruence ($p>0.3$), nor the interaction between semantic category and attribute prominence for the semantic category ($p>0.2$) reached the threshold of statistical significance. The only two-way interaction which turned out to be statistically significant was an interaction between attribute prominence for the semantic category and attribute congruence ($Chisq(1)=6.1431$, $p<0.05$).

The dataset was then divided into two subsets based on the levels of the factor congruence, and the resulting subsets of congruent and incongruent attributes were analyzed separately. Consequently, attribute congruence as a fixed factor and random slopes for attribute congruence were removed from the models, leaving two fixed factors (semantic category and attribute prominence for the semantic category), and participants, concepts, and attributes as random factors.

According to the results of the error rate analysis in congruent trials, there was neither a main effect of semantic category ($p>0.5$) nor an interaction between the factors semantic category and attribute prominence for the semantic category ($p>0.2$). However, there was a highly significant main effect of attribute prominence for the semantic category ($Chisq(1)=33.166$, $p<0.001$), such that there were less errors made in response to primary attributes than in response to secondary attributes.

As for the incongruent trials, there was a significant main effect of attribute prominence for the semantic category ($Chisq(1)=6.024$, $p<0.05$), showing that there were less errors made in trials with primary attributes than in trials with secondary attributes. The main effect of semantic category did not reach significance ($p>0.05$), however, pairwise comparisons between different semantic categories revealed that participants tended to make less erroneous responses in trials with concepts from the semantic category "food" than in trials with the semantic category "tools" ($p=0.026$), whereas other pairwise comparisons turned out not to be significant ($p>0.1$ in the comparison between the semantic categories "animals" and "tools" and $p>0.2$ in the comparison between "animals" and "food"). The interaction between the factors semantic category and attribute prominence for the semantic category did not reach the threshold of statistical significance ($p>0.5$).

To determine whether the pattern of results was different when attribute prominence for individual items was used instead of attribute prominence for the semantic category, a new model was computed with semantic category (animals, food, tools) and attribute prominence for individual items (primary attribute, secondary attribute) as fixed factors, as well as participants, concepts, and attribute values as random factors. This model was used for analyzing error rates in congruent and incongruent trials separately. The error rates for congruent and incongruent attributes of each level of attribute prominence for individual items from each semantic category are depicted in Figure 2.

For the congruent attributes, similarly to the results obtained from the analysis of the model including a different kind of attribute prominence, neither a main effect of semantic category ($p>0.4$) nor an interaction between the factors semantic category and attribute prominence for individual items ($p>0.9$) were detected. There was, however, a significant main effect of attribute prominence for individual items ($Chisq(1)=13.609$, $p<0.001$), such that trials with primary attributes were less prone to errors than trials in which secondary attributes were presented. Nevertheless, this effect was somewhat smaller than the main effect of attribute prominence for the semantic category.

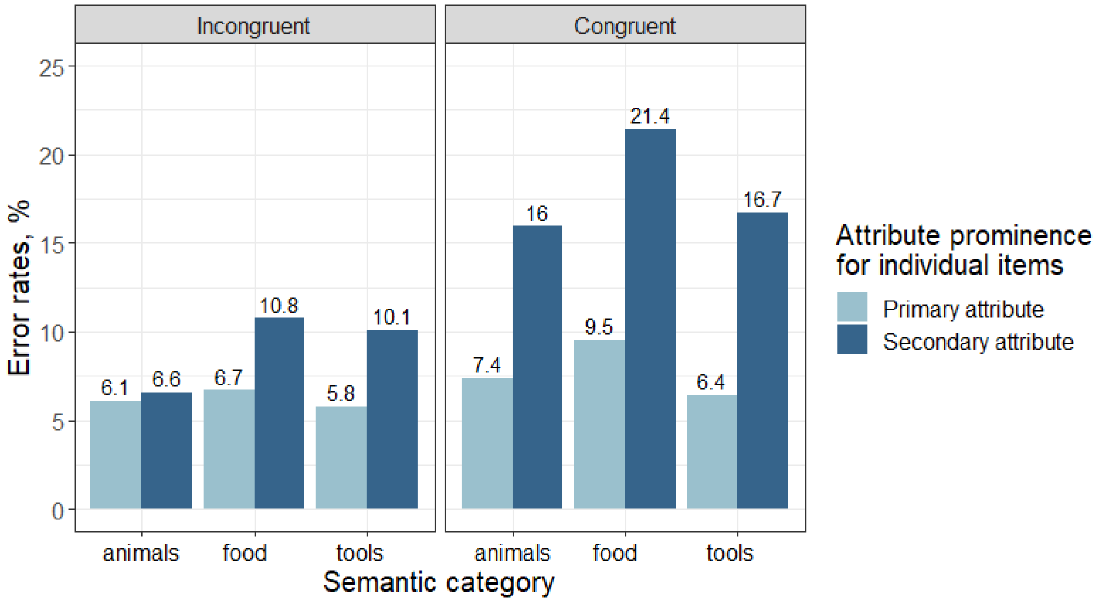


Figure 2. Error rates in Experiment 1 for each level of attribute prominence for individual items.

As for the incongruent condition, as opposed to the analysis of the model including attribute prominence for the semantic category, in this case none of the effects reached the threshold of statistical significance: $p>0.1$ for the main effect of semantic category, $p>0.1$ for the main effect of attribute prominence for individual items, and $p>0.9$ for the interaction between these two factors.

Reaction Times

For the analysis of reaction times, a linear mixed-effects model was computed, which had the same set of fixed and random factors as the model for the error rate analysis, namely semantic category (animals, food, tools), attribute prominence for the semantic category (primary attribute, secondary attribute), and attribute congruence (congruent attribute, incongruent attribute). In this model, the predicted variable was transformed using a logarithmic function to ensure normality of distribution of the reaction time data. The random factors included participants, concepts, and attribute values, with by-participant, by-concept and by-value slopes for attribute congruency. This model with attribute prominence for the semantic category as a whole was then compared to an analogous model, in which this factor was replaced by attribute prominence for individual items, using ANOVA and the AIC values of these models. Removing attribute prominence for the semantic category from the first model resulted in a significant difference between a model containing it and a model without it (Chisq(1)=20.162, p<0.001), whereas removing attribute prominence for individual items from the second model did not show a significant difference between the two variants of the model (p>0.2).

Using the AIC values to compare a model with attribute prominence for the semantic category and a model with attribute prominence for individual items revealed that attribute prominence for the semantic category was a significantly better predictor for the reaction times observed in this experiment, since the model containing it had a much lower AIC value. Table 3 provides information on the AIC values of both models.

Table 3. Comparison of linear mixed-effects models containing attribute prominence for the semantic category vs. attribute prominence for individual items in Experiment 1.

	K	AICc	Delta AICc	AICcWt	Cum.Wt	LL
exp1.rt_prominence.for.category	22	-18588.73	0.00	1	1	9316.41
exp1.rt_prominence.for.item	22	-18553.36	35.37	0	1	9298.73

Therefore, the model including attribute prominence for the semantic category among its fixed factors was used in the main analysis of reaction times.

The reaction time data for each level of the factors semantic category, attribute congruence, and attribute prominence for the semantic category is presented in Figure 3.

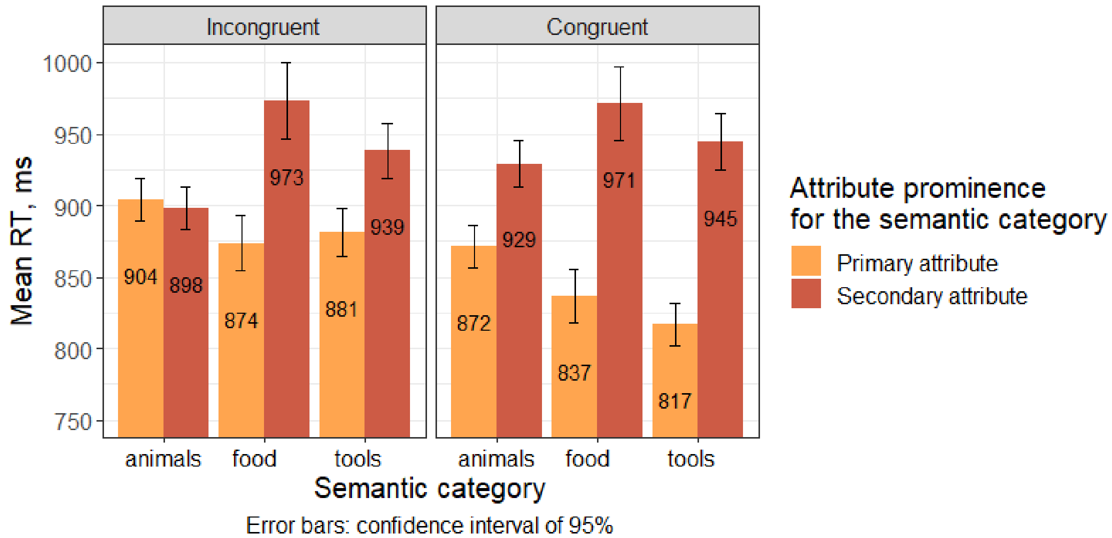


Figure 3. Reaction times in Experiment 1 for each level of attribute prominence for the semantic category.

The analysis of variance showed that there was a significant main effect of attribute prominence for the semantic category (Chisq(1)=20.162, p<0.001), such that reaction times to primary attributes were shorter (M=2.921, SD=0.1195, SE=0.0016) than reaction times to secondary attributes (M=2.953, SD=0.123, SE=0.00175).

As expected, there was also a significant main effect of attribute congruence (Chisq(1)=290.93, $p<0.001$) with shorter reaction times for congruent attributes (M=2.931, SD=0.125, SE=0.00176) than for incongruent ones (M=2.941, SD=0.1187, SE=0.0016). However, the main effect of semantic category did not reach the threshold of statistical significance ($p>0.06$).

As far as interactions between the fixed factors are concerned, there was an interaction between semantic category and attribute prominence for the semantic category (Chisq(2)=6.5091, $p<0.05$). The dataset was then divided into three subsets by the levels of the factor semantic category. Further analysis conducted for each of the subsets of the data revealed that the main effect of attribute prominence for the semantic category turned out to be significant for food items (Chisq(1)=12.786, $p<0.001$) with shorter reaction times for primary attributes (M=2.917, SD=0.1115, SE=0.0032) than for secondary attributes (M=2.969, SD=0.126, SE=0.004) and for tools (Chisq(1)=6.9735, $p<0.01$) with shorter reaction times for primary attributes (M=2.911, SD=0.1223, SE=0.0028) in comparison with secondary attributes (M=2.954, SD=0.129, SE=0.0031). For items from the semantic category “animals”, the main effect of attribute prominence for the semantic category did not reach significance ($p>0.9$).

Another two-way interaction that reached the threshold of statistical significance was the two-way interaction between attribute prominence for the semantic category and attribute congruence (Chisq(1)=17.998, $p<0.001$).

No significant interaction was found between semantic category and attribute congruence ($p>0.3$).

Finally, there was a significant three-way interaction between semantic category, attribute prominence for the semantic category and attribute congruence (Chisq(7)=30.109, $p<0.001$).

The dataset was then divided into two subsets based on the levels of the factor attribute congruence, such that congruent and incongruent attributes could be analyzed separately.

For congruent attributes, the analysis revealed a highly significant main effect of attribute prominence (Chisq(1)=35.285, $p<0.001$) with shorter reaction times for primary attributes (M=2.909, SD=0.122, SE=0.002) than for secondary attributes (M=2.9565, SD=0.124, SE=0.003). Neither the main effect of semantic category, nor the interaction between semantic category and attribute prominence for the semantic category were significant ($p>0.2$ in both cases).

As for the incongruent attributes, the main effect of semantic category failed to reach the threshold of statistical significance ($p>0.08$), but there was a highly significant main effect of attribute prominence for the semantic category (Chisq(1)=6.909, $p<0.01$), showing that reaction times were shorter for primary attributes (M=2.933, SD=0.116, SE=0.002) than for secondary attributes (M=2.9495, SD=0.121, SE=0.002). The two-way interaction between semantic category and attribute prominence for the semantic category also turned out to be significant (Chisq(2)=10.173, $p<0.01$). The dataset with incongruent attributes was then divided into three subsets by the levels of the factor semantic category and the main effect of attribute prominence for the semantic category was investigated for animals, food, and tools separately. This revealed that, for animals, there was no difference in reaction times between trials with primary attributes and trials with secondary attributes ($p>0.2$). For concepts from the semantic category “tools”, the reaction times tended to be shorter for primary attributes than for secondary attributes, but the main effect of attribute prominence for the semantic category did not reach the threshold of statistical significance ($p\approx0.09$). There was, however, a significant main effect of this factor for food items (Chisq(1)=7.666, $p<0.01$), such that it took less time for the participants to reject a primary attribute (M=2.9265, SD=0.111, SE=0.0045) than a secondary attribute (M=2.968, SD=0.129, SE=0.006).

In addition, an analysis was performed based on the model in which attribute prominence for the semantic category was replaced by attribute prominence for individual items, in order to determine whether substituting this factor would change the overall pattern of results. The data on reaction times for each level of the factors semantic category, attribute congruence, and attribute prominence for individual items is represented in Figure 4.

The analysis of the congruent trials showed that there was no significant main effect of semantic category ($p>0.5$) and no interaction between attribute prominence for individual items and semantic category ($p>0.8$). The only significant effect was the main effect of attribute prominence for individual items (Chisq(1)=8.505, $p<0.01$), such that the reaction times were shorter for primary attributes (M=2.917, SD=0.122, SE=0.002) than for secondary attributes (M=2.9465, SD=0.127, SE=0.003).

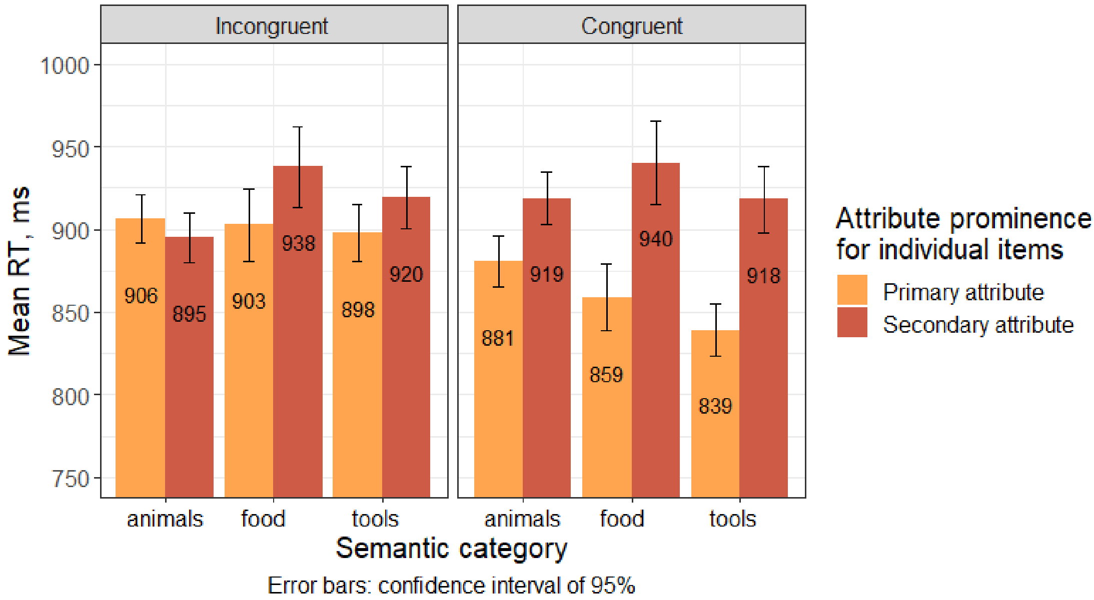


Figure 4. Reaction times in Experiment 1 for each level of attribute prominence for individual items.

As far as the incongruent trials are concerned, none of the effects reached the threshold of statistical significance: $p>0.3$ for the main effect of semantic category, $p>0.8$ for the main effect of attribute prominence for individual items, and $p>0.4$ for the interaction between these two factors.

Attribute Effects

Part attributes were present in all three semantic categories: for items from semantic categories “food” and “tools” part attributes were considered secondary, and for the semantic category “animals” part attributes were primary. Considering that for the semantic category “animals” the differences in reaction times and error rates were no longer observed in the incongruent condition, whereas the other two semantic categories showed reaction time and error rate patterns similar to the congruent condition, it might indicate that part attributes are inherently different from color and affordance attributes. To find out whether that is the case, it would be useful to introduce attribute as a factor. To make sure that the analysis would not be performed using rank deficient models, two semantic categories sharing the same attributes were selected, namely “animals” and “food”. All tools were filtered out from the datasets for reaction time analysis and for error rate analysis, and the main effect of attribute with the levels “part” and “color” was subsequently investigated. As a result of this analysis, it was established that there was a significant main effect of attribute (Chisq(1)=6.0199, $p<0.05$) with longer reaction times in trials with part attributes ($M=2.943$, $SD=0.123$, $SE=0.0021$) than in trials with color attributes (M=2.935, SD=0.114, SE=0.002). In case with the error rates, the main effect of attribute did not reach the threshold of statistical significance, but revealed a trend, according to which error rates were higher in trials with part attributes than in trials with color attributes ($p=0.064$).

3.1.3. Discussion

Experiment 1 was a property verification task designed to assess the influence of a range of factors including semantic category, attribute prominence for the semantic category, attribute prominence for individual items, and attribute congruence on the speed and accuracy of accessing attribute values. It was hypothesized that primary attributes, being more prominent in a given conceptual structure, should be available earlier than secondary attributes and should be retrieved more accurately due to a stronger connection to their concept nodes, leading to shorter reaction times and lower error rates for primary attributes in comparison with secondary attributes.

