Comparison of Different Lexical Resources With Respect to the Tip-of-the-Tongue Problem

Michael Zock, Chris Biemann

Comparison of Different Lexical Resources With Respect to the *Tip-of-the-Tongue* Problem

Michael Zock[1], Chris Biemann[2]

[1]*AMU (LIF, CNRS)163 Avenue de Luminy, 13288 Marseille, France*

[2]*Language Technology Group, Vogt-Kölln-Straße 30, 22527 Hamburg, Germany*

michael.zock@lif.univ-mrs.fr, biemann@informatik.uni-hamburg.de

Abstract

Language production is largely a matter of words which, in the case of access problems, can be searched for in an external resource (lexicon, thesaurus). When accessing the resource, the user provides her momentarily available knowledge concerning the target and the resource-powered system responds with the best guess(es) it can make given this input. As tip-of-the-tongue studies have shown, people always have some knowledge concerning the target (meaning fragments, number of syllables, ...) even if its precise or complete form is eluding them. We will show here how to tap on this knowledge to build a resource likely to help authors (speakers/writers) to overcome the Tip-of-the-Tongue (ToT) problem. Yet, before doing so we need a better understanding of the various kinds of knowledge people have when looking for a word. To this end, we asked crowd workers to provide some cues to describe a given target and to specify then how each one of them relates to it, in the hope that this could help others to find the elusive word. Next, we checked how well a given search strategy worked when being applied to differently built lexical networks. The results showed quite dramatic differences, which is not really surprising. After all, different networks are built for different purposes; hence each one of them is more or less well suited for a given task. What was more surprising though is the fact that the relational information given by the users did not allow us to find the elusive

Manuscript received 5 April 2020; Revised version received 28 May 2020; Accepted 5 June 2020

Journal of Cognitive Science 21-2:193-252, 2020

word in WordNet more easily than without relying on this information.

Keywords: *word access, tip of the tongue problem, indexing, knowledge states, meta-knowledge, mental lexicon, navigation, lexical networks*

1. The problem : word access in language production[1]

Language production (speaking, writing) consists largely in choosing and combining words that implies, of course, that we are able to retrieve them. To achieve this goal we can ask another person, rely solely on our brain, or resort to an external resource (dictionary). Obviously, the processes involved in each case are not at all the same.

In spontaneous discourse (on-line processing) word access is fast, automatic, and unconscious. During off-line processing (dictionary look-up) it is slow, controlled, and deliberate.[2] Obviously, words in books and our brains are not quite the same (Zock, 2015). In the first case (dictionaries) they are holistic entities, meaning and form being stored next to each other (Saussure, 1916). In the human brain (mental lexicon) they are decomposed. A word will reveal its form only if all its components (conceptual, syntactic, phonological) have been activated.

Our concern here is not really the human brain or the strategies people use to access the mental lexicon. We are more concerned with the qualities an external resource (electronic dictionary) must have to allow humans to reach naturally and quickly their goal, find the word they are looking for[3].

1 This is the revised version of a paper presented at the last CogALex workshop : https://sites.google.com/site/cogalex2016/home

2 Psychologists (Posner and Snyder, 1975) draw a clear line between automatic and attentional processes. The former are fast, parallel, and mandatory. They do not tax memory, neither do they interfere with other tasks, and they are not available to introspection (consciousness). On the other hand, attentional processes are slow, serial and they can be observed and controlled, i.e., stopped. Hence they tax memory, they are likely to interfere with other tasks, and their (intermediate) results may be accessible to our awareness, i.e., consciousness.

3 Being in variable cognitive states people have different needs, requiring different

More precisely, our goal is interactive lexical access, that is, finding a word with the help of a computer in an external resource. In sum, we consider word-finding as a dialogue between a human and a machine, both having some (partial) knowledge concerning words.

Word-finding can be considered as successful if the user has managed to reduce the entire set of words (stored in the lexicon), to one, the target [4]. Given the fact that we do this naturally, many times a day, and without consulting an external resource, one may wonder how the 'system' allowing for this is organized (neuronal architecture, information flow). The performance of the human brain is all the more remarkable as discourse is fast[5] and the lexicon is quite large. If an author had a vocabulary of 60-100,000 words[6], then he would have about 3-500 milliseconds to locate and retrieve a word[7]. Despite these constraints, humans succeed quite well most of the time, at least in their mother tongue. In addition, they make very few mistakes, according to psychologists (Butterworth, 1992; Levelt, 1989; Hotopf, 1983; Garnham, et al., 1982;), only one or two out of 1000.

---

kinds of resources : one for access based on meaning, another for access based on sound (rhymes), still another for access based on relations (associations)… Yet, most of the time people have access to information coming from various levels, i.e., that is, they have access to different retrieval patterns, which is why we should combine our resources in such a way as to allow for the optimal exploitation of (all) the available fragments of the different levels.

4 The assumption being that 'words' are 'stored' in a dictionary, assumption which cannot be taken for granted in the case of the mental lexicon (human brain). Note also that not all used words are stored. Just think of numbers, 357 or 1.586, complex noun phrases like the famous German word 'Donau-dampf-schiffarts-gesellschaft', or the possible combinations, i.e. 'words' in an agglutinative language like Turkish: 'ev' (house), 'evler (houses), 'evlerimiz' (our houses). According to (Hankamer, 1989), speakers of Turkish produce new words in almost every sentence, whereas speakers of English rarely do so.

5 According to Deese (1984), the speech rate in ordinary discourse is about 100-200 words per minute, that is, 2-3 words per second.

6 (Miller, 1991 ; Levelt, 1989 ; Oldfield, 1963)

7 Actually, humans seem to be able to achieve this much faster, getting from syntax to phonology in 40 milliseconds (van Turennout, et al. 1998).

This being said, human performance is not always perfect. People do make mistakes (Cutler, 1982, Fromkin, 1980), and their flow of words (hence, possibly also, their thoughts) may be disrupted, momentarily, yielding hesitations (Maclay and Osgood, 1959) or for good, discourse stops (Rastle and Burke, 1996). Not being able to access the desired word on the fly (instant recall), authors have to make a deliberate effort to find it. This is when they start to consider using a dictionary.

Alas, most dictionaries are not well suited for this task. Having been built with the reader in mind, alphabetically organized dictionaries are only of limited use for language producers. While they can be used for spell-checking, retrieval of grammatical information (e.g., gender in languages like German or French) or the definition of a word[8], they are of little use if the goal is to find a word on the basis of conceptual input (meaning, definition), the most frequent situation. Indeed, how to query such a dictionary in order to find the right word (form) expressing the notion of a 'large mammal living in Africa endowed with a trunk, and two large ivory tusks'?

Obviously, readers and writers have different needs (Humblé, 2001)[9], and while both provide words as input, they clearly pursue different goals. Readers start from word forms in the hope to get meanings, while authors start from meanings, their fragments, topical categories (food) or target-related words (associations, co-occurrences) in the hope to find the elusive word. We will be concerned here with this latter kind of search.

There are two major access modes, one automatic, and the other

8 This can be useful even in language production, for example, to avoid repetitions, or to check and highlight the differences between two words, say, caiman/crocodile or loneliness/solitude.

9 For excellent readers or overviews concerning lexicography, see (Durkin, 2015; Ostermann, 2015; Hanks, 2013; Svensén, 2009 ; Atkins & Rundell, 2008; Fontenelle, 2008; van Sterkenburg, 2003; Landau, 2001).

deliberate. The former relies solely on our brain when speaking[10] or writing (on-line processing), whereas the latter uses an additional, external resource, a paper or electronic dictionary / thesaurus (off-line processing). We resort to this second strategy only if spontaneous access fails. Alas, as mentioned already, most dictionaries are not very well suited for this task (see Section 3). Yet, in order to build such a dictionary and to make it work[11], we need to have a clear idea concerning a number of issues:

(a) what is a word (di Sciullo & Williams, 1987; Grefenstette & Tapanainen, 1994), or, what counts as lexical entry (hot, dog, hot-dog)?
(b) How many words and what kinds of words does a particular group of people know (Brysbaert et al. 2016 ; Zechmeister et al.1993)?
(c) Which words are particularly prone to yield word access problems (names, abstract words, technical terms, words having a specific phonological features)?
(d) What are typical retrieval patterns?
(e) What words to include in the resource? Are named entities good candidates (Sekine & Nobata, 2004 ; Nadeau & Sekine, 2007)?

Yet this is not all, we also need to address issues like the following: index creation, i.e., organization of the data[12], input (query), output, determination of search space, navigation .... More precisely,

---

10 For books, see (Bonin, 2004, Jescheniak 2002, Stemberger, 1985). For papers describing major systems or their components, see (Dell et al.2014 ; Goldrick, 2014; Griffin & Ferreira, 2011; Bock & Griffin, 2002; Meyer, 2002; Levelt, 2001; Levelt, et al. 1999 ; Schriefers et al., 1990; Butterworth, 1989). For neuronal accounts, see (Kemmerer, 2015; Indefrey & Levelt, 2000 ; Pulvermüller, 1999 and 2002).

11 Beware that our goal is not the creation of a dictionary, but the building of an index whose purpose is to help finding the intended word in the lexical resource. Even if we combined different dictionaries, we do not create a new resource, we only increase the power and flexibility of the access modes.

12 If texts are the territory, then the index is the map revealing how words (i.e., the lexical entries listed in a dictionary) are related. Searching words in a dictionary without an index is like exploring a city without a map.

(a) what information shall the user provide as input to allow the resource to guess the elusive word: one word, several words, a hint signaling his goal? (See also Section 4, meta-knowledge.) Obviously, not all query terms are alike in terms of quality[13]. Also, how to describe, i.e., what terms to use to tell the system the meaning of the specific word we are looking for? Imagine any of the following targets : 'junta, entropy, justice, black hole'? This is relevant if ever you use a reverse dictionary.

(b) How to avoid misunderstandings? Having received 'mouse' as input (query), the system needs to know whether you were referring to the *rodent* or the *computer device*, as in the absence of this information it may present the wrong set of candidates.

(c) Which words to show as output, and how to present them: in alphabetical order, as a ranked list, as categorial trees, in mixed order? Since a query (input), may trigger many outputs we may get drowned under a huge flat list[14] unless the resource builder has organized it according to some principle making sense to the dictionary user. Yet, there are many ways to organize data[15], semantically, topically (Roget, 1852), relationally (Miller, 1990), alphabetically, by frequency, etc. The latter two are definitely not very satisfying in the case of concept-driven, i.e., onomasiological search. Synonyms (big, great, tall) appear very far apart from each other in an alphabetically organized resource. The same is true for resources ordering their output solely on frequency. If you take a look at an association

13 For an experiment comparing the quality of various query terms see (Zock & Schwab 2011 and 2016).

14 Think of a search in a web search engine.

15 Some of the work on classification or library and information science is very relevant here and has been widely discussed in the literature (Bliss, 1935; Borges, 1984, Dewey, 1996; Eco, 1995; Foucault, 1989; Ranganathan, 1959 and 1967; Svenonius, 2000; Wilkins, 1668).

thesaurus like the E.A.T. (Kiss et al. 1973)[16] you will notice that categorially similar items (say, Persia and Afghanistan) are often quite far apart (Zock & Schwab, 2013). Also, if the resource within which search takes place is a hybrid semantic network —words being connected via typed and untyped relations— than we have different search scenarios. If the cue and the target are connected via a typed link, the output is fairly straightforward[17]. Either the user can provide the link-type with the input, say, 'horse+hypernym', in which case the system displays a single term (equid), or a small list as output (synonym: equine, stallion). In the opposite case, the system will present a set of lists, clustered by link-type (synonym, hypernym,…). If ever the words are connected via free associations (simple co-occurrence, hence, untyped links), then we get a flat, unstructured list, and if this list is long, the user may end up having a navigational problem. In order to avoid that (Zock, 2019) proposed to cluster this list by categories and to give names to each one of them (categorial tree). Hence, the words in the following list {Afghanistan, Africa, Asia, black, brown, China, cow, flies, Pakistan ….} could be subsumed under the following categories : animal, color, continent, country.

