Introduction

See cow. This is what is preserved of the stimulus sentence I can see a cow that is repeated by a two-year-old child in an experiment by Brown and Frase (Brown, 1973, p. 76), illustrating telegraphic speech. When we have limited resources, because of developmental constraints, or length limitations in old-fashioned telegrams or modern tweets, we resort to content words and mostly nouns and verbs as the main carriers of meaning. In studying language acquisition, we want to know how many words a language learner knows and can use, how these words are stored in and retrieved from the mental lexicon, and how often the language learner has encountered or will encounter the word in the everyday language environment.

Words, as part of linguistic systems, and lexical knowledge, as part of language users’ language proficiency, play an essential role in language use and language learning. Investigating this role requires language researchers to take careful account of the words they use in their research materials and the way they investigate the lexical skills of language users and language learners. The contributions in this handbook underline that lexis and lexical knowledge are not as simple as is sometimes suggested. In this chapter we will focus on a few of the considerations researchers have to make in designing and conducting their research on single-word items. When we refer to the language user in general terms, this includes (L1 and L2) language learners. In this chapter, we confine ourselves to the features and processing of content words, that is nouns, verbs, adjectives, and adverbs, and ignore the processing of function words. The divide between these two categories, content and function words, can be somewhat fuzzy, but these discussions fall outside of the scope of this chapter.

Our discussion of issues in researching single-word items firstly focuses on the many word characteristics that play a role in lexical processing, and therefore should ideally be considered when conducting any type of vocabulary-oriented research. However, as stated in the title, while this chapter focuses on single words (contrary to collocations and other multiword expressions), these single words derive their meaning and usability to a large extent from the connections they have with other words, and their level of interconnectedness: in this chapter they are single, but not unrelated. Therefore, in the second part of this chapter,
we will zoom in on tasks that can be used to target a special feature of knowledge of single-word items, namely, their place in the language user’s semantic network.

Critical Issues and Topics: Form and Meaning Characteristics

Theoretical Issues

Including single-word items in stimulus materials or tests of one’s research may initially seem quite straightforward. However, this impression is highly misleading. Content words in texts, stimulus sentences or vocabulary tests seldom represent unidimensional form-meaning combinations. A closer look at words as single linguistic units, and at the use of words in language interaction shows that words and word use are highly multidimensional or, in other words, have many facets that need to be taken into account when designing lexically oriented studies. Language users are sensitive to these word characteristics in processing lexical information, and therefore it is essential that researchers are aware of their possible effects. These characteristics or features not only relate to the form of the word, be it the acoustic or visual form, but also its meaning and its use. Multilingual language users find themselves in an even more complicated situation, because form, meaning and use of words in the different languages might interfere with each other, at all levels.

In this section we will introduce a number of word characteristics that may affect word processing and thus results of lexically oriented research. For practical reasons, we distinguish between characteristics relating to use, form and meaning, but obviously these facets are highly interconnected, as becomes evident when one takes into account derivational morphology of words. Our goal is to bring a number of important word characteristics to the readers’ attention, which may either facilitate or inhibit smooth word processing, and which are sometimes mediated by other factors, such as language modality. Some characteristics are inherent to the word form or meaning, such as morphological complexity (cf. “intra-lexical factors”; Laufer, 1997) and others are induced by the use of the word or the language user, such as word frequency and cognate status of words in a multilingual’s lexicon, respectively. Where possible, we will provide information on available databases targeting the various characteristics that are discussed. Our discussion of these issues is by no means comprehensive, but the reader will find additional sources in the other chapters of the book, or in Schmitt’s vocabulary research manual (Schmitt, 2010; see Further Reading).

Use, Form and Meaning Characteristics in Research

Word Use Characteristics

Word use characteristics play a prominent role in the processing of lexical items in everyday language use, but also, or especially, in the assessment of receptive lexical skills or verbal fluency in research. For words that we have frequently heard or seen before, it will be relatively easy to process the form and to comprehend the meaning. This makes word frequency an important feature in language processing and thus in lexical research. In most cases it needs to be taken into account when designing lexical materials for empirical studies (cf. Baayen, 2001; Balota, Yap, & Cortese, 2006; Breland, 1996). Word frequency in spoken and in written language are strongly correlated, and researchers, often for practical
reasons, use written word frequencies in their studies as these are more widely available. However, it will be more accurate to distinguish between the two types of word frequency and to choose the one that fits one’s research goals best. See, among others, Baayen, Feldman, & Schreuder (2006) for differential effects of spoken language and written language frequencies – as a written-to-spoken frequency ratio – on lexical decision and word naming latencies. At the same time, they showed that, across modality – spoken word frequency can be predictive of latencies in visual word recognition, with highly frequent spoken words leading to shorter latencies in both visual word recognition and naming. Remarkably, words that are far more frequent in written language than in spoken language showed longer latencies in both tasks (Baayen et al., 2006). An interesting alternative to written and spoken word frequencies is presented in work by Brysbaert, who studied the quality of a corpus based on subtitles. He showed that the frequencies based on subtitles are in many respects superior to frequencies based on other corpora, such as Kučera and Francis’s (1967) classical corpus of “present-day” American-English, CELEX (Baayen, Piepenbrock, & Van Rijn, 1993), and others (see Brysbaert & New, 2009). Word frequencies derived from subtitles are now available for a range of languages including American English, British English, Chinese, Dutch, German, Greek, Polish, and Spanish (see http://crr.ugent.be/programs-datasubtitle-frequencies).