The data analysis revealed that reaction times in the congruent condition were shorter than in the incongruent condition. There are several possible explanations for this. Firstly, participants gave positive responses using their right hand and, considering that the majority of them were righthanders (32 out of 40 people) and that responses made by the dominant hand are usually quicker than responses made by the non-dominant hand (Kerr, Mingay & Elithorn, 1963), one would expect to see shorter reaction times in congruent trials, in which participants had to confirm an object's property. Secondly, shorter reaction times in the congruent condition could have been due to priming effects produced by congruent attribute values. A third explanation could be that perhaps the participants had a bias, such that they were generally more inclined to give a positive response. However, the results of the error rate analysis revealed that more errors were made in trials with congruent attribute values than in trials with incongruent attribute values. Therefore, the latter explanation for shorter reaction times in the congruent condition has to be ruled out. Nevertheless, since neither the possible confounding effects of handedness nor priming effects (pre-activation of a concept by a congruent attribute value) were of interest for the current study, the incongruent condition was considered more representative of the role of attribute prominence in conceptual processing.

The hypothesis concerning attribute prominence for the semantic category was confirmed: according to the results of the analysis, shorter reaction times were observed in trials with primary attributes as compared to trials with secondary attributes, which is not surprising, considering that primary (more prominent) attributes have a stronger association with their respective concept labels (Barsalou, 1982, 1992), which should facilitate retrieval of such attributes from the semantic memory. Moreover, secondary attributes were also more prone to errors than primary attributes. These differences in reaction times and error rates were more pronounced for congruent attributes than for incongruent ones.

Since some items in each of the three semantic categories were deviating from the overall pattern of attribute prominence for their respective semantic categories in such a way that an attribute which was considered secondary for the semantic category as a whole was more prominent for their conceptual structure than an attribute which was considered primary for that semantic category, it created an opportunity to find out whether attribute prominence for individual items might be a better predictor for reaction times and/or error rates. For that purpose, two models were computed: one of them included attribute prominence for the semantic category and in the other model this factor was replaced by attribute prominence for individual items. These models were then compared based on their AIC values and based on the significance and strength of the main effect of each of the two kinds of attribute prominence. The comparison of these models showed that attribute prominence for the semantic category is a better predictor both for reaction times and for error rates in Experiment 1, which allows us to conclude that attribute prominence for the semantic category as a whole overrides (or has a priority over) attribute prominence for individual items. Nevertheless, it should be noted that the number of deviating items in each semantic category was relatively small (only 12-13 deviating items per category) and the difference in frequency of occurrence between primary and secondary attributes of deviating items was much smaller than that for regular items.

No statistically significant difference in reaction times was found between primary and secondary attributes of the semantic category “animals” in the incongruent condition, even though for the other two semantic categories, “food” and “tools”, reaction time patterns in the incongruent condition remained largely the same as in the congruent condition, albeit these differences had somewhat higher p-values in the incongruent condition than in the congruent one. Given that there was no significant main effect of semantic category, this might indicate that the part attribute is inherently different from the affordance and color attributes. An additional analysis revealed that reaction times were longer and error rates tended to be higher for part attributes than for color attributes. One could assume that part attributes generally take longer to access and cause more errors due to an additional processing step, i.e., in order to decide whether a particular element is present in an object, one needs to mentally disassemble the object. An alternative explanation for that could be that part attributes have a different way of organization within a frame structure as compared to color and affordance attributes. "Because values are concepts, they in turn can be attributes having still more specific values" (Barsalou, 1992: 32), which enables part attributes to be further specified by the number of elements in question, e.g., an attribute "has legs" can be found in conceptual structures "owl" and "dog", but these animals have different numbers of legs and it has to be reflected in their frame structures (or, as Barsalou argued, a type of legs can also be specified in the frame structure of an animal), whereas color attributes usually do not require any further specification beyond their value. This kind of issue can be hard to control for, considering that the feature database by McRae et al. (2005), which was used as a source material for the experimental stimuli, does not reflect the approach of the frame theory and, as a consequence, properties like "has legs" and "has 4 legs" are often mentioned in this database as two separate features for the same concept, each of them having its own production frequency.

Even though reaction times were predicted to be shorter for primary attributes than for secondary attributes even in the incongruent condition, at this point it is not yet clear whether this effect of attribute prominence indeed emerges on the conceptual level because of stronger connections between primary attributes and concepts, and, therefore, quicker retrieval of primary attributes in comparison with secondary attributes. One could argue that, for instance, if participants see the attribute value "IST ROT" and then the concept label "BANANE", they access the corresponding attribute of the concept in question and realize that the attribute value they have just seen on the screen does not belong to the object in question. Thus, no facilitation would have been expected in such cases. However, since reaction times were still significantly shorter for primary attributes than for secondary attributes, it would mean that attribute prominence for the semantic category can indeed influence the speed of accessing a certain attribute irrespective of attribute congruence. On the other hand, given that verbal stimuli were used in this experiment, an alternative explanation could be that this difference in reaction times between primary and secondary attributes in the incongruent condition could have been caused at the lexical level by the phenomenon known as mediated priming (see De Groot, 1983; Hill, Strube, Roesch-Ely & Weisbrod, 2002; Hill, Ott & Weisbrod, 2005), in which case priming effects are present, if a prime and a target are associated with each other not directly, but through another item, i.e., a "mediator". Given that every incongruent attribute value in the current experiment was taken from another member of the same semantic category, a residual priming effect might be observed in the incongruent condition, if there is a strong enough association between the concept that attribute belongs to and the concept it was re-assigned to. As a way to control for that, a similar verification task could be conducted with images instead of verbal stimuli. In that case, if there is still a significant effect of attribute prominence, it would mean that it is not a lexical, but rather a conceptual phenomenon. Before such an experiment can be conducted, though, it would be necessary to make sure, given the apparent complexity of part attributes in comparison with color attributes, that the color changes and part changes in the pictorial stimuli are comparable with each other, in order to avoid confounding the findings with the effects that can be explained by the difference in the perceptual complexity of the depicted attributes. In other words, a pre-test is needed to assess how quickly and accurately certain attribute changes can be detected in pictorial stimuli. It would be useful to conduct a pre-test both on the perceptual level (ensuring comparable discriminability of the attributes in isolation from their respective concepts) and on the conceptual level (by investigating how quickly and accurately the same changes can be detected when the whole concept is represented in the image). To this end, Experiments 2 and 3 were conducted.

3.2. Experiment 2

Experiment 1 has demonstrated that in an attribute verification task with verbal stimuli primary attributes are accessed more quickly and accurately than secondary attributes, and attribute prominence for the semantic category as a whole has a larger influence on reaction times and error rates in comparison with attribute prominence for individual items. However, it is not yet clear whether this effect emerged on the lexical level due to mediated priming, since incongruent attribute values were taken from items that are categorically related to the concepts they were re-assigned to, or whether it can be observed on the conceptual level as well. To investigate this question, a verification task with pictorial stimuli should be conducted, and the images would have to be modified accordingly to depict congruent and incongruent attribute values. Whereas color and part attributes can be relatively easily depicted in a static image, it is much harder to do it with affordance as an attribute of the semantic category "tools". Even if a corresponding image could be produced, it would most likely have to include not just the tool itself, but essentially a whole situation and perhaps other objects that reflect the use of that particular tool (e.g., a hammer and a nail), which would make the stimuli overly complicated. Items from the semantic categories "animals" and "food" are more comparable in terms of the level of image complexity, since concepts from these semantic categories can be easily depicted with no background or additional objects without any detriment to their recognizability. Furthermore, both animals and food items have two attributes in common, namely colors and parts, differing only in their prominence depending on the semantic category or on the concept in question. Therefore, the semantic categories "animals" and "food" were chosen to be used in Experiments 2 and 3.

It has been observed in Experiment 1 that color attributes seem to have an advantage that is reflected in the fact that reaction times were shorter and error rates tended to be lower for color attributes than for part attributes. To avoid this kind of confounding factor, one could create several degrees of color changes and see whether a more subtle degree of color change would be perceptually more comparable with part attribute changes in terms of reaction times and error rates. Before an experiment with images depicting whole objects can be conducted, perceptual comparability of the attributes in isolation from their respective images should be tested. Hence, the aim of Experiment 2 was to check for perceptual comparability between color and part attributes in isolation from their respective concepts using pictorial stimuli in a two-alternative forced choice task, in which participants had to match a target picture to one of two alternative images, one of which depicted the congruent attribute value of the object in question, whereas the other one depicted a changed version of it. A set of images depicting animals and food items was selected based on the verbal stimuli from Experiment 1, such that the corresponding attribute values were visible in those images. In addition, the selected images were modified, so that the original (congruent) attribute values were replaced by incongruent attribute values from other items from the same semantic category. The original and the modified attributes were then extracted from the images and used as stimuli in Experiment 2. Reaction times were measured, and error rates were calculated to determine how quickly and accurately participants could detect the differences between the original and the altered versions of the attributes depicted in those images. A more subtle color change was expected be recognized more slowly and less accurately than more obvious color changes, making it more comparable to part attribute changes.

3.2.1. Methods

3.2.1.1. Participants

Twenty adult German native speakers (11 female, 9 male; age: from 22 to 35 years old, mean age: 26.35 years old) took part in Experiment 2 and were paid for their participation. All participants reported to have normal or corrected-to-normal vision and normal color perception.

Since color perception is an important aspect of this experiment, bilingual participants with more than one mother tongue were not allowed to participate in this experiment, since, according to some studies (e.g., Winawer et al., 2007), their second mother tongue could have an influence on the speed of color recognition, making it easier or more difficult to recognize certain colors, in case they are categorized differently in the second mother tongue in comparison with color labels in the German language.

The study was approved by the local ethics committee of Heinrich Heine University Düsseldorf (study number 2019-436-Drittmittel).

3.2.1.2. Stimuli

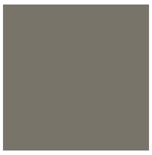
The verbal stimuli that were selected from the feature database by McRae et al. (2005) for Experiment 1 were also utilized to select their pictorial equivalents, which were subsequently used as stimuli in Experiment 2. In the semantic category "animals" 24 items with part attributes "HAT FELL", "HAT WOLLE", "HAT HAARE" or "HAT FEDERN" had to be omitted, since it was not possible to depict an attribute violation and to ensure its reliable recognition by participants (e.g., replacing feathers of a parakeet with feathers of another bird without changing their color or the overall shape of the animal would not necessarily result in a recognizable change in the image). The rest of the stimuli were kept, and the corresponding images were found using the Bank of

Standardized Stimuli, or BOSS (Brodeur, Dionne-Dostie, Montreuil & Lepage, 2010; Brodeur, Guérard & Bouras, 2014), and, if no image for a certain concept was available there or if the required attribute was not depicted in the provided image, the database "Freepng" (URL: https://www.freepng.ru) was used instead. All in all, pictures of 52 items from the semantic category "animals" and 40 items from the semantic category "food" were selected, based on the fact that their respective attributes were visible in the corresponding static images. Fourteen items from the semantic category "animals" and 8 items from the semantic category "food" were taken from the BOSS database; and 38 items from the semantic category "animals" and 32 items from the semantic category "food" were taken from "Freepng". All images were files in the PNG format with a transparent background.

For each item, 6 different images were created: one image depicting the original color of the object, one image with the original part attribute, three images involving a color attribute changed by 33%, 66%, and 100%, respectively, and one image with an altered part attribute. In altered color conditions, a color of another item from the same semantic category was used that was not typical for the given item. The altered image was then superimposed onto the original image and the opacity of the former was set to 33%, to 66%, and to 100% to create the corresponding conditions with various degrees of color changes. For the conditions involving part attributes, each concept's part attribute that was used in Experiment 1 was replaced by a corresponding element of another concept from the same semantic category while matching the color of the original picture. An element was considered a suitable replacement, if it represented an analogous structural part, e.g., an elephant's trunk was replaced by a pig's nose.

After all the images depicting concepts with original and changed color and part attributes had been prepared, those attributes were extracted from their respective pictures. First, the original picture of every item was used to obtain a patch of original color associated with the corresponding concept. For that purpose, each picture was scaled, such that an area of the color in question would fully occupy a canvas of 300×300 pixels and the function of pixelation was used to turn that area into a color patch. As for the three altered color conditions, the image of a concept with its color changed by 100% was used to obtain a patch of color using pixelation in the same way. The resulting color patch was later used as a stimulus for the condition "color 100%". The patch of that altered color was then superimposed onto the patch of the respective item's original color and the opacity of the former was set to 33% and to 66% to create patches of color for the conditions “color 33%” and “color 66%”, respectively. As far as part attributes are concerned, the original part associated with the concept and the altered part of the corresponding pictures were cut out and saved separately. All pictorial stimuli in Experiment 2 were calibrated to have the same size of 300×300 pixels and were saved in the PNG format on a transparent background using an open-source raster graphics editor “Krita”, version 4.3.0 (URL: https://krita.org). Examples of these pictorial stimuli are given in Table 4.[3]

Table 4. Examples of stimuli in Experiment 2.

	COLOR				PART	
	original	color 33%	color 66%	color 100%	original	part 100%
ANIMALS						
FOOD						

Once the pictorial stimuli were ready, combinations of stimuli for each trial were determined, such that the alternatives in each trial contained one unaltered image (an original part of the concept or a patch of its original color) and one image with a changed attribute (a replacement for the original part or a patch of color changed by 33%, 66%, or 100%). As a result, a set of stimuli was created, in which 4 pairs of alternatives were created for each item: original part vs. altered part, original color vs. color altered by 33%, original color vs. color altered by 66%, and original color vs. color altered by 100%. In these pairs of alternatives, the original attribute was supposed to be depicted on the left side of the screen and the target image depicted the original attribute, i.e., the expected response was “left”. Another set of stimuli was then created, such that the target image depicted the altered attribute, i.e., the expected response was “right”. To counterbalance the position of the stimuli, two additional sets of stimuli were created, in which the target

[3] A full set of the experimental stimuli for this experiment as well as other experiments described in the present thesis can be found in the supplementary materials in the OSF (Mnogogreshnova, 2023, URL: https://osf.io/3fbn2/?view_only=ee43eb8a793e408ab8d6b504bc4daaff). These materials can also be accessed by scanning a QR code (see Appendix).

pictures remained the same, but the positions of the two alternatives were switched. Trials from these sets were divided into 4 lists of stimuli. In each of these lists, all 4 attribute comparisons were presented for each concept: one trial with the comparison of the original and the altered part, and 3 trials with color patch comparisons with different degrees of color change; in 2 of the trials the original attribute was depicted on the left (with expected responses "left" in one of these trials and "right" in the other trial), in the other 2 trials the original attribute was presented on the right (with expected responses "left" in one of these trials and "right" in the other trial). These 4 lists of stimuli were pseudorandomized using the program "Mix" (Van Casteren & Davis, 2006) under the following constraints: attributes of the same concept had to be at least 10 trials away from each other; attributes of concepts from the same semantic category could not be presented in more than 3 trials in a row; the same attribute (part or color) could not be presented in more than 4 trials in a row; identical responses could not be given in more than 3 trials in a row. At the beginning of each list, 4 practice trials were added to enable participants to get used to the task. The stimuli for these practice trials were created the same way as experimental trials and belonged to the same semantic categories. The resulting lists of stimuli included 372 trials each.

3.2.1.3. Procedure

Experiment 2 was programmed using PsychoPy (Peirce et al., 2019) and, due to COVID restrictions, it was conducted online via Pavlovia (URL: https://pavlovia.org). All participants received an email containing detailed instructions concerning the experimental task, their personal participant number, and the link to the experiment. The aforementioned participant numbers were used to assign a particular list of stimuli to ensure that each list was used equally often and to simplify the replacement of participants, if necessary, by giving the same number to another participant, for example, when the data provided by a participant was incomplete or could not be saved due to technical problems.

Once a participant accessed the experiment using the link, they typed their participant number and email address in separate fields and were presented with a short version of instructions written in white font over a dark blue background. The color of the background was selected to ensure good visibility of all pictures, considering that they had a relatively wide range of colors. Upon reading the instructions, the participants could start the experiment by pressing the space bar. Each trial started with a fixation cross presented for 500 ms in the middle of the upper half of the screen. After that, three images were presented on the screen simultaneously: the target image was presented in the middle of the upper half of the screen, where the fixation cross was previously visible, and two alternatives were presented on the left and on the right side in the bottom half of the screen (see Figure 5 for a sample trial). The images remained on the screen for 2 seconds, during which participants had to press a left or a right arrow key to indicate which of the two alternatives looks the same as the target image. If participants did not give their response in time, or as soon as they pressed one of the arrow keys, an inter-trial interval of 500 ms was initiated, and the next trial began.

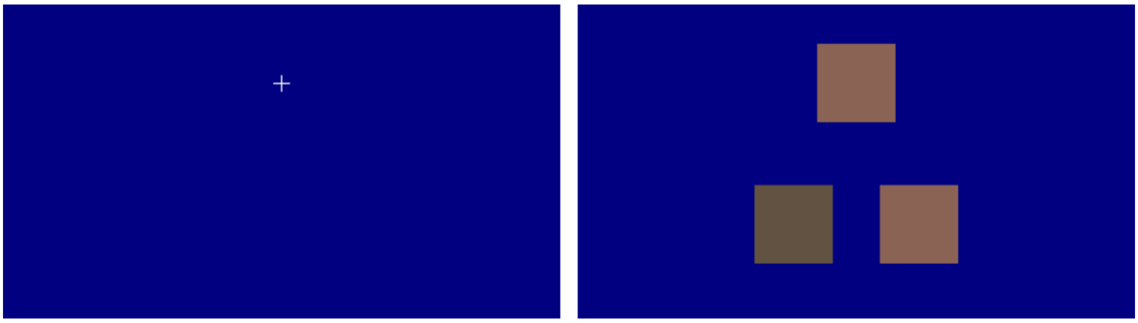


Figure 5. Sample trial from Experiment 2.

Thus, the total duration of a trial amounted to maximally 3 seconds. At the end of the experiment, participants were informed that the task was finished, and were asked to press the space bar and then close the browser window.

3.2.2. Results

A total of 7360 data points were collected in Experiment 2. To start with, error rates per participant and per item were checked. No participants made errors in more than 25% of the total number of trials and no items caused mistakes in more than 25% of trials, therefore, no participants or items were excluded from the analysis. At this stage the dataset was considered ready for the error rate analysis.

To prepare the dataset for the analysis of reaction times, the following steps of data filtration were implemented. First, all trials, in which erroneous responses were given, were filtered out, which lead to elimination of 176 trials (2.4%) from the dataset. Secondly, the data was filtered in such a way as to ensure that each participant's and each item's mean reaction time lay within 3 standard deviations from the overall mean reaction time. No trials were lost as a result of this step of filtration. Finally, 154 trials (2.1%) deviating more than 3 standard deviations from the respective participant's mean reaction time were discarded. After that, 7030 data points remained in the dataset, and the data was considered ready for the reaction time analysis.

