The way idioms are stored in the mental lexicon

Siri Brouwer
3001210
Thesis Research Master Linguistics
Supervisor: prof. dr. Martin Everaert
Second supervisor: prof. dr. Frank Wijnen
August 2011

1. Introduction

Idioms are an interesting part of our language. Idioms are expressions, words, or phrases that have a figurative meaning which is separate from the literal meaning of the words of which they are made. An example of an idiom is ‘wait to see which way the cat jumps’. This expression not only has a literal meaning, it also means ‘to delay making a decision until one knows what is going to happen’.

Within idioms little pieces of the history of our culture are contained. Every language is full of these expressions. We are raised with them and learn how to deal with them. Even when people don’t know the origin of an idiom they can still use it in daily life. We laugh when children don’t understand what we are talking about, and with a smile on our face we watch them struggle with the figurative meaning of these sentences. It reminds us of how we once had troubles with these so called idioms.

In this thesis the focus lies on how idioms are stored in the mental lexicon. The mental lexicon is the storage of a speaker’s knowledge of the vocabulary of a language. The vocabulary consists of words and expressions. However, the lexicon of a language is not like a dictionary. Words in the lexicon are not stored in alphabetical order. Idioms are special because they are composed from combinations of words that are already stored in the lexicon. Now, the interesting question emerges: How are these idioms stored in the mental lexicon? The question becomes even more interesting if we take into account the fact that there are idioms which contain words that only occur in the idiom itself and not outside the idiom, for example ‘voor iemand in de bres springen’ (to jump into the breach for someone).

Since the lexicon itself is unobservable, its study must be based on what experimental evidence can show us. In this thesis I will describe three experiments which will hopefully shed some light on the question of how idioms are stored in the mental lexicon. The first experiment investigated if it is possible to understand an idiom if you are missing a lexical entry in your mental lexicon. The second experiment investigated if idiom words are comparable to non-words or to real words. The third experiment investigated if a semantically related prime can speed up the production of an idiom and if it can cause a lower error rate in the production of idioms. These experiments will be discussed in relation to the experiments done by Sprenger et al (2006). They investigated

lexical access during the production of idiomatic phrases. The assumption of Sprenger et al (2006) is that idioms are compositional on the level of lemmas. Sprenger et al (2006) based this assumption on the results of their experiments. In this thesis the design of these experiments will be criticized and the assumption that idioms are compositional will be discussed.

I would like to thank the following people for their useful comments and helpful support:
Martin Everaert and Frank Wijnen
Leon Geerdink and Peter Coopmans
Maartje Schulpen and Marja Oudega
Theo Veenker and Iris Mulders
Anne van Leeuwen and Marjolein van Egmond
All the participants and Primary School de Sleutel in Schalkhaar

2. Theoretical framework

2.1 Definition of an idiom

Idioms are sentences that can or cannot have a literal meaning. Idioms can be sayings, proverbs, expressions or even simple words. 'Idiom' is defined in the Oxford English Dictionary as: 'A form of expression, grammatical construction, phrase, etc., used in a distinctive way in a particular language, dialect, or language variety; spec. a group of words established by usage as having a meaning not deducible from the meanings of the individual words' (OED, 2011).

In the literature on idioms there are many definitions. (Levorato, 1993) describes the difference between literal and figurative language by given characteristics of figurative language. 'The first characteristic is a gap between the speaker's words and his or her communicative intention. An idiom departs from its original meaning, the literal one, and acquires new meaning by means of strongly held conventions.' For this thesis, the following definition by (Sprenger et al 2006) is taken to be adequate; 'Fixed expressions are phrasal units, and they exist in many varieties (e.g., phrasal verbs, restricted collocations, idiomatic expressions, and sayings and proverbs). Idiomatic expressions or idioms are a particularly interesting variant of fixed expressions, because their meaning is partly or completely non-compositional.' In this thesis, the focus lies on proverbs and sayings that form a complete sentence. A proverb is 'A short, traditional, and pithy saying; a concise sentence, typically metaphorical or alliterative in form, stating a general truth or piece of advice; an adage or maxim.' (OED, 2011). For example, 'haste makes waste' is a popular English proverb.

2.1.2 Three forms of idioms

Idioms can be divided into three categories: normally decomposable phrases, abnormally decomposable phrases and non-decomposable phrases (Gibbs & Nayak 1989).

Normally decomposable phrases are idioms where the figurative and literal meaning are very similar, and it is therefore not difficult to grasp the figurative meaning. An example of a decomposable phrase is 'button your lips', which means close your mouth, or stay silent.

Abnormally decomposable phrases are idioms that require more analysis before one understands their figurative meaning. There is a so called metaphorical link that needs to be made between the literal meaning and the figurative concept. “To carry a torch”, which mean to secretly love someone who does not love you, is an example of an abnormally decomposable phrase. One can only deduce the meaning of the idiom if one understands that the torch is a metaphor for having a warm feeling.

Non-decomposable phrases are typically idioms that are based on the historical and cultural background of a language. It is very difficult to grasp the figurative meaning if one is unfamiliar with these idioms.

2.2 Comprehension and production of idioms

How idioms are learned by children might show us something about how people are able to comprehend and produce idioms. It can give us insight in the development of lexical knowledge. The ability to understand figurative meaning by children occurs simultaneously with the development of cognitive abilities in general and with linguistic development in particular. According to Levorato (1993, p. 102) children acquiring the figurative meaning of language have difficulties understanding: ‘(a) that a distinction may exist between what is said and what is meant, (b) that the conventional meaning may differ from the literal one, and (c) that one needs to make use of all the available contextual information, the linguistic and non-linguistic context as well as the relevant world knowledge, to identify the exact meaning of an expression.’ Levorato’s findings (1993:105) imply that idioms that have a figurative meaning that is very different from the literal meaning ensures that children only process the literal meaning. For very young children the literal strategy is the only one they can apply.

Children make many mistakes in producing idioms until a relatively late age. An example from Schaerlaekens and Gillis (Gillis & Schaerlaekens, 2002, p. 219) produced by an eight year old: ‘Hij sloeg maar wat in het wilde westen’ (He only hit somewhat in the Wild West). Instead of ‘in het wilde weg’ (in the wild way). Another example produced by a nine year old child: ‘Maar ik zie nog geen haartje voor ogen’ (But I don’t even see a hair before my eyes) instead of ‘Ik zie nog geen hand voor ogen’ (I don’t even see a hand before my eyes). A last example is from an eleven year old: ‘Anders ben ik

echt een moederzielskindje hoor' (Normally I am really a mother soul child) instead of 'moederskindje' (mothers child).

Cacciari investigated how children assign meaning to idioms (Cacciari, 1993) by giving ten year old children a questionnaire, questioning them on three lists consisting of ten familiar idioms each. The three lists were differentiated by the type of idiom they tested. Cacciari used the same division as has been discussed in 2.1.2, but used a different terminology. The first list consisted of transparent idioms (normally decomposable phrases), the second of quasi-metaphorical idioms (abnormally decomposable phrases) and the third of opaque idioms (non-decomposable phrases). Every child was asked to fill in for each idiom: (a) a paraphrase of the meaning of the idiom, (b) the reasons motivating the meaning of the idiom, (c) whether a six year old is able to understand this idiom, and (d) what a six year old could do to be able to understand the idiom (Cacciari, 1993:42). The results showed that ten year olds understand the meaning of idioms quite well. "69.9 % of correct paraphrases for the quasi-metaphorical ones, 51.9 % for opaque idioms and 47.9% for transparent idioms." (Cacciari, 1993:44). The children had the following reasons for motivating the answer for the correct meaning of the idiom: usage conditions, causal explanations, explanations based on the literal state or action expressed by the idiom, explanation based on the literal and figurative outcomes of the idiomatic action, explanation based on analogies, explanations based on perceived symbols or explanations based on other idioms (Cacciari, 1993:44). So these are very adult ways in which children are able to assign meaning to idioms. The results for the question whether six year olds are able to understand the meaning of these idioms were the following: for the quasi-metaphorical ones 38.6% thought that six year olds would understand them, 43.9% for the transparent ones and 40.6% for the opaque idioms (Cacciari, 1993:45). It was really interesting to see how the ten year olds answered the question what a six year old can do in order to be able to understand the idiom. Cacciari (1993:45) found that exemplification is an option, performing the idiomatic action, performing the literal action, asking adults, by reasoning, by a correct similitude, but some children also responded that a six year old is unable to understand the idiom. Later on in this thesis I will discuss these results from my experiment in comparison with the results of Cacciari (1993).

Levorato (Levorato, 1993) describes an experiment she conducted in 1989 with first and fourth grade children about the role that context plays. Her hypothesis was that: 'Comprehension would be better when the children could make use of contextual information in order to figure out what the idiomatic expression meant.' (Levorato, 1993:107) In order to test this hypothesis, Levorato tested children's comprehension of idioms under two conditions:' (a) with idioms embedded in a linguistic context consisting of a short narrative, and (b) with idioms presented alone and out of context. The idea was to use very common idioms with both a literal and a figurative meaning. An example of an idiom presented embedded in a linguistic context taken from Levorato (1993:107): 'A little boy named Paul moved to another town. It was winter so he had to change school. His mother suggested that he should try and get to know his new schoolmates. Once at school he lent them his crayons and that helped to break the ice. What did Paul do when he broke the ice? (a) He made friends with his schoolmates. (b) He broke a piece of ice. (c) He told his mummy everything.' The results showed that answer (a) was chosen more often if the idiom was presented in a context than without a context. Answer (a) was chosen more often by older children than by younger children.

Levorato (1993:113) also conducted an experiment consisting of a comprehension and a production task, using children from the first and fourth grades of elementary school, by comparing the role of context with the role of familiarity. The comprehension task consisted of two stories for each idiom, ending with an idiomatic expression, where familiar and unfamiliar idioms were used. Levorato (1993:114) used a strategy to make a difference in choice between the literal and figurative meaning: 'In the first story, which we refer to as the idiomatic context, the literal interpretation was inappropriate because of the incongruence between the semantic field of the linguistic context and the semantic field of the idiom. In the second story, hereafter the literal context, a literal interpretation also would have been possible because the literal meaning related to objects or events in the context, although the idiomatic interpretation was still the most appropriate one.' Levorato used a multiple-choice test with three possible answers consisting of an idiomatic interpretation, a paraphrase of the literal interpretation and an associate answer for the comprehension task. The children had to listen to the story and had to answer the multiple-choice test. The results showed that younger children in both contexts chose the

idiomatic interpretation less often than the older children. The results also showed that the role of context is important for unfamiliar idioms. Levorato (1993:116) also did a production task with children from the second and fifth grades of elementary school, using the same stories but without the last part of the idiom. The children had to complete the sentence. The results showed that completing familiar idioms was easier for the children than completing unfamiliar idioms. A literal completion occurred more often with younger children and also with unfamiliar idioms.

Thus, how easy an idiom is to produce depends on its familiarity and transparency and its context. The familiarity of an idiom is defined as how frequently an idiom occurs in the language. (Nippold & Taylor, 1995) The transparency of an idiom is based on how easy it is to analyze it..

There are two points in an idiomatic sentence which are important for the interpretation. The first one can be called 'the point of idiom uniqueness', the second one can be called 'the point of idiom recognition'. This terminology was proposed by Flores d' Arcais (Flores D'Arcais, 1993). The point of idiom uniqueness is the point at which the idiom becomes uniquely identifiable. This is often the last word of the idiom. The point of idiom recognition is the point where the reader retrieves the meaning of the idiom, this point depends on the context and the familiarity of the reader with the idiom.

2.3 The mental lexicon

According to Jackendoff (1995) the most likely place where fixed expression like an idiom are stored is the mental lexicon, given the linguistic properties of fixed expressions. According to Jackendoff people would know at least as many fixed expressions as single words.