Including word frequencies in one’s studies as they can be derived from large corpora serves at least two purposes. One is that infrequent words are expected to be more difficult and less well known than frequent words. This kind of information will predict item difficulty in, for example, the development of (adaptive) lexical knowledge tests (Breland, 1996). Breland (1996) found values of .70 and higher for the correlations between word difficulty and standard frequency indices from multiple corpora. Many vocabulary size tests are organized according to frequency bands from which the words in the corpora are sampled, such as Nation’s Vocabulary Size Test (Nation & Beglar, 2007; Coxhead, Nation, & Sim, 2014) or Nation’s more diagnostic Vocabulary Levels Test and its successors (Nation, 1983; Schmitt, Schmitt, & Clapham, 2001; Webb, Sasao, & Ballance, 2017). A second purpose for including word frequencies in one’s research is that these frequencies are considered to approximate the exposure an individual has had to the respective words. Frequent previous encounters and thus processing of a word most likely facilitates its processing upon new encounters, which may affect results in psycholinguistically oriented studies that use, for example, processing speed as measured by reaction times. These kinds of studies depend on subtle differences in reaction times between conditions, and poor control of word frequencies in the research design may confound research results. The extent to which word frequencies can be considered a valid approximation of an individual’s exposure to the words very much depends on the match between the source materials of the corpus from which the frequencies were derived and the individual’s acquaintance with such language materials. Obviously, word frequencies derived from a corpus of academic books will not be highly suitable for studies on language processing by primary school children. On the other hand, in many cases it will be difficult to find appropriate corpora for one’s research goals (see however, Anthony, this volume, for a list of available corpora). Furthermore, in using corpora, the absolute frequency may not be the only relevant parameter to consider: words can be highly frequent in certain text types or topical domains but virtually nonexistent in others. To account for this imbalance, the dispersion of words across subcorpora could provide valuable additional information (Gries, 2010). Using word frequencies presupposes a clear distinction between the units that count as a word. However, this might not always be that simple. Do we want to count the frequency
of specific form-meaning pairings, or just word form frequencies. For instance, the noun *work* as in *The work was done* will have a certain frequency, so will the verb *work* as in *We worked very hard*. The form *work* will have the combined frequency. In this example the part of speech, verb or noun, makes the difference, as the two word meanings are strongly related. This is very different in homonyms, such as *husky* referring to a type of dog and *husky* meaning hoarse. It will depend on the specific research questions whether we want to combine the two word frequencies or whether we are concerned about just one of them, especially in cases where there is a dominant, frequently used word meaning and a non-dominant, infrequently used one.

A final consideration regarding frequency concerns not just the frequency of the word as a whole, but of parts of the word, as encountered in written language. In studies on reading, highly frequent words may be easily processed because of the visual regularity of the letter pattern. The frequency of letter patterns can be a factor in speed of word recognition, and therefore the positional bigram frequency is often used as an index of the expected visual familiarity of the word. To calculate this measure, words are split up in all their bigrams (e.g., *mo-ot-th-he-er* for *mother*) and the frequency of the bigram at that particular position in a word is computed. For example, the bigram *mo* at the initial two positions in six-letter words has a frequency of 1,721, whereas its frequency would be 7 for the final two positions (Solso & Juel, 1980, using Kučera & Francis, 1967). The bigram frequencies can be summed and averaged according to word length (Ellis & Beaton, 1993). Assuming that language users are especially sensitive to unexpected patterns, researchers may want to include the lowest bigram frequency of a word as a special index in their analyses (Lemhöfer et al., 2008). This latter kind of index may also capture syllable boundaries in multisyllabic words, as bigrams surrounding a syllable boundary often show a relatively low bigram frequency (Balota et al., 2006). For example, a comparison of the bigram frequencies of *nt* as in *pint* vs. *tn* as in *fitness* in a Google corpus (Lydell, 2015) are very different: (rounded) $29.359 \times 10^9$ and $0.282 \times 10^9$, respectively.