Error Rates

Error rates in all conditions of Experiment 2 are shown in Figure 6.

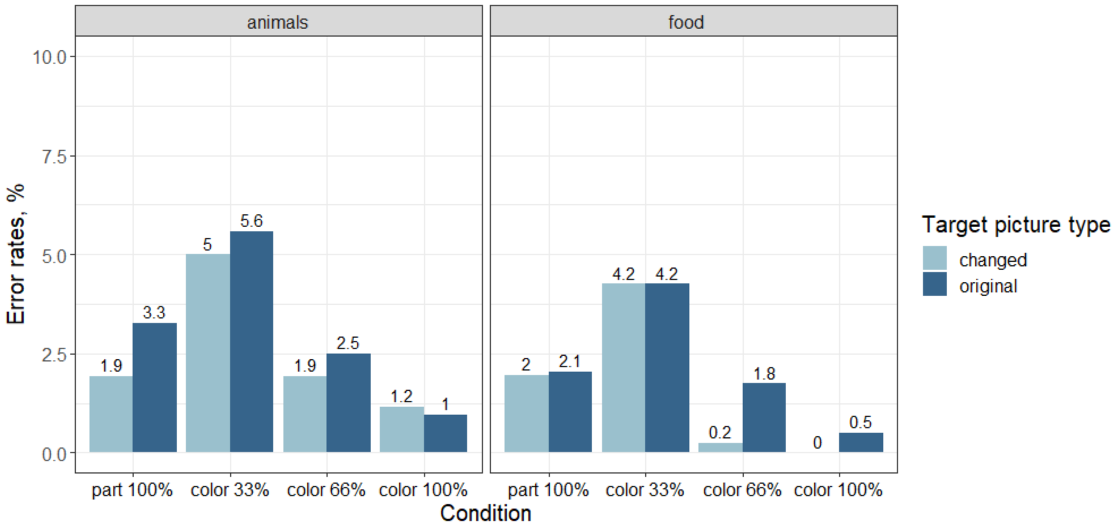


Figure 6. Error rates in Experiment 2.

For the error rate analysis, a binary logistic regression model was computed, which included participant and item as random intercepts, as well as the following fixed factors: semantic category (animals, food), condition (color 33%, color 66%, color 100%, part 100%), and type of the target picture (original, changed). Attribute prominence for the semantic category and attribute prominence for individual items were excluded from the analysis, since in trials with color patches participants could not know what concepts the colors belonged to, whereas "part 100%" was the only condition in which assumptions about certain objects could be made. Hence, any potential prominence effects would correspond to an interaction between semantic category and condition, particularly to a significant main effect of semantic category in the part condition, but not in the color conditions.

The results of the analysis showed that there was a significant main effect of condition ($Chisq(3)=74.841$, $p<0.001$) with a higher accuracy in the condition "color 100%" in comparison with the three other conditions: "color 33%" ($p<0.001$), "color 66%" ($p<0.005$), and "part 100%" ($p<0.001$); error rates in the condition "color 33%"

were higher than in the condition “part 100%” ($p<0.001$) or “color 66%” ($p<0.001$), but there were no significant differences in error rates between the conditions “part 100%” and “color 66%” ($p>0.1$). No other main effects or interactions between the fixed factors reached the threshold of statistical significance.

Reaction Times

Reaction times in all conditions of Experiment 2 are shown in Figure 7.

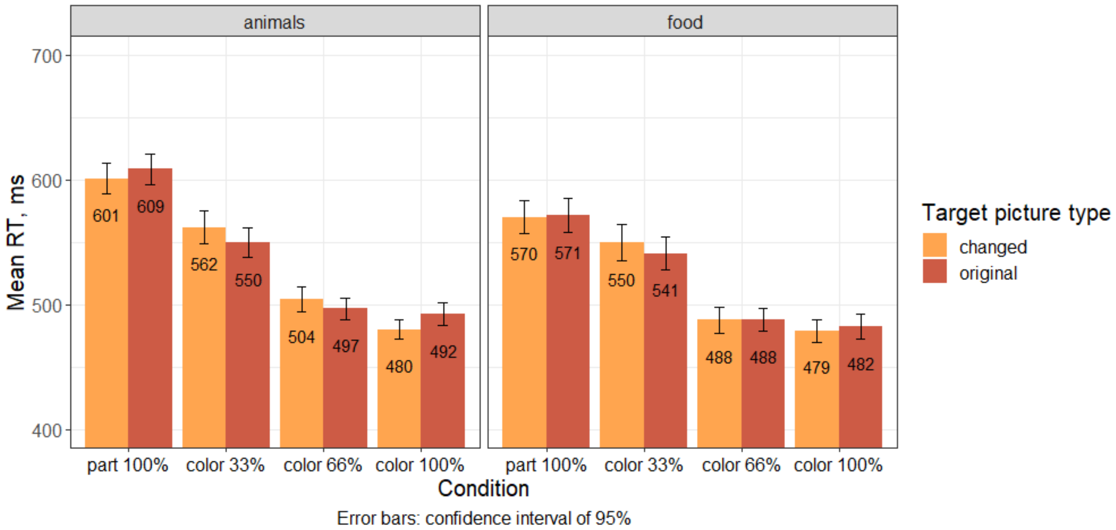


Figure 7. Reaction times in Experiment 2.

To analyze reaction times, using the “lme4” package (Bates et al., 2015) in R Studio (R Core Team, 2022), a linear mixed-effects model was built which included the same set of random and fixed factors as the model used in the error rate analysis, namely participant and item as random intercepts, as well as the following fixed factors: semantic category (animals, food), condition (color 33%, color 66%, color 100%, part 100%), and type of the target picture (original, changed). To ensure normality of distribution of reaction times, a base 10 logarithm was calculated from the original reaction time values, and it was then used as the dependent variable in the model.

The analysis of reaction times revealed a significant main effect of semantic category ($Chisq(1)=6.9139$, $p<0.001$) with slightly longer reaction times in trials involving animals ($M=2.717$, $SD=0.098$, $SE=0.0016$) than in trials involving food items ($M=2.706$, $SD=0.094$, $SE=0.0017$).

There was also a significant main effect of condition ($Chisq(3)=1230$, $p<0.001$), such that there were significant differences between all four conditions of the experiment: “color 100%” ($M=2.677$, $SD=0.0801$, $SE=0.0019$), “color 33%” ($M=2.7295$, $SD=0.099$, SE=0.0024), “color 66%” (M=2.686, SD=0.084, SE=0.002), and “part 100%” (M=2.7598, SD=0.097, SE=0.0023).

There was also a significant interaction between the factors semantic category and condition (Chisq(3)=18.249, $p<0.001$). To investigate this interaction further, the dataset was split into four subsets by the levels of the factor condition, and the main effect of semantic category was analyzed in each of those subsets separately. This main effect did not reach significance in any of the color conditions ($p>0.4$ for conditions “color 100%” and “color 33%”, and $p>0.08$ for the condition “color 66%”), whereas there was a significant difference in reaction times between the semantic categories in the condition “part 100%” (Chisq(1)=7.1564, $p<0.005$) with longer reaction times in trials with animal part attributes (M=2.771, SD=0.095, SE=0.003) than in trials with food part attributes (M=2.745, SD=0.097, SE=0.0035).

There was also a marginally significant interaction between the factors condition and target picture type (Chisq(3)=9.0681, $p<0.05$). Further analysis showed that the main effect of target picture type did not reach significance in conditions “color 100%” ($p>0.05$), “color 66%” ($p>0.3$), or “part 100%” ($p>0.1$), and there was only a marginally significant main effect of target picture type in the condition “color 33%” (Chisq(1)=4.9655, $p<0.05$) with somewhat longer reaction times in trials with changed pictures (M=2.733, SD=0.102, SE=0.0035) than with original pictures (M=2.726, SD=0.096, SE=0.0033).

However, there was no main effect of target picture type ($p>0.9$) or interaction between semantic category and target picture type ($p>0.7$).

3.2.3. Discussion

Experiment 2 utilized a two-alternative forced choice task, in which participants were presented with images depicting color and part attributes extracted from pictures of animals and food items: color attributes were presented in the form of color patches reflecting the original color of each concept or different degrees of color changes (i.e., colors altered by 33%, 66%, and 100%), whereas part attributes were presented in the form of original or altered structural elements of the concepts that were cut out of images depicting objects in the corresponding conditions. The attributes in question were the same attributes that were used in Experiment 1 based on the feature database by

McRae et al. (2005). In Experiment 2, participants had to match a target picture to one of two alternatives presented on the screen at the same time, which would allow to assess how quickly the depicted attributes could be distinguished from one another and to what extent each experimental condition was prone to errors. This experiment was designed to establish whether any of the degrees of color changes were comparable with the part changes in terms of reaction times and/or error rates.

As expected, a subtle color change by 33% was the most difficult of all color conditions in this experiment, causing longer reaction times as well as higher error rates in comparison with colors changed by 66% or by 100%. Trials with color patches reflecting 100% color changes elicited the shortest reaction times and the lowest error rates in this experiment, reflecting the easiness of detecting this degree of color change. Trials with 66% color changes elicited much shorter reaction times and much lower error rates than trials with 33% color changes, but the difference between 66% and 100% color changes was somewhat smaller, making reaction times in the condition “color 66%” slightly longer and error rates slightly higher than in the condition “color 100%”.

When it comes to trials with part attributes, reaction times in this condition were significantly longer than in any of the color change conditions, confirming that it takes longer to detect the difference between isolated part attributes, which can be explained by the visual complexity of this kind if stimuli as compared to simple color patches, in which no small objects are represented and thus no additional processing is demanded. However, even though participants needed more time to make a response in trials involving part attributes, the error rates in this condition are comparable with the error rates in the condition involving 66% color changes.

In trials with color patches, participants were unable to discern which concept a particular color attribute is assigned to, since there was no other conceptual information available for participants to draw any unequivocal conclusions about that, whereas part attributes might have provided enough conceptual information to draw conclusions about the objects in question. This is confirmed by the fact that an interaction between semantic category and condition was observed, such that there was a main effect of the category for part attributes, but not for color attributes.

Thus far, it can be concluded that the color advantage observed in Experiment 1 also arose on the perceptual level, allowing people to quickly detect color changes in comparison with changes in isolated parts. Nevertheless, this color advantage seems to be less salient in terms of error rates, since moderate color changes caused approximately the same percentage of errors as changes in part attributes. Meanwhile, according to the reaction time analysis, the condition "color 33%", while being significantly different from the condition "part 100%", is the closest match to it out of the three color change conditions in this experiment. Another pilot experiment is needed to definitively establish whether 33% or 66% color changes should be used in the subsequent verification task with pictorial stimuli. In addition, the question remains whether adding the rest of conceptual information (apart from color or part attributes) to the pictorial stimuli will change the pattern of results observed in Experiment 2. To investigate it, Experiment 3 was designed, in which images depicted whole objects either in their original version or with a modified part or color attribute.

3.3. Experiment 3

In Experiment 2, participants were presented with color patches and part attributes isolated from their respective concepts. Two color conditions were identified, which would be potentially comparable to the condition involving part attribute changes: 66% color changes had approximately the same error rates as the condition with part attributes, whereas 33% color changes were the closest to it in terms of reaction times. To determine which of these two color conditions could be used in the subsequent verification task with pictorial stimuli and how using images depicting every object as a whole would influence the pattern of results observed in the previous experiment, Experiment 3 was designed, in which the same two-alternative forced choice paradigm was implemented, but a different set of stimuli was used, namely the images from which the stimuli of Experiment 2 were extracted. Thus, in Experiment 3, participants were presented with visually more complex pictures and were asked to choose one of two alternatives matching a target picture in every given trial.

These changes in the stimuli were expected to cause the following changes in the findings of this experiment as compared to Experiment 2. First, the reaction times in general, irrespective of the conditions, were expected to become longer due to the increase in the visual complexity of the stimuli. Secondly, a main effect of attribute prominence was expected, i.e., shorter reaction times for primary attributes in comparison with secondary attributes across both semantic categories, since now participants would be able to recognize the depicted objects and that might provide them with additional cues as to which attributes should be accessed and processed first. Thirdly, the advantage of color attributes was expected to be replicated, however, now that whole objects were depicted in the images, effects of object knowledge on color perception might arise, for example, the effect of color constancy, meaning that subtle color changes might become even more difficult to detect. One of the ways in which these changes were expected to manifest was an increase in error rates for subtle color changes (for conditions “color 33%” and possibly “color 66%”) relative to other conditions in comparison with the pattern of results observed in the same conditions in Experiment 2. Object knowledge might make it is easier for the human brain to “compensate” subtle color changes and perceive them still as the original colors of the objects in question (e.g., Mitterer & de Ruiter, 2008). Therefore, incorrect responses should be observed more often in these conditions with subtle color changes in comparison with the other experimental conditions.

3.3.1. Methods

3.3.1.1. Participants

Forty adult German native speakers from 18 to 35 years old (mean age: 25.85 years) took part in Experiment 3 and were paid for their participation. All of them reported having German as their only mother tongue, having normal color perception as well as normal or corrected-to-normal vision. To prevent any effects caused by familiarity of the stimuli, it was made sure that none of these participants had taken part in Experiment 2.

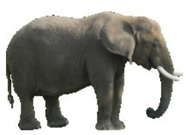
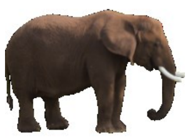
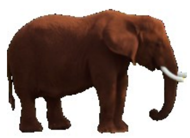
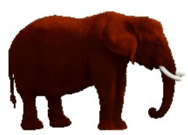
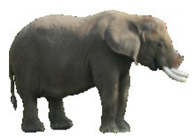
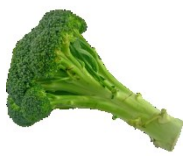
Since color perception was supposed to play an important role in this experiment, bilingual participants with more than one mother tongue were not allowed to participate in this experiment, since, according to some studies (e.g., Winawer et al., 2007), their second mother tongue could have an influence on the speed of color recognition, facilitating the recognition of certain colors, if they happen to be categorized differently in the second mother tongue in comparison with color labels in the German language.

The study was approved by the local ethics committee of Heinrich Heine University Düsseldorf (study number 2019-436-Drittmittel).

3.3.1.2. Stimuli

In Experiment 3, the same pictorial stimuli were used, from which color and part attributes had been extracted for Experiment 2 (see the information on the preparation of the stimuli for Experiment 2 in Section 3.2.1.2). For every concept, 5 different images were created: the original image in which the original color and part attributes were represented, one image with an altered part attribute, and three images with different degrees of color modifications (i.e., with the color attribute changed by 33%, 66%, and 100%). In the altered color conditions, a color of another item from the same semantic category was used, which was not typical for a given item, e.g., broccoli was changed from green as its original color to red as the color of a strawberry. The altered image was then superimposed onto the original image and the opacity of the former was set to 33%, to 66%, and to 100% to create the corresponding conditions with various degrees of color changes. For the conditions involving part attributes, a conceptual attribute used in Experiment 1 was replaced by an analogous structural part of another concept from the same semantic category while matching the color of the original picture. Similarly to the stimuli in Experiment 2, all pictorial stimuli of Experiment 3 were calibrated to the size of 300×300 pixels and saved in the PNG format on a transparent background using an open-source raster graphics editor "Krita", version 4.3.0 (URL: https://krita.org). Examples of these stimuli are provided in Table 5.[4]

Table 5. Examples of stimuli in Experiment 3.

	original	color 33%	color 66%	color 100%	part 100%
ANIMALS					
FOOD					

[4] A full set of stimuli for this experiment is available in the OSF (Mnogogreshnova, 2023, URL: https://osf.io/3fbn2/?view_only=ee43eb8a793e408ab8d6b504bc4daaff). The supplementary materials can also be accessed using a QR code (see Appendix).

Once the pictorial stimuli were ready, combinations of stimuli for each trial were determined, such that the alternatives in each trial were represented by one original image and one image with a changed attribute (a picture with a replaced part or with a color changed by 33%, 66%, or 100%). As a result, a set of stimuli was compiled, in which 4 pairs of alternatives were created for each item: an original image vs. an image with an altered part, an original image vs. an image with a color altered by 33%, an original image vs. an image with a color altered by 66%, and an original image vs. an image with a color altered by 100%. These stimuli were used to create 4 lists of stimuli, in each of them every concept was presented in 4 trials (once in each condition described above), the position of the images and the expected responses were counterbalanced within each list, i.e., in two of these trials the original image was on the left side of the screen and the altered image was on the right side of the screen (in one of these cases the participants were expected to choose the image on the left, in another one they were expected to choose the image on the right), while in the other two trials the position was reversed, whereas the expected responses were counterbalanced the same way. The total number of trials in each list of stimuli amounted to 372 and consisted of 368 experimental trials preceded by 4 warm-up trials with images that were prepared the same way as the experimental stimuli and belonged to the same semantic categories. The program “Mix” (van Casteren & Davis, 2006) was used to pseudorandomize these lists of stimuli under the following constraints: trials involving identical concepts had to be separated by at least 10 trials involving other objects; items from the same semantic category could not be used in more than 3 subsequent trials; the attribute (part or color) could not be repeated more than 4 times in a row; identical expected responses (e.g., “left” or “right”) could not be given more than 3 times in a row.

3.3.1.3. Procedure

Similarly to Experiment 2, Experiment 3 was programmed in PsychoPy (Peirce et al., 2019) and was conducted online via Pavlovia (URL: https://pavlovia.org). Participants received an email with detailed instructions and a participant number to ensure that each list of stimuli was used by an equal number of participants and to enable targeted replacement of participants (for example, if their data was incomplete or was not saved due to technical problems). Experiment 3 utilized the same type of task, namely the two-alternative forced choice task, and both the structure of the trials and the instructions for the participants were the same as in Experiment 2, the only difference being that a new set of stimuli was used in Experiment 3 (see a sample trial provided in Figure 8).