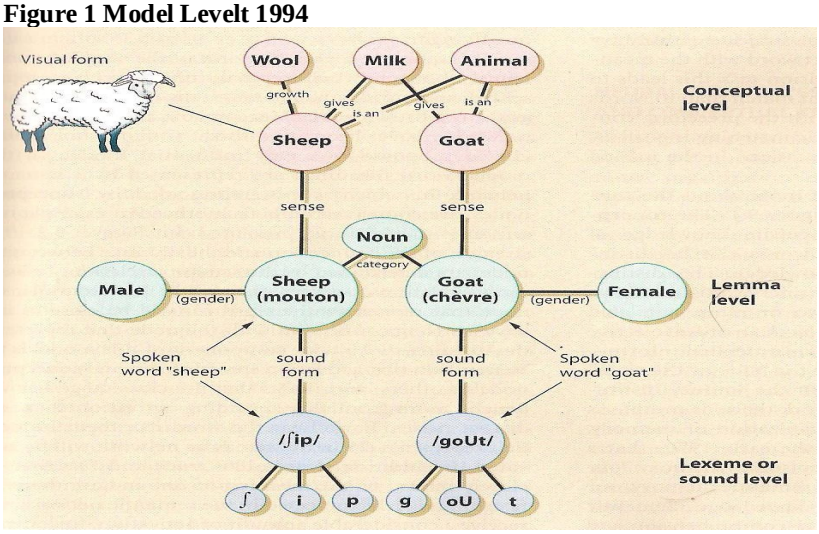
The mental lexicon is where the knowledge of the vocabulary of the speaker's language is stored. It contains all the words that the speaker knows. However, it does not contain all the forms of each word, e.g. it does not contain *run*, *runs* and *ran* separately. Only one abstract entity is stored in the mental lexicon, which is called a lexeme. A lexeme can be designated by more or less arbitrarily picking one of the forms as its standard, or canonical form. This canonical form of a lexeme is called a lemma. For instance, we could designate the lemma *run* to be the canonical form of all the forms of *to*

run, like *ran* and *is running*. We can then designate the lexeme corresponding to all these forms with the lemma *run*. In models of speech production and recognition, however, it is the abstract unit which carries the meaning of the lexeme, but not yet any specific form (it does contain information about possible forms), which is called the lemma. In the discussion below this second meaning of lemma is meant.

The mental lexicon contains information about the meaning of a word, the forms of the word and also the syntactic properties of the word. We must try to understand how people learn new words and how these words are stored in the mental lexicon. Waxman & Lidz (2006) propose the following steps of how people might learn new words, focusing mainly on infants: "infants must (1) identify the relevant conceptual units (e.g., an individual, a category of individuals, an event, and so on), (2) identify the relevant linguistic units (e.g., words), and (3) establish a mapping between them." Where a lexeme is situated in the mental lexicon, that is, how it relates to other lexemes in the lexicon, is based on its relationships of meaning.

Levelt (1994) proposes a model of how words are stored as an information specific network (see figure 1). This model presupposes three levels, the conceptual level, the lemma level and the lexeme or sound level. On the conceptual level concepts are stored, and concepts that are conceptually similar are related to each other within this conceptual network. For instance, according to figure 1, goat and sheep are related, because they are both animals and give milk. These concepts are also related to corresponding lemmas. On the lemma level the syntactic properties of the words whose meaning are the concepts become available, e.g. that 'sheep', which is the English word for sheep, is a noun. The lemmas are also stored in the form of a network. The connections of this network are based on the categories of the lemmas. For instance, all nouns are related to each other on the lemma level, as are all verbs, etc. The lemmas are also connected to the sound level. On this level the sounds that make up the words corresponding to the concepts are stored. These are also interrelated, for instance words that have the same number of syllables are related on the lexeme or sound level.

Figure 1 Model Levelt 1994



Evidence that words are stored in this way comes from semantic priming studies and lexical decision tasks. These studies show that participants respond faster and more accurate if the prime is related to the target word. There are a lot of different theories on the storage of words. The main problem with all these different theories is that it is difficult to find evidence for a proposed model of a stored word. The model by Levelt is a basic model and has found its evidence in lexical decision tasks and priming studies. For this thesis we take the model by Levelt as the basic theory for the storage of words.

2.4 Theories on the processing of idioms for comprehension

In the literature one can find three different models on processing idioms (Everaert et al 1995). The first model is called the idiom list hypothesis. The assumption here is that idioms are stored and accessed from a separate list. In this model the non-idiomatic meaning is processed first. Support for this hypothesis is given by the experimental fact that idioms take less time in processing than non-idiomatic sentences do (Everaert et al 1995: 10). The second model is called the lexical representation model. The assumption here is that idioms are stored in the lexicon as holistic entries, i.e. the parts of the idiom together form the entry of the idiom. In this model the idiomatic meaning is processed first. The third model is called the idiom decomposition model. The assumption here is that some idioms are decomposable. These decomposable idioms have a compositional

nature, i.e. that the meaning of the idiom is composed out of the meaning of its parts. In this model the non-idiomatic and idiomatic meaning are simultaneously processed.

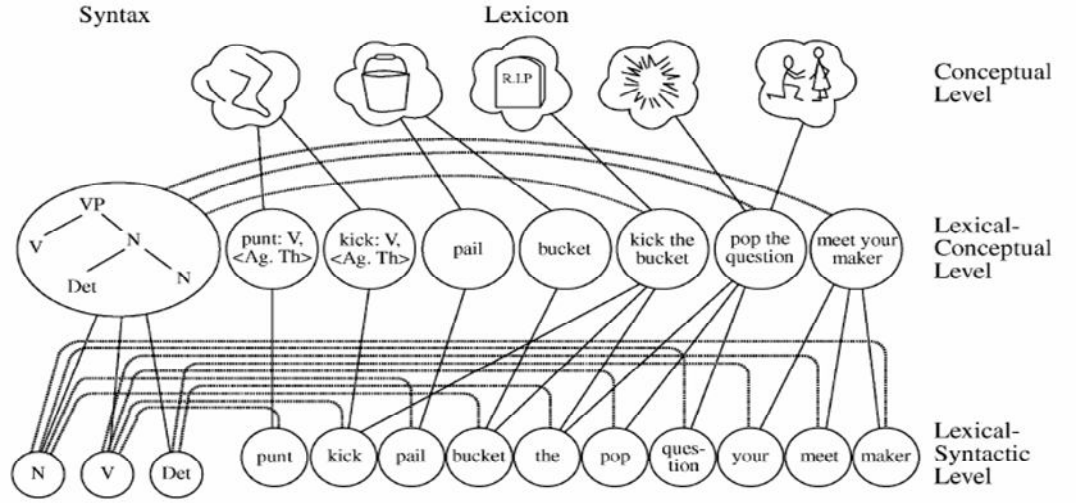
Cacciari (Cacciari 1995) argues that when we are processing idioms for comprehension, two types of meaning are available. The specific meaning that is assigned to the idiomatic expression is available, but the meanings of the linguistic constituents is also available. Cacciari conducted a study concerning the mental imagery of idioms. The experiment consisted of an image production task. The participants were given an idiom and were asked to form a detailed mental image and describe it verbally. The results showed that the number of literal pictures reported was significantly higher than the number of figurative pictures reported. The results of this study justify the decision to make use of literal pictures in the experiment that will be shown in part 5.

Gibbs (Gibbs 1995) assumes that many idioms are compositional or analyzable. The process of idiom comprehension is according to Gibbs as follows: 'the listener relies on a fast unconsciousness process, they seek to discover the independent meanings of the parts of the idiom and combine these to recognize what idioms mean as wholes.' A problem with this approach is that it cannot account for parts in the idiom that contain unknown words.

2.5 Super lemma theory

Sprenger et al (2006) investigated lexical access during the production of idiomatic phrases. Her idea is that "idioms must be retrieved from long-term memory [i.e. lexicon, SB]. Their mental representation must comprise at least the set of words, their syntactic idiosyncrasies, and their figurative interpretation." (Sprenger et al 2006: 162) Her experiments were designed in such a way that they would allow for a better understanding of how idioms are stored and produced. Sprenger et al (2006) tested the two core assumptions of the hybrid account of idiom production, proposed by Cutting and Block (1997).

Figure 2 Hybrid account of idiom production (Cutting and Block, 1997)



According to this theory, idioms are composed from single words on a lexical-syntactic level, but have their own representation on the conceptual level in the mental lexicon. When these single representation are activated, activation spreads to all their component parts.

The first experiment was a cued recall experiment. It was hypothesized that during the process of producing an idiomatic phrase the single words that make up the utterance were accessed separately. Participants were asked to produce idiomatic and literal phrases in response to a visually presented prompt word. Every phrase was linked to a prompt word that was intended to trigger the production of the phrase. For 13 of these pairs the prompt word was a common Dutch name. The same names were used for triggering the production of idiomatic phrases and literal phrase. Participants saw only one of these pairs, not both. For the remaining three items, the prompt consisted of a short phrase. During the learning phase participants were asked to memorize the idiomatic and literal phrases in such a way that they could produce the phrase fluently whenever they were presented with the prompt. After a fixation cross had appeared in the center of the screen, participants saw one of the prompt words. At the same time, a prime word was presented via headphones. A prime is a stimulus which influences the response of a later stimulus. In this case it was hypothesized that a phonologically or semantically related prime word would affect the time the participants needed for producing the prompted idiomatic or literal phrase, i.e. the response time. A phonologically related prime word is a prime which sounds similar to the noun in the prompted phrase. A

semantically related prime word is a prime which is similar in meaning to the noun in the prompted phrase. This prime was either identical to the noun in the phrase or to the phrase to be produced itself or unrelated. The participant's task was to react to the visually presented prompt word by producing the appropriate phrase as quickly as possible. An example: priming [road] in [clean the road] will result in shorter production latencies compared to an unrelated prime. The effect was stronger when an identity prime was presented with an idiomatic phrase than when an identity prime was presented with a literal phrase, for example priming [road] in [hit the road] would result in a stronger effect of reducing the response time then priming [road] in [clean the road]. This proved not be the case when using phonologically or semantically unrelated primes.

The second experiment was a completion task. The question was whether the production of idioms involves the same lemmas that are otherwise part of non-idiomatic language production and whether idioms have their own meaning and lexical concept. Completing well known idioms allows the study of idiom production without an initial learning phase, because reading the first part of the idiom provides easy access to its remaining parts. Participants were asked to produce the last word of an idiom. The first part of the sentence functioned as stimulus for producing the last word of the idiom. Acoustic primes were presented at different stimulus onset asynchronies, relative to the presentation of the visual stimulus. Within each of the four soas (stimulus onset asynchronies), each of the 16 items was presented in each of the 32 trials. Each item was presented subsequently during the following conditions: with a semantic prime, with a phonological prime, with an unrelated prime and without a prime. Each item was presented eight times under each of the four conditions. Participants made fewer errors when no prime word was presented than when an unrelated prime was presented. For example: The participants were presented with the phrase 'Jan viel door de…(Jan fell through the…)' and simultaneously presented with the phonological prime 'map' (file) with the intention that they produced 'mand' (basket). This same phrase was subsequently presented with the semantic prime 'korf' (basket).

The third experiment tested whether the literal meanings of the component words of the idiom were active during the production of an idiom. While planning to complete

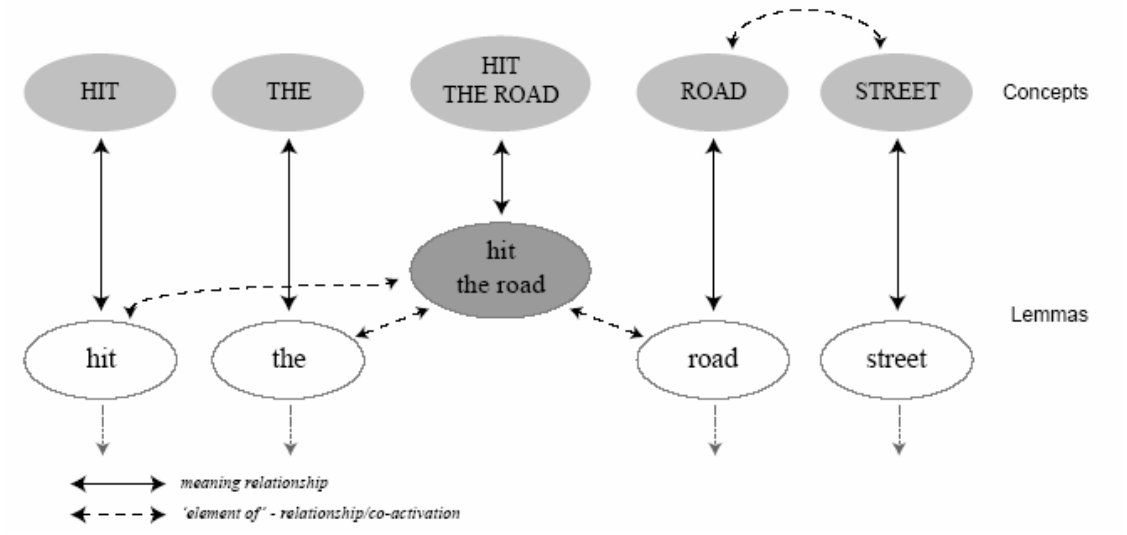
an idiomatic expression, the speaker had to switch task and read out loud a visually presented word that is semantically related to the literal meaning of the target word.