**Word Form Characteristics**

The word form is the first to be encountered in listening or reading, and a main word form feature is word length. Longer words generally take longer to process. This is obvious in spoken language where words are produced and processed linearly in time. The more phonemes or syllables a word consists of, the more time it is likely to take to be processed, if all else remains the same. However, language users do not always need to hear the full sound string for adequate word identification. For instance, the English word */trespass/* will usually be identified at the */p/*, because no other English word that has the same initial sounds, */tresp/*, concurs with */trespass/* (Marslen-Wilson, 1987). At this so-called recognition point or uniqueness point, the word can be identified and selected, and the initially concurrent words can be inhibited; they are no longer competing as candidates for recognition (see the Cohort model, Marslen-Wilson, 1987). In written language, similar word length effects can be expected, although reading is not necessarily linear. Our eyes jump across the text lines, as is known from eye-tracking studies, and in a single fixation, multiple graphemes can be processed. The span of a fixation may thus cover full words, and this suggests that in a visual word recognition task one single fixation may be enough to recognize words of different lengths. This is in line with a study by Juhasz and Rayner (2003), in which participants read target words in a sentence and it was found that word length did not affect first fixation duration, nor fixation duration for words fixated only once. However,
the likelihood of refixating on a word increased with its length, as did total fixation time, which includes backtracking and refixations of the word. These findings suggest that longer words require refixations in order to be processed properly, which implies that word length is a factor to take into account in both visual and aural word recognition. This conclusion is restricted to language perception, i.e., word recognition. The production of words as in naming follows an incremental process which can start almost immediately, irrespective of the length of the word to be produced. In an experimental study of naming action pictures, Shao, Roelofs, and Meyer (2014) found only a weak correlation between naming latency and word length. One can imagine that the naming latency is largely determined by the interpretation of the picture and the lexical word selection, as a speaker can start naming based on the partial representation of the word form, for example the initial syllable (Shao et al., 2014).

A word form characteristic that partially interacts with word length, is a word’s morphological derivation. Some basic word forms reoccur in language in various derived forms and these derivations usually make the words longer, but in a structured way that will be familiar to the language user – though not necessarily to the second language learner. Moreover, the appearance of the base word in different derived forms could add to the word frequency of (at least) that part of the word. For example, *use* can be recognized in *usage*, *usable*, *usability*, *user*, and so on. These words belong to the family of *use*, and words with many of such family members are likely to be processed more easily. Family size of a word can be defined as the number of derivations and compounds of the base word involved; words with large families are associated with shorter response latencies in lexical decision tasks (Schreuder & Baayen, 1997; Lemhöfer et al., 2008), and this effect seems to exist independent of the cumulative frequency of the family members, that is a word’s family size was a better predictor of the lexical decision reaction time than the cumulative frequency of these family members. However, the family size effect did not show in a progressive demasking task, which taps into the early stages of perceptual word recognition, suggesting that the family size effect occurs at a later stage in visual processing of words (Schreuder & Baayen, 1997).

That various words, i.e., family members, can influence the recognition and processing of a particular single target word, underscores that single words are not processed in isolation. These kinds of effects are not only shown in the family size effect, but also in the so-called neighborhood size effect. Whereas family members are linguistically related through derivation or compounding, neighbors are close to each other at a more superficial level, that is, by virtue of their visual or auditory similarities. Characteristics of the neighborhood also impact the recognition of target words. Neighborhood size or density refers to the number of words that differ orthographically in just one grapheme from the target word, preserving letter position (Grainger, O’Regan, Jacobs, & Segui, 1989), or to the size of the set of words that differ minimally in sound from the target word (Vitevitch & Luce, 2016). A large neighborhood implies that there are many competitors that need to be excluded for unique word identification. And it is not just the number of competitors that we should be concerned about, but also their relative word frequencies. A target word with a highly frequent neighbor will take more time and effort to be recognized (Grainger et al., 1989). Conversely, in word production the opposite seems to happen. Words with sparse neighborhoods are produced more slowly and less accurately, suggesting that activation spreads back and forth, thus reinforcing the production of the word form selected to match the intended lemma. Sparse neighborhoods provide little reinforcement (Vitevitch & Luce, 2016). The effects of neighborhood size are mediated by neighborhood spread,
that is the distribution across the word of the replaceable graphemes or phonemes that form the neighborhood. For example, three letter word *peg* can have a grapheme replacement at all three positions: *beg*, *pig*, and *pen*, and thus has a P-metric of 3; for *hen* no replacement of the medial grapheme is possible, only at the initial and final position (*den* and *hem*, respectively) and thus its P-metric is 2 (Vitevitch, 2007). Vitevitch found – among other things – that in an auditory lexical decision task words with metric P = 2 were recognized more quickly than those with metric P = 3.