(d) How to determine the optimal search space? Dictionaries are generally quite large, hence the question, how to reduce this set to one, the target[18]? Put differently, how to reduce the entire lexicon (initial set) to a subset, which is neither too big nor too small? While we do not want

16 Since the University of Edinburgh does not host the E.A.T. any more, the following link being broken (http://www.eat.rl.ac.uk), we suggest to go to the following site (http://rali.iro.umontreal.ca/rali/?q=en/XML-EAT), created by Guy Lapalme of the University of Montreal.

17 This has been nicely exploited by WordNet.

18 While this question makes sense for the computer scientist, it does not really for the psychologist, who considers word access to be a matter of activation rather than search. According to psychologists words are synthesized rather than located.

to drown the user, we also want to avoid eliminating potential target words by making the search space too small. This assumes, of course, a reasonable input, ideally a word closely related to the target (direct neighbor).

Of course, there are many more questions, but in order to find answers to at least some of them, it may be good to start by taking a look at well attested situations, like the ***T**ip-of-the-**T**ongue* problem, henceforth ToT.

2. The Tip-of-the-tongue problem

There are different sorts of impairment (aphasia, anomia ....), and different reasons causing word-finding problems[19]. Yet, probably the best known and most intensively studied one is the ToT problem (Brown, 1991, Brown, 2012, Nickels, 2002). Someone is said to be in this state when he knows what to say, he also knows the corresponding form, but for some reason he simply is not able to access it in time, at the very moment of speaking or writing.

One may wonder what causes such a state. Actually, there are many possible reasons: (a) proximity at the semantic or phonological level (left/right – historical/hysterical), (b) low frequency of the target word; (c) lack of usage (decay), (d) age (loss of neurons or destruction of neuronal pathways), (e) distraction (lack of attention), (f) interference. The last two points can have quite a dramatic effect, though the causes are rarely obvious. Interference may occur at various levels and it is generally due to formal similarities with the target, possibly yielding the production of a cognate, i.e., false friend[20], or an otherwise similar-looking or sounding

19 Zock (2002) describes a method of how to solve name-finding problems due to phoneme- or syllable scrambling.

20 For example, the Spanish word *'libreria'* evokes the English form *'library'*, yet when Brits talk about a 'library' they want the Spaniards to understand that they are referring to 'biblioteca', the corresponding Spanish word. Likewise, when

word: hungry/angry; emigrate/immigrate ; anagrams : cheater/teacher[21]. The fact that the speaker has produced one form instead of the other ('heel' instead of 'heal') may have as side effect to evoke the meaning of the intrusive form ('foot'), which considerably reduces the chances to revive the desired target form. See how the misproduction of one form, say 'heel' instead of 'heal' or 'hell' may lead you off the path to a completely different set of thoughst: foot, medicine, or even 'heaven' (heel-hell). Associations can be fatal, distracting the user to a point that he cannot remember anymore his goal. Yet, they may also be salvatory, i.e., helping us to find the right form.

To get a better understanding of the problem at hand, let us consider word access in its natural context, language production[22]. After all, words are generally used in this situation[23], even if this is not always the case.

---

a Spanish speaker uses the word 'libreria' he wants the British to understand 'bookstore' and not 'library', the similar-sounding word.

21 Suppose you wanted to refer to an *animal* having the following features {woolly usually horned ruminant mammal related to the goat}, then there could be a competition between any of the following words, i.e., lexical concepts, at the semantic level: {mutton, ram, ewe, lamb, sheep, goat, bovid, ovis}. If we had chosen the lemma *"sheep"*, which at this stage is still a phonologically empty form, then any of the following phonological neighbors {cheap, jeep, schliep, seep, sheep, sleep, steep, streep, sweep}may come to our mind, while obviously we were looking only for /ʃiːp/. Note that confusion or interference can occur not only at the vertical (paradigmatic axis), but also on the horizontal level. For example, in the attested example, 'glear plue sky' there has been a switch of the voiceless form of one element (the 'b' from 'blue') to the voiced of the other ('c' stemming from'clear'), yielding in each case the wrong form: 'g' instead of 'c' and 'p' instead of 'b'.

22 For a broad view from a psycholinguistic or neuroscientist's perspective, see (Levelt, 1989; Rapp & Goldrick, 2006; Goldrick, et al. 2014). The equivalent, but from an engineering point of can be found in (Dale & Reiter, 2000, Krahmer & Theune, 2010). For a state-of-the-art approach, see (Bateman & Zock, 2016 ; Gatt & Krahmer, 2018).

23 For an excellent survey for word access in language understanding, a task which is very different from language production, see (Altmann, 1997). For a discussion concerning research devoted to each one of them, see (Roelofs, 2003).

Language production involves three major tasks (Bock, 1996; Levelt, 1989), most of which apply not only for sentence generation, but also for the production of words. Given some goal (need, visual input), we activate a concept, idea or message ($level_1$) which at that point is only a more or less abstract entity devoid of linguistic information. We certainly do not have yet the lexical form, neither do we have grammatical information (part of speech, gender) nor the word's sound form (type and number of syllables, intonation, etc.). All we have is something abstract and underspecified like 'MOVE' / 'REPTILE', or something more elaborate like the incomplete description (definition) of a word : <domesticated four legged MAMMAL kept for its meat or its thick wooly coat>. Yet, in order to produce the concrete lexical forms we must either add information (which applies in the case of underspecified inputs: MOVE/REPTILE), or we must choose among a set of alternatives ('sheep, ram, goat') which becomes necessary if ever we provided only an incomplete conceptual input. While the first may yield 'walk, limp, run' or 'alligator, caiman, crocodile', the latter may yield 'sheep'.

Alas, the conversion of meaning to sound is not straightforward, it requires several steps. During the first we converge towards a lexical concept, i.e. *lemma* ($level_2$)[24], say 'sheep' rather than 'goat'. During the

24 The lemma specifies the part of speech and some general information concerning the phonological form (number of syllables, intonation, …). Please beware though, that different communities associate different meanings with the term 'lemma'. While for linguists it refers to the word's citation- or dictionary form (nouns in singular, verbs in the infinitive, ...), for psychologists (Levelt, 1989; Roelofs et coll. 1998) it refers to some kind of abstract entity encoding semantic and syntactic information (part of speech, gender). Since lemma do not contain the phonological form (lexeme), we cannot produce at that stage their physical form (spoken or written form). All we have at that moment is an abstract entity (lexical concept). Put differently, the *lemma* 'sheep' is not the same as the corresponding sound form or lexeme: /□i□p/ .This is in sharp contrast to our traditional view of dictionary entries where meaning and form appear next to each other. This holistic view is inherited from Saussure (1916) who considered language as a system of signs, a set of organic entities composed of a *signifier*

second step (level$_3$) our brain specifies the phonological form, yielding a *lexeme*, the word's concrete form. This allows us then to compute the required motor program to carry out the necessary steps to produce the final written or spoken form, say 'sheep' or /ʃiːp/. A tip-of-the-tongue state occurs if there is an interruption between level$_2$ and level$_3$. The figure below illustrates this process for the word 'sheep'. It is inspired by the work done by Levelt and his colleagues (1999).

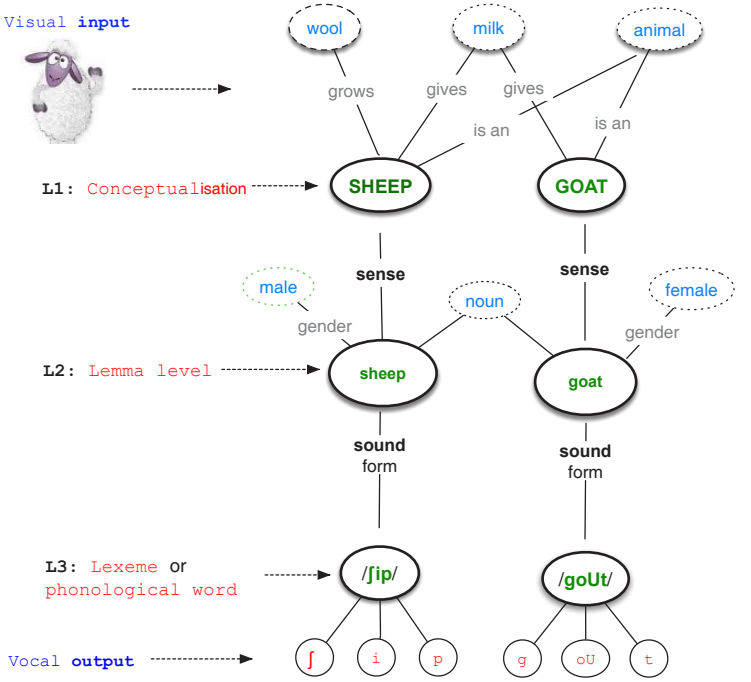

Figure 1. Word access based on (Levelt et al. 1999)[25]

(form) and the *signified* (meaning), the two appearing together, one being simply the reverse side of the other.

25 Levelt's word production model (Levelt et al., 1999) is actually quite a bit more sophisticated. It requires the following six steps : (1) conceptual preparation → lexical concept ; (2) lexical selection (abstract word) → lemma; (3) morphological encoding → morpheme ; (4) phonological encoding

ToT problems can be seen as a puzzle which can be solved by providing or priming the missing elements. This can be done indirectly (cf. Abrams et al., 2007). James and Burke (2000) designed a protocol to do precisely this. They presented some pictures or definitions asking their subjects to find the corresponding word. Those who failed, but knew the word, i.e., those who were in a ToT state, were used for the main part of the experiment. This group was then divided in two equal parts. Half of the participants were asked to read aloud a list of words that cumulatively contained all of the syllables of the ToT word. Suppose someone failed to retrieve the target *abdicate*, in this case he would be asked to read the following list of ten words, *ab*stract, *in*digent, *trunc*ate, *tradition* and *loc*ate, each of which contains a syllable of the target. The other half was also given a list of 10 words, but phonologically unrelated. Having done this exercise, participants were asked to try again to retrieve the target. And this time most of the members of the group being exposed to phonologically related words succeeded, while the other group did not.

Obviously, in a natural situation we cannot expect the phonological primes to occur, neither can we provide them as the psychologists did in their experiment, as this would require knowledge of the target. Yet if we knew the target then we would give it, since this is what the author is looking for. To conclude, we cannot provide the missing parts or offer form-related cues (for example, phonological cues), what we can do though is to provide semantically related words, associations, i.e., words related to the user input.

Actually, we can do more than that, as studies have shown that people having word-finding problems always know something concerning the target word: meaning, sound, origin, etc. This being so, we could start from

(syllabification) → phonological word; (5) phonetic encoding → phonetic gestural code ; (6) articulation → sound wave. Note that it postulates two knowledge bases: the mental lexicon, vital for lemma retrieval, and the syllabary, important for phonetic encoding.

this knowledge, incomplete as it may be, and build a bridge between this point, the *source word* (s_w) and the goal, i.e., *target word* (t_w).

Indeed, there are at least three things an author usually knows before opening a dictionary: (a) the word's *domain* (e.g., tourism, sport, food), (b) its meaning (its definition), or at least part of it, and (c) its relation to other concepts or words: [X is $_{more\ general}$ than Y]; [X is the $_{equivalent\_of}$ Y]; [X is the $_{opposite\_of}$ Y]. In other words; X being the hypernym/synonym or antonym of Y, etc. In all these cases x and y could be either the *source*- or *target word*, that is, the word coming to someone's mind (access word), or the term she is looking for. This is basically conceptual and encyclopedic knowledge.