The experiments by Sprenger et al (2006) show that: (1) The production of idiomatic and literal phrases can be primed successfully by means of priming one of their content words. For instance, there is a relationship between the word [road], the lemma [road] and the lexical entry [hit the road]. This is interpreted as supporting the view that in the lexicon representations of idiomatic expressions are compositional on some level of description. (1b) The effect of priming is stronger for idiomatic phrases then for literal phrases. This is interpreted as supporting the view that the different components of an idiom are bound together by one common entry in the mental lexicon. (2) The production of nouns as parts of idiomatic expressions can be sped up by both phonologically related and semantically related acoustic distractors. This is interpreted as supporting the view that the representation of an idiom activates simple lemmas that are lexical entries on their own. (3) The results indicate that the preparation of a word as part of an idiom can affect the production latencies of words that are phonologically or semantically related to this word. The presence of a semantic effect is interpreted as evidence for the activation of literal word meanings during the production of idioms.

Sprenger et al (2006) proposes a revised model of the earlier discussed hybrid account of idiom production by (Cutting and Bock, 1997) for the lexical representation of an idiom. The revised model was based on the results of the three experiments. It was also based on a model in which the same lexical-conceptual and the lexical-syntactical processing levels are responsible for both comprehension and production (Kempen, 1997). According to this view it should be possible to read a language processing model as a production model by going from concepts to lemmas and as a comprehension model by going from lemmas to concepts. To make sure that the hybrid account of idiom production fulfills this requirement Sprenger et al introduced the notion of ‘superlemma’. Superlemmas are a separate representation of idiomatic phrases on the lexical-syntactic level. The superlemma is a representation of the syntactic properties of the idiom, and is connected to its building blocks, the simple lemmas. The purpose of the superlemma is to fine tune the lexical frames that are associated with individual lemmas. According to this view, the syntactic features of the superlemma’s elements form the building blocks of its

structure. This structure is reduced in its syntactic potential, making the idiom syntactically less flexible.

Figure 3 Model Superlemma (Sprenger, 2006)



Sprenger et al's experimental results provide an indication that words in an idiom are connected to semantically and phonologically related words. The question if activation can occur from the idiom to a related word and backward has been answered. But, the way in which this is demonstrated is disputable, because in the first experiment the use of a learning phase and the way in which the stimulus is presented is problematic. The learning phase itself has a priming effect on the results. If a subject has to memorize a sentence like 'hit the road' or 'clean the road' together with a prompt then, because of memorization, these phrases are treated as a single unit. Now, when we move from the learning phase to the experiment, this connection to the single unit is still active and if we now present the prompt, there is no need for another prime. That means that we have a double prime, which problematizes the validity of the experiment. With this double prime we cannot simulate a real language situation. If we want to investigate how an idiom is stored in the mental lexicon then we have to come up with an experiment that comes as close to a real language situation as possible. As a consequence association learning should be avoided. It is therefore important to use well-known idioms that can easily be elicited. The prompt words used by Sprenger et al are common Dutch names. These common names are stored together with the sentence in a single memory unit. Again, this is not part of a real language situation. In daily life, an idiom is based on a particular

situation and used as such. An idiom that is stored in the mental lexicon will in daily life become activated by a particular situation or by a related content word. Therefore, in the experiment it is important to use a stimulus that is close to daily life situations. For example, a picture of a well-known idiom could function as a stimulus. If it is presented together with a prime we have a stable basis.

Sprenger et al (2006) found a significant difference in response time between related and unrelated primes in the second experiment. But, there were 16 items and every item was presented eight times with a semantic prime, eight times with a phonological prime, eight times with an unrelated prime and eight times without a prime. The response was prompted by the first part of the idiom. The experimenters also tested the participants' knowledge of the idioms before the experiment. All these factors influence one another. Every time an idiom is presented again, there is an effect from the previous presentation. It is not possible to control for all these factors.

A replication of the experiments by Sprenger et al with the just mentioned adjustments will be discussed in part 6. We are justified to conclude that words that are related to the idiom become activated. But this doesn't yet show anything about how an idiom is stored. It only shows something about the connections around the idiom, but nothing about the idiom itself. The necessity for the notion of a superlemma, i.e. a representation of the syntactic properties of the idiom that is connected to its building blocks, cannot be concluded based on these three experiments.

If we assume that an idiom is stored in the form of a sum of component lemmas in the lexicon then it is necessary that a person has these lemmas in their lexicon and therefore should know the meaning of the words connected to these lemmas. But an argument against this assumption is that it is possible to use an idiom correctly and nevertheless not know the meaning of each separate word that occurs in the idiom. This can be shown in an experiment and will be shown in part 4. Another counterargument against the view that idioms are stored in the form of a sum of its component lemmas is that some patients with Alzheimer or aphasia are capable of using idioms while they are having problems with grammatical constructions. This suggests that idioms are independently stored. A study by (Papagno, 2001) showed that a decline of figurative language is not an early symptom of dementia. It can occur independent of an impairment

in propositional language. A case study by (Nenonen et al 2002) showed that an aphasic patient was able to read idiomatic noun phrases. Nenonen's result suggest that noun phrase idioms are stored like holistic long words and that verb phrase idioms must be decomposed.

Furthermore, children remember text often as a sequence of sounds. They pay less attention to the meaning of a word and more to sounds (Wray, 2002). When a child learns a rhyme or a song it doesn't take an analytic approach. If we watch, for example, the German Idols show, we can see people singing in English even though they do not understand the language and therefore do not understand the words they are singing. These two facts show that it is possible to produce language or words that have no meaning for us. These words might be stored only with the sound mapping and not with a meaning. If we produce an idiom that contains an unknown word then we can assume that this unknown word is an empty category in the lexicon. The unknown word is not linked to a meaning. It is interesting to see how people in a lexical decision task under time pressure would judge a strange word that comes from an idiom. Would they judge idiom words to be non-words as often as non-words or would they more often judge them to be an existing word like real words. This experiment will be shown in part 5. Support for this can be found in Cruse (1986), who indeed claimed that words in idioms can be semantically empty.

3. Research framework

By asking the first main question: ‘is it possible to understand an idiom if you are missing a word in your mental lexicon?’ we can investigate how an idiom is built. We can think of an idiom that contains an unfamiliar word. The suggestion in the main question is that people might not know the meaning of this word. The unfamiliar word belongs to one of the following three categories: Hapax legomenon: a word or phrase that appears only once in a manuscript, document, or particular area of literature. (dictionary.com). Archaic word: of or belonging to the distant past (= a long time ago); from an ancient period in history. (Cambridge dictionary, Press, 2008). Jargon: special words and phrases which are used by particular groups of people, especially in their work. (Cambridge dictionary, Press, 2008)

It is interesting to see if there are cases in which the meaning of idiom is available while the meaning of the unfamiliar word is not available. Can we still construct an idiom if we are missing a word in our lexicon? And can we still correctly produce and use such an idiom? To answer these questions I have designed and run an experiment that can help me solve these issues. We can try to understand if an idiom is compositional or non compositional. If idioms are composed of single words that have their own representation in the mental lexicon, then it would be difficult to build an idiom if a lexical entry is missing. The hypothesis is that it is possible to understand an idiom if you are missing a lexical entry. The prediction is that people are able to understand and use idioms even if they don’t know what the meaning is of one of the words that the idiom contains. The prediction is based upon observations in daily speech. Therefore, idioms might not be compositional.

By asking the second main question: ‘are idiom words more similar to non-words or to real words?’ we can investigate if words that most of the time occur only in idioms exist as a separate unit in our lexicon or only exist in our lexicon as they are captured in the idiom itself. By missing a word in our lexicon we must define at what level the word is missing. For example, it might be the case that we don’t know the meaning of the word ‘loef’ (luff), but we might know all the syntactic properties of this word and recognize this word as an existing Dutch word. This would imply that at the conceptual level there is no information stored about the semantic knowledge of the word ‘loef’. But that there

is information stored about the grammatical properties and the word form at the lemma and lexeme level. This would suggest that there is no meaning relationship between the unit ‘loef’ at the lemma/lexeme level and the conceptual level. To understand if words like ‘loef’ are more similar to non-words or to real words I have designed and run an experiment that will hopefully give more insight in this complicated matter. The hypothesis is that idiom words are more similar to non-words. The prediction is that participants will decide that a lot of the idiom words are non-words. The prediction is based upon observations in daily speech of people who claimed that they never heard the word ‘loef’.

The third main question: ’Does a semantically related prime word speed up the production of an idiom as compared to an unrelated prime word?’ will tell us more precisely how the idiom is connected to other lexical items. We can imagine a network in which a node represents an idiom such as in (1a), with the word for word translation in (1b).

1. a. de kat uit de boom kijken
   b. the cat out of the tree look
   ‘wait to see which way the cat jumps’

We can imagine that this node might have connections to separate nodes of ‘kat’, ‘boom’ and ‘kijken’ but we can also imagine that the node of the idiom has a direct connection to related words like ‘kater’ (tomcat), ‘poes’ (pussycat), ‘tak’ (branch), ‘blad’ (leaf), ‘zien’ (see). If we present a related word like ‘poes’ together with a stimulus that can trigger the production of the idiom then the idea is that activation of the node ‘poes’ will result in spreading activation through this part of the network and will also activate the node that contains ‘kat’, or it might directly activate the node that contains the idiom ‘de kat uit de boom kijken’ and this might decrease the production time of the idiom. Another question we must ask is ‘Does a semantically related prime cause a lower error rate in the production of an idiom?’ If related nodes in a network have become activated then we could expect a lower error rate in the production of the idiom. But, this last assumption is dangerous because activating related nodes might also cause a speech error more easily. It could also be the case that people get distracted by the prime word and try to think of an idiom that contains the prime word. To gain more insight in what is really going on

here I have designed an experiment based on the earlier discussed experiments by Sprenger et al (2006). The hypothesis is that idioms are connected to semantically related content words. The predictions are that a semantically related prime will speed up the production of the idiom and that a semantically related prime might cause a lower error rate. The predictions are based upon the results by Sprenger et al (2006).

4. Word meaning in idioms and the meaning of idioms: Questionnaire

4.1 Method

The research question of this experiment is: is it possible to understand an idiom if you are missing a lexical entry in your mental lexicon? The hypothesis is that it is possible to understand an idiom if you are missing a lexical entry. The prediction is that people are able to understand and use idioms even if they don't have a conceptual mental representation of one its contained words. Idioms might not be compositional.

Participants

Group 1: A total of 63 children from Primary school OBS de Sleutel in Schalkhaar, Overijssel, Netherlands participated in this experiment. The children were from 7th and 8th grade. The age range of the children in this group varied from 11 to 13 years old. The group consisted of 30 boys and 33 girls. There were four bilingual children and 59 monolingual children. The bilinguals were not excluded from the dataset. They were Turkish/Dutch, Arabic/Dutch and Frisian/Dutch. All of the children were born in the Netherlands.

Group 2: Adults

A total of 50 Students from Utrecht University participated in the experiment. 36 students got paid for their participation and 14 volunteered. The age range of the participants varied from 20 to 48 years old. The group consisted of 40 women and 10 men. There were three bilingual and 47 monolingual participants. The bilingual participants were not excluded from the dataset. They were Italian/Dutch, Frisian/Dutch and Surinamese/Dutch. All of the participants were born in the Netherlands. The participants were born in different areas.