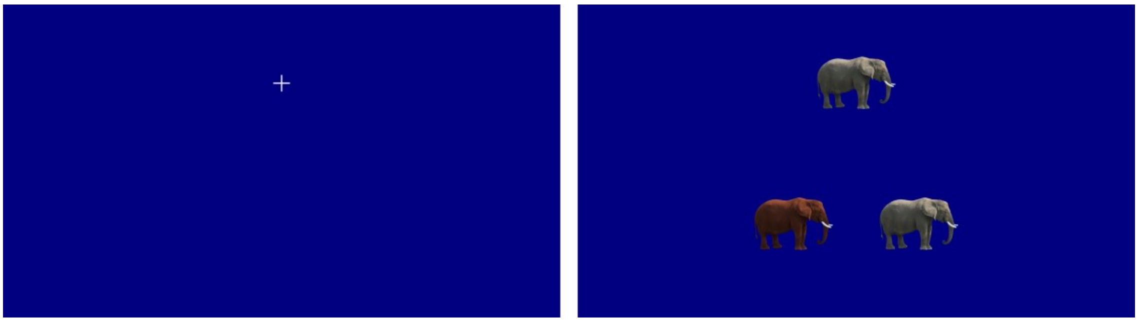


Figure 8. Sample trial from Experiment 3.

3.3.2. Results

The total number of data points collected in Experiment 3 was 14720. To make sure that the dataset is suitable for the error rate analysis, error rates per participant and per item were checked, so that there would be neither participants who gave wrong responses in more than 25% of trials nor items which caused errors in more than 25% of trials. No participants were eliminated as a result of this step of data filtration, however, 2 items (pictorial equivalents of items "Hahn" and "Pflaume") were filtered out, because they had error rates of 25% or more. For the item "Pflaume", errors were detected in 40 out of 160 trials: 25 errors were made in trials with 33% color changes, 11 errors in trials involving 66% color changes, 3 errors in trials with 100% color changes, and 1 error in trials involving part changes. In this case, that might have been caused by the fact that the skin of a plum, which was described by the original color attribute "ist lila" ("is purple"), was dark, making color manipulation less recognizable. The item "Hahn" caused errors in 41 out of 160 trials: 22 errors were detected in trials with part changes, 15 errors in trials with 33% color changes, 3 errors in trials with 66% color changes, and 1 error in trials involving 100% color change. Thus, for this item the part changes turned out to be less noticeable, which could be explained by the fact that the original part attribute was "hat Füße" ("has feet"), which describes a very small part of the image, so the participants could have missed that level of details. After filtering out these two items, the dataset with the remaining 14400 trials was considered suitable for the error rate analysis.

In order to prepare the dataset for the analysis of reaction times, further steps of filtration were necessary. First, 1138 trials, in which wrong responses or no responses were given, were filtered out. Next, it was necessary to make sure that the mean reaction time of each participant as well as the mean reaction time to each item were within 3 standard deviations from the overall mean reaction time. No trials were deleted based on this criterion. Finally, reaction time in each trial had to be within 3 standard deviations from the corresponding participant's mean reaction time (this led to elimination of 165 trials, which constitutes 1.1% of the dataset). As a result of the whole process of data filtration, the final dataset contained 13097 trials, which amounts to 89% from the overall number of trials.

Error Rates

Error rates in trials with original and changed target pictures in every experimental condition for both semantic categories are represented in Figure 9.

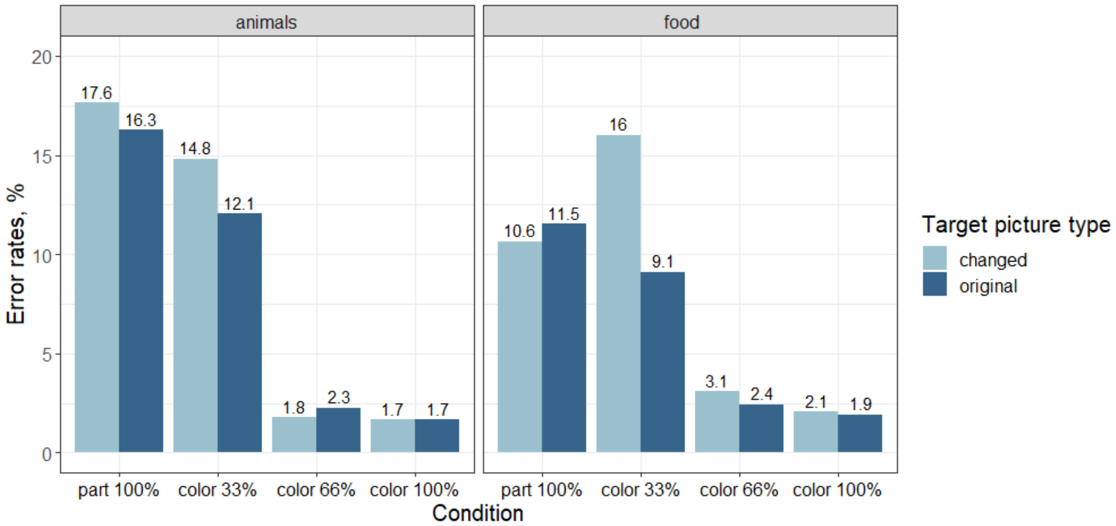


Figure 9. Error rates in Experiment 3.

To analyze the error rates in the data obtained in Experiment 3, a binary logistic regression analysis was performed based on a model that included the following fixed factors: semantic category (animals, food items), condition (color 33%, color 66%, color 100%, part 100%), type of the target picture (original, changed), and attribute prominence for the semantic category (primary, secondary), as well as participants and items as random intercepts.

Since there are some items with deviating attribute prominence, this model was compared to another model, in which attribute prominence for the semantic category was replaced with attribute prominence for individual items, but the rest of the structure remained the same. These two models were compared using ANOVA and AIC values. Removing attribute prominence for the semantic category from the model containing this factor resulted in a significant difference (Chisq(1)=15.755, $p<0.001$), whereas removing attribute prominence for individual items from the other model showed that this factor had no significant main effect in terms of error rates ($p>0.06$). To find out, which of these models better fits the data, a function computing the Akaike information criterion of each model was used. The subsequent comparison of the AIC values of both models showed that the model including attribute prominence for the semantic category had a lower AIC score than the model with attribute prominence for individual items, which means that attribute prominence for the semantic category was a better predictor of error rates. Thus, it was decided to use the model including attribute prominence for the semantic category in the following error rate analysis. The results of the comparison between AIC scores of the two models are presented in Table 6:

Table 6. Comparison of binary logistic regression models containing attribute prominence for the semantic category vs. attribute prominence for individual items in Experiment 3.

	K	AICc	Delta AICc	AICcWt	Cum.Wt	LL
exp3.er_prominence.for.category	9	6734.47	0.00	1	1	-3358.23
exp3.er_prominence.for.item	9	6746.70	12.23	0	1	-3364.34

According to the results of the error rate analysis, there was a marginally significant main effect of semantic category (Chisq(1)=3.897, $p<0.05$), such that there were slightly less errors made in trials with food items than with animals.

There was also a highly significant main effect of condition (Chisq(3)=745.31, $p<0.001$), indicating that there were more errors made in the condition “color 33%” than in “color 100%” ($p<0.001$), more errors in “part 100%” than in “color 100%” ($p<0.001$); the error rates were also higher for “part 100%” condition than for “color 66%” ($p<0.001$), and the error rates in the condition “color 33%” were higher than in the condition “color 66%” ($p<0.001$). There was, however, no significant difference in error rates between the conditions “color 66%” and “color 100%” ($p>0.1$) or between the conditions “color 33%” and “part 100%” ($p>0.3$).

There was also a significant main effect of the target picture type (Chisq(1)=8.536, $p=0.0035$), such that the error rates were higher, when the target picture depicted the changed version of an object than when the original (unaltered) target picture was presented ($p<0.01$).

Another main effect which turned out to be significant was the main effect of attribute prominence (Chisq(1)=15.755, $p<0.001$), showing that lower error rates were observed in trials with secondary attributes in comparison with trials involving primary attributes ($p<0.001$).

The only two-way interaction that turned out to be significant was the interaction between the factors condition and target type (Chisq(3)=9.097, $p<0.03$). The dataset was then divided into four subsets by the levels of the factor condition, and further analysis showed that target type had a significant effect on error rates only in the condition “color 33%” (Chisq(1)=18.244, $p<0.001$), such that participants made more errors in trials with altered target pictures than with original pictures ($p<0.001$).

A three-way interaction between semantic category, condition and target type was also significant (Chisq(9)=18.779, $p<0.03$). To investigate this interaction, the dataset was divided into two parts by the levels of the factor semantic category. Comparison of the interaction between the factors condition and target type in these two subsets showed that this interaction was only significant for food items (Chisq(3)=12.297, $p<0.01$). Further subdivision of the dataset by experimental conditions revealed that target type had a significant effect on error rates only in the condition “color 33%” (Chisq(1)=18.695, $p<0.001$), in which case a higher error rate was observed in trials with altered target pictures than in trials with original target pictures ($p<0.001$). Similar comparisons in other experimental conditions for the same semantic category did not approach the threshold of statistical significance ($p>0.3$ in all cases). Even though the interaction between the factors condition and target picture type did not reach significance for the semantic category “animals”, dividing the corresponding subset of the data by the levels of the factor condition revealed that, for the condition “color 33%”, there was a marginally significant main effect of the target picture type (Chisq(1)=3.858, $p<0.05$), showing that participants also tended to make more errors, when a changed picture appeared as a target than when an original picture was used as a target. There was no significant main effect of the target picture type for any other condition: “color 66%” ($p>0.4$), “color 100%” ($p>0.9$), or “part 100%” ($p>0.3$).

Reaction Times

Reaction times in trials with original and changed target pictures in every condition for both semantic categories are presented in Figure 10.

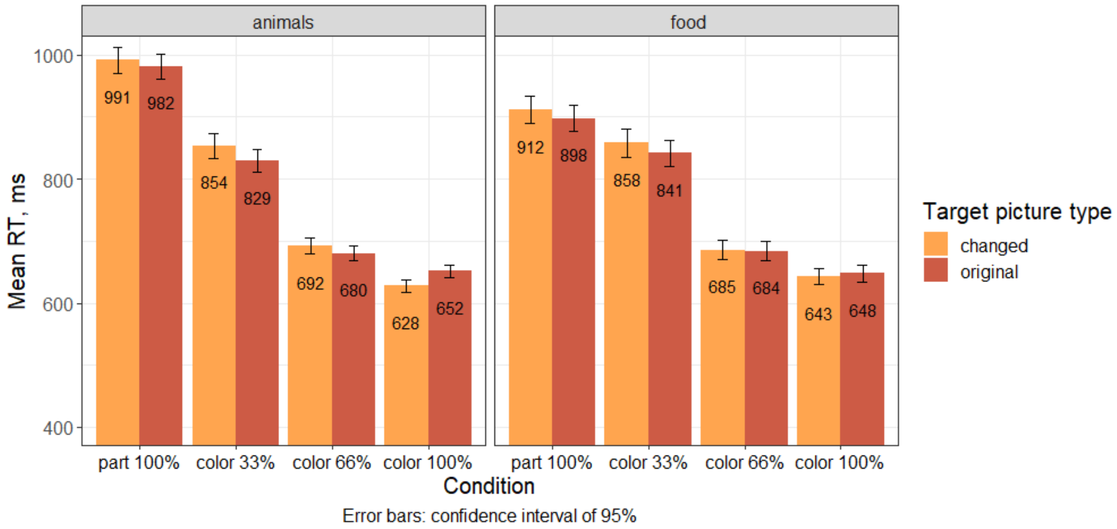


Figure 10. Reaction times in Experiment 3.

To analyze the reaction time data, the "lme4" package (Bates et al., 2015) was used in R Studio (R Core Team, 2022). For this purpose, a mixed-effects logistic model was built that included the same factors as the model for the error rate analysis presented above: semantic category (animals, food), condition (color 33%, color 66%, color 100%, part 100%), type of the target picture (original, changed), and attribute prominence for the semantic category (primary, secondary) as fixed factors; whereas the random factors were participant and item. Since reaction times were not normally distributed, a base-10 logarithmic function was computed based on reaction times in each trial and used as the dependent variable in the model.

A second model with attribute prominence for individual items instead of attribute prominence for the semantic category was compiled to estimate which of these factors would be a better predictor for the reaction times in this experiment. These two models were then compared using ANOVA and AIC values of the models. Removing attribute prominence for the semantic category from the model containing it resulted in a significant difference (Chisq(1)=81.276, p<0.001). Removing attribute prominence for individual items revealed that this factor also has a significant main effect, albeit of a smaller magnitude (Chisq(1)=36.392, p<0.001). Comparing AIC scores of the two models in question showed that attribute prominence for the semantic category was a better predictor, since the model containing it had a much lower AIC value. Therefore, the model containing this factor was subsequently used in the reaction time analysis. The results of this comparison are presented in Table 7 below:

Table 7. Comparison of linear mixed-effects models containing attribute prominence for the semantic category vs. attribute prominence for individual items in Experiment 3.

	K	AICc	Delta AICc	AICcWt	Cum.Wt	LL
exp3.rt_prominence.for.category	10	-21532.02	0.00	1	1	10776.02
exp3.rt_prominence.for.item	10	-21487.14	44.88	0	1	10753.58

According to the results of the analysis of variance performed based on the observed reaction times, there was a significant main effect of condition (Chisq(3)=4189.6, p<0.001) with statistically significant differences between all four conditions: “color 33%” (M=2.902, SD=0.146, SE=0.0026), “color 66%” (M=2.819, SD=0.117, SE=0.002), “color 100%” (M=2.794, SD=0.106, SE=0.0018), and “part 100%” (M=2.956, SD=0.138, SE=0.0025).

There was also a significant main effect of attribute prominence for the semantic category (Chisq(1)=81.276, p<0.001) with pairwise comparisons showing longer reaction times in trials with primary attributes (M=2.873, SD=0.149, SE=0.0019) than in trials with secondary attributes (M=2.855, SD=0.136, SE=0.0016).

Apart from that, there was a significant interaction between the factors condition and target picture type (Chisq(3)=15.75, p=0.0013). In order to obtain more detailed information about the interaction, the dataset was divided into four subsets based on the four conditions in question. Further analysis showed that the main effect of target picture type was only significant in the conditions “color 33%” (Chisq(1)=7.4797, p<0.01) showing shorter reaction times in trials with original target pictures (M=2.898, SD=0.141, SE=0.00355) as compared to trials with changed target pictures (M=2.907, SD=0.151, SE=0.004), and “color 100%” (Chisq(1)=11.421, p<0.001) with longer reaction times in trials with original target pictures (M=2.799, SD=0.1097, SE=0.003) than in trials with changed target pictures (M=2.7895, SD=0.102, SE=0.0024). In the remaining two conditions, “color 66%” and “part 100%”, there was no significant difference between reaction times in trials with original and changed target pictures (p>0.1 in both cases).

There was also a significant three-way interaction between the factors semantic category, condition, and target picture type (Chisq(9)=22.711, p<0.01). The dataset was subsequently divided into two subsets based on the semantic category of items and another analysis of variance was conducted on each subset, revealing that the interaction between condition and target picture type was only significant for the semantic category "animals" (Chisq(3)=13.528, $p<0.01$). Further division of the subset based on the levels of the factor condition showed that the main effect of target picture type was only significant in the condition "color 100%" (Chisq(1)=16.268, $p<0.001$) with longer reaction times in trials with original target pictures (M=2.801, SD=0.1045, SE=0.0033) than with changed target pictures (M=2.786, SD=0.097, SE=0.003).

3.3.3. Discussion

Experiment 3 utilized a two-alternative forced-choice task with pictorial stimuli that represented either an original image of an animal or a food item, or an image with an altered color or part attribute. In each trial of this experiment, the participants had to match one of the two alternatives presented on the screen to a target picture presented on the screen at the same time. Whereas Experiment 2 contained color and part attributes isolated from their respective images, Experiment 3 aimed to investigate whether the fact that images depicted objects as a whole could influence processing of color and part attributes and whether a processing advantage of color attributes found in Experiments 1 and 2 could also be observed at the conceptual level. Apart from that, this experiment would also help to establish which color change could be considered comparable with the condition involving changes in part attributes, so that it would be possible to use those two conditions in a subsequent verification task with pictorial stimuli.

As far as reaction times are concerned, the overall pattern of results was largely similar to the pattern of reaction times observed in Experiment 2, meaning that the longest reaction times in Experiment 3 were measured in trials involving part attributes, followed by somewhat shorter reaction times in the condition with a 33% color change and even shorter reaction times in the condition involving a 66% color change, while the shortest reaction times were observed in the condition with a 100% color change. Nevertheless, while the relative length of reaction times among these four conditions remained similar to that of Experiment 2, reaction times in all conditions of Experiment 3 were generally longer in terms of their absolute values. This could be explained by the fact that images in Experiment 3 were no longer simple color patches or separate parts of concepts, but rather whole images of concrete objects, which elicited longer reaction times due to their higher visual complexity.

Counterintuitively, reaction times to primary attributes happened to be longer than to secondary attributes. It might be the case that the participants in Experiment 3 spent longer periods of time comparing the two alternatives to make sure which one matches the target picture, and, apparently, stimuli with manipulations in their primary attributes turned out to be more prone to this process. Similar re-checking processes accompanied by a P600 effect are also induced in experiments involving strong violations of expectations in the form of orthographic errors in high-cloze sentences (e.g., Vissers et al., 2006) or semantically completely incongruent critical stimuli (e.g., van de Meerendonk et al., 2008). Since primary attributes are more prominent for a given semantic category, it would be reasonable to assume that they represent a stronger expectation than secondary attributes do. A somewhat similar conclusion was also reached by Võ and Wolfe (2013), who investigated semantic and syntactic scene processing. In their study, they used images depicting semantic inconsistencies (i.e., scenes with objects that do not belong to the environment they are presented in, like a soap bar near a laptop) and mild or extreme syntactic violations (mild violations represented misplaced objects in an otherwise appropriate environment, while extreme syntactic violations represented objects violating the laws of physics in an otherwise semantically consistent environment). In their article, the authors note that mild syntactic violations "may trigger scene reanalysis marked by a P600 response" (Võ & Wolfe, 2013: 1822). No such effect was found for the extreme syntactic violations, indicating that, in order for re-analysis processes to be initiated, the violation in question has to be correctable. Scenes without a critical object and scenes with a critical object were presented one at a time, but the participants received a visual cue to indicate where to expect a consistent or an inconsistent critical object to appear in the image, which made it possible for them to compare the images with and without the critical objects. Unfortunately, they used a different kind of task (their participants had to press a button, if they detected a repetition of a scene), which renders it impossible to compare their reaction time data to the data from Experiment 3. Nonetheless, their definition of a mild syntactic violation could be compared to an attribute violation in the current experiment.