Yet, people seem to know more than that. For example, they seem to know many details concerning the lexical *form*: number of *syllables; beginning/ending* of the target word; *part* of *speech or lexical category* (Brown & McNeill, 1996; Burke *et al*., 1991), and sometimes even the *gender* (Vigliocco *et al*. 1997).

The same holds when people look for a specific target, say 'mocha'. In this case the user may have access to the following fragments, which reveal his current knowledge state: *domain* (food) ; *meaning fragments* (coffee mixed with chocolate; Arabian coffee); *part-of-speech* (noun); *rhymes_ with* (soccer, coca); *evokes* (Starbucks, coffee beans, Yemen, dark brown); *related terms* (latte, espresso, cappuccino).

The fact that authors have access to different fragments at query time implies that they have different information needs. Since the available fragments to belong to different layers (meaning, sound, etc.), we need to create different resources, one for each specific need (Zock et al. 2010), and we should combine them[26], as otherwise our search space will be too large,

26 This is precisely what Castro & Stella (2018) have done by introducing the notion of multiplex networks, which allowed them to achieve excellent results for picture naming. It should be noted though that combining different layers, databases or indexes is not trivial, as in the absence of sound principles, it may lead to an unnecessarily large search space. This is why we suggest to have the

or it will not contain the target.

As one can see, there are many possibilities to facilitate word-finding. Yet, we will deal here only with one of them, associations[27], or, more precisely, the relations between concepts or words. Knowing the nature of a relation and using it are vital for many tasks : understand the (physical/mental) world, creation of coherent discourse, etc. With respect to word-finding relations are important for orientation and search space reduction (see Section 4, knowledge states).

3. Related Work

Concerning lexical access, several communities are concerned: engineers from the natural language generation community (NLG), psychologists[28],

user to provide with the input (query) additional information concerning the kind of word he is looking for (word related by meaning, sound, association, domain), as this is something the computer cannot guess (see Section 4: knowledge states). Obviously, given the following input ('harbor') a user may expect quite different proposals from the system depending on his goal : does he want the system to find him a *synonym* (anchorage), a *hypernym* (seaport) or a *sound-related* word (barber)'.

27 For more on this topic, see : Hees, 2018, Rapp, 2017; de Deyne & Storms, 2015; Hörmann (2013, chapter 6), Meara, 2009; Miller, 1991 and 1969; Schvaneveldt, 1989; Strube, 1984; Findler, 1979 ; Jenkins, 1970; Cramer, 1968, Deese, 1966 ; Freud, 1901; Jung, 1910 and Galton, 1880, the first one to prove scientifically the validity of this fundamental notion.

28 The dominant psycholinguistic theories of word production are all activation-based, multilayered network models. Most of them are implemented, and their focus lies on modeling human performance: *speech errors* or the *time course* (latency distribution) as observed during the access of the mental lexicon. The two best-known models are those of (Dell, et al., 1997) and (Levelt, et al., 1999), which take opposing positions concerning *conceptual input* (conceptual primitives vs. holistic lexicalized concepts) and *activation flow* (unidirectional vs. bidirectional). For a comparison and evaluation of five major theories or models, see (Rapp & Goldrick, 2000). Note that there is also quite some work where other techniques are used (chronometry, eye tracking, electrophysiology and neuroimaging). For more details concerning the various theories, methods and implementations, see the work cited in footnote 10.

computational linguists and lexicographers. Space constraints prevent us from referring to all this work, but a summary and discussion of the various models can be found in (Zock et al. 2010). Hence, we will focus here mainly on the work done in lexicography. Note though, that the problem addressed by the NLG community deals only with 'lexical choice'[29], but not with 'lexical access'. Yet, before choosing a word, one must have accessed it.

How words are stored and processed in the human mind has extensively been dealt with by psychologists[30]. Yet, while there are many papers dealing with the *tip-of-the-tongue phenomenon* (Brown, 1991), or the problem of lexical access[31], they do not consider the use of computers for helping people in their task (our goal).

Lexicographers have tried their best to bridge this gap. Unfortunately, until recently most of their tools have been devoted to the language receiver. Yet, to be fair, one must admit that great efforts have been made recently to address also the problems of the language producer.

3.1 Onomasiological Resources

In fact, there are quite a few *onomasiological dictionaries* (van Sterkenburg, 2003) or lexical resources allowing search based on meaning. For example, Roget's thesaurus (Roget, 1852) or its modern incarnation built with the help of corpus linguistics (Dornseiff, 2003). There are also analogical dictionaries (Boissie`re, 1862; Robert et al., 1993), Longman's Language Activator (Longman, 1993), Fontenelle's semantic networks (Fontenelle, 1997), the Oxford Learner's Wordfinder Dictionary (Trappes-Lomax, 1997), and various network-based lexical resources: *WordNet*, henceforth WN (Miller,1990), *Framenet* (Fillmore et al. 2003), *PropBank*

29 For excellent surveys see (Robin,1990; Stede, 1995; Wanner, 1996).

30 Aitchinson, 2003; See also footnote 11.

31 For example, Levelt et al. 1999; Pulvermüller, 1999; Rapp & Goldrick, 2006; Griffin & Ferreira, 2011; Dell et al. 2014; Goldrick, 2014.

(Palmer et al. 2005), *VerbNet* (Kipper-Schuler 2005), *MindNet* (Richardson et al., 1998), *HowNet* (Dong & Dong, 2006) and *Pathfinder* (Schvaneveldt, 1989). Other proposals are 'The active vocabulary for French' (Mel'cˇuk & Polguère, 2007), and ANW, a dictionary of Dutch whose author (Moerdijk, 2008, http://anw.inl.nl) introduces the interesting notion of Semagrams (see below).

There are also various collocation dictionaries (Benson et al., 2010), the Oxford Learner's Wordfinder (Trappes-Lomax, 1997), and web-based resources like Lexical FreeNet[32] (Beeferman,1998), that mixes semantic relations and phonetically derived relations, and OneLook[33]. Like Yago (Suchanek et al. 2007) and BabelNet (Navigli & Ponzetto, 2012) OneLook combines a dictionary (WN) and an encyclopedia (Wikipedia), though putting the emphasis on onomasiological search, i.e., lexical access. Finally, there is MEDAL (Rundell and Fox, 2002), a thesaurus produced with the help of Kilgariff's Sketch Engine (Kilgarriff et al., 2004), and various reverse dictionaries (Bernstein, 1975; Kahn, 1989; Edmonds, 1999, Thorat & Choudhari, 2016) built by hand (Bernstein, 1975) or with the help of machines[34]. In both cases, one draws on the words occurring in the definition. Thorat and Choudhari (2016) try to extend this idea by introducing a distance-based approach to compute word similarity. Given a small set of words they compare their approach with OneLook and with dense-vector similarity. While we adopt part of their methodology in our

32 http://www.lexfn.com

33 http://onelook.com

34 For details, see (Liu et al. 2019 ; Pilehvar, 2019 ; Reyes-Magaña et al. 2019 ; Zhang et al., 2019 ; Hill, et al., 2016 ; Thorat & Choudhari, 2016 ; Shaw et al., 2011 ; Dutoit and Nugues, 2002). Note that there are also several commercial versions of reverse dictionaries, 'Dictionary.com' (www.dictionary.com) being one of them. It reveals candidates based on the user's word description. OneLook takes a different approach. It searches within the index of the dictionaries it has access to (more than 1000), mainly all major freely available lexical resources (WN, Wiktionary, etc.). The same can be achieved in French via *JeuxdeMots* whose data are free (http://www.jeuxdemots.org//diko.php).

evaluation scheme, we are more reserved with respect to their architecture. Since it requires a fully computed similarity matrix for the entire vocabulary, their work cannot scale up: it is unreasonable to assume that the lexicon is stored in a fully connected similarity matrix, which grows quadratically in the size of the vocabulary.

3.2 Vector-space Models, Distributional Semantics and (complex) Graph Theory

While not being directly connected to lexicography, *word-space* and *vector-space models* (Sahlgren, 2006; Turney & Pantel, 2010; Widdows, 2004), *distributed semantics* (Baroni & Lenci, 2011) and graph theory, i.e., *complex graphs* (Mihalcea & Radev, 2011) are very relevant with respect to the construction of the dictionary of tomorrow so many authors have been dreaming of[35]. *Graph models*[36] allow to reveal the topology of the mental lexicon (Vitevitch, et al. 2014), that is, they can show the hubs (local density and number of connections), the position of a word in the network, its relative distance and connectedness to other words (co-occurrences and direct neighbors), etc. *Vector space models and distributional semantics* rely on words in context. Hence both can reveal hidden information, and allow, at least in principle, to build applications capable of brainstorming, reading between the lines, and much more (Steyvers et al. 2002 ; Steyvers & Tenenbaum, 2005 and 2010, Vitevitch, 2008). Yet, more is to come, psychologists offering methods to organize knowledge (concepts, words) in line with the pathways of the human brain (Sporns, 2011; Lamb, 1999; Spitzer, 1999).

As one can see, a lot of progress has been made during the last three decades, yet more can be done especially with respect to indexing

35 (Atkins and coll.,1994; de Schryver, 2003; Leech & Nesi, 1999; Rundell, 2007; Sinclair, 1987).

36 (Siew et al. 2018; Kenett et al. 2014; Zortea et al. 2014; Baronchelli et al. 2013, Bieman, 2012; Morais et al. 2013; Motter et al. 2002)

(organization of the data) and navigation.

4. Navigation, a Fundamentally Cognitive Process

As we will show in this section, navigation in a lexical resource is above all a knowledge-based process. Before being able to use a word, we must have acquired it. It is only then that it has become part of our knowledge. Yet, storage or knowledge do not guarantee word-finding, i.e., access (Zock & Schwab, 2011; Zock & Schwab, 2016). This fact has not received the attention it deserves by lexicographers. Also, there are several kinds of knowledge (declarative, meta-knowledge and knowledge states) of which the last two at least need to be taken into account (Zock & Tesfaye, 2015; Zock, 2019) if we want to build a user-friendly tool, allowing to find even needles in a haystack.

4.1 Declarative Knowledge

Declarative knowledge is what we acquire when learning words (meaning, form, spelling, usage). This is the information we generally find in dictionaries (meaning, form, grammatical information, usage). Obviously, in order to find a word or to find the information associated with it, they must be stored, though this is not enough. Also, when speaking or writing we generally start from meanings, though this is not necessarily all. Suppose, you were looking for a word expressing the following ideas: *domesticated animal, producing milk, suitable for making cheese.* Suppose further that you knew that the target word was neither *cow* nor *sheep.* While none of this information is sufficient to guarantee the access of the intended word *goat*, the information at hand (fragments of the definition) could certainly be used. Yet, next to definitional information, people have other kinds of knowledge concerning the *target word*. For example, they know how it relates to other words. They also know that *goats* and *sheep* are somehow

connected in that both of them are *animals*, that *sheep* are appreciated for their wool and meat, that they tend to follow each other blindly, while *goats* manage to survive, while hardly eating anything, etc. In sum, people have in their mind a whole encyclopedia concerning the knowledge evoked by words.

4.2 Meta-Knowledge

Next to knowledge we need to acquire *meta-knowledge*, which plays a fundamental role in cognition. Being generally unavailable for in(tro) spection, meta-knowledge reveals itself in various ways. For example, via the information available when we *fail to access a word* (Schwartz, 2006), or via the *query* we provide at the onset of the search. As mentioned already, and as word association experiments have shown (Aitchison, 2003) words always evoke something. Since this is true for all words one can conclude that all words are connected in our mind, which implies that all words are accessible from anywhere like in a fully connected graph[37]. All we have to do is to provide some input (source word, available information) and follow then the path linking this input to the output (target). Interestingly, people hardly ever start from words remotely related to the target. Quite to the contrary, they tend to start from a more or less direct neighbor of the target, the distance between the two, exceeding rarely the distance of 2[38]. Also, dictionary users often know the type of relationship

37 Note that this does not hold for WN, as WN is not a single network, but a set of networks. There are 25 for nouns, and at least one for all the other parts of speech.