**Word Meaning Characteristics**

Characteristics of lexical meaning may affect the processing of words as well, especially when these characteristics are not represented in the word form per se. Transparent compounds (*travel agency*) and derivations (*friendliness*) may be long and less frequent, but the constituting parts may still be processed relatively easily, and thus the multi-morphemic word may not cause serious problems to the language learner. Still, some mono-morphemic words have features that make them harder to process than others. One such feature is abstractness or lack of concreteness. Abstract words (e.g., *virtue* and *regret*) are, generally speaking, harder to learn than concrete words (e.g., *needle* and *pirate*) in paired-associate learning tasks (cf. De Groot, 2006; De Groot & Keijzer, 2000). The differences are assumed to be related to differences in mental representations of concrete and abstract words; concrete words are supposed to have at least a visual and a verbal representation, whereas abstract words may only have a verbal one (De Groot, 2006). The dimension of concrete to abstract is highly correlated with the related dimension of imageability, with some scholars using one for the other. Balota et al. (2006) define concrete words as words that can be the object of a sense verb, such as *touch*, *see*, and *hear*, which would make *music* a concrete word; they also state that imageability usually is operationalized by gathering the subjective ratings of participants. Due to the high correlations between the two dimensions, researchers often do not make a distinction between them (Balota et al., 2006; De Groot & Keijzer, 2000). Effects of abstractness or imageability have been shown in lexical decision and naming tasks (Balota et al., 2006). In recent years, databases with concreteness and imageability ratings have become available for research purposes, such as Schock et al.’s (2012) imageability estimates for 3,000 disyllabic words, and Cortese and Fugett’s (2004) imageability ratings for 3,000 monosyllabic words. Brysbaert and colleagues collected concreteness ratings for 40,000 words, including some two-word expressions (Brysbaert, Warriner, & Kuperman, 2014). Concreteness and frequency can coincide, but also go in opposite directions. A relatively frequent word such as *although* belonged to the least concrete words, whereas a less frequent words such as *walrus* was among the most concrete words.

Matters get more complicated, when the relationship between lexical form and meaning is not one-to-one, but one-to-multiple as is often the case in natural languages. This feature is captured in meaningfulness, which often quite literally refers to the number of meanings a word represents. Meaningfulness of stimulus words can be assessed straightforwardly as the number of entries in a dictionary (Balota et al., 2006), but more advanced approaches are used by, among others, Baayen and colleagues (2006). They counted the number of synsets a word is listed in; synsets being a set of semantically related words according to Wordnet (Miller, Beckwith, Fellbaum, Gross, & Miller, 1990). Words with high meaningfulness ratings show a facilitatory effect in lexical decisions and naming, but when semantic decision-making gets involved in the response, as is the case in semantic categorization or semantic
relatedness judgment, “meaningful” words show inhibitory effects, probably indicating that the different meanings are in competition (Balota et al., 2006).

Interlanguage Effects

As mentioned earlier, in research on multilingual language users, features of words in the different languages interact with each other. This concerns form as well as meaning features. Language learners may benefit from similarities between languages, for example because of a common ancestor of the languages involved, such as Greek and Latin for many western languages. Dutch medicijn, coming from Latin (res) medicina, will not cause much problems for English, German, French, or Spanish learners of Dutch, as the translation in their first language is medicine, Medizin, médecine, medicina, respectively, whereas the Dutch synonym with a partial Gothic origin geneesmiddel will be much harder to recognize and to learn. These forms of overlap are helpful to the language learner, but can also lead to unexpected and maybe unwanted effects in psycholinguistic experiments. Both cognates – words that are similar both in form and meaning across two languages – and “false friends” – which are similar in form but misleadingly not in meaning (e.g., French/English: attendre/attend and plus/plus) – influence the processing of words in a target language. In a multilingual language learning setting, both the target word and its translation equivalent bring their characteristics to the table (cf., Dijkstra, Timmermans, & Schriefers, 2000). As a consequence, word frequency or family size in one language may affect processing speed of words in the other (Van Heuven, Dijkstra, & Grainger, 1998).

Word Use, Form and Meaning Characteristics: Concluding Remarks

Whether one wants to investigate effects of the aforementioned word features on language acquisition or language processing, or whether one wants to control for them or not, largely depends on the research context, research design, and resources. In most lexically oriented studies, it is important to pay careful attention to these characteristics in the compilation of the testing materials. An important distinction may be whether the research is geared towards auditory or visual recognition or towards semantic classification or interpretation. In the former case, form characteristics may be more important to control for, and in the latter meaning features may play a larger (additional) role. In the next section, we will zoom in on another characteristic of individual words, namely the position they occupy in the semantic network, connected to other words via various semantic relations.