When it comes to error rates, the percentage of errors in trials involving 66% color changes turned out to be comparable to the percentage of errors in trials with 100% color changes, while the percentage of errors made in the condition involving part attributes was similar to that in the condition with 33% color changes. The percentage of errors in the condition “color 33%” was also much higher than in either of the other two color-related conditions. Had it been merely due to the fact that 33% color changes are generally harder to detect, it would not have made any difference whether the target picture in this condition was an original image or an image with a slightly altered color, and the error rates would be similar irrespective of the target picture type, similar to the result of Experiment 2. However, in the condition “color 33%” of Experiment 3 participants also tended to make significantly more errors in response to altered target pictures in comparison with the original target pictures, i.e., the participants were more likely to choose the original version of the image when presented with a target picture having a slightly changed color than to choose a picture with an altered color when presented with a non-modified target picture. In other words, participants seem not to have detected a slight color change in target pictures. This effect was particularly strong for food items, for which color was a primary attribute. It is likely that this bias is connected to the influence of object knowledge on color perception, due to object knowledge including information about the typical color associated with an object. In other words, color as an attribute in the conceptual frame representation of the object is able to affect the participants’ response. Our result is in line with the results of a study by Mitterer and de Ruiter (2008) who presented their participants with objects in different hues between yellow and orange and asked them to classify the color that they saw in the picture either as yellow or as orange. Ambiguous hues were more often classified as yellow, when they were presented on prototypically yellow objects (e.g., a banana), and as orange, when they were presented on prototypically orange objects (e.g., a carrot), whereas no such bias was observed with neutral objects that had no typical color (e.g., a sock).

All in all, even though there was a significant difference in reaction times between conditions involving part changes and 33% color changes, the error rates reveal that these conditions are comparable with one another. In addition, they are similarly affected by depicting an object as a whole as opposed to depicting only a color patch or a part attribute. Taking this finding into consideration, it would be appropriate to use 33% color changes in Experiment 4 along with the condition involving part changes, since “color 33%” proved to be more comparable to the condition involving part changes than any of the other degrees of color changes used in Experiments 2 and 3.

3.4. Experiment 4

Experiment 1 utilized a verification task with verbal stimuli, however, it was not clear whether the effects observed in that experiment arose purely on the lexical level or if they could also be detected on the conceptual level. Experiment 4 aimed to resolve that issue by replicating the verification task using pictorial stimuli. For this purpose, it was necessary to determine a degree of color change that would be most comparable to the changes in part attributes. According to the findings of Experiments 2 and 3, it turned out to be the condition involving 33% color changes. Thus, Experiment 4 included the same two semantic categories that were used in Experiments 2 and 3 (animals and food), and three conditions: an original image, an image with a 33% color change, and an image with a changed part attribute. The task for the participants was to press a button indicating whether the object in the image was depicted correctly or whether there was something wrong with its depiction. According to our hypothesis, if the most prominent attributes were accessed first, then a violation in a more prominent attribute (a body part of an animal or color of a food item) would be detected earlier than a violation in a less prominent attribute (color of an animal or a part of a food item). On the other hand, it was also possible that part attributes are more complex in terms of processing, leading to generally longer reaction times in the condition involving part changes irrespective of the prominence of this attribute for a given semantic category and, consequently, to the absence of the main effect of attribute prominence for this type of attribute, given that a similar finding was observed in Experiment 1.

Since only one picture was presented in each trial of this experiment, the participants were expected to have relatively high error rates in comparison with Experiments 2 or 3, where they had other images to compare the target image to.

3.4.1. Methods

3.4.1.1. Participants

Sixty-three German native speakers (age: 18 to 38 years old, mean age 22.95 years) took part in Experiment 4. All participants reported to have normal or corrected-to-normal vision, normal color perception, and no second mother tongue besides German.

None of these participants took part in Experiments 2 or 3. The study was approved by the local ethics committee of Heinrich Heine University Düsseldorf (study number 2019-436-Drittmittel).

3.4.1.2. Stimuli

The same pictures as in Experiment 3 were also used as stimuli in Experiment 4, however, the conditions “color 66%” and “color 100%” were left out, since, according to Experiments 2 and 3, reaction times and error rates in these conditions were not comparable with reaction times and error rates in the condition “part 100%”. As a result, for every concept there were three types of pictures: an original picture, in which none of the attributes were altered, a picture with a 33% change in color, and a picture in which an object had an altered part attribute. In the process of stimuli preparation, the pictures in question were saved as PNG files with transparent background and had the size of 500×500 pixels.[5]

Three lists of stimuli were organized in such a way that every concept was presented in each list only in one of the three conditions (an original image, an image with an altered color, or an image with an altered part), leading to the total number of experimental trials of 92 per list (52 animals and 40 food items). To equalize the number of trials with original pictures and trials with altered pictures (i.e., number of trials with two different expected responses), filler pictures were used, which matched the format, size and semantic categories of the experimental stimuli, while also making the number of food items equal to the number of animal items in each list. The total number of trials in each list amounted to 132, including 8 warm-up trials created in the same way as the experimental stimuli.

3.4.1.3. Procedure

The experiment was programmed using PsychoPy (Peirce et al., 2019) and conducted online via Pavlovia (URL: https://pavlovia.org). Before the experiment started,

[5] A full set of stimuli for Experiment 4 is available in the OSF (Mnogogreshnova, 2023, URL: https://osf.io/3fbn2/?view_only=ee43eb8a793e408ab8d6b504bc4daaff). These materials can also be accessed via a QR code (see Appendix).

the participants were asked to fill out three text fields. In the first field, they indicated their month of birth using a number between 1 and 12, which was then used to assign a certain list of stimuli to each participant: list 1 corresponded to numbers from 1 to 4, list 2 – to numbers from 5 to 8, and list 3 – to numbers from 9 to 12. For example, if a participant was born in August, number 8 would be typed in the text field and the program would select the second list of stimuli. In the second field, participants wrote their email address, which would help to filter out the data from people who participated in Experiments 2 and 3, so that we could ensure, as far as possible, that the remaining participants were not familiar with the pictorial stimuli. In the third text field, they indicated their age. Next, the participants were presented with instructions clarifying the task and the buttons which they would have to use. The participants were supposed to look at the target picture in each trial and decide whether that picture depicted an object correctly or if there was something wrong in the picture (the ways in which the stimuli had been modified were not explicitly stated). The participants had to press one of two buttons as quickly as possible to express their decision: an arrow to the left for pictures with attribute violations or an arrow to the right for correct depictions of objects. As soon as they read and understood the instructions, they pressed a space bar on their keyboard and the experiment began, starting with 8 warm-up trials followed by experimental items from the list of stimuli selected by the program based on the participant's month of birth.

Each trial began with a fixation cross shown in the center of the screen for 1 second followed by a target picture presented in the center of the screen. The target picture remained on the screen for 2 seconds or until a participant gave a response by pressing one of the two buttons. If no response was given within 2 seconds, the target picture disappeared, an inter-trial interval of 500 ms was initiated, and after that the next trial started (see Figure 11 for a sample trial). At the end of the experiment, the participants were informed that the experiment was finished, they were asked to press the space bar and then could close their browser window.

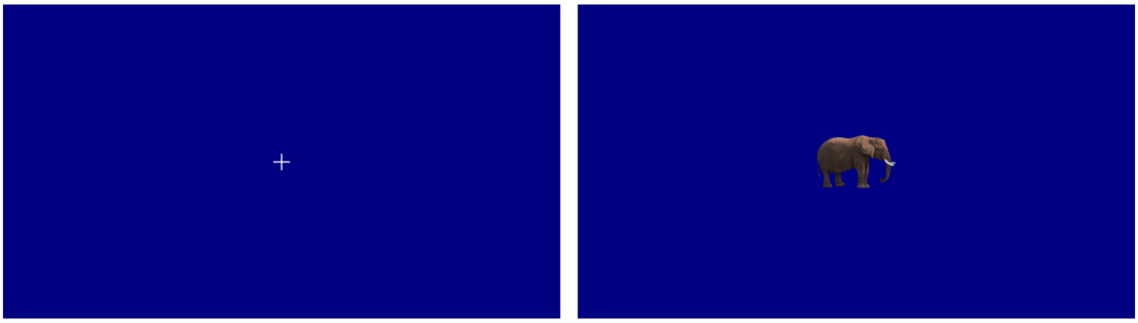


Figure 11. Sample trial in Experiment 4.

The data obtained from the participants who did not fill out the three fields prior to starting the experiment or who failed to perform the task throughout the whole experiment (e.g., if they had technical problems preventing them from using the buttons properly or if they stopped giving any responses in the middle of the experiment) was discarded. To ensure that each image was seen by an equal number of participants, it was planned that each list of stimuli would be used by 20 participants. If the expected amount of data was obtained using a particular list of stimuli, it was replaced in such a way as to ensure that other lists of stimuli are used instead until they, too, would have been used by 20 participants.

3.4.2. Results

The total number of data points collected in Experiment 4 was 7812. Since high error rates were expected in this experiment due to the absence of reference pictures, error rates per item and per participant were checked, but no items were eliminated based on that factor. Error rates per item varied from 3.18% to 68.25%, mean error rates: 36.62% (excluding fillers: from 15.87% to 68.25%, mean error rates: 44%); error rates per participant ranged from 24.19% to 70.16%, mean error rates: 36.62% (excluding fillers: from 28.26% to 66.3%, mean error rates: 44%). Lower error rates in trials with filler items could be explained by the fact that fillers were unmodified pictures, therefore, participants had no difficulty confirming that objects in those images were represented correctly.

To prepare the dataset for the error rate analysis, filler items were eliminated from the dataset, leaving a total of 5796 trials with experimental items in the dataset. For the reaction time analysis, further data filtration steps were needed. Firstly, trials with incorrect responses were filtered out, after which 3246 trials remained in the dataset. Secondly, it was made sure that mean reaction times of each item and mean reaction times of each participant were within 3 standard deviations from the overall mean reaction time. Moreover, reaction times in each trial also had to be within 3 standard deviations from the respective participant's mean reaction time. As a result, a dataset with a total of 3234 trials (41.4% from the total number of trials in the raw dataset) was considered ready for the reaction time analysis.

Error rates

Percentage of errors in each condition for both semantic categories is represented in Figure 12.

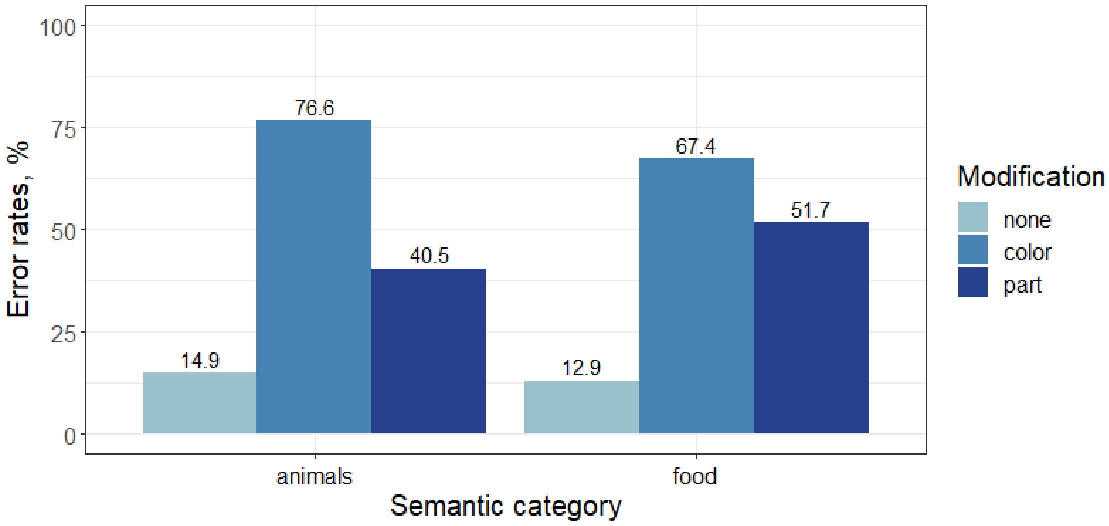


Figure 12. Error rates in Experiment 4.

To analyze the error rates, a binary logistic regression was used. The model for this type of analysis included the following parameters: semantic category (animals, food) and picture modification (none, color, part) as fixed factors, as well as participants and items as random intercepts.

According to the results of this analysis, there was no significant main effect of semantic category ($p>0.9$). However, a significant main effect of picture modification was found ($Chisq(2)=1574.3$, $p<0.001$), showing that there were significantly more errors in trials with color modifications than in trials with part modifications or in trials with non-modified pictures, as well as significantly more errors in trials with part modifications than in trials with non-modified pictures ($p<0.001$ in all pairwise comparisons).

There was also a significant interaction between factors semantic category and picture modification ($Chisq(2)=43.802$, $p<0.001$). The subsequent division of the dataset into two subsets by the values of the factor modification revealed that there was a significant main effect of semantic category for color modifications ($Chisq(1)=5.089$, $p=0.024$) with less errors made in trials with color-modified food items than in trials with color-modified animals. The participants also tended to make less errors in trials with pictures of animals with modified parts, but this effect did not reach the threshold of statistical significance (p=0.08). However, for non-modified pictures, no significant main effect of semantic category was found (p>0.2).

Reaction Times

The reaction time data from Experiment 4 is represented in Figure 13.

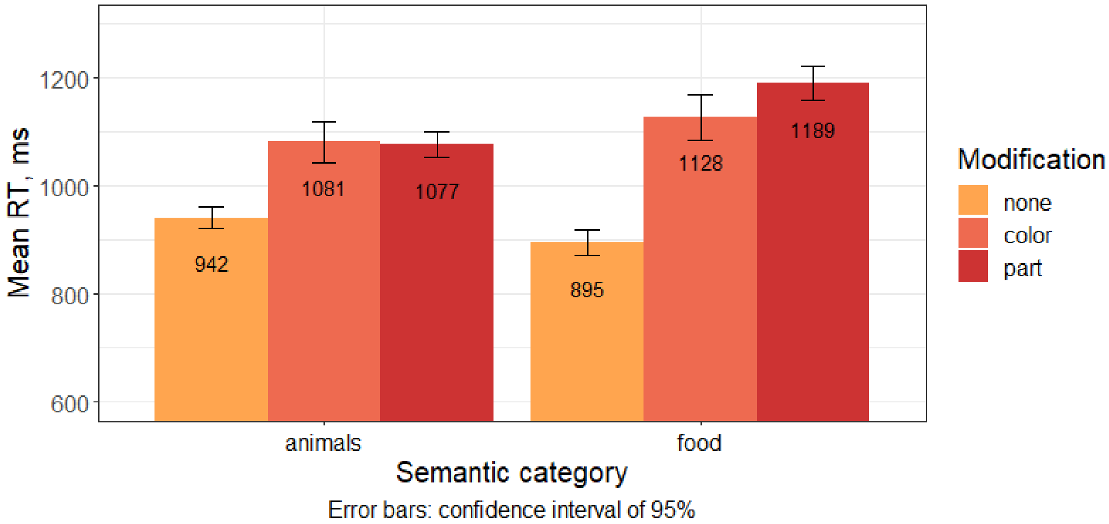


Figure 13. Reaction times in Experiment 4.

To analyze the reaction time data, the “lme4” package (Bates et al., 2015) was used in R Studio (R Core Team, 2022). The optimal random effects structure was determined with the help of the Akaike information criterion and the resulting model included semantic category (animals, food) and modification (none, color, part) as fixed factors, as well as participant identification numbers and items as random intercepts. To make sure that the reaction times were normally distributed, a logarithmic function (base 10 logarithm) was computed based on reaction times in each trial and used as the dependent variable in the model.

According to the results of the reaction time analysis, there was a significant main effect of picture modification (Chisq(2)=387.38, p<0.001). Pairwise comparisons between the experimental conditions revealed that there was a significant difference (p<0.001) between original, i.e., non-modified, pictures (M=2.941, SD=0.142, SE=0.0035) and pictures with altered colors (M=3.025, SD=0.127, SE=0.0056), as well as between original pictures and pictures with altered parts (M=3.033, SD=0.121, SE=0.004). There was, however, no statistically significant difference between the condition with altered colors and the condition with altered parts (p=0.059).

The main effect of semantic category on reaction times was not statistically significant (p>0.6).

Also, an interaction between picture modification and semantic category proved to be significant (Chisq(2)=42.589, $p<0.001$). The dataset was subsequently divided into two subsets based on the values of the factor semantic category and the main effect of picture modification was investigated in both subsets. As a result, the main effect in question was significant for items from the semantic category “animals” (Chisq(2)=123.82, $p<0.001$), showing that the reaction times were significantly shorter in trials with non-modified pictures (M=2.952, SD=0.138, SE=0.005) than in trials with altered colors (M=3.017, SD=0.1196, SE=0.005) or altered parts (M=3.016, SD=0.119, SE=0.005) with $p<0.001$ in both cases, but there was no significant difference in reaction times between the two latter conditions ($p>0.4$). For food items, the main effect of picture modification was significant as well (Chisq(2)=249.3, $p<0.001$), with shorter reaction times for non-modified pictures (M=2.927, SD=0.147, SE=0.005) than for pictures with altered colors (M=3.032, SD=0.134, SE=0.008) or altered parts (M=3.0596, SD=0.119, SE=0.006), however, there was also a significant difference between the condition with altered colors and the condition with altered parts ($p<0.001$), showing that reaction times were shorter in trials with color modifications than in trials with part modifications.

When the original dataset was split into three subsets based on the values of the factor modification, the main effect of semantic category was significant for pictures with part modifications (Chisq(1)=10.838, $p<0.001$) with longer reaction times in trials with pictures of food items with modified part attributes (M=3.0596, SD=0.119, SE=0.006) than in trials with pictures of animals with modified body parts (M=3.016, SD=0.119, SE=0.005), whereas there was no main effect of semantic category for non-modified pictures ($p>0.1$) or pictures with modified color attributes ($p>0.4$).