4.2 Materials

The participants were given a questionnaire that consists of two parts. Part 1 is called *word meaning in idioms*. Part 2 is called *meaning of idioms.* The content of the questionnaire is enclosed as appendix I-materials. The idioms were selected on the basis of the following criteria:

- The idiom was in the Van Dale dictionary; idioms for children.
- The idiom contained a word that seemed to be unusual. The criteria are that the word belongs to one of the following three categories:
    - Hapax legomenon: a word or phrase that appears only once in a manuscript, document, or particular area of literature. (dictionary.com)
    - Archaic word: of or belonging to the distant past (= a long time ago); from an ancient period in history. (Cambridge dictionary, Press, 2008)
    - Jargon: special words and phrases which are used by particular groups of people, especially in their work. (Cambridge dictionary, Press, 2008)
- The idiom is well known; the idiom is used frequently. The criteria are that the idiom appeared more than 25 times in a year in the LexisNexis database. "The LexisNexis database – is updated on a daily basis to include content from 35,000 reliable international sources, such as newspapers, magazines (including news and opinion magazines), trade publications, market and country reports." (LexisNexis.nl, 2011)
- The idiom did not contain a word that came from another language.
- The idiom did not contain a word that was a name.

Part 1 consists of the selected words that seem to be unusual or unknown in general. To make sure that the participants give a real answer and pay attention to the question, well known words are included. In part 1 the participant is asked to answer the following question for every word:

- Have you seen or heard this word before?
- What do you think that the meaning is of this word?

In part 2 the participant is asked to answer the following question for every idiom:

- Have you seen or heard this idiom before?
- What do you think that the meaning is of this idiom?

4.3 Procedure

As to the order of the experiment it was better to start with the separate words, this is to avoid a bias, the participants could be influenced if they have just seen the target word in the idiom and in that case they won't be able to answer the first question with no.

The children were told that they would have to fill in two questionnaires, one on the meaning of words and one on the meaning of idioms, they were told that it is not a problem if they did not know a word. The children from each grade all started to fill in the questionnaire at the same time. The children were not allowed to talk with each other and they were not allowed to use a dictionary or something else that could help them. The questionnaire required approximately 30 minutes. The correct answers were discussed with each group after the experiment. In this way, the experiment was a kind of lesson on idioms for the children. The children were rewarded with a treat in exchange for their participation.

The procedure for the adults was as follows: the participants were tested separately. All participants participated in three experiments, the second and third experiment are reported below. The questionnaire was the last experiment. Every participant was asked to fill in the questionnaire. They were not allowed to use a dictionary or something similar. The questionnaire required approximately 20 minutes.

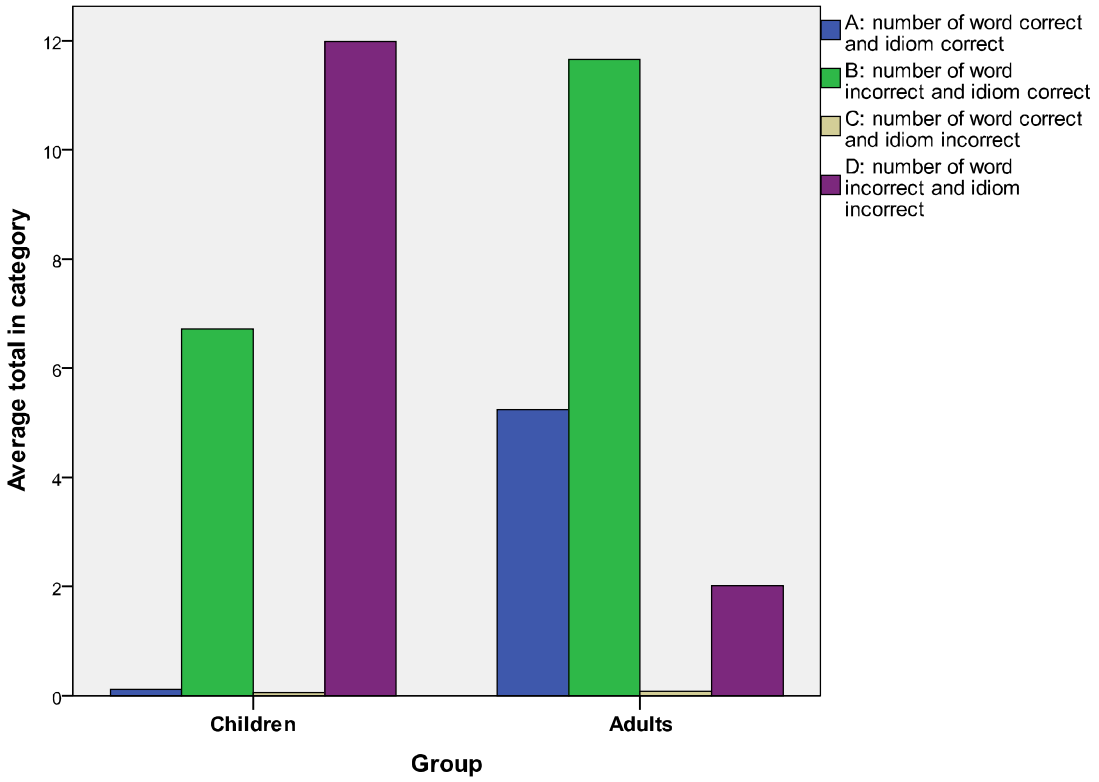
4.4 Results

The questionnaires were scored. If an answer was judged to be correct than this was indicated by 1. If an answer was judged to be wrong than this was indicated by 0. There are four possible categories. For example: if an adult or child wrote down the correct meaning of the target word, for instance ‘loef’, and if he wrote down the correct meaning of the idiom that contained the word, than this answer was valued as A.

Category A = number of word correct and idiom correct.
B = number of word incorrect and idiom correct.
C = number of word correct and idiom incorrect.
D = number of word incorrect and idiom incorrect.

The research question is as follows: Is it possible to understand an idiom if you are missing a word in your mental lexicon? The question is whether category B does exist in this dataset. And if category B does exist, then we want to know if the value of category B is higher than any of the other categories.

Figure 4 Overview Results Experiment 1



In the above figure we can see the average total sorted by group for each category. We can see a difference between the performance of the adults and the children. Overall, the adults performed better than the children, because their total average in category A and B far exceeds that of the children. Most of the adults' answers (578 out of 950) belonged to category B, i.e. they mostly had the word incorrect and the idiom correct. This implies that their knowledge of the idioms was high, whereas their knowledge of the target words was not very high. The children had difficulties with the idioms and the words, thus most of the children's of the answers (752 out of 1197) belonged to category D, i.e. they mostly had the word incorrect and the idiom incorrect.

If idioms are compositional then this would imply that we have all the building blocks of an idiom in our lexicon. Therefore, Sprenger et al (2006) is able to say that category B (word wrong and idiom correct) should not exist and if it does exist then its value should not be very high. The null hypothesis would be that category B has the lowest value. We can test this prediction with use of a chi square goodness of fit test. We

will first look at the adults. There were 19 possible combinations and 50 adults, therefore we have a total of 19*50=950 to divide among the four categories.

Table 1 chi square goodness of fit test adults

	Observed N		**Expected N**	**Residual**
A. word correct and idiom correct	28,11 %	= 267	25 %	3,11 %
B. word incorrect and idiom correct	60,84 %	= 578	25 %	35,84 %
C. word correct and idiom incorrect	0,42 %	= 4	25 %	-24,58 %
D. word incorrect and idiom incorrect	10,63 %	= 101	25 %	-14,37 %
Total	100 %	= 950	100 % = 950	

We can see that the values for the observed n and the expected n differ. The prediction that adults know a lot of the idioms is correct. But, the value for category B is extremely high. The test statistic is 2781 and the df is 3, this implies that the frequency distribution of the variable is not equal to the theoretical distribution. The Asymp. Sig. is 0,000. Based on this data, we can reject the null hypothesis. The values for the observed data differ from the expected data. The value for category B is very high. It follows that we can reject the assumption that idioms are compositional as was proposed by Sprenger et al and that we can accept the hypothesis that it is possible to understand an idiom if you are missing a lexical entry.

We already know that the children did not perform as well as the adults did. We can therefore expect a higher value for category D.

Table 2 chi square goodness of fit test children

	Observed N		Expected N	Residual
A. word correct and idiom correct	0,59 %	= 7	25 %	-24,4 %
B. word incorrect and idiom correct	36,26 %	= 434	25 %	11,26 %
C. word correct and idiom incorrect	0,33 %	= 4	25 %	24,67 %
D. word incorrect and idiom incorrect	62,82 %	= 752	25 %	37,82 %
Total	100 %	=1197	100 % = 1197	

We can see that the values for category B and D are very high. The test statistic is 1322 and the df is 3, this implies that the frequency distribution of the variable is not equal to the theoretical distribution The Asymp. Sig. is 0,000. All the idioms that were used in the experiment came from the Van Dale dictionary; idioms for children. This dictionary is especially for children from the age of ten. It turned out that a lot of the chosen idioms from this dictionary were too difficult for the children. The value of category A is very low. If children knew the idiom, then the chance that they knew the word was not high. The values for the observed data differ from the expected data. It follows that the assumption that all the building blocks of an idiom are in our lexicon cannot be justified.

The interesting question now is if our hypothesis still holds if we take the adults and children together as a group.

Table 3 chi square goodness of fit test both groups

	Observed N		Expected N	Residual
A. word correct and idiom correct	12,76 %	= 274	25 %	-262,8
B. word incorrect and idiom correct	47,14 %	= 1012	25 %	475,3
C. word correct and idiom incorrect	0,37 %	= 8	25 %	-528,8
D. word incorrect and idiom incorrect	39,73 %	= 853	25 %	316,3
Total	100 %	= 2147	100 % = 2147	

We can clearly see that if we take the adults and children together, category B is the biggest category. The test statistic is 1256 and the df is 3, this implies that the frequency distribution of the variable is not equal to the theoretical distribution. The Asymp. Sig. is 0,000. The values for the observed data differ from the expected data. It follows that it is not necessary to have all the words of an idiom in our lexicon at the conceptual level. The results of this experiment are enclosed as appendix I-results. The results of this experiment will be discussed in part 7 (discussion).

5. Experiment 2: Lexical decision task

5.1 Method

The research question is: are idiom words comparable to non-words or to real words? The hypothesis is that idiom words are comparable to non-words. The predictions are that participants will decide that a lot of the idiom words are non-words. Idiom words will require a higher reaction time than real words.

Participants

A total of 50 students from Utrecht University participated in the experiment. The participants also took part in the first and third experiment. 36 students were paid for their participation and 14 students volunteered. The age range of the participants varied from 20 to 48 years old. The group consisted of 40 women and 10 men. There were three bilingual and 47 monolingual participants. The bilingual participants were not excluded from the dataset. They were Italian/Dutch, Frisian/Dutch and Surinamese/Dutch. All of the participants were born in the Netherlands.

5.2 Materials

For this experiment a Lexical Decision task was used. The idiom words from the first experiment were used again. 25 non-words were generated with a non-word generator. (Duyck, W., Desmet, T., Verbeke, L., & Brysbaert, M. (2004)). The standard parameter settings that were used in the Non-Word generator were as follows: the constrain number of neighbors: min 2- max 10. There were no other standard constraints selected.

Since the number of letters in the idiom words and real words differs, the parameter setting for number of letters for creating the non-words differs. Some non-words were made into verbs by adding a suffix. The number of letters in the non-words was similar to the number of words in the idiom words and real words, namely between 3 and 7.

The list of words that were used in this experiment is enclosed as appendix II-materials.