Critical Issues and Topics: Connections in a Semantic Network

Theoretical Issues

Even when focusing on single words, it is important to note that knowledge of such items is never isolated: all words are connected in the mental lexicon through various types of semantic, phonological, and orthographic ties. These links form part of our knowledge of individual lexical items, which is a topic of interest in its own right: Which types of ties are prominent? In what order do they develop? How do they affect language use? In this section, we will focus on semantic links specifically, as they form part of individual words’ meanings. We will discuss the different types of semantic ties that exist within the semantic network and various tasks used to measure knowledge and activation of these links.
Importance of Semantic Relations: Language Acquisition and Processing

When studying single-word items, what is the relevance of the semantic ties they hold with other words? Firstly, as with all other vocabulary knowledge, the semantic networks surrounding words need to be acquired. Researchers may therefore be interested in how the network develops and which relations are important throughout acquisition. For example, a developmental path from a focus on context-dependent knowledge, i.e., contextual semantic relations such as squirrel – forest, to the gradual abstraction towards more context-independent knowledge, i.e., categorical relations such as squirrel – animal, has been identified for L1 children (cf. Nelson, 2007, inter alia).

Secondly, many studies have shown that the processing of single-word items is affected by the processing of other items within their networks. During language use, the activation of individual words encountered in discourse spreads through the network to related words, which may then be accessed more quickly for comprehension and production during the immediately following discourse. This effect of spreading activation (Collins & Loftus, 1975) can be studied by means of semantic priming tasks which may take on various forms (McNamara, 2005).

Finally, the degree to which network knowledge of lexical items is developed may in turn affect other language skills. For example, a number of studies have found that there is a relation between reading comprehension and various measures of semantic network quality (cf. Betjemann & Keenan, 2008; Swart et al., 2017; Tannenbaum, Torgesen, & Wagner, 2006).

Types of Relations

The semantic network has been studied extensively and a multitude of relations that may exist between words in the network have been identified. Although many different classifications have been offered, most are centered around a dichotomy that distinguishes relations based on shared intrinsic characteristics, on the one hand, and concepts related through context, on the other hand. Table 32.1 provides examples and the most commonly used overarching terms. Typically, concepts considered to be related through shared characteristics include category relations such as sub-, super-, and coordination, and often also synonyms, antonyms, and part-whole relations. Contextual relations may include a wide variety of connections, for example between objects and their locations, typical events at which they occur, and instruments that may be used to manipulate them. Finally, the dichotomy may also be conceptualized as a continuum from context-independent relations to context-dependent relations, where some types of connections may be more in the middle of the scale, such as intrinsic characteristics or behaviors of objects, animals, or people.

In addition to these semantic classifications of word relations, there are two other important types of relations distinguished in the literature, which may or may not overlap with the semantic types discussed previously. Most notably, associative relationships are those identified through large-scale studies of people’s word association responses. An example of a highly common associative stimulus – response pair is blood – red. Associative links may also include form-based connections, but these are typically far less common than meaning-based associations (De Deyne & Storms, 2008). Secondly, words may be related through collocation, i.e., common co-occurrence in speech and text (or in the outside world), such as pencil – paper (example from COCA, Davies, 2008). All three types of relations (semantic, associative, and collocation) may overlap, as is the case in the word pair salt – pepper.
<table>
<thead>
<tr>
<th>The paradigmatic/syntagmatic distinction (i.e., Cruse, 2000): a syntactically defined categorization system</th>
<th>The thematic/taxonomic distinction (i.e., Estes, Golonka, &amp; Jones, 2011): an object-focused categorization system</th>
<th>The context-independent to context-dependent continuum (from Spätzgens &amp; Schoonen, submitted; based on De Deyne &amp; Storms, 2008; Cremer, Dinghoff, de Beer, &amp; Schoonen, 2011): a semantically focused categorization system</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Paradigmatic</strong> (between words of the same syntactic category): including coordinates, superordinates, subordinates, partonyms, and synonyms</td>
<td><strong>Taxonomic</strong> (hierarchically related objects sharing many features): coordinates, superordinates, subordinates.</td>
<td><strong>Taxonomic</strong> (category relations) including coordinates, superordinates, subordinates, and synonyms.</td>
</tr>
<tr>
<td><strong>Syntagmatic</strong> (between words of a different syntactic category): typically, syntagmatic relations are not specifically defined other than by the difference in syntactic category</td>
<td><strong>Thematic</strong> (concepts playing complementary roles in the same event), including relations that are spatial: jungle – bird; temporal: summer – holiday; causal: wind – erosion; functional: fork – knife; possessive: police – badge; and productive: cow – milk.</td>
<td><strong>Feature</strong> (necessary or prototypical characteristics), including physical characteristics: banana – yellow; internal characteristics: snail – slow; behavior: dog – bark; functions: cot – sleep; and partonyms.</td>
</tr>
<tr>
<td><strong>Situational</strong> (concepts related through possible contexts), including co-occurrence: princess – hat; contexts: ocean – rocks; actions: bakery – choose; context-dependent attributes: car – blue.</td>
<td><strong>Subjective</strong> (dependent on personal context), including emotion/evaluation: cheese – eew; personal: hair dresser – my mother.</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Overlap between the classification systems is not always complete, since definitions are not similarly detailed.
However, they may also exist separately, such as the pair *dolphin – elephant*, which are both mammals and therefore semantically related, but which are neither associates nor collocates.