38 This is probably one of the reasons why we would feel estranged if someone provided as cue 'computer', while his target is 'mocha'. The two are definitely not directly connected, though, there is a path between them, even though it is not obvious (the chosen elements being underlined.): computer → (*Java, Perl, Prolog ; mouse, printer ; Mac, PC*) ; (1) Java → (*island, programming language*) ; (2) Java (island) → (coffee; Kawa Igen); (3) coffee → (Cappuccino, Mocha, Latte). Note that 'Java' could activate 'Java beans', a notion inherent to JAVA, the programming language. In this case it would lead the user directly to

holding between the input (prime) and the target. These two observations clearly support our idea that people have a considerable amount of (meta-) knowledge concerning the organization of words in their mind, i.e., their mental lexicon. The idea of relationship has been nicely exploited by WN, which due to this feature keeps the search space, i.e., a set of candidates among which the user has to choose, quite small. The idea of relatedness has led lexicographers already in the past to build thesauri, collocation- and synonym dictionaries. Obviously, an input consisting only of a simple word is hard to interpret. Does the user want a more general/specific word, a synonym or antonym? Is the input semantically or phonetically related to the target, or is it part of the target word's definition (dog-animal)? In each case the user is expecting a different word (or set of words) as output. Hence, in order to enable a system to properly interpret the users' goals we need this kind of metalinguistic information (neighbor of the target, i.e., s_w + relation to the t_w) at the input[39]. If ever the user cannot provide it, the system is condemned to make a rough guess, presenting all directly connected words. Obviously, such a list can become quite large. This being so, it makes sense to provide the system this kind of information so that it can produce the right set of words, while keeping the search space small.

4.3 Knowledge States

Finally, *Knowledge states*, refer to the knowledge activated at a given point in time, for example, when launching a search. What has been primed? What is available in the user's mind? Obviously, in order to find a word or to find the information associated with it, they must be stored.

---

the class (hypernym) containing the desired target word (mocha).

39 This has, of course, consequences with respect to the resource. To be able to satisfy the different user needs (goals, strategies) we probably need to create different databases: Obviously, to find a target on the basis of sound (rhymes), meanings (meaning fragments) or related words (co-occurrences), requires in each case networks encoding different kinds of information.

Yet this is not enough, as not all stored information is equally available or prominent anytime. Knowledge states (KS) are highly fluctuant, varying from person to person and from moment to moment. In the case of word-finding, knowledge states are characterized by the value of two parameters : the s_w and its *relation* to the t_w. The first one is the specific item coming to the author's mind when looking for a word, say 'black' (cue, query), while the eluding terms (t_w) may be 'white', 'dark' or 'coffee'. Since the s_w and the t_w can be related in many ways (in our case: opposite_of, similar_ to, color), and since the author generally knows the type of relation, while the system does not, it is good that the former reveals it to the latter. This is precisely what is achieved via the second kind of information. It tells the system how the sw is related to the target. Note that even though the author fails to retrieve the target, he generally knows the relation type, for example, B $_{sounds\_like}$ A, B $_{part\_of}$ A; B $_{used\_for}$ A, where X and B are respectively the source and the target. This latter kind of information is also precious as it can be used to signal the system the seeker's goal, namely, that he is looking for a semantically, phonologically (sounds_like, rhymes_with) or otherwise related word (free association; co-occurrence). In sum, both types of information are vital for the system, as they signal where to start the search from (i.e., what is a reasonably close neighbor of the target?), and what the user's goal is. In the absence of his information, the system has no clue what data to present to the user. Hence the output may be irrelevant, or too numerous, the search space having become too large.

Obviously, the fact that peoples' knowledge *states* vary is important, as it determines to a large extent the users' search strategies. This is why it should be taken into consideration by the system designer. Actually, a system can become truly useful only if it is 'aware' of the user's knowledge state (and meta-knowledge) and his goal. This is what allows it to determine the optimal search space, reducing the scope only to reasonable outputs, while in the absence of this information the scope could be the entire lexicon.

The example here below illustrates to some extent these facts with regard to word-finding in an electronic resource. Suppose you are looking for a word conveying the idea of a *large black-and-white herbivorous mammal of China.* Yet, for some reason you fail to retrieve the intended form, Panda, even though you know a lot concerning the target. People being in this ToT state would definitely appreciate if a system could help them to find the target, by taking as input the information they currently have access to. Figure 2 illustrates the process of getting from a visual stimulus to its expression in language via a lexical resource. Given an external stimulus (A) our brain activates a set of features (B) that ideally allow us to retrieve the target form. If our brain fails, we use a fallback strategy and give part of the activated information to a lexical resource (C) expecting it to filter its base (D) in the hope to find the target (panda) or a somehow related word (E). As one can see, we consider lookup basically as a two-step process. At step one the user provides some input (current knowledge) to which the system answers with a set of candidates, at the step two the user scans this list to make her choice.

A	*perceptual input*, i.e., target	
B	*associated features* in the mental lexicon (brain)	*type* : bear *lives_in* : China *features* : black patches *diet* : eats bamboo
C	***input*** to lexical resource	bear China
D	lexical resource	**a**ardvark ... ... *panda* ... **z**ygote
E	***output*** of lexical resource	*panda* polar bear

Figure 2. Lexical access a two-step process mediated by the brain and an external resource (lexicon).

5. A Framework for Dictionary Navigation

Lexical access can be viewed as a dialogue or guessing game between a human and a machine, where the former provides a cue (the momentarily available knowledge concerning the target) and the latter responds with the best guess(es) it can make given this input. Obviously, all things being equal, the more precise the input, the better output.

Imagine the following situation, where the user's query is simply 'fish'. Without any further ado (information), there is no way for the system to know what the user wants, hence what to do with this input. Is he looking for the word's *meaning* (aquatic vertebrate with gills)? Is he searching for one of its *instances* (barracuda), a *more general* word (aquatic vertebrate, seafood), or some *associated* term ('chips', 'sign of the zodiac')? Also, did he mean 'fish' the animal, i.e., the noun, or 'to fish' the activity (verb), etc.?

Since the user's knowledge states vary considerably, he experiences different kinds of deficits, and depending on the level concerned (conceptual, phonological) he is able to provide different information: a word related in terms of meaning, sound, etc. This implies that we need to build different kinds of resources in order to address the various user needs. For example, if he can recollect only the first and last syllable of a word, we need a syllable dictionary. If in addition he can give us some clues concerning the word's meaning, we could also draw on another resource (reverse dictionary) and produce as output the combination of the two. While one can think of many kinds of dictionaries or lexical resources, we will consider here only the one where words are connected via associations, that is, an association thesaurus.

In the remainder of this section, we will try to answer briefly the following three questions : What should a resource look like to allow for the search described in the figure here above? How to use and how to build it[40]?

40 For a more detailed description, see (Zock & Schwab, 2011 ; Zock, 2019). We

5.1 What should a Resource Look like to Allow for this?

We would need a fully connected graph, or, more precisely, an association thesaurus (AT) containing typed and untyped links. Both kinds of links are necessary for filtering, i.e., to ensure that the search space is neither too big (typed links), nor too small (untyped links). Untyped links are a necessary evil: they are necessary to address the fact that two words evoke each other even though we are not able to qualify the nature of the link.

5.2 How to Use it?

Imagine an author wishing to convey the name of a beverage typically served in coffee shops. Failing to evoke the desired form ('mocha'), he reaches for a lexicon. Since dictionaries are too huge to be scanned from cover (letter A) to cover (Z), we will try to reduce the search space incrementally. Having received some input from the user, say 'coffee' — which is the word coming to his mind while failing to access the target— the system answers with a set of words among which the user chooses. If the input and the target are direct neighbors in the network, and if the user knows the link between the two (source + target), then the search space is generally quite small. In the opposite case, that is, if the user cannot specify the link, then the system is condemned to make an exhaustive search, retrieving all direct neighbors of the input. However, the system could cluster the words by affinity and give names to these categories, so that the user, rather than navigating in a huge flat list navigates in a categorial tree, which avoids scanning long lists.

This last step is very important in particular if ever the list gets very long. Suppose, we looked for the name of a famous spiritual leader (Gandhi), by providing 'India' as input, then we certainly would not want to get a flat list,

proposed then the following steps : (a) building a semantic network based on a corpus, (b) type the links (indexing), (c) rank words in terms of frequency, (d) cluster the system's output and name the categories (in the case of free associations, i.e., untyped links).

as this is the case with the Edinburgh Association Thesaurus, but rather a categorial tree, containing the retrieved words' named clusters, for example (**COUNTRY** ; Pakistan, China, Afghanistan) ; (**CONTINENT** : Africa, Asia) ; (**COLOR** : black, brown) ; (**ANIMAL** : cow, flies), ...

5.3 How to Build it?

While there are quite a few resources, in particular, association thesauri, they are too small to allow us to solve the ToT problem. The projected resource would still have to be built, and while one could imagine the use of combined resources, like BabelNet (Navigli and Ponzetto, 2012), or the combination of WN with other resources like topic maps (Agirre et al. 2001), Roget's Thesaurus (Mandala, 1999) or ConceptNet (Liu and Sing, 2004; Speer et al., 2017), it is not easy to tell which combination is best, all the more as besides encyclopedic knowledge, we also need episodic knowledge (Tulving, 1983).

One straightforward solution might be co-occurrences (Wettler & Rapp, 1993; Lemaire & Denhière, 2004; Schulte im Walde & Melinger, 2008). The problem with them is that they are too powerful, even after application of appropriate significance measures. While, they link many words, most of them are not the ones we need, that is, those coming to our mind when we think of a specific target. To convince yourself, you may want to take a look at Wikipedia, and choose an entry, say Panda (https://en.wikipedia.org/wiki/Giant_panda). Try to figure out which co-occurrences are equivalent to the associations you, or most people would have with respect to this kind of bear. Of all the words contained in this document, we are interested only in ["bear, zoo, cute, Himalaya, bamboo, herbivorous mammal, black and white patches, furry, Nepal, Tibet, China, diplomatic gift; ...."], as they are the most likely ones to evoke the target, *panda*, the rest being simply noise[41]. As

41 It is interesting to see the results you get by going to De Deyne's or Lafourcade's resources : (https://smallworldofwords.org/en/project/explore) and (http://www.

one can see, only very few words are really relevant for our task, and the challenge is to make sure to get precisely those. What is worse, is the fact we cannot generalize, as, what seems relevant for Pandas is not necessarily relevant for other, similar animals (bears). This being said, we believe that there are other solutions.

One being semagrams (Moerdijk, 2008) which are reminiscent of Fontenelle's (1997) lexical-semantic networks, resulting from the combination of the Collins-Robert dictionary and lexical functions (Mel'čuk, 1996). Semagrams represent the knowledge associated with a word in terms of attribute-values. Each semantic class has its type template and corresponding slots. For instance, the type template for *animals* contains the slots 'parts, behavior, color, sound, size, place, appearance, function', etc., whereas the one for *beverages* has slots for 'ingredients, preparation, taste, color, transparency, use, smell, source, function, 'composition', etc. Here below is an example for '*cow*'.

UPPER CATEGORY	is an animal
CATEGORY	is a bovine
SOUND	moos/lows, makes a sound that we imitate with a low, long-drawn 'moo'
COLOR	is often black and white spotted, but also brown and white spotted, black, brown or white
SIZE:	is big
BUILD	is big-boned, bony, large-limbed in build
PARTS	has an udder, horns and four stomachs: paunch, reticulum, third stomach, proper stomach
FUNCTION	produces *milk* and (being slaughtered) meat
PLACE	is kept on a *farm*; is in the *field* and in the winter in the *byre*

jeuxdemots.org//AKI.php).