5.3 Procedure

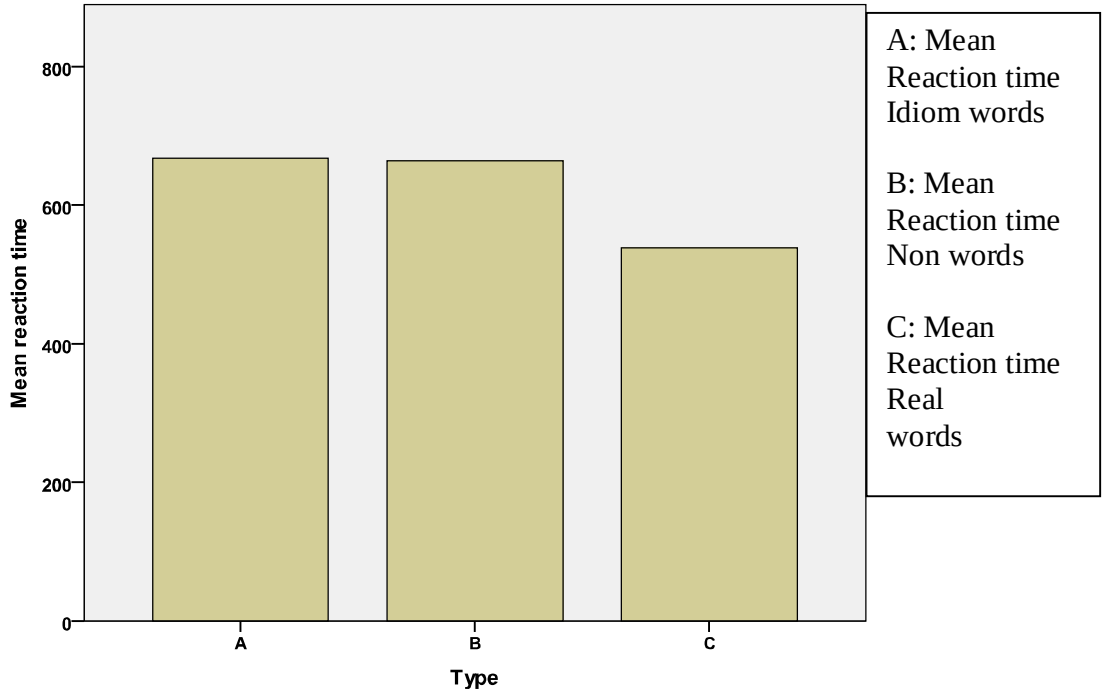
The computerprogram ZEP (Veenker, 2011) was used to run the experiment. Participants saw a letterstring, this could be a word from the list of real words, from the list of idiom

words or from the list of non-words. The word was presented in the middle of a computer screen. The order of presented words was randomized for every participant. The task was to decide whether the word is an actual existing word or not. If the word was according to the participant an actual existing Dutch word they had to press the yes button and otherwise the no button. The participants were under time pressure, the limited time was 1500 ms. The reaction times were measured while using Linux-real time.

5.4 Results

A value is indicated as a missing value if the participant did not press any button or if the participant did not pressed a button within 1500 ms. There were 25 missing values for the idiom words. There were 16 missing values for the non-words and there were 5 missing values for the real words. The idiom words did seem to cause some troubles at this point. The first question we need to investigate is: 'is there any difference in reaction times between the three groups of words?' The null hypothesis would be that there is no difference between the three types of words. Only the values of the correct answers were scored. A value was judged to be false if the participant gave an incorrect answer. The mean reaction times for each category are presented in figure 5.

Figure 5 Mean Reactiontime Lexical Decision task



The mean in milliseconds for the idiom words is 662. The mean for the non-words is 675. The mean for the real words is 535. These results indicate that the real words took on average the least time for the participants to judge. The non-words took on average the most time for the participants to judge. The difference between the non-words and idiom words is very small. With use of a Repeated Measurements test we can calculate what is going here. There was a significant effect, Wilks' Lambda = 0.11, F (2.48) =194.05 p < 0.001. This implies that for this dataset the differences in reaction times between the three types of groups are significant within subject. Thus, we can reject the null hypothesis. With use of a posthoc test with Bonferroni correction we can test which word categories differ significantly from each other.

Table 4 Posthoc Bonferroni Reaction times Lexical Decision task

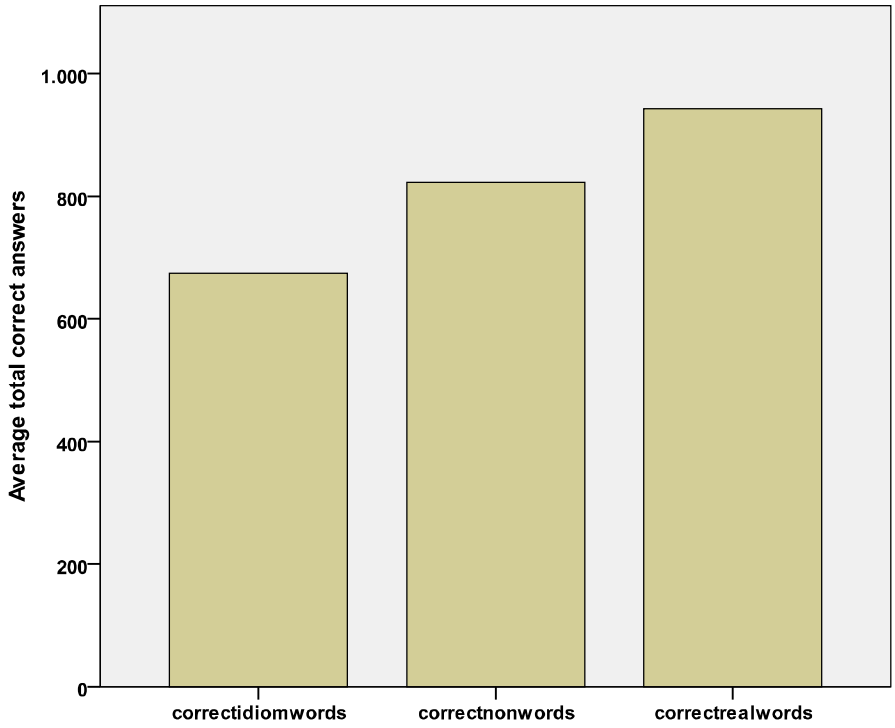
(I) type	(J) type	Mean Difference (I-J)	Std. Error	Sig.
1	2	-22,521	8,817	0,032
1	3	118,757*	8,561	0,000
2	3	141,278*	8,096	0,000

*. The mean difference is significant at the 0.01 level.

Type 1 is the category of idiom words. Type 2 is the category of non-words and type 3 is the category of real words. In table 6 we can read that the difference between the idiom words and the non-words is not significant. The difference between the idiom words and the real words is significant. The difference between the non words and the real words is significant. The research question of this experiment was: 'are idiom words comparable to non-words or to real words? The hypothesis was that idiom words are comparable to non-words. If we look at the results of the reaction times, we can see that the data is in accordance with the hypothesis.

The question now is how well did the participants performed on the three word categories. The mean correct answers for each category are presented in figure 6. The total of category A, B and C is 3000. There were 1000 answers for each category.

Figure 6 Mean correct answer Lexical Decision task



In 67% of the cases participants judged the idioms to be existing Dutch words. That means that in 33 % of the cases participants judged the idioms to be non-words. In 82 % of the cases participants judged the non-words to be non-words. In 18 % of the cases participants judged the non-words to be existing Dutch words. In 94 % of the cases participants judged the real words to be real words. Still, in 6 % of the cases real words were judged to be non-words. In figure 6 we can see that idiom words in the way they behave are more similar to non-words than to real words. The results of this experiment will be discussed in part 7 (discussion).

6. Experiment 3: The Sprenger et al experiment

6.1 Method

The research question is: does a semantically related prime speed up the production of an idiom as compared to an unrelated prime? And: 'does a semantically related prime cause a lower error rate in the production of an idiom?' The hypothesis is that idioms are connected to semantically related content words. The predictions are that a semantically related prime will speed up the production of the idiom. And that a semantically related prime might cause a lower error rate.

Participants

A total of 50 Students from Utrecht University participated in the experiment. The participants also took part in the first and second experiment. 36 students were paid for their participation and 14 volunteered. The age range of the participants varied from 20 to 48 years old. The group consisted of 40 women and 10 men. There were three bilingual and 47 monolingual participants. The bilingual participants were not excluded from the dataset. They were Italian/Dutch, Frisian/Dutch and Surinamese/Dutch. All of the participants were born in the Netherlands.

6.2 Materials

The experiment was a combination of a picture naming task and a priming experiment. The idioms were selected from the van Dale junior idiom dictionary (2007). The pictures were selected from Goldberg (2007) and from van Dale junior idiom dictionary (2007). The criteria are:

- The idiom is well known. the idiom is used frequently. The criteria are that the idiom appeared more than 10 times in a year in the LexisNexis database. "The LexisNexis database – is updated on a daily basis to include content from 35,000 reliable international sources, such as newspapers, magazines (including news and opinion magazines), trade publications, market and country reports." (LexisNexis.nl, 2011)
- The picture is able to stimulate the production of the idiom.

The prime word is either semantically related to the focus word of the idiom or semantically and phonologically unrelated to the focus word of the idiom. The materials that were used in this experiment are enclosed as appendix III-materials.

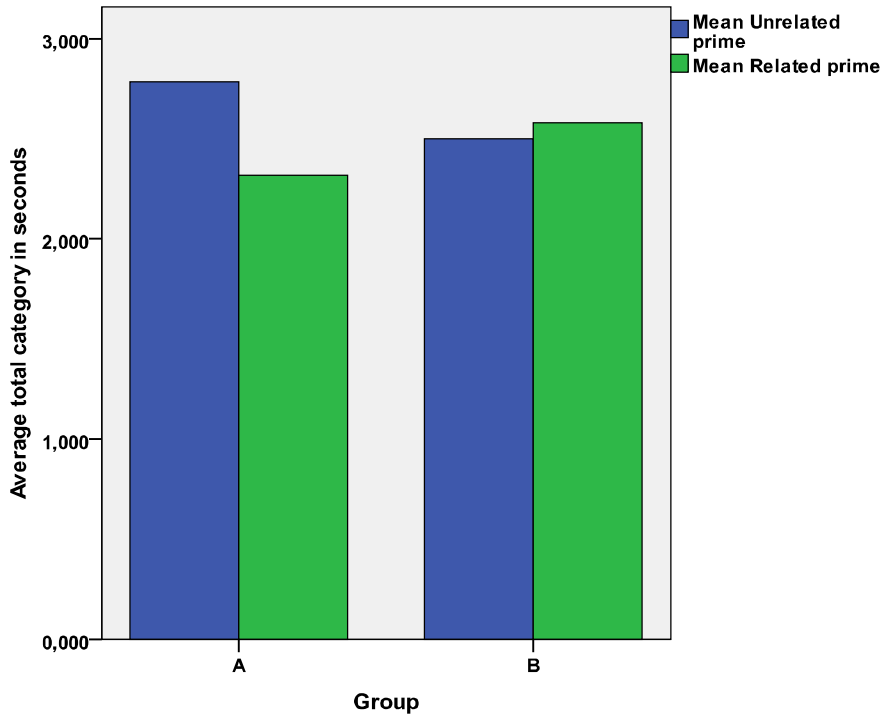
6.3 Procedure

The experiment is a variation of two experiments that were done by Sprenger et al (2006). This experiment will have a picture functioning as the stimulus. The picture is a literal image of the idiom. An interesting observation while investigating the school materials on idioms is that teachers often make use of literal pictures if they are using pictures. This finding justifies the use of literal pictures in an experiment. The picture will be presented on the screen for 5 seconds. The prime will be introduced together with the picture. The prime will be presented on the screen for 1.5 seconds. The prime can either be semantically related to one of the focus words of the idiom or semantically and phonologically unrelated. The participant saw a picture (the stimulus) and a word (the prime). Every idiom was presented only once. Therefore, the participants were divided between group A and group B. Thus, there is a split-design. The participants in Group A and B were given the same practice trails. For every participant the order of pictures was randomized. The task was to produce the appropriate idiom. The participant had to produce the idiom within a certain period of time, if the participant failed to produce the idiom, then the next idiom was presented. Every participants was given two opportunities to practice the task. The answers were recorded and reaction times were measured with the audio program Audacity. The reaction time is indicated in seconds and milliseconds. A value is indicated as a missing value if the participant gave an incorrect answer. The participant was not required to answer within 5 seconds. The participant was not required to produce a full sentence with articles. An answer was indicated as an incorrect answer if the nouns and verbs did not correspond to the idiom.

6.4 Results

The first question that we want to answer is if there is a difference in reaction time between a related and an unrelated prime. The total of unrelated and related primes is 544, this means that there are 456 missing values, because we have 50 participants and 20 questions. The mean reaction times are presented in figure 6.

Figure 7 Mean Reaction Time



The mean in reaction time of the related primes lies lower than the mean of the unrelated primes. Overall reaction times are faster with a related prime than with an unrelated prime. The results of this experiment are enclosed as appendix III-results. The mean of the related primes in group A is 2,3 seconds and in group B 2,6. The mean of the unrelated primes in group A is 2,8 seconds and in group B 2,5. The question now is if this difference between related and unrelated primes is significant. The null hypothesis would be that the difference is not significant. With use of a Repeated measurements test we can test if the difference between related and unrelated primes in reaction time is significant within subjects and between groups. There was not a significant effect for prime type, Wilks' Lambda = 0.92, F (1.46) = 3.99, p = 0.052. However, there was a significant effect for prime type * group, Wilks' Lambda = = 0.85, F (1,46) = 8.02, p = 0.007. Thus, the difference for prime type is not significant, there is a tendency that related primes

have a positive influence on the reaction time, but we cannot reject the null hypothesis. If we look at prime type * between groups, we can see that there is something going on here. It is remarkable that group has an influence on the reaction time. It is not yet clear what the cause of this difference is.