**Operationalizing the Semantic Network**

In this section, we will discuss a number of methodologies commonly used to study the semantic network, and the types of results these methodologies yield.

**Free Word Association Tasks**

As was discussed earlier, word association studies provide us with information on commonly linked word pairs based on people’s spontaneous responses. At the individual level however, word associations can also be used to uncover speakers’ preferences for certain types of semantic links. The basics of the procedure are very simple: participants are provided with a stimulus word and asked to respond with the first word, or first few words, that come to mind. The task can be performed either orally or in writing.

Traditionally, only a single response is required in most word association studies. However, eliciting multiple responses appears to be more compatible with the fact that any given word likely has many different connections to other words in an individual’s mental lexicon. Furthermore, recent research by De Deyne and Storms (2008) has shown that different response patterns occur when three responses are elicited. For example, they demonstrated that secondary and tertiary responses were less likely to be taxonomically related to their stimuli than first responses. Therefore, eliciting multiple responses appears to provide a more comprehensive view of a word’s semantic network and of individuals’ associative behavior.

Similar to the different types of classifications of semantic relations in general, there are many different coding schemes that have been used to classify word association responses, and these will of course affect the conclusions that can be drawn. Most early studies used a word class-based definition of the syntagmatic/paradigmatic distinction (see Table 32.1). A number of recent studies have provided more detailed classification schemes, defining not only specific context-independent relations but also context-dependent relations and sometimes form-related types (Cremer et al., 2011; De Deyne & Storms, 2008; Fitzpatrick, 2006; Spätgens & Schoonen, submitted). The choice for a specific coding scheme will of course affect the potential outcomes of any word association study: using a classification system that is detailed enough to distinguish the categories relevant to the research questions is therefore vital.

Besides the semantic classification of stimulus-response pairs, researchers have also used databases as a reference point to weigh participants’ associations against (Fitzpatrick, Playfoot, Wray, & Wright, 2013; Swart et al., 2017; Zareva & Wolter, 2012). In this way, the degree of similarity between, for example, an L1 child’s or an L2 adult’s associative behavior can be compared to that of L1 adults, to study and compare the development of the semantic network in these groups.

In sum, the relatively spontaneous word association task can provide us with a qualitative mapping of individual language users’ semantic networks, their potential preferences for certain kinds of semantic relations, and the way they compare to other groups of language users in this regard. Due to the spontaneous nature of the task however, it is difficult to use it as an assessment of semantic knowledge. Rather, it expresses the preferences or relative ease of access to certain kinds of semantic relations that language users may exhibit.
Tasks Assessing Declarative Knowledge of Semantic Relations

Various tools aimed at the assessment of knowledge of semantic relations have been developed, focusing on the ability to identify relations. In these tasks, participants are required to select semantically related words, distinguishing them from unrelated or differently related distractor words. For example, the Deep Vocabulary Knowledge (DVK) task developed by Qian (2002) based on the word associates format by Read (1993) requires participants to select synonyms and collocates for adjective stimuli. Similarly, a standardized task for knowledge of word relations is the CELF Word Classes subtest (Semel, Wiig, & Secord, 2003), which has been normed for children from 5 to 18 years of age. In this task, participants are required to select two related words from sets of three or four words. Semantic relations between these words include many different types, such as functional relations, contextual relations, and coordination, meaning that the task is suited for testing general knowledge, but not for targeting specific semantic relations or assessing the ability to distinguish them from other types of semantic relations.

The latter option has been explored by Schoonen and Verhallen (2008), who based their test on the finding that context-independent semantic knowledge develops later than context-dependent semantic knowledge. In their Word Association Task (WAT), developed in Dutch for children between 9 and 12 years of age, participants are required to select three target words that are context-independently related to a stimulus, with three unrelated or context-dependently related distractor words. Here then, not only the awareness of a semantic relation is tested, but also the ability to distinguish relations that hold true independent of context (i.e., *dog – poodle*) from context relations (*dog – jump*).