AGE	is an adult, has calved
PROPERTY	is useful and tame; is considered as a friendly, lazy, slow, dumb, curious, social animal
GENDER	is female
BEHAVIOR	grazes and ruminates
TREATMENT	is milked every day; is slaughtered
PRODUCT	produces milk and meat
VALUE	is useful

Semagrams are built by hand, and while it is unlikely that we can infer or mine semagrams automatically, chances are that we can populate them mechanically, which could then be seen as an alternative route of building an association thesaurus, but in a fairly controlled way. Note that this is a suggestion has been made already some time ago (Zock, 2015 ; Zock & Biemann, 2016), yet it seems, that there is now some concrete work going in this direction (Leone et al. 2020), though the focus is not on word access, but on the word's definition.

6. Experimental Setup

In this section, we describe the experimental setup to answer the following research questions:

a) When being in the ToT state what cues do people provide to help the system find the target?
b) How good are existing lexical resources for retrieving the targets by using these cues?
c) How big is the added value of knowing the relationship between the cue (source word) and the target? Put differently, does it enhance retrieval precision and speed?

6.1 Lexical Graphs as Dictionaries

For our experiments we used three different lexical networks: WN, distributional semantic models using word similarity and word co-occurrence. They were chosen deliberately to cover different structural aspects, different amounts of effort to construct them manually, and different degrees of language dependence. Note, that we could have chosen other resources, for example, the E.A.T. (Edinburgh Association Thesaurus), but, being based on free associations, it lacks typed relations. In addition, it is quite old (Kiss et al. 1973)[42], covering only a subset of the words used in our experiment.

- *WordNet:* WN 3.0 (Fellbaum, 1998) is a high-coverage, manually built lexical-semantic network of English. Words are organized in terms of synsets, i.e. sets of synonyms, which are linked in various ways depending on the part of speech. We used a subset of these links (synset, hyponymy, derivation, etc.) and domain categories in the hope to be able to retrieve the target.
- *Word Similarity:* We used the JoBimText distributional semantic model, its similarity score being based on common dependency parse contexts, which requires a language-specific parser. The JoBimText distributional thesaurus[43] (Biemann and Riedl, 2013) contains in ranked order the 200 most similar terms of a newswire corpus of 100 million sentences

42 For example, if you provide 'terrorism' as key, you will get the following list of ranked words as answer : Guerilla, Gun, Soldier, War, Guerrilla, Anarchist, Evil, Fear, Fighting, Rebel, Tyrant, Vandal, Vietnam, Abroad, Activities, Activity, Arab, Arson, Bandit, Blood, Bomb, Che, Che Guevara, Congo, Czech, Fight, Fighter, Gangster, Gorilla, Greek, Guerillas, Guns, Hooligan, Kill, Killer, Madness, Man, Mao, Maoist, Mexico, Night, Police, Regime, Revolution, Revolutionary, Rioter, Russian, Shoot, Terror, Tourist, Tree, Trotsky, Vietcong, Vietnamese, Wog. As one can see 'associations' change over time. The words we would associate nowadays with 'terrorism' are not the same as the ones people had associated in the seventies, the moment of history where this resource was built.

43 Available at www.jobimtext.org

in English. We expect this resource to be suitable for most associative queries, that is to help us find words occurring in contexts like "X is somehow like a Y or a Z" (e.g. "*a panda* is somehow like a *koala* or a *grizzly*"). This example illustrates 'co-hyponymy', a relation not directly encoded in WordNet. Similarities (for example, panda/koala vs. panda/dog) are ranked by context overlap.

- *Word Co-occurrence:* We compute statistically significant sentence-based word co-occurrences using the same corpus as here above, and following the methodology of (Quasthoff et al., 2006)[44]. We expect this resource to be suited for free associations, i.e., cue words whose link to the target cannot be specified. This resource has by far the highest rate of relations across different word classes, as they may occur in patterns like "With Xs, especially with Y ones, you can Z and W" (e.g. "with mochas, especially with iced ones, you can chill and have cookies"). Co-occurrences are ranked by the log-likelihood significance measure (Dunning, 1993).

6.2 Network Access

Given the structural differences of our resources, our networks are accessed with different query strategies. The general setup is to query the resource via a cue and to insert then the retrieved terms into a ranking. As long as the system has not found all the desired words, it will keep going by querying with words according to their rank, inserting previously un-retrieved terms below the ranking.

- *WordNet*: Having noticed that people tend to use hypernyms (flower) as cues to find the *hyponym* (rose, the target), we defined a heuristic supporting query using this relation. We start by querying for 'synonyms' of the cue, putting results first in the ranking. Next, we proceed along the sense numbers, senses being ordered by frequency in WN, which

44 Available at http://corpora.informatik.uni-leipzig.de

ensures that we start with the most common senses. Third, we add (in this order) direct 'hyponyms', 'meronyms' and 'domain members'. This order seems to be justified by the fact that most people tend to go from general to specifics, starting by a more general term when launching a search. Finally, we add other relations like 'similar', 'antonyms', 'hypernyms', 'holonyms, 'domains', etc. For example, for the cue "pronouncement", the target "affirmation" is found by first checking the cue's 'synonyms' ("dictum", "say-so"), before checking the direct *hyponym* and *hypernym* (directive, declaration). Next we navigate through directly related words of "dictum", synonym of "pronouncement", to find then the target as a direct *hypernym* of "say-so" in its first sense, resulting in rank 12.

- *Word Similarity*: We retrieve the most similar terms per query, ranked by their similarity. Note that due to structure limitations of the resource only 200 similar words can be retrieved per query.
- *Word Co-occurrence*: Having filtered out the 200 most frequent stop words, we retrieve terms co-occurring at least twice with a minimum log-likelihood score of 6.63.

Each cue returns a ranking of the full vocabulary. Working with three cues per target (see Section 6.3), we explore two different combinations of target ranks (minimum rank and merged rank) from querying with the three cues. Regarding *minimum rank*, the rationale is that for each cue, a retrieval process is started in parallel, terminating when the ToT target is encountered for the first time. Actually, only the rank of the 'best' cue is used. For *merged rank*, the rationale is as follows: we use all cues and merge the three rankings by a) adding the ranks per word and sorting by sum or b) multiplying the ranks and sorting byproduct. For more details, see Section 6.4.

6.3 Data Set

Since it is not trivial to put people in the ToT state, we have reformulated the problem in the following way: we ask people to describe a given target to other people who may not know the word (e.g., language learners), by providing three cues. Crowd workers were asked to provide single-word cues rather than descriptions or definitions. Note that the idea was not the creation of a resource, but rather the creation of a set of data to see how well they would behave with respect to our three resources (Section 6.1). Also, in order to get a clearer picture concerning our third question, i.e., the added value of the relation between cue and target, we asked subjects to also specify the relationship between the target and each one of the given three cues. Relations were defined indirectly, i.e., via examples. They comprise synonyms, hypernyms/hyponyms, meronyms/holonyms, typical properties, typical roles (verb-subject, verb-object) and free associations.

Data acquisition was done via the Crowdflower crowdsourcing platform[45]. In order to check whether crowd workers had given the right answer and understood the target, we presented the latter together with three definitions. For our experiment we used only trials that the crowd workers had fully understood, that is, for which they had picked the correct definition. After data collection, we excluded data from crowd workers that deliberately had ignored our instructions. For the targets and definitions, we used the 208 common nouns listed in (Abrams et al., 2007; Harley and Bown, 1998), who examined the ToT state from a psychological angle. Full data, instructions and judgments are available online.[46]

45 www.crowdflower.com, now called Figure Eight.

46 A full description of the crowdsourcing interface can be found here: https://www.inf.uni-hamburg.de/en/inst/ab/lt/resources/data/cogalex16-tot.html

Table 1. Crowd-sourced data collection in a nutshell.

Task

Given a *target* word, say, 'affidavit' you are expected to

(a) choose among a set of definitions to pick the one describing it best;

(b) describe the target in terms of other words (your associations); constraint: use only single words!

(c) specify the *relation* between the *target* and your *associated term*(s).

Here is a list of *possible relations*: (a) mean (about) the_same_thing ; (b) is_a_kind_of ; (c) part_of ; (d) property_of ; (e) typical_action ; (f) typically_used_for ; (g) somehow_associated_with

and here an illustration of valid terms and some of the relations

WALLET and PURSE $\underline{\text{mean about the same}}$ thing

HORSE $\underline{\text{is a kind of}}$ ANIMAL

BRICK is an important $\underline{\text{part of}}$ WALL

GREEN is a typical $\underline{\text{property of}}$ GRASS

Your turn

1° Which of the *definitions* applies to 'abacus' (target) ?

(a) What do you call the tubular viewing toy that produces symmetrical designs through an arrangement of colored chips and mirrors?

(b) What do you call an instrument for performing calculations by sliding beads along rods or grooves?

(c) What is the science of investigating the weather?

2° Please think of a term that you would use to describe abacus'(target), write it in the box below, and specify its relation.

Term$_1$: count

Relation : made_of

3° Do the same for all other terms (Term$_2$ and Term$_3$) describing your *target* (here : 'abacus')

Data collection yielded a total of 1186 cue triplets, provided by 65 participants, who worked on 3 to 132 targets. After manual correction of typos and lemmatization, cue triplets were filtered by eliminating words outside of the vocabulary of the respective resource used in the experiments. Inspection of the data revealed that crowd workers generally chose the cues quite well, but many of them had a hard time to assign the appropriate relation, which is not all that surprising, as this requires quite a bit of metalinguistic knowledge. It is also possible that some participants had chosen the relation without taking the needed care since we did not perform any quality checks during the task. We probably need a different kind of experiment to validate this or measure the extent to which linguistically innocent users can accurately classify semantic relations.

Table 2 below shows the distribution of relations expressed in the first 200 cue triplets (target range 'a-c', i.e., abacus – calisthenics, in alphabetical order) also containing some manually assigned relations. The results show the importance of taxonomic relations, a fact well exploited by WN. Representing nearly 46% of the relations, they confirm the intuition that paradigmatic associations are an important means to access the desired word. However, the next largest class is *syntagmatic*, i.e., untyped, associations (37%). Note that about 17% of the cues come from a different word class than the targets.

Table 2. Distribution of relations between target and cue, as well as typical part of speech (POS) for the cue (N: *Noun*, V: *Verb*, A: *Adjective*), manually assigned by the authors.

Relation	**associated**	**hyponym**	**synonym**	**quality**
Ex.: cue - target	tea - afternoon	story - anecdote	horoscope - astrology	white - albatross
Typ. POS	N	N	N	A
%	36.8%	23.5%	13.3%	8.2%

object	**meronym**	**holonym**	**subject**	**hypernym**
share - anecdote	letters - anagram	day - afternoon	cheer - audience	zombie - cadaver
V	N	N	V	N
5.2%	4.3%	4.2%	3.8%	0.6%

6.4 Evaluation Methodology

Our methodology is very similar to the one of Thorat and Choudhari (2016): we query the lexical network with cues and retrieve then a ranked list of potential ToT targets. With more appropriate cues and better lexical resources, our targets will probably get a boost, appearing higher in the list.

Our vocabulary of WN comprises 139,784 terms, including multi-words, which can be mutually reached through the query procedure. This vocabulary was used in the first experiment. The intersection of the vocabulary of the three networks consists of 34,365 terms, all of them being single words, just as the ones used in the second experiment. Here below are the criteria used in our evaluation:

- *Minimum rank per cue (MinRank)*: if all cues were processed strictly in parallel when would the target appear for the first time?
- *Target rank in sum of ranks (+Rank)*: if the retrieval time depends on the average rank per cue, we sum up the ranks of the three cues and sort the list of terms in ascending order, reporting the position of the target.