3.4.3. Discussion

Experiment 4 utilized a property verification task with pictorial stimuli. The participants were instructed to classify pictures they were presented with based on whether an object depicted there looked normal or whether something in its depiction was wrong. In some of the stimuli, a change in color or a change in a part attribute had been implemented, however, the participants were not warned about the types of manipulation involved in the experiment. The effects of picture modification and semantic category on reaction times and error rates were assessed. In Experiment 4, attribute prominence was not introduced as a fixed factor in the mixed-effects model. Instead, if any effects of attribute prominence on reaction times and/or error rates should arise, it would be evident from an interaction between the remaining two factors: the semantic category and the type of picture modification.

The analysis of the data obtained in this experiment showed that, while reaction time and error rate patterns in both semantic categories (animals and food) were generally similar (no main effect of semantic category was observed), there was a main effect of the type of picture modification both in reaction time and error rate analyses as well as an interaction between semantic category and type of picture modification. Reaction times to original (unaltered) pictures were consistently shorter than to pictures with altered colors or part attributes in both semantic categories. For animals, there was no difference in reaction times to images with altered colors in comparison with images with altered parts, however, for food items, reaction times to images with altered colors turned out to be significantly shorter than to images with altered parts. This might reflect the fact that part attributes are secondary for food items and, therefore, participants needed somewhat more time to process them. A similar result was observed in Experiment 1, which utilized a property verification task with verbal stimuli.

Another argument in favor of this interpretation would be the fact that pictures of food items with modified part attributes were also more prone to errors than pictures of animals with modified part attributes (within the experimental condition that included images with modified part attributes, erroneous responses were given in 51.7% of trials with food items and in 40.5% of trials with animals). This could not be explained by the fact that one of these semantic categories was more prone to errors, since error rates in both of them were comparable, i.e., there was no significant difference between the semantic categories “animals” and “food” in terms of error rates in general across all types of picture modifications. Moreover, similar effects of attribute prominence could be observed in color attributes: this type of attributes is considered secondary for the semantic category “animals”, and, indeed, the percentage of errors in trials with pictures of animals with modified colors was higher than the percentage of erroneous responses given in trials in which pictures of food items with modified colors were presented (76.6% vs. 67.4.%). However, reaction times in trials with color-modified food items were comparable to reaction times in trials with color-modified animals.

The fact that error rates were higher for pictures of animals and food items with color modifications than for pictures in either of the other two experimental conditions could be explained by the fact that, unlike in Experiments 2 and 3, this time only pictures with subtle color changes of 33% were used, which made it somewhat harder to detect those color attribute violations. In addition to that, only one image was presented in each trial of Experiment 4, which means that participants had no other image to compare it to, making their mental representations of the objects the only reference to rely on in making the decision on whether a picture provided a truthful depiction of the object in question. Such high error rates could also account for the absence of a significant reaction time effect between color-modified animals and color-modified food items: removing numerous trials with erroneous responses from the dataset to prepare it for the reaction times analysis reduced the statistical power of the analysis, making it harder to find an interaction between the factors semantic category and picture modification, which could have been indicative of an attribute prominence effects with respect to color.

All in all, the results of Experiment 4 indicate that attribute prominence for the semantic category does play a role, when it comes to error rates, making an attribute more prone to errors, if it is secondary for the given semantic category. As for reaction times, the pattern of results there is less clear. Part attributes as secondary attributes of food items caused significantly longer reaction times than part attributes as primary attributes of animals. However, reaction times for color attributes in those two semantic categories remained similar, in spite of their prominence for each semantic category, which might be connected with a relatively large loss of data as a result of filtering out erroneous trials.

4. GENERAL DISCUSSION

4.1. Summary of the Results

The aim of the current dissertation was to investigate the effects of prominence of certain attributes for specific semantic categories on time and accuracy of retrieval of those attributes. For that purpose, a series of four experiments was conducted involving stimuli from three semantic categories: animals, food, and tools. Attribute prominence was determined based on how frequently a given attribute is mentioned in response to a concept label, according to the feature database compiled by McRae et al. (2005).

In Experiment 1, a property verification task with verbal stimuli was used, in which the participants were presented with attributes and congruent or incongruent conceptual labels belonging to one of three semantic categories (animals, food, or tools) and had to indicate whether the concept they were presented with had the property shown prior to it. For animals, the primary attribute was a part attribute and the secondary attribute was a color attribute; for food items, color attributes were primary and part attributes were secondary; for tools, the primary attribute was affordance and the secondary attribute was a part attribute. Since confirmation of a congruent attribute would be prone to priming effects due to direct connections between congruent attribute values and their respective concept labels, negative responses were of more interest in that experiment. According to the results of the analysis, overall incongruent primary attributes were rejected more quickly than incongruent secondary attributes. The responses were also less accurate in trials with secondary attributes than in trials with primary attributes. No significant difference in reaction times or error rates was found between primary and secondary incongruent attributes of animals. A follow-up analysis involving part and color attributes of animals and food items revealed an effect of attribute, such that it took longer to reject part attributes than color attributes, and part attributes also tended to cause somewhat higher error rates than color attributes, probably due to their higher complexity. Given that verbal stimuli were used in this experiment and that incongruent attributes were taken from other members of the same semantic category, the fact that primary attributes were rejected more quickly and more accurately than secondary attributes could have an alternative explanation in the form of mediated priming effects. Thus, a property verification task with pictorial stimuli was necessary to resolve this issue.

Considering that color attributes were found to have an advantage over part attributes in terms of processing, it was decided that a more subtle color change should be used in the images, to compensate for this advantage and to find a degree of color change that would be comparable to the changes in part attributes. To determine the optimal degree of color change, Experiment 2 was conducted in the form of a two-alternative forced choice task aiming to test perceptual discriminability of part changes as well as several degrees of color changes (33%, 66%, and 100%). Pictorial equivalents of the verbal stimuli from Experiment 1 that belonged to the semantic categories "animals" and "food" were used to obtain isolated attribute values: color patches of original and altered colors as well as original and altered parts of the objects. In each trial, the participants saw a target picture in the upper half of the screen and two alternatives below (one of them was always an original attribute value and the other one represented an altered attribute value), from which they had to choose the one that corresponded to the target picture. The results of Experiment 2 revealed that the condition involving 33% color changes was the closest match to the condition involving part changes in terms of reaction times, but when it came to error rates, the condition with 66% color changes was a better match to the condition with part changes. To determine, which of these conditions could be used in a verification task with pictorial stimuli, another experiment was designed using the same part changes and the same degrees of color changes in images depicting not isolated attribute values, but whole objects.

In Experiment 3, the two-alternative forced choice task was replicated using pictorial stimuli, in which whole objects (animals or food items) were visible, presented in the same conditions: in each trial, a target picture was presented in the middle of the upper half of the screen simultaneously with two alternatives in the bottom half of the screen (one of them was always the original image and the other was an altered version of it, i.e., a picture in which a part of the object was changed, or its color was changed by 33%, 66%, or 100%). The participants had to choose the alternative corresponding to the target picture. The results of Experiment 3 showed that the condition involving 33% color changes was still the closest match to the condition with part changes in terms of reaction times (even though there was still a significant difference in reaction times between these two conditions), and that the error rates in trials with 33% color changes and in trials with part changes were comparable, making this condition a suitable color degree change condition to be used in the subsequent property verification task with pictorial stimuli.

Furthermore, the error rate analysis revealed a color correction bias with regard to food items in the condition "color 33%", which will be addressed later in this chapter.

Experiment 4 was a property verification task with pictorial stimuli depicting whole objects belonging to the semantic categories "animals" and "food". The participants saw a single target picture in each trial and had to indicate via a button press whether the target picture was a truthful depiction of an object. Each object was presented in one of three conditions: an original image, an image with an altered part attribute, or an image with an altered color. The analysis revealed a pattern of reaction times similar to that of Experiment 1: there was no difference between reaction times to images of animals with modified colors or modified parts, whereas for food items, color violations were detected significantly faster than part violations. While this might be an indication that both attribute prominence and visual or processing complexity of an attribute influence the speed of verification, a relatively large loss of data as a result of high error rates in Experiment 4, however, led to a reduction of statistical power, which was why the results of the error rate analysis were taken to be more reliable indicators of attribute prominence effects. And, indeed, the error rate pattern showed that both color and part attributes were less prone to errors when they were primary attributes for a given semantic category than when they were secondary attributes.

4.2. Attribute Prominence

While the aim of the current study was to investigate the role of attribute prominence in conceptual processing, it was discovered in the course of the preparation of the stimuli for Experiment 1 that for individual items the attribute prominence does not necessarily coincide with the attribute prominence for the semantic category as a whole. Both parameters – attribute prominence for individual items and attribute prominence for the semantic category – were used as fixed factors in the analysis of reaction time data as well as error rates in Experiments 1 and 3 described in the present thesis. This factor was omitted in Experiment 2, since color patches do not provide any information about the objects the colors belong to, therefore, no access to information on attribute prominence for the semantic category or for individual items is possible in such cases. In Experiment 4, attribute prominence effects for the semantic category were evident from the interaction between the factors semantic category and modification type, which rendered inclusion of this factor into the model unnecessary. As a result, it was consistently observed throughout Experiments 1 and 3, in which attribute prominence for the semantic category and attribute prominence for individual items were compared, that attribute prominence for the semantic category as a whole was a better predictor in the analysis of reaction times as well as in the analysis of error rates. A possible explanation for that finding could be that the differences in attribute frequencies between primary and secondary attributes of deviant items, i.e., of the items with an attribute prominence different from the attribute prominence for the respective semantic category, were much smaller than the differences between attribute frequencies of primary and secondary attributes of regular items, for which attribute prominence corresponds to the pattern expected for their semantic category. However, even in that case a model including attribute prominence for individual items instead of attribute prominence for the semantic category as a whole should have resulted in a slightly better fit to the data, which was not the case. Essentially, this result suggests that attribute prominence for the semantic category as a whole has a priority over attribute prominence for individual items, which could mean that some attributes and/or information about their prominence are stored at the level of the category rather than at the level of individual concepts. This organization of semantic memory would make retrieval of such attributes more efficient, since the information that is true of all or most members of a category does not have to be replicated for each conceptual node, instead it can be stored at the level of a category and retrieved via the fact that a given concept is a member of a particular category. An example of this kind of semantic memory organization was described by Collins and Quillian (1969): “the fact that a canary can fly can be inferred by retrieving that a canary is a bird and that birds can fly” (Collins & Quillian, 1969: 240). When, however, a member of a category is encountered that does not have that kind of an attribute value (e.g., an ostrich cannot fly), this information is then stored at the level of that particular concept, in order to prevent the wrong inference about the concept at an early stage of processing. McCloskey and Glucksberg (1979), while discussing the issue of typicality of certain concepts within a semantic category, also considered the possibility that some attributes are stored at the level of the category itself rather than at the level of an individual concept: “Typical category exemplars are those whose values for many attributes are among the values stored with the category itself. Atypical exemplars, on the other hand, are those which have, for many attributes, values that are not among those stored for the category” (McCloskey & Glucksberg, 1979: 13). Nevertheless, given a wide variety of colors for food items or for animals, it would not be particularly efficient to store all of them at the category level. However, information stored at the category level is more likely to include the weight of a certain type of attributes, e.g., the fact that something is a fruit could provide us with the information that color is of higher importance than parts of that object. In other words, information about prominence of a particular attribute is stored at the level of a category, even if the attribute itself is stored at the level of the concept node. For example, even though red is a very frequent color when it comes to fruit (it is the typical color of cherries, strawberries, tomatoes, etc.), not all kinds of fruit are red, so it would not be necessary to store this attribute value at the level of the whole category, since otherwise there would be too many exceptions (e.g., cucumbers, lemons, oranges, etc.). And yet, we do know that color is an important characteristic for fruit, because knowledge about the color that a particular fruit is supposed to have when it is ripe helps us to determine whether a particular exemplar of the fruit category is edible or not. Thus, it would be essential to pay attention first of all to color attributes, when it comes to fruit, and the information about the order of importance of at least some attributes might be stored at the level of the semantic category that a given concept belongs to. This information can then be used, for example, to guide our attention to the respective attributes in a verification task or when we are trying to judge the edibility of a fruit in real life.

The fact that there are differences between representations of different semantic categories in terms of prominence of certain attributes is also supported by evidence from category-specific semantic deficits, i.e., language impairments affecting certain semantic domains more than others. Such deficits can be explained by modular accounts hypothesizing that there are distinct areas of the brain responsible for mental representations of objects from different semantic categories in general or for mental representations of attributes that belong to different modalities, whereas items within one and the same category are represented by neural structures that have more overlap due to the items' similarity within a given category. In order to explain the vulnerability of certain categories and properties to language impairments, the Conceptual Structure Account, or CSA (Moss, Tyler & Taylor, 2007), takes into consideration such aspects of mental representations as the total number of features, their distinctiveness, intercorrelation, and types of features in question. According to the CSA, the type of an attribute defines whether the attribute in question describes a particular perceptual property of a concept (e.g., shape, color, sound, etc.), a functional property (e.g., use or behavior), or an abstract characteristic of a concept. The authors of this account also agree that the salience of certain attributes differs across semantic categories, which not only gives rise to differences in category-specific language impairments (i.e., the fact that a lesion affecting processing of a particular modality has a larger impact on the semantic category for which the corresponding attribute is more prominent), but also explains some phenomena observed in psycholinguistic studies with healthy participants (Moss, Tyler & Taylor, 2007). Thus, it is not the difference between categories per se, but rather the difference in prominence of certain attribute types that gives rise to category-specific impairments of semantic memory representations (Farah & McClelland, 1991).

4.3. Color Correction Bias

Experiment 3 yielded an unexpected finding concerning color attributes that was also connected to attribute prominence. Specifically, the participants made significantly more errors in trials involving images of food items with 33% color changes, when presented with an altered target picture (16%) than when an original target picture was presented (9.1%). In other words, even though they had a matching altered target image as a reference, they still chose the original alternative over the altered one more often than they chose an altered version as a match to an original target picture. In trials involving animals with the same degree of color changes, a difference in error rates between trials with altered targets (14.8%) and trials with original targets (12.1%) did not reach significance, and no such effect was observed for the conditions involving 66% or 100% color changes or part changes in either semantic category.

As opposed to 66% and 100% changes in color, which are more easily detectable, subtle 33% color changes in Experiment 3 could be compared to the ambiguous hues from the experiments by Mitterer and de Ruiter (2008), in which they presented prototypically yellow, prototypically orange, or color-neutral objects in a range of hues: yellow, orange, or one of 5 ambiguous hues between those two values, and after the exposure phase asked their participants to categorize an ambiguous color on a color-neutral object. As expected, their participants developed a bias towards the prototypical color of the objects they were presented with during the exposure phase, whereas exposure to color-neutral objects created no such bias, allowing the authors to conclude that object knowledge could influence the way ambiguous hues were categorized.

Therefore, a plausible explanation for this color correction bias would be that the participants' knowledge about the color a given concept should have had caused them to recalibrate their color perception and see the color altered by 33% as being the same as the "right" (original) color of the object, in spite of the simultaneous presence of all stimuli (the target and both alternatives) on the screen.

A similar result was obtained in a study by Vandenbroucke et al. (2016), who utilized line drawings of typically red, typically green or nonsense objects filled with ambiguous colors between red and green in two experiments. In the first experiment, they used fMRI and presented their participants with 16 objects from one of the sets of stimuli rotating around a fixation cross, while the participants were instructed to watch the fixation cross and to press a button whenever the cross turned into a circle. This allowed to determine brain areas involved in processing objects and their colors. Vandenbroucke et al. (2016) found that object knowledge shifted the neural representation of colors towards the expected color category and concluded that "subjective experience at least partly overrides the representation of physical stimulus properties at a relatively early stage of the visual processing hierarchy" (Vandenbroucke et al., 2016: 1406). According to their findings, dorsolateral prefrontal cortex might play a role in the involvement of object knowledge in color perception along with some of the visual areas, like V3 and V4, involved in the early stages of the visual processing. The fMRI experiment was followed by a behavioral study, in which the same line drawings in various ambiguous hues were presented one at a time with color noise masks shown between the experimental trials, and the participants had to classify the ambiguous color they were presented with as either red or green. They found that typically red objects filled with an ambiguous color were classified as being red more often than typically green objects were. According to the authors, behavioral findings correlated with the neuroimaging results obtained in the previous experiment. This essentially means that the physiological data also reflects behavioral findings concerning the color correction bias, such as the one discovered in one of the experiments described in the present thesis.

The fact that this color correction bias was only significant for food items and not for animals is likely to be linked to the previously described phenomenon of attribute prominence. For food items, color is the primary attribute, whereas for animals, color attributes are secondary. Thus, apparently, such a color correction bias only arises in trials with subtle color changes when the color attribute has a high level of prominence within a conceptual structure of a given semantic category.

4.4. Future Research Possibilities

One of the findings of the property verification tasks both with verbal stimuli (Experiment 1) and to some extent with pictorial stimuli (Experiment 4) was the fact that incongruent part attributes were rejected earlier when they represented the primary attribute of the category than when they represented the secondary attribute of the category. A possible reason for that could be that participants simply paid attention to part attributes earlier, if they represented a more prominent attribute for the semantic category in question. Whether this is actually the case, could be investigated with the help of eye-tracking, which has been successfully used for research questions concerning concepts and their attributes (e.g., Huettig & Altmann, 2011). To find out whether more prominent part attributes are examined first and/or for longer periods of time, one might mark the corresponding attributes as interest areas and investigate at what point in time participants would start to fixate on those interest areas and how long the dwelling time would be, and then compare the results between the semantic categories, for which part attributes have different prominence levels.

Further research in this area could be facilitated, for example, by creating a database of images that would be similar to SCEGRAM (Öhlschläger & Võ, 2017) which comprises images without any violations, images with semantic violations (e.g., a scene with an object that does not belong to the environment), and images with syntactic violations (e.g., a scene with an appropriate object placed in an unlikely or physically impossible location). In the new database, various image parameters could be controlled for and attributes of concepts from different semantic categories could be manipulated. Such a database could help to conduct a more consistent and detailed research on concepts, the role of certain attributes within conceptual structures and semantic violation detection using images of various objects, while keeping technical parameters of the images under control. A reliable and comprehensive feature database compiled in the target language to avoid confusion and loss of stimuli due to translation-related problems could be created and taken as the basis for such a project.