These tasks provide a measure of declarative knowledge of semantic relations, which can target knowledge of certain types of relations in a more principled way, with the potential of eliciting knowledge that is available but not accessed in the more spontaneous free association task.

Categorization Tasks

Differences in preferences for types of semantic organization can be tested using categorization tasks, in which participants are presented with a set of words or pictures and required to select the items that are related to a target. The set typically contains items that bear different types of semantic relations to the target, resulting in a measure of preference for different types of semantic relations.

For example, Mirman and Graziano (2012) tested adults (mean age 66) in a task which first presented a target picture (e.g., *finger*) followed by a thematically related (*ring*) and a taxonomically related picture (*thumb*) and required participants to choose the picture that “goes best” with a target picture. They found that most of their participants did not have a consistent preference for either type of semantic relation, but that participants who did show a taxonomic preference, also showed more competition of taxonomic items in a spoken word recognition task, as indicated by their fixations of taxonomic distractors within the array of visual response options. Conversely, Lin and Murphy (2001) did find many of their young adult participants exhibited consistent taxonomic or thematic preferences in a similar task using words instead of pictures, with a majority (62%) preferring thematic relations. This may be an indication of differences in organization preferences across the different age groups, but the use of different stimuli (pictures vs. words) may also play a role.
In contrast to tasks forcing participants to distinguish a certain type of semantic relation from unrelated or differently related word pairs, the categorization task allows for a more spontaneous measure of semantic preference, similar to word association tasks. However, by juxtaposing specific semantic relations, these preferences can be studied in a more principled and controlled fashion. An important methodological consideration in these types of tasks is the relationship strength between the target and the different answer options. A pretest establishing this strength is therefore a necessity and may be done using rating studies. For example, Mirman and Graziano (2012) asked a separate group of participants to indicate which two words in their triad stimuli belonged to the same category or were likely to occur in the same situation or scene, for taxonomic and thematic ratings, respectively.

Category Generation

In category generation tasks, knowledge of semantic relations within specific domains is targeted. Participants are asked to produce as many items as possible to fit a certain category, such as “tools”, “animals” or “occupations”. It can also be seen as a guided, instead of free, association task. Again, a normed and standardized version of such a task is available in the CELF test battery, namely the Word Association subtest (Semel et al., 2003). By setting a time limit (60s in the case of the CELF-WA subtest, but for example only 10s in a study by Tannenbaum et al., 2006), the response can be speeded, meaning that a measure of semantic fluency can be derived. Scores on this type of task have been found to be associated with reading comprehension skill (Nation & Snowling, 1998; Tannenbaum et al., 2006).

Different types of categories may be used to assess different types of semantic knowledge. Peña, Bedore, and Zlatic-Giunta (2002), for example, compared fluency in taxonomic categories such as “animals” with so-called slot-filler categories, such as “animals one would find at a farm”. The latter type of category is supported by co-occurrence in experience, meaning that it is more context-dependent than purely taxonomic categories. Peña et al. studied bilingual children’s ability to generate items within each of these types of categories in both of their languages and at different ages. They found that the bilinguals produced similar numbers of items across categories in each of their languages, but that older children produced more taxonomic items than slot-filler items, compared to a younger age group which produced similar numbers of concepts in each category.

The category generation task is thus an example of a guided production task that measures the fluency of access and availability of certain types of semantic relations, essentially by constraining the word association task. Careful definition of goal categories can allow for detailed comparisons of various semantic domains.

Semantic Priming

Online measures of the spreading activation that is occurring in individuals’ semantic networks during language use are also available, in the form of a range of different types of semantic priming tasks. In these tasks, participants are traditionally presented with single-word stimuli and required to provide a response, which is timed. By including related and unrelated prime – target pairs in the stimulus lists, for example animal – dog and bicycle – dog, the effect of relatedness on response time can be measured. If participants are faster to respond to targets preceded by related primes compared to targets preceded by unrelated primes, a priming effect is present. Provided the experiment is set up to elicit automatic rather than strategic processing, the effect is considered to be due to spreading activation.
Researching Single-Word Items

within the semantic network (cf., McNamara, 2005). In addition to single-word studies, the principle of semantic priming can be extended to the sentence level, for example by measuring reading times to related words in sentences using self-paced reading or eye tracking (Camblin, Gordon, & Swaab, 2007). Finally, some studies include neurological measures such as EEG (Landi & Perfetti, 2007) or fMRI (Sachs et al., 2008) to study the neural correlates of different types of semantic effects.