We can now take a look at the error rate. All incorrect answers were indicated as a missing value. The hypothesis is that unrelated primes might cause a higher error rate than related primes do. The mean of the missing values for the related primes is 4.40 with a total of 220 cases and the mean of the missing values for the unrelated primes is 4.56 with a total of 228 cases. This difference is very small. With use of a Repeated measurements test we can test if the difference in missing values between related and unrelated primes is significant within subjects and between groups. There was not a significant effect for prime type, Wilks' Lambda = 0.99, F (1,48) = 0.24, p = 0.63. There was a significant effect for prime type * group, Wilks' Lambda = 0.82, F (1.48) = 10.92, p = 0.002. Thus, the difference for prime type is not significant, there is a tendency that related primes have a positive influence on correct answers, but we cannot reject the null hypothesis. It seems that group has an influence on the missing values as well.

7. Discussion

Pattern of knowledge

The results of Cacciari (1993) showed a success rate of 51,9 % for this type of idioms. These were ten year old children. The performance of the children of the first experiment on the knowledge of idioms was if we take category A and B together a score of 37 % correct answers. The children were between 11 and 13 years old. The idioms that were used in this questionnaire were opaque idioms. This difference might be due to the fact that the idioms were chosen because they contained an unfamiliar word. If we look at the difference in performance between the children and the adults we clearly see a pattern of development. Children are still acquiring the meaning of a lot of words whereas the adults already have a lot of these meanings stored. The van Dale idioms dictionary for children is very useful and will help in understanding the meaning and heritage of a lot of idioms.

The way idioms are stored in the mental lexicon

In this thesis the focus was the storage of idioms in our mental lexicon. Sprenger et al (2006) proposed the superlemma. A superlemma is a representation of the syntactic properties of the idiom that is connected to its building blocks. The assumption of Sprenger et al is that idioms are compositional in building and that they are non-compositional in meaning. The first main question of this thesis was: *is it possible to understand an idiom if you are missing a word in your mental lexicon?* Yes, we can still correctly produce and use such an idiom if we are missing a word in our lexicon. Therefore, idioms cannot be compositional, not in meaning and not in building. If we assume that an idiom is stored as a sum of the words from the lexicon then it is necessary that a person has these words in its lexicon and knows what the meaning is of these separate words. But an argument against this assumption is that it is possible to use an idiom correctly and nevertheless not be able to know the meaning of a separate word in the idiom. It has just been proven in the first experiment that this is a possibility. In fact, it is not only a possibility. There were lots of cases of people knowing the idiom and not knowing the word. With this experiment it has been proven that it is possible to understand and use an idiom correctly, even if one of the building blocks of the idiom is

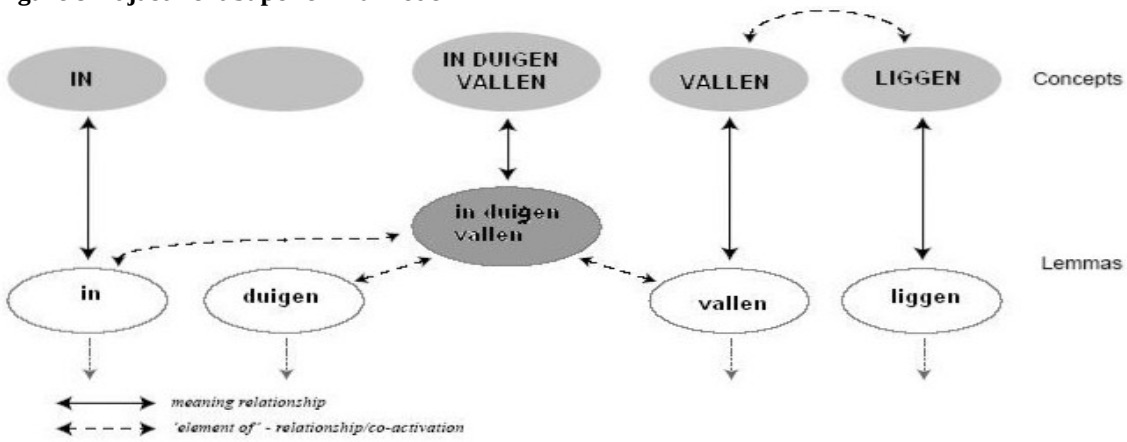
missing in the lexicon. Thus, if we take into account the fact that it is common that people can use and understand idioms correctly and fail to assign meaning to one of building blocks of the idiom, then this might be a problem for superlemma theory. Sprenger et al (2006) mentioned the compositional nature of idiomatic expressions. The word 'compositional' refers to combining distinct elements to form a whole. But what if one of these elements is missing? Can we still add up all the elements that are available and create the idiom? The most important question now is: do we have a word like 'loef' in our lexicon and simply fail to assign a correct meaning? This would imply that we do have the grammatical properties and the information about the word form of 'loef' stored at the lemma/lexeme level. Or does a word like 'loef' not exist as a single unit at any level in our lexicon but only in combination with the idiom? The results of experiment 2 might give us the answer.

The second main question was: *'are idiom words more similar to non-words or to real words?* Idiom words are more similar to non-words than to real words. If an idiom word it is not recognized as a real word then there is no information stored at the conceptual and also not at the lemma/lexeme level. It is also possible that there is no semantic knowledge stored but the idiom word has been recognized as an existing word. In this case, there is information stored at the lemma/lexeme level. We said that by missing a word in our lexicon we must define at what level the word is missing. If we do not know the meaning of the word 'duigen' (staves), then this would imply that at the conceptual level there is no information stored for this word. Because we can only know the meaning of the word 'duigen' if we have this knowledge stored at the conceptual level. If the participants judge the word 'duigen' to be an existing Dutch word, then this would imply that we do know all the syntactic properties and the word form of the word 'duigen'. If the participants judge the word 'duigen' to be a non-word, then this would imply that this word is not in their mental lexicon at all. Therefore, there can't be any information that is stored. We know that children and adults once they have acquired the grammatical knowledge are able to modify non-words to the grammatical tense that is asked. Evidence for this assumption comes from the famous WUG test (Berko, 1958). This experiment showed that even if a person does not have the conceptual knowledge, it is still possible to have some sort of grammatical knowledge. But what about idiom

words? If we look at the idiom words and see them as non-words then we can modify the words to any grammatical tense. But if we look at what they are, namely idiom words, then it doesn't feel right to adjust the grammatical tense. An interesting experiment would be to ask a new set of participants to modify a set of idiom words and also give them the option 'not possible'.

The results of the questionnaire for the adults showed that 16 % knew the meaning of the word 'duigen'. 96 % knew the meaning of the idiom 'in duigen vallen' (falling in staves). If we look at the results of the lexical decision task then we can see that 76 % of the participants gave a correct answer and judged that the idiom word 'duigen' is an existing Dutch word. This would imply that even if we do not have 'duigen' at the conceptual level, we can have 'duigen' at the lemma/lexeme level, because people recognize the idiom word as an existing Dutch word. In the model we can see how this should work. Cruse (1986) would say that such a word is semantically empty.

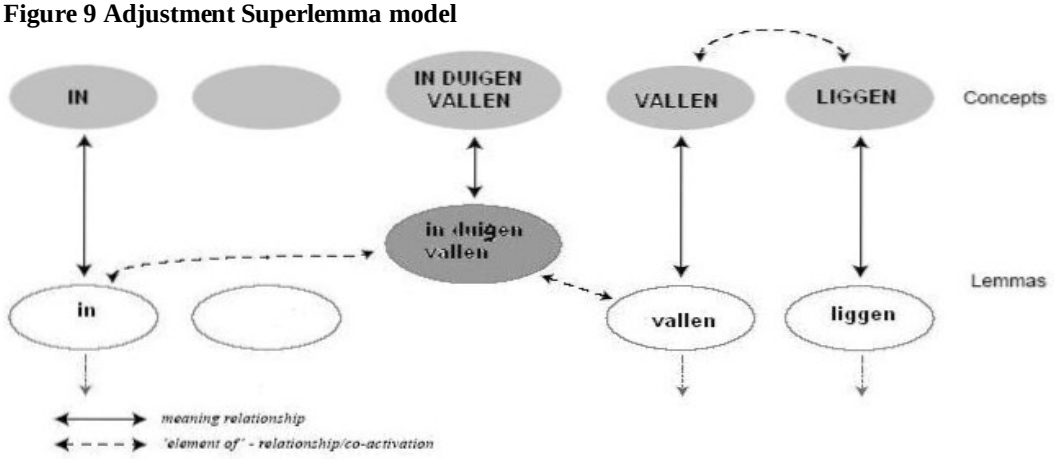
Figure 8 Adjustment Superlemma model



The results of the questionnaire for the adults showed that 0 % knew the meaning of the word 'falie' (a type of headscarf). 98 % knew the meaning of the idiom 'op je falie krijgen' (to get a sound beating). If we look at the results of the lexical decision task then we can see that 42 % of the participants gave a correct answer and judged that the idiom word 'falie' is an existing Dutch word. We can see here that it can be the case that the idiom word is not recognized as an existing Dutch word. This implies that the idiom word is not in our lexicon at all. The idiom word is now comparable to a non-word. It would be

ridiculous to say that every single non-word has its own lemma. Therefore, there must be some general grammatical knowledge stored for non-words that share properties with each other or with real words. Thus, idiom words that are seen as non-words do not have information stored at the lemma/lexeme level. In the model we can see how this should work.

Figure 9 Adjustment Superlemma model



The third main question was: *'does a semantically related prime word speed up the production of an idiom as compared to an unrelated prime word?'* We saw that this difference was not significant. The experiment in this thesis was a different version of experiments that were done by Sprenger et al (2006). The differences between this experiment and the experiment from Sprenger et al (2006) were that there was no training, there was a different set of idioms, there was a picture functioning as a stimulus, and the amount of primes was reduced to two, but the participant saw only one idiom together with the prime.

We are now interested in the related nodes. In the above model we can see a relationship between 'vallen' (fall) and 'liggen' (lie). We would suspect that if we present someone with a word that is semantically related to a focus word or another word of the idiom the speed of the production of the idiom will increase. The hypothesis of the third experiment was that idioms are connected to semantically related content words. The prediction was that a semantically and phonologically unrelated prime will slow down the production of the idiom and that it might also cause a higher error rate. The results of this experiment showed that for this group of participants the semantically related primes

indeed speed up the production of the idiom. However, the difference between a semantically related prime and an unrelated prime was not high and the difference was not significant, therefore we cannot draw the conclusion that semantically related primes speed up the production of idioms. The unrelated primes indeed cause a higher error rate for this group of participants, but the difference with related primes was very small and the difference was not significant, therefore we cannot draw the conclusion that unrelated primes cause a higher error rate. In general, the error rate was very high. This can have several causes. Without making use of a learning phase, there is a risk that people do not know the idioms. However, asking the participants revealed that they did know the idioms but they just could not come up with the idiom at that moment. A lot of the participants indicated that they got distracted by the words. Some participants indicated that they even tried to ignore the words. Ignoring words might work consciously but we do register the word unconsciously. Thus, a simplification of the Sprenger et al experiment showed that there is no significant difference between related an unrelated primes, but there is a tendency.

Bibliography

Berko, J. (1958). The Child's Learning of English Morphology. *Word*, 14, 150-177.

Cacciari, C. (1993). The place of idioms. In C. Cacciari, & P. Tabossi, *Idioms, Processing, structure and interpretation*. New Yersey: Lawrence Erlbaum Associates.

Cacciari, C., Glucksberg, S. (1995) Imagining Idiomatic expressions: Literal or figurative meanings? In Everaert, M. E.-J. van der Linden, A. Schenk, R. Schreuder (eds.) (1995). *Idioms: Structural and Psychological Perspectives*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Cambridge, U. (2008). *Cambridge dictionary*. Cambridge: Cambridge University Press.