CONCLUSION

The present dissertation investigated the influence of attribute prominence on the speed and accuracy of semantic access in a series of four behavioral experiments focusing on the following attributes and semantic categories: parts and colors of animals, colors and parts of food (fruit and vegetables), affordance and parts of tools. The main research question was whether primary (more prominent) attributes of the aforementioned semantic categories were accessed before their secondary attributes became available and whether attribute prominence also had an influence on the error rates. It was hypothesized that, for each semantic category, primary attributes would be accessed more quickly and more accurately than secondary attributes, due to stronger connections between primary attribute values and their respective concept labels. It was also established in the course of stimuli preparation that prominence of certain attributes within conceptual structures of some members of those semantic categories deviated from attribute prominence for the semantic category as a whole. Hence, an additional research question was whether attribute prominence for the semantic category or attribute prominence for individual items had a larger influence on the order and accuracy of retrieval of the respective conceptual attributes.

The hypothesis concerning attribute prominence for the semantic category was confirmed: primary attributes were accessed more quickly and accurately than secondary attributes. Moreover, attribute prominence for the semantic category turned out to be a better predictor of reaction times and error rates than attribute prominence for individual items. It means that, even though a secondary attribute of a particular semantic category might be more distinctive for an individual item than a primary attribute of the category, the primary attribute of the semantic category was accessed first, i.e., attribute prominence for the semantic category as a whole still had a priority over attribute prominence for individual items, suggesting that information about attribute prominence is stored at the level of the semantic category.

However, attribute prominence is not the only factor influencing the speed and accuracy of semantic access, and the complexity of attributes can also affect reaction times and error rates. For example, part attributes appeared to be more complex in terms of processing than color attributes, which also affected the relative speed and accuracy of their retrieval. Thus, while investigating color and part attributes using pictorial stimuli, the resulting color advantage had to be taken into consideration.

Another phenomenon that was discovered in the experiments described in the present thesis was the color correction bias. When comparing original images of objects to the images with subtle color changes in a two-alternative forced choice task, participants made more errors in trials with altered target pictures than in trials with original (non-modified) target pictures. In other words, they often failed to recognize subtle color changes in the corresponding images, categorizing the slightly altered color hue as being the same as the original color of the object, in spite of the simultaneous presence of all stimuli (a target picture and both alternatives) on the screen. This finding is in line with other studies that found evidence for the influence of object knowledge on color perception and color categorization (e.g., Mitterer & de Ruiter, 2008; Vandenbroucke et al., 2016).

All in all, the findings described in the current thesis indicate the importance of attribute prominence for the semantic category in determining the speed and accuracy of semantic retrieval during conceptual processing and confirm the unique status of color in comparison with other conceptual attributes.

REFERENCE LIST

1. Abad, M. J. F., Noguera, C., Ortells, J. J. (2003) Influence of Prime-Target Relationship on Semantic Priming Effects from Words in a Lexical Decision Task. Acta Psychologica, 113, pp. 283-295.

2. Abdel Rahman, R., Sommer, W. (2003) Does Phonological Encoding in Speech Production Always Follow the Retrieval of Semantic Knowledge? Electrophysiological Evidence for Parallel Processing. Cognitive Brain Research, 16, pp. 372-382.

3. Abdel Rahman, R., van Turennout, M., Levelt, W. J. M. (2003) Phonological Encoding Is Not Contingent on Semantic Feature Retrieval: An Electrophysiological Study on Object Naming. Journal of Experimental Psychology: Learning, Memory, and Cognition, 25 (5), pp. 850-860.

4. Adlington, R. L., Laws, K. R., Gale, T. M. (2009) The Hatfield Image Test (HIT): A New Picture Test and Norms for Experimental and Clinical Use. Journal of Clinical and Experimental Neuropsychology, 31 (6), pp. 731-753.

5. Anderson, J. R. (1983) A Spreading Activation Theory of Memory. Journal of Verbal Learning and Verbal Behavior, 22, pp. 261-295.

6. Ashcraft, M. H. (1976) Priming and Property Dominance Effects in Semantic Memory. Memory & Cognition, 4 (5), pp. 490-500.

7. Balota, D. A., Chumbley, J. I. (1984) Are Lexical Decisions a Good Measure of Lexical Access? The Role of Word Frequency in the Neglected Decision Stage. Journal of Experimental Psychology: Human Perception and Performance, 10 (3), pp. 340-357.

8. Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H. (2004) Visual Word Recognition of Single-Syllable Words. Journal of Experimental Psychology: General, 133 (2), pp. 283-316.

9. Balota, D. A., Yap, M. J., Cortese, M. J., Watson, J. M. (2008) Beyond Mean Response Latency: Response Time Distributional Analyses of Semantic Priming. Journal of Memory and Language, 59, pp. 495-523.

10. Barsalou, L. W. (1982) Context-Independent and Context-Dependent Information in Concepts. Memory & Cognition, 10 (1), pp. 82-93.

11. Barsalou, L. W. (1992) Frames, Concepts, and Conceptual Fields. In: Lehrer, A., Kittay, E. F. (Eds.) Frames, Fields, and Contrasts: New Essays in Semantic

and Lexical Organization, pp. 21-74. Lawrence Erlbaum Associates, Publishers. Hillside, New Jersey.

12. Bates, D., Mächler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67 (1), pp. 1-48, DOI: 10.18637/jss.v067.i01.

13. Benikos, N., Johnstone, S. J., Roodenrys, S. J. (2013) Varying Task Difficulty in the Go/Nogo Task: The Effects of Inhibitory Control, Arousal, and Perceived Effort on ERP Components. International Journal of Psychophysiology, 87, pp. 262-272.

14. Bevans, R. (2022) Akaike Information Criterion | When & How to Use It (Example). Scribbr. URL: https://www.scribbr.com/statistics/akaike-information-criterion (Accessed on 16.11.2022).

15. Bonin, P., Chalard, M., Méot, A., Fayol, M. (2001) Age of Acquisition and Word Frequency in the Lexical Decision Task: Further Evidence from the French Language. Current Psychology of Cognition, 20 (6), pp. 401-443.

16. Bouaffre, S., Frédérique, F.-A. (2007) Hemispheric Differences in the Time-Course of Semantic Priming Processes: Evidence from Event-Related Potentials (ERPs). Brain and Cognition, 63, pp. 123-135.

17. Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., Lepage, M. (2010) The Bank of Standardized Stimuli (BOSS), a New Set of 480 Normative Photos of Objects to Be Used as Visual Stimuli in Cognitive Research. PLoS ONE, 5 (5): e10773, DOI: 10.1371/journal.pone.0010773.

18. Brodeur, M. B., Guérard, K., Bouras, M. (2014) Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos. PLoS ONE, 9(9): e106953, DOI: 10.1371/journal.pone.0106953.

19. Brown, C., Hagoort, P. (1993) The Processing Nature of the N400: Evidence from Masked Priming. Journal of Cognitive Neuroscience, 5:1, pp. 34-44.

20. Brown, C. M., Hagoort, P., Chwilla, D. J. (2000) An Event-Related Brain Potential Analysis of Visual Word Priming Effects. Brain and Language, 72, pp. 158-190.

21. Brown, M., Besner, D. (2002) Semantic Priming: On the Role of Awareness in Visual Word Recognition in the Absence of an Expectancy. Consciousness and Cognition, 11, pp. 402-422.

22. Buchanan, E. M., Holmes, J. L., Teasley, M. L., Hutchison, K. A. (2013) English Semantic Word-Pair Norms and a Searchable Web Portal for Experimental Stimulus Creation. Behavior Research, 45, pp. 746-757.

23. Buchanan, E. M., Valentine, K. D., Maxwell, N. P. (2019) English Semantic Feature Production Norms: An Extended Database of 4436 Concepts. Behavior Research Methods, 51, pp. 1849-1863.

24. Carbon, C.-C., Faerber, S. J., Augustin, M. D., Mitterer, B., Hutzler, F. (2018) First Gender, Then Attractiveness: Indications of Gender-Specific Attractiveness Processing via ERP Onsets. Neuroscience Letters, 686, pp. 186-192.

25. Chancel, M., Ehrsson, H. H. (2020) Which Hand Is Mine? Discriminating Body Ownership Perception in a Two-Alternative Forced-Choice Task. Attention, Perception & Psychophysics, 82, pp. 4058-4083.

26. Cheesman, J., Merikle, P. M. (1986) Distinguishing Conscious from Unconscious Perceptual Processes. Canadian Journal of Psychology 40 (4), pp. 343-367.

27. Chiu, Y. S. (2012) The Time Course of Semantic and Phonological Encoding in Taiwan Sign Language: An ERP Study. International Journal of Psychophysiology, 85 (3), p. 364.

28. Chomsky, N., Halle, M. (1968) The Sound Pattern of English. New York, Harper & Row.

29. Chumbley, J. I., Balota, D. A. (1984) A Word's Meaning Affects the Decision in Lexical Decision. Memory & Cognition, 12 (6), pp. 590-606.

30. Coleman, L., Kay, P. (1981) Prototype Semantics: The English Word *Lie*. Language, 57 (1), pp. 26-44.

31. Collins, A. M., Loftus, E. F. (1975) A Spreading-Activation Theory of Semantic Processing. Psychological Review, 82 (6), pp. 407-428.

32. Collins, J., Pecher, D., Zeelenberg, R., Coulson, S. (2011) Modality Switching in a Property Verification Task: an ERP Study of What Happens When Candles Flicker after High Heels Click. Frontiers in Psychology, 2. URL: https://www.frontiersin.org/articles/10.3389/fpsyg.2011.00010 (Accessed on 19.01.2023).

33. Collins, A. M., Quillian, M. R. (1969) Retrieval Time from Semantic Memory. Journal of Verbal Learning and Verbal Behavior, 8, pp. 240-247.

34. Coltheart, M., Davelaar, E., Jonasson, J. T., Besner, D. (1977) Access to the Internal Lexicon. In: Dornič, S. (Ed.) Attention and Performance VI, pp. 535-555. Hillsdale, NJ: Erlbaum.

35. Cortese, M. J., Khanna, M. M. (2007) Age of Acquisition Predicts Naming and Lexical-Decision Performance Above and Beyond 22 Other Predictor Variables: An Analysis of 2,342 Words. The Quarterly Journal of Experimental Psychology, 60 (8), pp. 1072-1082.

36. De Groot, A. M. B. (1983) The Range of Automatic Spreading Activation in Word Priming. Journal of Verbal Learning and Verbal Behavior, 22, pp. 417-436.

37. Den Heyer, K., Briand, K., Dannenbring, G. L. (1983) Strategic Factors in a Lexical Decision Task: Evidence for Automatic and Attention-Driven Processes. Memory & Cognition, 11 (4), pp. 374-381.

38. Devereux, B. J., Tyler, L. K., Geertzen, J., Randall, B. (2014) The Centre for Speech, Language and the Brain (CSLB) Concept Property Norms. Behavior Research Methods, 46 (4), pp. 1119-1127.

39. Dumay, N., Gaskell, M. G. (2007) Sleep-Associated Changes in the Mental Representation of Spoken Words. Psychological Science, 18 (1), pp. 35-39.

40. Enriquez-Geppert, E., Konrad, C., Pantev, C., Huster, R. J. (2010) Conflict and Inhibition Differentially Affect the N200/P300 Complex in a Combined Go/Nogo and Stop Signal Task. Neuroimage, 51, pp. 877-887.

41. Falkenstein, M., Hoormann, J., Hohnsbein, J. (1999) ERP Components in Go/Nogo Tasks and Their Relation to Inhibition. Acta Psychologica, 101, pp. 267-291.

42. Farah, M. J., McClelland, J. L, (1991) A Computational Model of Semantic Memory Impairment: Modality Specificity and Emergent Category Specificity. Journal of Experimental Psychology: General, 120 (4), pp. 339-357.

43. Fillmore, C. J. (1982) Frame Semantics. Linguistics in the Morning Calm. Selected Papers from SICOL-1981, pp. 111-137. Hanshin Publishing Company, Seoul, Korea.

44. Fischler, I. (1977) Semantic Facilitation without Association in a Lexical Decision Task. Memory & Cognition, 5 (3), pp. 335-339.

45. Forster, K. I. (1994) Computational Modelling and Elementary Process Analysis in Visual Word Recognition. Journal of Experimental Psychology: Human Perception and Performance, 20 (6), pp. 1292-1310.

46. Foster, E. D., Deardorff, A. (2017) Open Science Framework (OSF). Journal of the Medical Library Association: JMLA, 105(2), pp. 203–206. DOI: https://doi.org/10.5195/jmla.2017.88 (Accessed on 13.12.2023).

47. Freepng.ru. URL: https://www.freepng.ru (Accessed on 15.01.2021).

48. Gajewski, P. D., Falkenstein, M. (2013) Effects of Task Complexity on ERP Components in Go/Nogo Tasks. International Journal of Psychophysiology, 87, pp. 273-278.

49. Gamerschlag, T., Gerland, D., Osswald, R. (Eds.) (2014) Frames and Concept Types: Applications in Language and Philosophy. Springer International Publishing AG.

50. Garman, M. (1991) Psycholinguistics. (Repr.) Cambridge [u.a.]: Cambridge Univ. Press.

51. Gere, A., Héberger, K., Kovács, S. (2021) How to Predict Choice Using Eye-Movements Data? Food Research International, 143. DOI: https://doi.org/10.1016/j.foodres.2021.110309 (Accessed on 13.11.2022).

52. Gerhand, S., Barry, C. (1999) Age of Acquisition, Word Frequency, and the Role of Phonology in the Lexical Decision Task. Memory & Cognition, 27 (4), pp. 592-602.

53. Glaser, W. R. (1992) Picture Naming. Cognition, 42, pp. 61-105.

54. Grainger, J., Jacobs, A. M. (1996) Orthographic Processing in Visual Word Recognition. Psychological Review, 103 (3), pp. 518-565.

55. Guo, T., Peng, D. (2007) Speaking Words in the Second Language: From Semantics to Phonology in 179 ms. Neuroscience Research, 57, pp. 387-392.

56. Gussenhoven, C., Jacobs, H. (1998) Understanding Phonology. Arnold, London, New York, Sydney, Auckland.

57. Harley, T. A. (2014) The Psychology of Language from Data to Theory. Fourth edition. Psychology Press, London and New York.

58. Hauk, O., Pulvermüller, F. (2004) Effects of Word Length and Frequency on the Human Event-Related Potential. Clinical Neurophysiology, 115, pp. 1090-1103.

59. Hayward, W. G. (1998) Effects of Outline Shape in Object Recognition. Journal of Experimental Psychology: Human Perception and Performance, 24 (2), pp. 427-440.

60. Hill, H., Strube, M., Roesch-Ely, D., Weisbrod, M. (2002) Automatic vs. Controlled Processes in Semantic Priming – Differentiation by Event-Related Potentials. International Journal of Psychophysiology, 44, pp. 197-218.

61. Hill, H., Ott, F., Weisbrod, M. (2005) SOA-Dependent N400 and P300 Semantic Priming Effects Using Pseudoword Primes and a Delayed Lexical Decision. International Journal of Psychophysiology, 56, pp. 209-221.

62. Hinton, J. (2005) Lilac Chaser. URL: https://michaelbach.de/ot/col-lilacChaser (Accessed on 23.02.2023).

63. Hirshman, E., Durante, R. (1992) Prime Identification and Semantic Priming. Journal of Experimental Psychology: Learning, Memory and Cognition, 18 (2), pp. 255-265.

64. Huettig, F., Altmann, G. T. M. (2011) Looking at Anything That Is Green When Hearing “Frog”: How Object Surface Color and Stored Object Color Knowledge Influence Language-Mediated Overt Attention. The Quarterly Journal of Experimental Psychology, 64 (1), pp. 122-145.

65. Hurlbert, A. (2007) Color Constancy. Current Biology, 17 (21): R906-R907. URL: https://www.cell.com/current-biology/pdf/S0960-9822(07)01839-8.pdf (Accessed on 13.11.2023).

66. Hurlbert, A. (2021) Colour and Vision. In: Fabian, A., Gibson, J., Sheppard, M., Weyand, S. (Eds.), Vision (Darwin College Lectures, pp. 57-106). Cambridge: Cambridge University Press. DOI: 10.1017/9781108946339.004.

67. Jakobson, R., Fant, C. G. M., Halle, M. (1952) Preliminaries to Speech Analysis. The Distinctive Features and Their Correlates. The MIT Press, Cambridge.

68. Jogan, M., Stocker, A. A. (2014) A New Two-Alternative Forced-Choice Method for the Unbiased Characterization of Perceptual Bias and Discriminability. Journal of Vision, 14 (3): 20, pp. 1-18. DOI: 10.1167/14.3.20.

69. Kahan, T. A., Neely, J. H., Forsythe, W. J. (1999) Dissociated Backward Priming Effects in Lexical Decision and Pronunciation Tasks. Psychonomic Bulletin & Review, 6 (1), pp. 105-110.

70. Kassambara, A. (2020) ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0, URL: https://CRAN.R-project.org/package=ggpubr (Accessed on 21.02.2023).

71. Katz, A. N. (1981) Knowing about the Sensory Properties of Objects. The Quarterly Journal of Experimental Psychology, Section A, 33 (1), pp. 39-49. URL: https://doi.org/10.1080/14640748108400767 (Accessed on 05.05.2023).

72. Kerr, M., Mingay, R., Elithorn, A. (1963) Cerebral Dominance in Reaction Time Responses. A British Journal of Psychology, 54 (4), pp. 325-336.

73. Konečná, K., Vaňkátová, K. (2017) A Database of Semantic Features for Chosen Concepts (Attested in 8- to 10-Year-Old Czech Pupils). Topics in Linguistics, 18 (1), pp. 71-97.

74. Lam, K. J. Y., Dijkstra, T., Rueschemeyer, S. A. (2015) Feature Activation During Word Recognition: Action, Visual, and Associative-Semantic Priming Effects. Frontiers in Psychology. URL: https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00659/full (Accessed on 26.09.2019).