In all, an extensive number of studies have been conducted, demonstrating priming effects for many types of context-independent and context-dependent relations both with and without the presence of an associative relation. Many excellent reviews on different types of priming effects are available, for example on differences between associative, categorical, and thematic relations (Jones & Estes, 2012), semantic priming with and without association (Lucas, 2000), and priming effects in discourse (Lédonu, Camblin, Swaab, & Gordon, 2006).

Semantic priming studies thus allow for the study of a range of semantic relations and their effects in both the aural and written modality, with single words and in (short) discourse contexts. Careful construction of both the related and unrelated word pairs is crucial to allow for a fair comparison, which involves controlling for a wide range of factors, such as word frequency, length, and phonological similarities, as discussed earlier. Some authors perform norming studies to ensure the relationship strength between different conditions is similar, especially when comparing different types of semantic relations (Hare, Jones, Thomson, Kelly, & McRae, 2009; Spätzgens & Schoonen, 2018). Depending on the goal of the study, the semantic relation in focus may also need to be carefully disentangled from potential associative relations. English word association data are currently available from various sources, e.g., the Edinburgh Associative Thesaurus (Kiss, Armstrong, Milroy, & Piper, 1973), the University of South Florida association norms (Nelson, McEvoy, & Schreiber, 2004) and notably, from the multiple association task used in the Small World of Words project (De Deyne & Storms, 2008; English data available from De Deyne via https://ppw.kuleuven.be/apps/concat/simon/?datasets). McNamara (2005) provides an excellent reference work for the many methodological considerations in priming studies. An important note here is that researchers have found that although priming effects are generally very robust at the group level, they may be less reliable within individuals. Especially when automatic processing is tapped, test-retest reliability may be low within participants, suggesting that such priming effects may be relatively unstable (Stolz, Besner, & Carr, 2005; Yap, Hutchison, & Tan, 2016).

Operationalizing the Semantic Network: Concluding Remarks

In sum, a host of tasks exist that can be used to measure semantic knowledge, each with their own specific characteristics and targeting knowledge of semantic relations in more or less guided ways, assessing both active and passive knowledge, both online and offline. Previous studies have shown that the semantic network evolves during language development, both in the L1 and the L2, and that differences in network structure and knowledge of semantic relations are associated with other aspects of language proficiency.

Future Directions

It will be obvious that word knowledge is multifaceted and that our language processing is affected by all these facets to a smaller or larger extent. It is virtually impossible to count with all these facets at the same time. In that respect we should conclude that each word is unique, in its form, its meaning, frequency, associations, and spot in the semantic network.
Nevertheless, we should strive for reasonable comparability when we select word stimuli for the different conditions in our lexical (training) experiments. Therefore, it is necessary that we continue our investigations into the relevant facets of single-word items and how they are interrelated and affect language processing, both language understanding and language production. Here we have focused on single-word items, but these cannot be completely isolated from their linguistic contexts, that is, the multiple word phrases they occur in, such as collocations and idiomatic expressions. These lines of research should be integrated.

Furthermore, most of the research reviewed discusses the effects of word characteristics at the participant group level, often young adults or language learning children. However, it is important to take this a step further to the level of the individual language user. How does an individual’s preference for certain semantic relations affect his or her reading fluency? Or, how does one’s sensitivity to priming relate to one’s listening skills? Is a dense semantic network beneficial to fluent language production? These kinds of questions have so far only been scarcely touched upon. But then, addressing these questions requires a very good understanding of what facets of single (and multiple) word expressions need to be taken into account. Fortunately, more and more information comes available about the facets that really matter in lexical and semantic tasks, and also more corpora and lexical databases are compiled and made available for research purposes. It only takes a few clicks to find out that Brown’s two-year-old’s *cow* has a word frequency of 25.51 per million and occurs in 8.68% of the more than 8,000 subtitled films (cf. Brysbaert & New, 2009) and *milk* is its strongest associate.

**Further Reading**


Schmitt provides an extensive manual for vocabulary research, including many references to resources.


Balota et al. conducted detailed research of factors involved in (visual) word recognition. They provide definitions and operationalizations of those factors. See also Lemhöfer et al. (2008) for another study on this topic.


McNamara is an excellent introduction to (general) issues concerning semantic priming. More specific reviews of priming research can be found in Jones & Estes (2012), Lucas (2000), and Ledoux et al. (2006).

**Related Topics**

Different aspects of vocabulary knowledge, the mental lexicon, vocabulary and proficiency, processing single-word and multiword items, measuring depth of vocabulary knowledge, sensitive measures of vocabulary knowledge, key issues in measuring vocabulary knowledge, resources for researching vocabulary
References


Tessa Spärgens and Rob Schoonen