Note that this score is strongly influenced by negative outlier cues.

- *Target rank in multiplication of ranks (*Rank)*: To model a multiplicative instead of an additive combination, we multiply the target ranks per cue, sort the list of terms by this score in ascending order, and report then the position of the target. This score is less sensitive to negative outliers.
- *Average Precision*@100 (P@100) measures the fraction of trials containing the target among the first 100 hits, for each of the above. While 100 is an arbitrary number, it seems a reasonable wordlist size to allow for the quick retrieval of a target.

Note that the *minimum rank* is not necessarily lower than the other two scores. It is possible, and it even happens in our data, that a target gets a low rank because all three cues rank it consistently low, while the targets preferred by single cues are ranked much less favorably than others. For example, the target "agnostic" was retrieved from WN (untyped) by its three cues "believer, God, atheist" with ranks 170, 890, respectively 25. Minimum Rank is thus 25, but ranking via sum of ranks lists the target at position 14, while the multiplicative combination results in rank 15.

In the next section, we will qualitatively assess the differences in rankings from our different semantic networks.

7. Results and Discussion

We ran two experiments. In the first we tried to find out whether the knowledge and usage of WN relations produce some added value in terms of retrieval. The goal of the second experiment was to compare the retrieval performance of our three dictionary resources.

7.1 Retrieval Along Semantic Relations

To answer the question whether the usage of relations improves word

access, i.e., retrieval, we used WN, as it is highly structured and our relations can be directly mapped to it. For incorporating relations, we adopted the following query procedure (cf. Section 6.2): we first query for the target relation and then for all the others. For example, for the target ”abacus” and the clue “bead” of type *meronym*, we would first retrieve the *holonyms* of “bead”, then all other relations in the order given in Section 6.2, for initial and subsequent queries. If the supplied relation between the cue and the target is directly given in WN, retrieval is quick. Since the WN hierarchy is quite fine-grained, and since a hyponym relation might be contained over several transitive steps, we keep this order throughout the entire query process.

Table 3. Scores for target retrieval in WordNet by using or ignoring relational information for 200 cue triples on a vocabulary of 139,784 terms.

strategy \ score	Min-Rank	P@100	+Rank	% top100	*Rank	P@100
WordNet untyped	12352.7	40.5%	22403.2	7.5%	17993.5	21.5%
WordNet relations	11733.2	42.0%	22722.7	9.5%	17786.0	22.5%
Random Baseline (STDEV)	35480.7 (514.8)	0.2%	70264.7 (636.0)	0.1%	70438.5 (777.1)	0.1%

Both settings perform much better than the random baseline, which returns the vocabulary in random order irrespective of the dictionary’s structure. The random baseline was obtained by running simulations over the same size of the dataset; we also provide the standard deviation on 10 runs in parentheses where applicable. Since more than 40% of the targets are among the first 100 retrieved words in the MinRank setting, we conclude that WN is indeed suitable. A manual analysis statistically confirmed our intuition: WN is very good for retrieving targets based on

taxonomically related cues (e.g., calculator – abacus), while it does not perform well at all for syntagmatically related words or for noun-noun cues (e.g., beads – abacus, gluten – allergies).

Regarding the added value of relations for retrieval, we conclude that typed relations only help to a small extent, if at all. Our data show fluctuations in the range of a relative -2% to +5% between the settings. Note that this may be a side effect of the sample size, which is quite small. Interestingly, the differences decreased when repeating the experiment with the smaller vocabulary from Experiment 2. Clearly, more work is needed here.

7.2 Comparison of the Three Resources

In order to assess differences between our dictionary resources, we consider the 964 cue triplets per target matching, the common vocabulary of our three resources (see Table 4 below).

Table 4. Scores for target retrieval in our resources for 964 cue triples based on a common vocabulary of 34,365 words.

dictionary \ score	Min-Rank	P@100	+Rank	P@100	*Rank	P@100
Word Similarity	**523.6**	**61.0%**	**1945.9**	**40.6%**	**1040.2**	**55.7%**
Word Co-occurrence	1748.0	44.2%	4205.6	27.2%	3226.9	33.6%
WordNet	2615.4	51.2%	6132.9	13.0%	4247.2	30.3%
Random Baseline (STDEV)	8543.0 (189.7)	0.9%	17156.6 (260.7)	0.2%	17113.8 (252.0)	0.3%

All dictionaries allow for a much better retrieval than the random baseline. The results provide a clear picture: the *word similarity* resource achieves the lowest average ranks on all scores. In 61% of the cases, the

target is among the top 100 retrieved words if we consider only the most effective cue (MinRank). Note that more than half of the targets are found in the top 100 for the multiplicative combination (*Rank). This is surprising, as the relations between the cues and the target are quite diverse (see Section 6.3), and Word Similarity mostly contains direct and indirect taxonomic relations, such as co-hyponyms. The second-best resource in this evaluation is the word co-occurrence network, which outperforms WN on all metrics except the P@100 of MinRank scores.

We also analyzed the differences qualitatively and looked at cue-target-pairs where the three networks perform very differently. As our findings show, different networks have different potential with respect to the retrieval of ToT targets based on a given cue:

- *WordNet* good, Co-*occurrence* poor: Synonyms or near-synonyms, like javelin – spear, cadaver – corpse. These do not co-occur in sentences, also cf. (Biemann et al., 2012).
- *WordNet* poor, *Co-occurrence* good: associations, like hospital–doctor or hospital–sick. They are not encoded in WordNet, its associative relations are very spotty. Note that placing them first in the order of relations did not increase performance.
- *WordNet* good, *Similarity* poor: meronyms/holonyms, such as door–knob, road–asphalt. These are not similar at all from a distributional point of view.
- *WordNet* poor, *Similarity* good: relations that should be in WN, but for some reason are missing, e.g., torpedo–missile, calligraphy–art, gazebo–pavilion.
- *Co-occurrence* good, *Similarity* poor: associations, part-of and cross-POS relations, such as orthodontist–braces, hospital–ER and growth–economy. Though being related, these words are not similar.
- *Co-occurrence* poor, *Similarity* good: (near) synonyms, such as mercenary–warrior, lampoon–caricature, orthodontist–dentist. Again,

they rarely co-occur in the same sentence.

8. Summary, Comments and Future Work

In this article, we have examined the use of lexical semantic networks to overcome the ToT problem. After an analysis of possible causes and a survey of existing work, we have evaluated and analyzed three lexical networks meant to overcome the ToT problem: WordNet, a word similarity network and a word co-occurrence network. Our setup was to query the network with a cue and check whether this would allow us to retrieve the target. To see its relative efficiency, we measured the rank of the ToT target over the retrieved vocabulary.

A ToT state can be induced by describing a given target to another person by providing some cues and ask him then to name it. Something similar can be achieved via crowdsourcing. We assumed that the cues retrieved via this technique, are similar to the ones humans typically use for the target retrieval. In order to determine the added value of a cue, we asked subjects to specify also the relationship between the target and each one of the given three cues. It turned out that traditional X-'onym' relations (hyponym, hypernym, ...) represent about half of the relations, while the remainder are mainly free associations, i.e., untyped relations.

While we could not successfully exploit relational information to enhance retrieval, we could show the relative efficiency of different lexical semantic networks with respect to word access. As expected, *WordNet* is very good for retrieving targets on the basis of synonyms or taxonomically related cues, but it scores much lower when it comes to syntagmatically related words[47]. *Word co-occurrence* excels in associations, qualities and typical

47 Similar conclusions have been reported elsewhere (Zock & Schwab, 2011 and 2016) whose authors describe an experiment comparing WN to a lexical resource created on the basis of Wikipedia. The latter performed a lot better than WN in terms of wordfinding, which is due to its far richer pool of syntagmatic links.

actions. Yet, the best network in our experiment was the one based on *word similarity*, as, apart for meronym/holonym relations, it combines the advantages of the other two. Hence, it covers basically the same aspects as WN, but it is more complete. Like the co-occurrence network, it contains many more syntagmatically associated terms.

The fact that WN does not perform well for syntagmatically related words is well known by the WN community who refers to it as the 'tennis problem' (Fellbaum, 1998)[48]. Actually, serious efforts have been made to enrich WN by adding syntagmatic links (Bentivogli et al., 2004) and various kinds of encyclopaedic information: topic signatures (Agirre et al. 2001), domain-specific information…[49]. Alas none of them seems to be integrated in the version accessible via the web[50]. Yet this is the one accessed by the ordinary language. Of course, one could also think of other solutions, for example, lexical functions (Mel'čuk, 2006). Actually, Mel'čuk's Explanatory Combinatory Dictionary (ECD) is probably better suited for our task than WN, all the more as it is a language production model, called 'Meaning-Text Model' (Mel'čuk, 2012), The ECD captures a much larger range of lexical relations (50+ lexical functions) than WN. Alas, the problem with the ECD is its coverage and availability. Though being extremely fine-grained the ECD covers so far only a subset of the words normally found in a lexicon. Also, it is not available in digital form.

Other potentially interesting alternatives would be association networks. Unfortunately, these resources are either not free (Gavagai)[51], too old (Kiss, et al. 1973), not rich enough in terms of coverage (de Deyne, et al. 2016; Nelson, et al. 2004), or not in the needed language, English. This holds in

48 For more on this and related problems, see (Polguère, 2014 ; Hanks, 2013 ; Wilks, et al. 1996).

49 Boyd-Graber et al., 2006; Gliozzo & Strapparava, 2008; Fernando, 2013, as well as : http://wndomains.fbk.eu/hierarchy.html

50 http://wordnetweb.princeton.edu/perl/webwn

51 https://explorer.gavagai.se and https://lexicon.gavagai.se

particular for JeuxDeMots (JdM) (Lafourcade, 2007, 2015), probably the largest, and arguably the best association thesaurus at this moment. JdM is a crowdsourced resource created via a game, hence its name 'word games' [52]. At the moment it has more than 4 million terms, and many more relations than WN, actually more than 80, falling into four broad categories: lexical, ontological, associative and predicative (Chatzikyriakidis et al., 2017). Alas, JdM's coverage of English is very small, and its website is in French which are probably two of the reasons why, alas, it is so little known 'abroad'.

Note that there is a particular class of association networks that might have interest for our work, free associations, or, more precisely, normed free associations[53]. Free word associations (WA) have been collected for decades and for many languages[54] by applying the following strategy: the experimentalist provides a stimulus word (cue) asking the participants to produce the first word coming to their mind. By doing so for a larger group, he will get an idea of what are typical answers, i.e., associations for a specific cue or stimulus word. For example, "light" was produced in more than 70% of the cases (Palermo and Jenkins, 1964) to the cue "lamp", followed in decreasing order by « shade, table, bulb » etc. whose associative strength is weaker.

While one may collect associations in the wild, i.e., from anyone, one may as well do so only for a population corresponding to a given norm (age,

52 www.jeuxdemots.org/jdm-accueil.php; and www.jeuxdemots.org/AKI.php

53 They are usually referred to as WAN, standing for 'Word Association Norm'. Yet, it would probably be better to call them NWAN, that is, **n**ormed **w**ord **a**ssociation **n**etworks. Actually, the tables capturing the relative associative strength of words (with respect to some input) can also be presented as directed graphs whose words are connected via untyped links, hence the term 'free association'.

54 For summaries and related word, see (Bel-Enguix, et al., 2019 ; Church & Hanks, 1990 ; De Deyne, et al. 2019; Ferrand & Alario, 1998 ; Im Walde, et al., 2008; Jenkins & Palermo, 1964 ; Jenkins, J.J. 1970 ; Lubaszewski, et al. 2017 ; Moss, et al. 1996; Nelson, et al. 1998 ; Postman &. Keppel, 1970; Reyes-Magaña, et al. 2019 ; Wettler, et al. 2005)

sex, nationality, …), and rank then the words in terms of some criteria, say frequency. The fact that a resource is normed has many advantages, and in our case, it could be used to parameterize the output a dictionary provides for a specific kind of user (children, adult, student, etc.). The problem is that we need to build the corresponding WANs, as the existing ones are probably too small to be representative[55]. Most of them have been built by hand. The problem with automatically created resources, or resources created via crowd-sourcing is control. Who are the contributors (Reyes-Magaña et al. 2019)? Last, but not least, the existing resources do not always correspond to our target groups. In sum, more work is needed.