Cutting, J.C., & Bock, K. (1997). That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory & Cognition*, 25 (1), 57-71.

Daniels, W. (2007). Junior Spreekwoordenboek. Wat betekenen onze spreekwoorden? Vanaf 10 jaar. Van Dale Uitgevers, Utrecht/Antwerpen.

Duyck, W., Desmet, T., Verbeke, L., & Brysbaert, M. (2004). Wordgen: A Tool for Word Selection and Non-Word Generation in Dutch, German, English, and French. *Behavior Research Methods, Instruments & Computers*, 36 (3), 488-499.

Everaert, M. E.-J. van der Linden, A. Schenk, R. Schreuder (eds.) (1995). *Idioms: Structural and Psychological Perspectives*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Everaert, M., E.-J. van der Linden, A. Schenk, R. Schreuder, (1994). Introduction, In: M. Everaert et al. *Idioms: Structural and Psychological Perspectives*, pp 1-13

Flores D'Arcais, G. (1993). The comprehension and semantic interpretation of idioms. In C. Cacciari, & P. Tabossi, *Idioms, Processing, structure and interpretation*. New Yersey: Lawrence Erlbaum Associates.

Gibbs, R. (1995). Idiomaticity and human cognition. In Everaert, M. E.-J. van der Linden, A. Schenk, R. Schreuder (eds.) (1995). *Idioms: Structural and Psychological Perspectives*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Gillis, S., & Schaerlaekens, A. (2002). *Kindertaalverwerving*. Groningen: Martinus Nijhoff.

Jackendoff, R. (1995). The boundaries of the lexicon. In M. Everaert, E.J. VanderLinden, & R. Schreuder (Eds.), *Idioms. Structural and psychological perspectives* (pp.133-165). Lawrence Erlbaum Associates, Publishers, Hillsdale, New Yersey.

Kempen, G. (1996). Computational models of syntactic processing in human language comprehension. In T. Dijkstra & K. DeSmedt (Eds.), *Computational linguistics: Symbolic and subsymbolic models of language processing* (pp. 192–220). London: Taylor and Francis.

Levelt, W.J.M. (1994). The skill of speaking. In P.Bertelson, P. Eelen, and G.d'Ydewalle (Eds.), International Perspectives on Psychological Science. Vol.1: Leading themes (pp.89-103). Hove, England:Lawrence Erlbaum Associates. In M. Gazzaniga, R. Ivry, G.Mangun. (2002). Cognitive neuroscience, the biology of the mind. W.W. Norton & company, New York.

Levorato, M. (1993). The acquisition of idioms and the development of figurative meaning. In C. Cacciari, & P. Tabossi, *Idioms, processing, structure and interpretation* (p. 101). New Yersey: Lawrence Erlbaum Associates.

Nenonen. (2002). Representation and processing of idioms: evidence from aphasia. *Journal of Neurolinguistics 15* , 43-58.

Nippold, M., & Taylor, C. (1995). Idiom understanding in youth: Further examination of familiarity and transparency. *Journal of Speech and Hearing Research 38* , 426-433.

OED. (2011). *Oxford English dictionary (3rd ed.)*. Oxford: Oxford University Press.

Papagno, C. (2001). Comprehension of metaphors and idioms in patients with Alzheimer's disease. *Brain* , 1450-1460.

Sprenger, S., Levelt, W., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language 54* , 161-184.

Veenker, Theo. (2011) The Zep Experiment Control Application.
http: //www.let.uu.nl/~Theo.Veenker/personal/zep/

Waxman, S., Lidz, J. (2006). Early word learning. In D. Kuhn & R. Lerner, *Handbook of child psychology* (p. 299). John Wiley & Sons, Inc.

Appendix I

Materials Experiment 1 Word meaning in idioms and the meaning of idioms

A. Words used in the questionnaire

loef	rapport	biezen	vleet	baat
wolf	kladden	cirkel	vuur	politie
kijf	metten	kaf	heft	ornaat
bloem	koffer	grif	klippen	bres
kielzog	duigen	falie	lachen	ijs
blaam	stempel	heil	heinde	gareel

B. Idioms used in the questionnaire

1. in het gareel lopen.
Hendrik heeft zijn kinderen keurig in het gareel lopen.
2. van heinde en verre.
Ze kwamen van heinde en verre naar het festival.
3. ergens baat bij hebben.
Mijn moeder heeft veel baat bij deze kruidenthee.
4. iemand bij de kladden grijpen.
De politie greep Ron meteen bij zijn kladden.
5. mij treft geen blaam.
De lerares trof in ieder geval geen blaam.
6. het heft in handen nemen.
Jan nam bij de organisatie het heft in handen.
7. in duigen vallen.
Sarah zag haar droom in duigen vallen.
8. op je falie krijgen.
Erik kreeg op zijn falie van zijn vader.
9. het kaf van het koren scheiden.
De redactie was druk bezig het kaf van het koren te scheiden.
10. grif van de hand gaan.
De koekjes van Roos gingen grif van de hand.
11. ergens geen heil in zien.
Janneke zag er geen heil in.
12. in het kielzog van iemand.
We liepen in het kielzog van tante Beppie.
13. op de klippen lopen.
Het huwelijk van zijn ouders is op de klippen gelopen.
14. iemand de loef afsteken.
Het lukte Henk om Klaas de loef af te steken.
15. korte metten met iets maken.
De directeur van de school maakte korte metten met het spijbelen.
16. je biezen pakken.
Het kassameisje werd ontslagen en moest onmiddelijk haar biezen pakken.
17. in vol ornaat.
De burgemeester kwam in vol ornaat de school openen.
18. bij de vleet.
Sarah heeft geld bij de vleet.
19. voor iemand in de bres springen.
De brandweerman sprong voor haar in de bres.
20. dat staat buiten kijf.
Het staat buiten kijf dat veel snoepen ongezond is.

Results Experiment 1 Word meaning in idioms and the meaning of idioms

Answer

Item			Frequency	Percent	Valid Percent	Cumulative Percent
baat	Valid	word correct and idiom correct	46	40,7	40,7	40,7
		word incorrect and idiom correct	5	4,4	4,4	45,1
		word incorrect and idiom incorrect	62	54,9	54,9	100,0
		Total	113	100,0	100,0	
biezen	Valid	word correct and idiom correct	3	2,7	2,7	2,7
		word incorrect and idiom correct	100	88,5	88,5	91,2
		word incorrect and idiom wrong	10	8,8	8,8	100,0
		Total	113	100,0	100,0	
blaam	Valid	word correct and idiom correct	35	31,0	31,0	31,0
		word incorrect and idiom correct	11	9,7	9,7	40,7
		word correct and idiom incorrect	2	1,8	1,8	42,5
		word incorrect and idiom incorrect	65	57,5	57,5	100,0
		Total	113	100,0	100,0	
bres	Valid	word correct and idiom correct	2	1,8	1,8	1,8
		word incorrect and idiom correct	73	64,6	64,6	66,4
		word incorrect and idiom incorrect	38	33,6	33,6	100,0
		Total	113	100,0	100,0	
duigen	Valid	word correct and idiom correct	8	7,1	7,1	7,1

		word incorrect and idiom correct	83	73,5	73,5	80,5
		word incorrect and idiom incorrect	22	19,5	19,5	100,0
		Total	113	100,0	100,0	
falie	Valid	word correct and idiom correct	2	1,8	1,8	1,8
		word incorrect and idiom correct	102	90,3	90,3	92,0
		word incorrect and idiom incorrect	9	8,0	8,0	100,0
		Total	113	100,0	100,0	
gareel	Valid	word correct and idiom correct	1	,9	,9	,9
		word incorrect and idiom correct	56	49,6	49,6	50,4
		word incorrect and idiom incorrect	56	49,6	49,6	100,0
		Total	113	100,0	100,0	
grif	Valid	word correct and idiom correct	5	4,4	4,4	4,4
		word incorrect and idiom correct	67	59,3	59,3	63,7
		word incorrect and idiom incorrect	41	36,3	36,3	100,0
		Total	113	100,0	100,0	
heft	Valid	word correct and idiom correct	24	21,2	21,2	21,2
		word incorrect and idiom correct	66	58,4	58,4	79,6
		word incorrect and idiom incorrect	23	20,4	20,4	100,0
		Total	113	100,0	100,0	
heil	Valid	word correct and idiom correct	20	17,7	17,7	17,7
		word incorrect and idiom correct	54	47,8	47,8	65,5
		word incorrect and idiom incorrect	39	34,5	34,5	100,0

		Total	113	100,0	100,0	
heinde	Valid	word correct and idiom correct	5	4,4	4,4	4,4
		word incorrect and idiom correct	37	32,7	32,7	37,2
		word incorrect and idiom wrong	71	62,8	62,8	100,0
		Total	113	100,0	100,0	
kaf	Valid	word correct and idiom correct	25	22,1	22,1	22,1
		word incorrect and idiom correct	33	29,2	29,2	51,3
		word incorrect and idiom incorrect	55	48,7	48,7	100,0
		Total	113	100,0	100,0	
kielzog	Valid	word correct and idiom correct	10	8,8	8,8	8,8
		word incorrect and idiom correct	29	25,7	25,7	34,5
		word incorrect and idiom incorrect	74	65,5	65,5	100,0
		Total	113	100,0	100,0	
kijf	Valid	word correct and idiom correct	8	7,1	7,1	7,1
		word incorrect and idiom correct	62	54,9	54,9	61,9
		word incorrect and idiom incorrect	43	38,1	38,1	100,0
		Total	113	100,0	100,0	
klippen	Valid	word correct and idiom correct	34	30,1	30,1	30,1
		word incorrect and idiom correct	40	35,4	35,4	65,5
		word correct and idiom incorrect	4	3,5	3,5	69,0
		word incorrect and idiom incorrect	35	31,0	31,0	100,0
		Total	113	100,0	100,0	

loef	Valid	word correct and idiom correct	3	2,7	2,7	2,7
		word incorrect and idiom correct	37	32,7	32,7	35,4
		word correct and idiom incorrect	1	,9	,9	36,3
		word incorrect and idiom incorrect	72	63,7	63,7	100,0
		Total	113	100,0	100,0	
metten	Valid	word correct and idiom correct	2	1,8	1,8	1,8
		word incorrect and idiom correct	89	78,8	78,8	80,5
		word incorrect and idiom incorrect	22	19,5	19,5	100,0
		Total	113	100,0	100,0	
ornaat	Valid	word correct and idiom correct	26	23,0	23,0	23,0
		word incorrect and idiom correct	19	16,8	16,8	39,8
		word correct and idiom incorrect	1	,9	,9	40,7
		word incorrect and idiom incorrect	67	59,3	59,3	100,0
		Total	113	100,0	100,0	
vleet	Valid	word correct and idiom correct	15	13,3	13,3	13,3
		word incorrect and idiom correct	49	43,4	43,4	56,6
		word incorrect and idiom incorrect	49	43,4	43,4	100,0
		Total	113	100,0	100,0	

Appendix II

Materials Experiment 2 Lexical decision task

A. List of idiom-words

loef	falie	klippen	heinde	kladden
kijf	grif	heft	ornaat	duigen
kielzog	kaf	baat	metten	bres
blaam	biezen	vleet	heil	gareel

B. List of non-words

tupten	makkel	huggen	zeffen	tirts
zomig	zoug	plump	kieg	kwif
stunk	spen	sleef	driep	spletten
schouf	mamp	tubel	oraat	huts

C. List of real words

koffer	stempel	lachen	ijs	zinken
bloem	rapport	vuur	politie	avond
wolf	cirkel	zingen	koek	vliegen
bezem	krant	sneeuw	koorts	weiland