75. Levelt, W. J. M. (1991) Speaking: From Intention to Articulation. A Bradford Book. Cambridge, Massachusetts. [u.a.]: MIT Press.

76. Löbner, S. (2013) Understanding Semantics. Understanding Language Series. Second Edition, pp. 228-236. Routledge.

77. Löbner, S. (2015) Functional Concepts and Frames. In: Gamerschlag, T., Gerland, D., Osswald, R., Petersen, W. (Eds.) Meaning, Frames, and Conceptual Representation, pp. 35-62. Düsseldorf University Press, Düsseldorf.

78. Lucas, M. (2000) Semantic Priming without Association: A Meta-Analytic Review. Psychonomic Bulletin & Review, 7 (4), pp. 618-630.

79. Mayall, K., Humphreys, G. W. (1996) Case Mixing and the Task-Sensitive Disruption of Lexical Processing. Journal of Experimental Psychology: Learning, Memory and Cognition, 22 (2), pp. 278-294.

80. Mazerolle, M. J. (2020) AICcmodavg: Model Selection and Multimodel Inference Based on (Q)AIC(c). R package version 2.3-1. URL: https://cran.r-project.org/package=AICcmodavg (Accessed on 16.11.2022).

81. McClelland, J. L., Rumelhart, D. E. (1981) An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings. Psychological Review, 88 (5), pp. 375-407.

82. McCloskey, M., Glucksberg, S. (1979) Decision Process in Verifying Category Membership Statements: Implications for Models of Semantic Memory. Cognitive Psychology, 11, pp. 1-37.

83. McKoon, G., Ratcliff, R. (1989) Assessing the Occurrence of Elaborative Inference with Recognition: Compatibility Checking vs Compound Cue Theory. Journal of Memory and Language, 28, pp. 547-563.

84. McNamara, T. P. (2005) Semantic Priming: Perspectives from Memory and Word Recognition. Essays in Cognitive Psychology. Psychology Press. New York and Hove.

85. McRae, K., Cree, G. S., Seidenberg, M. S., McNorgan, C. (2005) Semantic Feature Production Norms for a Large Set of Living and Nonliving Things. Behavior Research Methods, Instruments & Computers, 37 (4), pp. 547-559.

86. Meyer, D. E., Schvaneveldt, R. W. (1971) Facilitation in Recognizing Pairs of Words: Evidence of a Dependence between Retrieval Operations. Journal of Experimental Psychology, 90 (2), pp. 227-234.

87. Mitterer, H., de Ruiter, J. P. (2008) Recalibrating Color Categories Using World Knowledge. Psychological Science, 19 (7), pp. 629-634.

88. Mnogogreshnova, S. (2023, December 13). The Role of Attribute Prominence in Conceptual Processing. Supplementary Materials. Open Science Framework (OSF). URL: https://osf.io/3fbn2/?view_only=ee43eb8a793e408ab8d6b504bc4daaff (Accessed on 14.12.2023).

89. Montefinese, M., Ambrosini, E., Fairfield, B., Mammarella, N. (2013) Semantic Memory: A Feature-Based Analysis and New Norms for Italian. Behavior Research Methods, 45 (2), pp. 440-461.

90. Morton, J. (1969) Interaction of Information in Word Recognition. Psychological Review, 76 (2), pp. 165-178.

91. Moss, H. E., Tyler, L. K., Taylor, K. I. (2007) Conceptual Structure. In: Gaskell, M. G. (Ed.), Altmann, G., Bloom, P., Caramazza, A., Levelt, P. (Consulting Eds.) The Oxford Handbook of Psycholinguistics, pp. 217-234. Oxford University Press.

92. Murdock, B. B. (1982) A Theory of Storage and Retrieval of Item and Associative Information. Psychological Review, 89 (6), pp. 609-626.

93. Murphy, G. L. (2002) The Big Book of Concepts. Cambridge, Massachusetts. MIT Press.

94. Naor-Raz, G., Tarr, M. J. (2003) Is Color an Intrinsic Property of Object Representation? Perception, 32, pp. 667-680.

95. Neely, J. H., VerWys, C. A., Kahan, T. A. (1998) Reading “Glasses” Will Prime “Vision,” but Reading a Pair of “Glasses” Will Not. Memory & Cognition, 26 (1), pp. 34-39.

96. Neumann, Y., Obler, L. K., Shafer, V., Gomes, H. (2008) Segmental and Syllabic Processing in Healthy Younger and Older Adults: An Electrophysiological Study. CUNY Conference on the Syllable, January 2008. URL: https://www.researchgate.net/publication/237768249_Title_Segmental_and_Syllabic_Processing_in_Healthy_Younger_and_Older_Adults_An_Electrophysiological_Study (Accessed on 09.02.2019).

97. New, B., Ferrand, L., Pallier, C., Brysbaert, M. (2006) Re-Examining Word Length Effects in Visual Word Recognition: New Evidence from the English Lexicon Project. Psychonomic Bulletin & Review, 13, pp. 45-52.

98. Noguera, C., Ortells, J. J., Abad, M. J. F., Carmona, E., Daza, M. T. (2007) Semantic Priming Effects from Single Words in a Lexical Decision Task. Acta Psychologica, 125, pp. 175-202.

99. Ortells, J. J., Abad, M. J. F., Noguera, C., Lupiáñez, J. (2001) Influence of Prime-Probe Stimulus Onset Asynchrony and Prime Precuing Manipulations on Semantic Priming Effects with Words in a Lexical Decision Task. Journal of Experimental Psychology: Human Perception and Performance, 27 (1), pp. 75-91.

100. Ortells, J. J., Kiefer, M., Castillo, A., Megías, M., Morillas, A. (2016) The Semantic Origin of Unconscious Priming: Behavioral and Event-Related Potential Evidence during Category Congruency Priming from Strongly and Weakly Related Masked Words. Cognition, 146, pp. 143-157.

101. Öhlschläger, S., Võ, M. LH. (2017) SCEGRAM: An Image Database for Semantic and Syntactic Inconsistencies in Scenes. Behavior Research Methods, 49, pp. 1780–1791. URL: https://doi.org/10.3758/s13428-016-0820-3 (Accessed on 02.03.2023).

102. Pecher, D., Zeelenberg, R., Barsalou, L. W. (2004) Sensorimotor Simulations Underlie Conceptual Representations: Modality-Specific Effects of Prior Activation. Psychonomic Bulletin & Review, 11 (1), pp. 164-167.

103. Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. K. (2019) PsychoPy2: Experiments in Behavior Made Easy. Behavior Research Methods, 51. pp. 195-203.

104. Perea, M., Gotor, A. (1997) Associative and Semantic Priming Effects Occur at Very Short Stimulus-Onset Asynchronies in Lexical Decision and Naming. Cognition, 62, pp. 223-240.

105. Perea, M., Rosa, E. (2000) The Effects of Orthographic Neighborhood in Reading and Laboratory Word Identification Tasks: A Review. Psicológica, 21, pp. 327-340.

106. Perea, M., Rosa, E. (2002) The Effects of Associative and Semantic Priming in the Lexical Decision Task. Psychological Research, 66, pp. 180-194.

107. Perfetti, C. (2007) Reading Ability: Lexical Quality to Comprehension. Scientific Studies of Reading, 11 (4), pp. 357-383.

108. Petersen, W. (2007) Representation of Concepts as Frames. The Baltic International Yearbook of Cognition, Logic and Communication, 2, pp. 151-170.

109. Price, C. J., Humphreys, G. W. (1989) The Effects of Surface Detail on Object Categorization and Naming. The Quarterly Journal of Experimental Psychology, 41 (4), pp. 797-828.

110. R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org (Accessed on 04.07.2022).

111. Ratcliff, R., McKoon, G. (1988) A Retrieval Theory of Priming in Memory. Psychological Review, 95 (3), pp. 385-408.

112. Redmann, A., FitzPatrick, I., Hellwig, F., Indefrey, P. (2014) The Use of Conceptual Components in Language Production: an ERP Study. Frontiers in Psychology, 5, pp. 1-13.

113. Revelle, W. (2022) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA. Version = 2.2.5. URL: https://CRAN.R-project.org/package=psych (Accessed on 28.11.2022).

114. Rodriguez-Fornells, A., Schmitt, B. M., Kutas, M., Münte, T. F. (2002) Electrophysiological Estimates of the Time Course of Semantic and Phonological Encoding during Listening and Naming. Neuropsychologia, 40, pp. 778-787.

115. Rosch, E., Mervis, C. B. (1975) Family Resemblances: Studies in the Internal Structure of Categories. Cognitive Psychology, 7, pp. 573-605.

116. Rothe, J., Schulte-Körne, G., Ise, E. (2014) Does Sensitivity to Orthographic Regularities Influence Reading and Spelling Acquisition? A 1-Year Prospective Study. Reading and Writing, 27 (7), pp. 1141-1161.

117. Rumelhart, D. E., Hinton, G. E., McClelland, J. L. (1986) A General Framework for Parallel Distributed Processing. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, 1, pp. 45-76.

118. Ruts, W., De Deyne, S., Ameel, E., Vanpaemel, W., Verbeemen, T., Storms, G. (2004) Dutch Norm Data for 13 Semantic Categories and 338 Exemplars. Behavior Research Methods, Instruments & Computers, 36 (3), pp. 506-515.

119. Sánchez-Casas, R., Ferré, P., Demestre, J., García-Chico, T., García-Albea, J. E. (2012) Masked and Unmasked Priming Effects as a Function of Semantic Relatedness and Associative Strength. The Spanish Journal of Psychology, 15 (3), pp. 891-900.

120. Schmitt, B. M., Münte, T. F., Kutas, M. (2000) Electrophysiological Estimates of the Time Course of Semantic and Phonological Encoding during Implicit Picture Naming. Psychophysiology, 37, pp. 473-484.

121. Schmitt, B. M., Schiltz, K., Zaake, W., Kutas, M., Münte, T. F. (2001) An Electrophysiological Analysis of the Time Course of Conceptual and Syntactic Encoding during Tacit Picture Naming. Journal of Cognitive Neuroscience, 13 (4), pp. 510-522.

122. Seidenberg, M. S. (1985) The Time Course of Information Activation and Utilization in Visual Word Recognition. In: Besner, D., Waller, T. G., MacKinnon, E. M. (Eds.) Reading Research: Advances in Theory and Practice, pp. 199-252. New York: Academic Press.

123. Smith, E. E., Shoben, E. J., Rips, L. J. (1974) Structure and Process in Semantic Memory: A Featural Model for Semantic Decisions. Psychological Review, 81 (3), pp. 214-241.

124. Snodgrass, J. G., Vanderwart, M. (1980) A Standardized Set of 260 Pictures: Norms for Name Agreement, Image Agreement, Familiarity, and Visual Complexity. Journal of Experimental Psychology: Human Learning and Memory, 6 (2), pp. 174-215.

125. Solomon, K. O., Barsalou, L. W. (2004) Perceptual Simulation in Property Verification. Memory & Cognition, 32 (2), pp. 244-259.

126. Tanaka, J. W., Presnell, L. M. (1999) Color Diagnosticity in Object Recognition. Perception & Psychophysics, 61 (6), pp. 1140-1153.

127. Taylor, I. (1990) Psycholinguistics: Learning and Using Language. Englewood Cliffs, NJ: Prentice Hall.

128. Taylor, J. R. (2003) Linguistic Categorization. Third edition. Oxford University Press, New York.

129. Therriault, D. J., Yaxley, R. H., Zwaan, R. A. (2009) The Role in Color Diagnosticity in Object Recognition and Representation. Cognitive Processing. Springer-Verlag. URL: https://link.springer.com/article/10.1007/s10339-009-0260-4 (Accessed on 11.01.2018).

130. Thompson-Schill, S. L., Kurtz, K. J., Gabrieli, J. D. E. (1998) Effects of Semantic and Associative Relatedness on Automatic Priming. Journal of Memory and Language, 38, pp. 440-458.

131. Tyler, L. K., Moss, H. E. (1997) Functional Properties of Concepts: Studies of Normal and Brain-Damaged Patients. Cognitive Neuropsychology, 14 (4), pp. 511-545.

132. Valdés, B., Catena, A., Marí-Beffa, P. (2005) Automatic and Controlled Semantic Processing: A Masked Prime-Task Effect. Consciousness and Cognition, 14, pp. 278-295.

133. Van Casteren, M., Davis, M. H. (2006) Mix, a Program for Pseudorandomization. Behavior Research Methods, 38 (4), pp. 584-589.

134. Van de Meerendonk, N., Kolk, H. H. J., Vissers, C. Th. W. M., Chwilla, D. J. (2008) Monitoring in Language Perception: Mild and Strong Conflicts Elicit Different ERP Patterns. Journal of Cognitive Neuroscience, 22 (1), pp. 67-82.

135. Van Turennout, M., Hagoort, P., Brown, C. M. (1997) Electrophysiological Evidence on the Time Course of Semantic and Phonological Processes in Speech Production. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23 (4), pp. 787-806.

136. Vandenbroucke, A. R. E., Fahrenfort, J. J., Meuwese, J. D. I., Scholte, H. S., Lamme, V. A. F. (2016) Prior Knowledge about Objects Determines Neural Color Representation in Human Visual Cortex. Cerebral Cortex, 26, pp. 1401-1408.

137. Vinson, D. P., Vigliocco, G. (2008) Semantic Feature Production Norms for a Large Set of Objects and Events. Behavior Research Methods 40 (1), pp. 183-190.

138. Vissers, C. Th. W. M., Chwilla, D. J., Kolk, H. H. J. (2006) Monitoring in Language Perception: The Effects of Misspellings of Words in Highly Constrained Sentences. Brain Research, 1106, pp. 150-163.

139. Vivas, J., Vivas, L., Comesaña, García Coni, A., Vorano, A. (2017) Spanish Semantic Feature Production Norms for 400 Concrete Concepts. Behavior Research Methods, 49 (3), pp. 1095-1106.

140. Võ, M. L.-H., Wolfe, J. M. (2013) Differential Electrophysiological Signatures of Semantic and Syntactic Scene Processing. Psychological Science, 24 (9), pp. 1816–1823. DOI: https://doi.org/10.1177/0956797613476955.

141. Whaley, C. P. (1978) Word-Nonword Classification Time. Journal of Verbal Learning and Verbal Behavior, 17, pp. 143-154.

142. Wickham, H. (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

143. Wickham, H., François, R., Henry, L., Müller, K. (2022). dplyr: A Grammar of Data Manipulation. URL: https://dplyr.tidyverse.org, https://github.com/tidyverse/dplyr (Accessed on 17.11.2022).

144. Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., Boroditsky, L. (2007) Russian Blues Reveal Effects of Language on Color Discrimination. Proceedings of the National Academy of Sciences, 104 (19), pp. 7780-7785.

145. Witzel, C., Gegenfurtner, K. R. (2018) Color Perception: Objects, Constancy, and Categories. Annual Review of Vision Science, 4, pp. 475-499. URL: https://doi.org/10.1146/annurev-vision-091517-034231 (Accessed on 21.02.2023).

146. Witzel, C., Gegenfurtner, K. (2020) Memory Color. In: Shamey, R. (Ed.) Encyclopedia of Color Science and Technology. Springer. Berlin, Heidelberg. URL: https://doi.org/10.1007/978-3-642-27851-8_58-9 (Accessed on 24.02.2023).

147. Wurm, L. H., Legge, G. E., Isenberg, L. M., Luebker, A. (1993) Color Improves Object Recognition in Normal and Low Vision. Journal of Experimental Psychology: Human Perception and Performance, 19 (4), pp. 899-911.

Appendix: Supplementary Materials

Supplementary materials including full sets of verbal and pictorial stimuli for each of the experiments described in the present thesis, as well as scripts containing all commands used for the statistical analysis of the data and creation of graphs in R Studio are available via the Open Science Framework (OSF):

Mnogogreshnova, S. (2023, December 13). The Role of Attribute Prominence in Conceptual Processing. Supplementary Materials. Open Science Framework (OSF). URL: https://osf.io/3fbn2/?view_only=ee43eb8a793e408ab8d6b504bc4daaff

For the sake of convenience, a QR code was created to simplify access to the supplementary materials for readers of the printed version of the present dissertation:

Eidesstattliche Versicherung

Ich versichere an Eides statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Ordnung über die Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Ort, Datum: Düsseldorf, 19.12.2024   Unterschrift:

Svetlana Mnogogreshnova

Lebenslauf

Svetlana Mnogogreshnova

Berufserfahrung

04/2023 – 09/2024	Russischdozentin Heinrich-Heine-Universität Düsseldorf, Sprachenzentrum
10/2018 – 11/2023	Wissenschaftliche Mitarbeiterin Heinrich-Heine-Universität Düsseldorf, Institut für Linguistik
06/2017 – 12/2018	Wissenschaftliche Hilfskraft Heinrich-Heine-Universität Düsseldorf, Institut für Sprache und Information
03/2014 – 09/2015	Managerin-Übersetzerin (Sprachen: Russisch, Englisch, Deutsch) Übersetzungsbüro „Neo Translating“, Krasnojarsk
02/2014	Praktikantin Übersetzungsbüro „Neo Translating“, Krasnojarsk

Akademische Ausbildung

2018 – 2024	Promotion Fach: Allgemeine Sprachwissenschaft Heinrich-Heine-Universität Düsseldorf
2015 – 2018	Masterstudium Fach: Linguistik Heinrich-Heine-Universität Düsseldorf
2010 – 2014	Bachelorstudium Fach: Linguistik Sibirische föderale Universität Krasnojarsk
bis 2010	Internatsschule Nr.1, Krasnojarsk

Zusatzqualifikationen

10/2022 – 07/2023	Zusatzqualifikation „Deutsch als Fremdsprache“ Heinrich-Heine-Universität Düsseldorf
07/2019	Sommerkurs Finnisch A1 Åbo Akademi, Turku (Finnland)
12/2014	Seminar des Goethe-Instituts Moskau für Literaturübersetzer/innen
09/2014	TestDaF (5544)
2014	Seminar „Methoden des Englischunterrichts“, British Council
10/2013	Seminar des Goethe-Instituts Moskau für angehende Literaturübersetzer/innen
08/2013	Sommerkurs Deutsch als Fremdsprache C1 Sprachen- und Dolmetscher Institut München