One last word concerning 'relations'. Since we do believe in the virtues of relational information —they are a critical component of the input— we plan to revisit the problem of navigation in lexical graphs, but on the basis of cues enriched with relational information. Relations provide a context for the input. Revealing the users' goal, they tell the information provider (human or system) what to do with the input: provide a synonym, hypernym, etc. Obviously, a user expects quite different outputs for the following inputs : ['similar_to' (knife)], ['more general' than (knife)], or ['part_of' (knife)].

While typed relations are extremely important, we still need to keep untyped relations, as the user is not always able to tell the system what links the source to the target. While ignorance of the link type increases the search space, throwing untyped relations (free associations) over board risks to cut the branch we are sitting on, i.e., eliminate a whole set of words, possibly containing the target.

Concerning relations, we may also consider thematic roles, all the more as some of them are frequently used as cue words especially for named entities (typically found_at <location>; comes_from <country>, is used_for

55 But, see but see, Sinopalnikova & Smrz. 2006; Kwong, 2015; Reyes-Magaña et al. 2019.

<action>, etc.)[56]. Actually, a lot of work has been done on this topic since the seminal work of (Fillmore et al. 2003; Gildea & Jurafsky, 2002). For example, (Palmer, et al. 2005 ; Shen & Lapata, 2007; Kaisser & Webber, B. 2007 ; Young & Mitchell, 2017). The latter introduced a neural network approach enriched with word-word dependencies to predict the words' roles directly from a text.

Since our ultimate goal is the creation of a resource helping people to overcome the ToT problem, we plan to combine different types of corpora, possibly include named entities[57] to build then a hybrid semantic network, that is, an association thesaurus containing typed and untyped relations. The first to keep the search space small, the second to make it large enough to include potentially relevant words, possibly even our target.

As mentioned, knowledge states are highly volatile, varying from person to person and from moment to moment. In addition, when searching for a word, a user may have access to information coming from various levels[58]. This implies that we create different resources, one for each level, which, once combined allow us to capitalize on the various knowledge fragments in order to filter then the respective knowledge bases. Obviously, in order to do so, we must have access to the needed lexical resources. As one can see (again), there is still quite some work ahead of us.

---

56 Imagine that you were looking for 'Mozart', then the following associations could be useful : ACTIVITY: composer; ORIGIN: Austria; BORN: Vienna.

57 There exists already a large structured database, freely available (Nadeau & Sekine, 2007 ; Sekine & Nobata, 2004 ; http://nlp.cs.nyu.edu/ene/).

58 Let the target be 'incarnation', and the available information at the semantic and phonological level be the following: (a) *semantic level*: 'embodiment in a previous life'; (b) *phonological level*: word being composed of three segments of which the first and the last one are known [<in> <???><nation>]).

Acknowledgments

The authors are greateful to Benjamin Heinzerling, Simon de Deyne, and Massimo Stella for their valuable comments on a preliminary version of this document. Being carried out within the Labex BLRI (ANR-11-LABX-0036) and the Institut Convergence ILCB (ANR-16-CONV-0002), this work has been supported by the French National Agency for Research (ANR) and the Excellence Initiative of Aix-Marseille University (A*MIDEX).

References

Abrams, L., Trunk, D. L., and Margolin, S. J. 2007. Resolving tip-of-the-tongue states in young and older adults: The role of phonology. *In L. O. Randal (Ed.), Aging and the Elderly: Psychology, Sociology, and Health* (pp. 1-41). Hauppauge, NY: Nova Science Publishers, Inc.

Agirre, E., Ansa, O., Hovy, E., and Martinez, D. 2001. Enriching WordNet concepts with topic signatures. *In: SIGLEX workshop on WordNet and Other Lexical Resources: Applications, Extensions and Customizations*, Pittsburgh, PA, USA. http://arxiv.org/abs/cs.CL/0109031

Aitchison, J. 2003. *Words in the Mind: an Introduction to the Mental Lexicon*. Oxford, Blackwell.

Altmann, G.T.M. 1997. Words, and how we (eventually) find them. *The Ascent of Babel: An Exploration of Language, Mind, and Understanding*. Oxford: Oxford University Press. pp. 65–83.

Atkins, B. S., & Rundell, M. 2008. The Oxford guide to practical lexicography. Oxford University Press.

Atkins, B., Fillmore, C. J., Lowe, J. B., and Urban, N. 1994. The dictionary of the future: A hypertext database. In Presentation and on-line demonstration at the Xerox-Acquilex Symposium on the Dictionary of the Future. Uriage.

Baronchelli, A., Ferrer i Cancho, R. Pastor-Satorras, R. Chater, N. and Christiansen, M. H. 2013. Networks in Cognitive Science. Trends in Cognitive Sciences, 17.7: 348–360.

Baroni, M. and Lenci, A. 2010. Distributional memory: A general framework for corpus-based semantics. *Computational Linguistics*, 36(4):673–721.

Bateman J. and M. Zock. 2016 Natural Language Generation. In: *R. Mitkov (Ed.) Handbook of Computational Linguistics (2nd edition)*, Oxford University Press.

Beeferman, D. (1998). Lexical discovery with an enriched semantic network. In *Usage of WordNet in Natural Language Processing Systems.*

Bel-Enguix, G., Gómez-Adorno, H., Reyes-Magaña, J., & Sierra, G. 2019. Wan2vec: Embeddings learned on word association norms. *Semantic Web, 10(6)*, 991-1006.

Benson, M., Benson, E., and Ilson, R. 2010. The BBI Combinatory dictionary of

English. John Benjamins, Philadelphia.

Bentivogli, L. and Pianta, E. 2004. Extending WordNet with Syntagmatic Information. Sojka, P., Pala, K., Smrz, P., Fellbaum, C. & Vossen, P. (Eds.): GlobalWor(l)dNet Conference, Proceedings, pp. 47-53. Masaryk University, Brno

Bernstein, T. 1975. *Bernstein's Reverse dictionary.* Crown, New York.

Biemann, C. 2012. Structure discovery in natural language. Springer.

Biemann, C. and Riedl, M. 2013. Text: Now in 2D! A Framework for Lexical Expansion with Contextual Similarity. *Journal of Language Modeling* 1(1):55-95.

Biemann, C., Roos, S. and Weihe, K. 2012. Quantifying Semantics Using Complex Network Analysis. In: *Proceedings of COLING-12*, Mumbai, India, pp. 263-278.

Bock, J.K. 1996. Language production: Methods and methodologies. *Psychonomic Bulletin & Review*, 3:395-421.

Bock, K. & Griffin, Z. M. 2002. Producing words: How mind meets mouth. In Wheeldon, L. (Ed.). *Aspects of language production.* Psychology Press. 7-47.

Boissière, P. 1862. Dictionnaire analogique de la langue française : répertoire complet des mots par les idées et des idées par les mots, Paris

Bonin, P. 2004. Mental Lexicon: Some Words to Talk about Words. Nova Science Publishers

Boyd-Graber, J., Fellbaum, C., Osherson, D. & Schapire, R. 2006. Adding dense, weighted connections to WordNet. In Sojka, P., Choi, K.S., Fellbaum, C. & Vossen, P. (Eds.): Proceedings of the Third International WordNet Conference, GWC 2006, South Jeju Island, Korea, Masaryk University, Brno, pp. 29–35.

Brown, A. S. 1991. A review of the Tip-of-the-Tongue Experience. *Psychological Bulletin*, 109:204 – 223.

Brown, A. S. 2012. *The tip of the tongue state.* Taylor & Francis.

Brown, R. and Mc Neill, D. 1966. The tip of the tongue phenomenon. *Journal of Verbal Learning and Verbal Behaviour*, 5:325-337.

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. 2016. How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in*

psychology, 7, 1116.

Burke, D.M., MacKay, D.G.; Worthley, J. S. and Wade, E. 1991. On the Tip of the Tongue: What Causes Word Finding Failures in Young and Older Adults?. In: Journal of Memory and Language 30, 542-579.

Butterworth, B. 1989 Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical representation and process*. Cambridge, MA: MIT Press: I08-135.

Butterworth, B. 1992. Disorders of phonological encoding. *Cognition*, 42, 261–286.

Castro, N. and Stella, M. 2018. The multiplex structure of the mental lexicon influences picture naming in people with aphasia. *Journal of Complex Networks*, 7(6), 913-931.

Chatzikyriakidis, S., Lafourcade, M., Ramadier, L., and Zarrouk, M. 2017. Type Theories and Lexical Networks: Using Serious Games as the Basis for Multi-Sorted Typed Systems. Journal of Language Modeling, Vol. 5, No2, pp. 229–272, http://jlm.ipipan.waw.pl/index.php/JLM/issue/view/16

Church, K.W. & Hanks, P. 1990. Word association norms, mutual information, and lexicography. *Computational Linguistics* 16 (1), 22–29.

Cramer, P. 1968. *Word Associations*. New York: Academic Press.

Cutler, A. (Ed.) 1982. Slips of the Tongue and Language Production. Amsterdam: Mouton.

de Deyne, S. and Storms, G. 2015. Word associations. In J. R. Taylor (Ed.), *The Oxford Handbook of the Word*. Oxford University Press, Oxford, UK.

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., and Storms, G. 2019. The "Small World of Words" English word association norms for over 12,000 cue words. *Behavior research methods*, 51(3), 987-1006

de Deyne, S., Verheyen, S. and Storms, G. 2016. Structure and organization of the mental lexicon: A network approach derived from syntactic dependency relations and word associations. In: *Towards a Theoretical Framework for Analyzing Complex Linguistic Networks* (pp. 47-79). Springer Berlin.

Deese, J. (1984). Thought into Speech. Prentice Hall. Englewood Cliffs. NJ.

Deese, J. (1966). *The structure of associations in language and thought*. Johns Hopkins University Press.

de Saussure, F. 1972 [1916]. Cours de linguistique générale. Paris, Payot.

de Schryver, G.-M. 2003. Lexicographers dreams in the electronic-dictionary age. International Journal of Lexicography 16, 2, 143–199.

Dell, G. S., Nozari, N. and Oppenheim, G. M. 2014. Word production: Behavioral and computational considerations. Goldrick, M. A., Ferreira, V. S. & Miozzo, M. (Eds.). (2014). *The Oxford handbook of language production.* Oxford University Press, 88-104

Deuter, M. (Ed.). 2008. Oxford collocation dictionary: for students of English. Oxford University Press.

Di Sciullo, A. M. and Williams, E. 1987. On the definition of word (Linguistic Inquiry Monographs 14). *Cambridge, Massachusetts/London.*

Dong, Z. and Q. Dong. 2006. *HOWNET and the computation of meaning.* World Scientific, London.

Dornseiff, F. 2003. *Der deutsche Wortschatz nach Sachgruppen.* Berlin & New York: W. de Gruyter.

Dunning, T. 1993. Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1):61–74.

Durkin, P. (ed.). 2015. The Oxford Handbook of Lexicography. Oxford University Press.

Dutoit, D. and P. Nugues 2002. A lexical network and an algorithm to find words from definitions. In: *van Harmelen, F. (ed.): Proceedings of the 15th European Conference on Artificial Intelligence*, pp. 450-454 Lyon, France.

Fellbaum, C. (Ed.) 1998. *WordNet: An electronic lexical database and some of its applications.* MIT Press.