Results Experiment 2 Lexical decision task correct answers

Correct answers idiom words Lexical Decision Task

word			Frequency	Percent	Valid Percent	Cumulative Percent
baat	Valid	false	8	16,0	16,0	16,0
		true	42	84,0	84,0	100,0
		Total	50	100,0	100,0	
biezen	Valid	false	10	20,0	20,0	20,0
		true	40	80,0	80,0	100,0
		Total	50	100,0	100,0	
blaam	Valid	false	15	30,0	30,0	30,0
		true	35	70,0	70,0	100,0
		Total	50	100,0	100,0	
bres	Valid	false	15	30,0	30,0	30,0
		true	35	70,0	70,0	100,0
		Total	50	100,0	100,0	
duigen	Valid	false	12	24,0	24,0	24,0
		true	38	76,0	76,0	100,0
		Total	50	100,0	100,0	
falie	Valid	false	29	58,0	58,0	58,0
		true	21	42,0	42,0	100,0
		Total	50	100,0	100,0	
gareel	Valid	false	9	18,0	18,0	18,0
		true	41	82,0	82,0	100,0
		Total	50	100,0	100,0	
grif	Valid	false	27	54,0	54,0	54,0
		true	23	46,0	46,0	100,0
		Total	50	100,0	100,0	
heft	Valid	false	11	22,0	22,0	22,0
		true	39	78,0	78,0	100,0
		Total	50	100,0	100,0	
heil	Valid	false	3	6,0	6,0	6,0
		true	47	94,0	94,0	100,0
		Total	50	100,0	100,0	
heinde	Valid	false	31	62,0	62,0	62,0
		true	19	38,0	38,0	100,0

		Total	50	100,0	100,0	
kaf	Valid	false	17	34,0	34,0	34,0
		true	33	66,0	66,0	100,0
		Total	50	100,0	100,0	
kielzog	Valid	false	9	18,0	18,0	18,0
		true	41	82,0	82,0	100,0
		Total	50	100,0	100,0	
kijf	Valid	false	25	50,0	50,0	50,0
		true	25	50,0	50,0	100,0
		Total	50	100,0	100,0	
kladden	Valid	false	10	20,0	20,0	20,0
		true	40	80,0	80,0	100,0
		Total	50	100,0	100,0	
klippen	Valid	false	10	20,0	20,0	20,0
		true	40	80,0	80,0	100,0
		Total	50	100,0	100,0	
loef	Valid	false	24	48,0	48,0	48,0
		true	26	52,0	52,0	100,0
		Total	50	100,0	100,0	
metten	Valid	false	27	54,0	54,0	54,0
		true	23	46,0	46,0	100,0
		Total	50	100,0	100,0	
ornaat	Valid	false	16	32,0	32,0	32,0
		true	34	68,0	68,0	100,0
		Total	50	100,0	100,0	
vleet	Valid	false	18	36,0	36,0	36,0
		true	32	64,0	64,0	100,0
		Total	50	100,0	100,0	

Appendix III

Materials Experiment 3 Another version of the Sprenger experiment

List of used idioms, pictures and prime words.

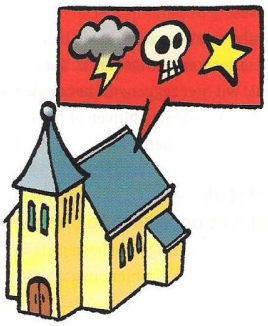
p1. iemand om de tuin leiden

prime word: computer

p2. van een mug een olifant maken

prime word:vlieg

1. vloeken in de kerk



related prime word: schelden
unrelated prime word: trein

2. de bloemetjes buiten zetten

related prime word: plantjes
unrelated prime word: hoes

3. de hond in de pot vinden

related prime word: blaffen
unrelated prime word: regen

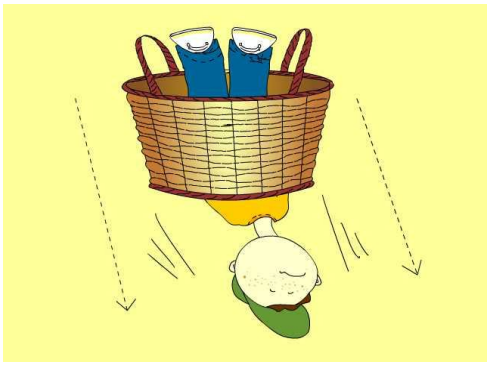
4. de kat uit de boom kijken

related prime word: poes
unrelated prime word: scherm

5. de kop in het zand steken

related prime word: hoofd
unrelated prime word: bus

6. door de mand vallen



related prime word: korf
unrelated prime word: boek

7. een geschenk uit de hemel

related prime word: gift
unrelated prime word: huis

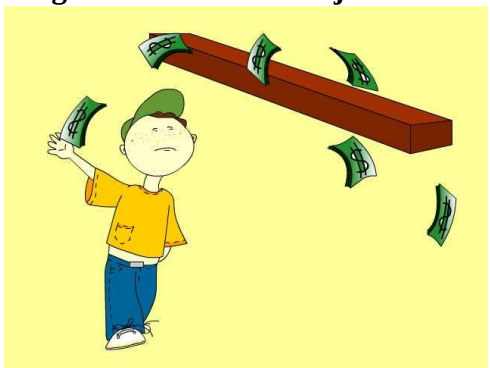
8. een heilig boontje

related prime word: peultje
unrelated prime word: schoen

9. een speld in een hooiberg zoeken

related prime word: draad
unrelated prime word: krant

10. geld over de balk smijten

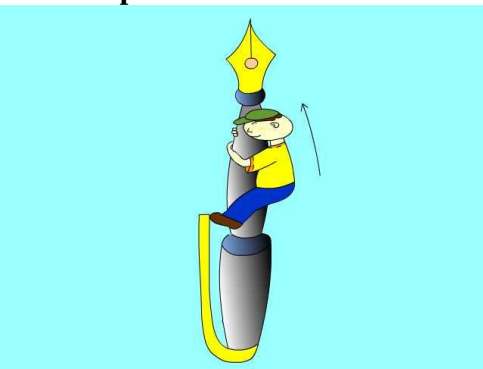


related prime word: plank
unrelated prime word: kers

11. het hart op de tong hebben

related prime word: gevoel
unrelated prime word: stoel

12. in de pen klimmen



related prime word: stift
unrelated prime word: brood

13. in de wolken zijn

related prime word: lucht
unrelated prime word: fiets

14. met het verkeerde been uit bed stappen

related prime word: teen
unrelated prime word: plein

15. naast je schoenen lopen

related prime word: pantoffels
unrelated prime word: geit

16. op de kast zitten

related prime word: lade
unrelated prime word: auto

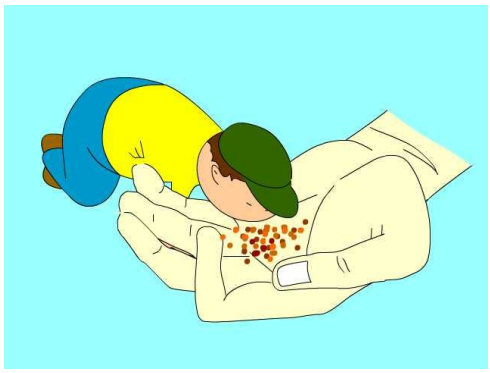
17. tegen de lamp lopen

related prime word: licht
unrelated prime word: kaas

18. uit de band springen

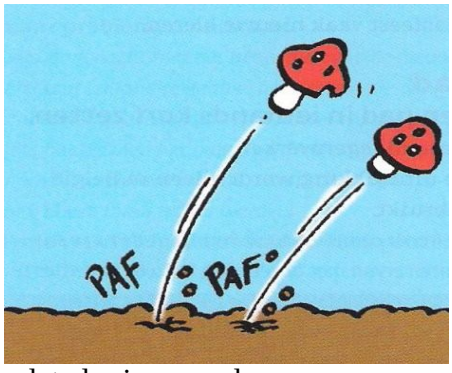
related prime word: rond
unrelated prime word: muur

19. uit de hand eten



related prime word: palm
unrelated prime word: flat

20. als paddestoelen uit de grond schieten.



related prime word: zwammen
unrelated prime word: pleister

Results Experiment 3 Another version of the Sprenger experiment

Statistics

Reactiontime wih unrelated prime

balk	N	Valid	19
		Missing	6
	Mean		2,43326
	Minimum		1,242
	Maximum		4,411
band	N	Valid	4
		Missing	21
	Mean		2,7745
	Minimum		0,963
	Maximum		5,143
been	N	Valid	19
		Missing	6
	Mean		2,33489
	Minimum		1,056
	Maximum		5,863
bloemetjes	N	Valid	18
		Missing	7
	Mean		2,08611
	Minimum		0,975
	Maximum		4,144
boontje	N	Valid	15
		Missing	10
	Mean		2,26507
	Minimum		1,195
	Maximum		4,481
geschenk	N	Valid	4
		Missing	21
	Mean		3,6945
	Minimum		2,774
	Maximum		4,435
hand	N	Valid	15
		Missing	10
	Mean		3,19613
	Minimum		1,23
	Maximum		5,816
hart	N	Valid	11
		Missing	14
	Mean		3,207
	Minimum		1,451
	Maximum		5,514
hond	N	Valid	14
		Missing	11
	Mean		2,71214
	Minimum		1,219
	Maximum		5,456
kast	N	Valid	6
		Missing	19
	Mean		3,022
	Minimum		1,88

Statistics

Reactiontime with related prime

balk	N	Valid	11
		Missing	14
	Mean		1,90055
	Minimum		1,091
	Maximum		3,401
band	N	Valid	6
		Missing	19
	Mean		3,2135
	Minimum		2,519
	Maximum		4,493
been	N	Valid	21
		Missing	4
	Mean		2,12024
	Minimum		1,126
	Maximum		4,086
bloemetjes	N	Valid	22
		Missing	3
	Mean		1,78795
	Minimum		0,847
	Maximum		3,065
boontje	N	Valid	19
		Missing	6
	Mean		2,12711
	Minimum		0,986
	Maximum		4,922
geschenk	N	Valid	6
		Missing	19
	Mean		2,71617
	Minimum		1,404
	Maximum		3,633
hand	N	Valid	10
		Missing	15
	Mean		2,6918
	Minimum		1,486
	Maximum		5,224
hart	N	Valid	17
		Missing	8
	Mean		2,92794
	Minimum		1,439
	Maximum		4,899
hond	N	Valid	7
		Missing	18
	Mean		2,80086
	Minimum		1,3
	Maximum		4,864
kast	N	Valid	11
		Missing	14
	Mean		3,61755
	Minimum		2,171

	Maximum		4,4
kat	N	Valid	20
		Missing	5
	Mean		2,80915
	Minimum		1,822
	Maximum		4,574
kop	N	Valid	19
		Missing	6
	Mean		2,14311
	Minimum		1,137
	Maximum		4,446
lamp	N	Valid	14
		Missing	11
	Mean		2,53886
	Minimum		0,835
	Maximum		4,051
mand	N	Valid	17
		Missing	8
	Mean		2,30794
	Minimum		1,068
	Maximum		4,62
paddestoel en	N	Valid	7
		Missing	18
	Mean		2,751
	Minimum		1,207
	Maximum		3,97
pen	N	Valid	9
		Missing	16
	Mean		2,55356
	Minimum		1,288
	Maximum		4,028
schoenen	N	Valid	15
		Missing	10
	Mean		2,36487
	Minimum		0,905
	Maximum		5,212
speld	N	Valid	17
		Missing	8
	Mean		2,61512
	Minimum		1,184
	Maximum		4,98
vloeken	N	Valid	10
		Missing	15
	Mean		2,3272
	Minimum		1,335
	Maximum		3,343
wolken	N	Valid	12
		Missing	13
	Mean		2,67467
	Minimum		0,882
	Maximum		5,317

	Maximum		5,537
kat	N	Valid	15
		Missing	10
	Mean		3,0956
	Minimum		1,242
	Maximum		5,712
kop	N	Valid	16
		Missing	9
	Mean		1,73163
	Minimum		0,812
	Maximum		3,239
lamp	N	Valid	17
		Missing	8
	Mean		1,78406
	Minimum		0,87
	Maximum		4,771
mand	N	Valid	17
		Missing	8
	Mean		2,36712
	Minimum		0,928
	Maximum		4,249
paddestoelen	N	Valid	5
		Missing	20
	Mean		3,0204
	Minimum		1,648
	Maximum		5,189
pen	N	Valid	15
		Missing	10
	Mean		2,57307
	Minimum		1,195
	Maximum		4,69
schoenen	N	Valid	18
		Missing	7
	Mean		2,05511
	Minimum		1,137
	Maximum		4,794
speld	N	Valid	22
		Missing	3
	Mean		2,59782
	Minimum		1,648
	Maximum		5,166
vloeken	N	Valid	8
		Missing	17
	Mean		2,4695
	Minimum		0,824
	Maximum		5,421
wolken	N	Valid	16
		Missing	9
	Mean		2,12919
	Minimum		0,94
	Maximum		4,562