Fernando, S. 2013. *Enriching Lexical Knowledge Bases with Encyclopedic Relations* (Doctoral dissertation, University of Sheffield).

Ferrand, L., & Alario, F. X. (1998). Normes d'associations verbales pour 366 noms d'objets concrets. L'Année psychologique, 98(4), 659-709.

Fillmore, C., Johnson, C. and Petruck, M. 2003. Background to FrameNet. *International Journal of Lexicography* 16:235–250.

Findler N. (Ed.) 1979. Associative Networks: Representation and Use of Knowledge by Computers. Academic Press, Orlando.

Fontenelle, T. (Ed.). 2008. Practical lexicography: a reader. Oxford University Press.

Fontenelle, T. 1997. Using a bilingual dictionary to create semantic networks,

International Journal of Lexicography, Vol.10, n4, Oxford University Press, pp. 275-303

Freud, S. (1901). Zur Psychopathologie des Alltagslebens (Vergessen, Versprechen, Vergreifen) nebst Bemerkungen über eine Wurzel des Aberglaubens. pp. 1–16. *European Neurology*, 10(1), 1-16. (1966). The psychopathology of everyday life. WW Norton & Company

Fromkin V. (ed.). 1980. Errors in linguistic performance: Slips of the tongue, ear, pen and hand. San Francisco: Academic Press.

Galton, F. 1880. Psychometric experiments. *Brain*, 2, 149-162. DOI: 10.1093/brain/2.2.149

Garnham A., Shillock R. S., Brown G. D., Mill A., and Cutler A. 1982. Slips of the tongue in the London-Lund corpus of spontaneous conversations. in A. Cutler (Ed.). Slips of the tongue and language production. Berlin Mouton, 251-263.

Gatt, A., & Krahmer, E. (2018). Survey of the state of the art in natural language generation: Core tasks, applications and evaluation. *Journal of Artificial Intelligence Research, 61*, 65-170.

Gildea, D. and D. Jurafsky. 2002. Automatic labeling of semantic roles. *Computational Linguistics*, 28(3):245–288.

Gliozzo, A. and Strapparava, C. 2008. Semantic domains in computational linguistics. Springer.

Goldrick, M. 2014. Phonological Processing: The Retrieval and Encoding of Word Form Information in Speech Production. In Goldrick, M. A., Ferreira, V. S. & Miozzo, M. (Eds.). The Oxford handbook of language production. OUP, 228-244.

Goldrick, M. A., Ferreira, V. and Miozzo, M. 2014. *The Oxford handbook of language production*. Oxford University Press.

Grefenstette, G. and Tapanainen, P. 1994. What is a word, what is a sentence?: problems of Tokenisation.

Griffin, Z. M. and Ferreira, V. S. 2011. Properties of spoken language production. In Traxler, M. & Gernsbacher, M. A. (Eds.). *Handbook of psycholinguistics* (pp. 21-59). Academic Press.

Hankamer, J. 1989. Morphological parsing and the lexicon. In W. Marslen-Wilson (Ed.). Lexical representation and process (pp. 392-408). Cambridge, MA: MIT Press.

Hanks P., and Pustejovsky, J. 2005. A Pattern Dictionary for Natural Language Processing. Revue Française de linguistique appliquée Vol. 10 (2), 2005: 63-82.

Hanks, P. 2012. The corpus revolution in lexicography'. In *International Journal of Lexicography* 25 (4).

Hanks, P. 2013. Lexical analysis: Norms and exploitations. MIT Press.

Harley, T. A. and Bown, H. E. 1998. What causes a tip-of-the-tongue state? Evidence for lexical neighborhood effects in speech production. *British Journal of Psychology*, 89:151–174.

Hees, J. 2018. Simulating Human Associations with Linked Data – End-to- End Learning of Graph Patterns with an Evolutionary Algorithm, PhD dissertation, TU, Kaiserslautern

Hill, F., Cho, K., Korhonen, A. & Bengio, Y. 2016. Learning to understand phrases by embedding the dictionary. *Transactions of the Association for Computational Linguistics*, 4, 17-30.

Hörmann, H. 2013. Psycholinguistics: an introduction to research and theory. Springer Science & Business Media.

Hotopf, W.H.N. 1983. Lexical slips of the pen and tongue. In B. Butterworth (Ed.), *Language production, Vol. 2.* San Diego: Academic Press.

Humblé, P. 2001. *Dictionaries and Language Learners*, Haag and Herchen, Frankfurt am Main.

Im Walde, S. S., Melinger, A., Roth, M., & Weber, A. 2008. An empirical characterisation of response types in German association norms. *Research on Language and Computation, 6*(2), 205-238

Indefrey, P., & Levelt, W.J.M. 2000. The neural correlates of language production. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 845–865). Cambridge, MA: MIT Press.

James, L. and Burke, D. 2000. Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26:1378-1391.

Jenkins, J. J. & Palermo, D. S. 1964. Word Association Norms. Minn.:University of Minnesota Press.

Jenkins, J.J. 1970. The 1952 Minnesota word association norms. In: L. Postman; G. Keppel (eds.): *Norms of Word Association.* New York: Academic Press, 1-38.

Jescheniak, J. D. 2002. *Sprachproduktion. Der Zugriff auf das lexikale Gedächtnis beim Sprechen.* Göttingen: Hogrefe

Jung, C.G. 1910. The Association Method. In: *The American Journal of Psychology*, 21.2, pp. 219–269.

Kahn, J. 1989. *Reader's Digest Reverse Dictionary.* Reader's Digest, London.

Kaisser, M., & Webber, B. 2007. Question answering based on semantic roles. In *ACL 2007 Workshop on Deep Linguistic Processing* (pp. 41-48).

Kemmerer, D. 2015. Cognitive Neuroscience of Language. New York: Psychology Press.

Kenett, Y. N., Anaki, D., & Faust, M. (2014). Investigating the structure of semantic networks in low and high creative persons. *Frontiers in human neuroscience*, 8, 407, 1-16

Kilgarriff, A., P. Rychly, P. Smrz and D. Tugwell. 2004. The Sketch Engine. In Proceedings of the Eleventh EURALEX International Congress, pages 105–116, Lorient, France.

Kipper-Schuler, K. 2005. VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon. Ph.D. dissertation. University of Pennsylvania, Philadelphia, PA.

Kiss, G.R., Armstrong, C., Milroy, R. and Piper, J. 1973. An associative thesaurus of English and its computer analysis. In: *A. Aitken, R. Beiley and N. Hamilton-Smith (eds.): The Computer and Literary Studies*. Edinburgh: University Press. pp. 153-165

Krahmer, E. and Theune, M. (Eds.). 2010. *Empirical Methods in Natural Language Generation-Data-oriented Methods and Empirical Evaluation.* Series: Lecture Notes in Computer Science, Vol. 5790, Springer Berlin Heidelberg.

Kwong, O.Y. 2015. Enhancing word access in Chinese dictionaries: Implications from word association norms. *International Journal of lexicography*, 28(1), 62-85.

Lafourcade, M. 2007. Making people play for Lexical Acquisition with the JeuxDeMots prototype. In: *Proceedings of the 7th International Symposium on Natural Language Processing*, Pattaya, Chonburi, Thailand.

Lafourcade, M. and Joubert, A. 2015. TOTAKI: A help for lexical access on the TOT Problem. In Gala, N., Rapp, R. et Bel-Enguix, G. (Eds). *Language Production, Cognition, and the Lexicon.* Festschrift in honor of Michael Zock. Series Text, Speech and Language Technology XI. Dordrecht, Springer, pp. 95-

112

Lamb, S. M. 1999. Pathways of the brain: The neurocognitive basis of language. John Benjamins Publishing.

Landau, S. I. 2001. Dictionaries: The Art and Craft of Lexicography. 2nd ed. Cambridge University Press.

Leech, G., and Nesi, H. 1999. Moving towards perfection: the learners' (electronic) dictionary of the future. In T. Herbst and K. Popp (Eds). The Perfect Learners' Dictionary (?). Tubingen: Max Niemeyer Verlag. pp: 295-306

Lemaire, B. and Denhière, G. 2004. Incremental construction of an associative network from a corpus. *Proceedings of the 26th Annual Meeting of the Cognitive Science Society (CogSci'2004)*, 825-830

Leone, V, Siragusa, G., Di Caro, L. & Navigli, R. 2020. Building Semantic Grams of Human Knowledge. *Proceedings of the 12th Conference on Language Resources and Evaluation*, pages 2991–3000

Levelt W., Roelofs A. and A. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22:1-75.

Levelt, W. 1989. Speaking: *From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. 2001. Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences, 98*(23), 13464-13471.

Levelt, W. J. M. 1992. Accessing words in speech production: Stages, processes and representations. *Cognition*, 42:1-22.

Liu, H. and Singh, P. 2004. ConceptNet: A practical commonsense reasoning tool-kit. *BT Technology Journal* 22(4):211–226

Liu, X., Rafal Rzepka, R. and K. Araki. 2019. Creating Reverse Dictionary of English Idiomatic Expressions by Mapping Word Embeddings to Singular Vectors. Technical Report of JSAI Special Interest Group for Interactive Information Access and Visualization, SIG-AM 7, Yokohama, Japan.

Longman, A. 1993. *Language Activator: the world's first production dictionary.* Longman, London.

Loper, E., Yi, S.T., and Palmer, M. 2007. Combining lexical resources: Mapping between PropBank and VerbNet. In: J. Geertzen et coll. (eds.). Proceedings of the 7th International Workshop on Computational Semantics (= IWCS). Tilburg: ITK, Tilburg University, 118–129.

Lubaszewski, W., Gatkowska, I., & Godny, M. 2017. How a Word of a Text Selects the Related Words in a Human Association Network. In In Sharp, B., Sedes, F., & Lubaszewski, W. (Eds.). *Cognitive Approach to Natural Language Processing* (pp. 41-62). Elsevier.

Maclay, Y. and Osgood, C.E. 1959. Hesitation phenomena in spontaneous speech. *Word*, IS, 19-44.

Mandala, R., Tokunaga, T. and Tanaka, H. 1999. Complementing WordNet with Roget's and Corpus-based Thesauri for Information Retrieval. In: *Proceedings of EACL 99*, pp. 94-101, Bergen, Norway

Meara, P. 2009. Connected Words: word associations and second language lexical acquisition. Amsterdam: John Benjamins.

Mel'čuk I. , and Polguère, A. 2007. *Lexique actif du français : l'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. Champs linguistiques. De Boeck, Bruxelles.

Mel'čuk, I. 1996. Lexical Functions: A Tool for the Description of Lexical Relations in the Lexicon. In L. Wanner (ed), Lexical Functions in Lexicography and Natural Language Processing. Language Companion Series 31, Amsterdam/ Philadelphia: John Benjamins, 37–102.

Mel'čuk, I. 2006. Explanatory Combinatorial Dictionary. In G. Sica (ed), Open Problems in Linguistics and Lexicography. Monza: Polimetrica, 225–355.

Mel'čuk, I. 2012. Semantics: From meaning to text. Volume 1, Studies in Language Companion Series 129, Amsterdam/Philadelphia: John Benjamins.

Meyer, A. S. 2002. Form representations in word production. In Wheeldon, L. (Ed.). *Aspects of language production.* Psychology Press. (pp. 61-82).

Mihalcea, R., and Radev, D. 2011. Graph-based natural language processing and information retrieval. Cambridge University Press

Miller, G.A. (ed.). 1990. WordNet: An On-Line Lexical Database. *International Journal of Lexicography*, 3(4):235-244.

Miller, G.A. 1969. The organization of lexical memory: Are word associations sufficient. Talland, G. A., Waugh, N. C., & Adams, R. D. (Eds.). *The pathology of memory.* Academic Press Inc, 223-237.

Miller, G.A. 1991. The science of words. Scientific American Library.

Moerdijk F. 2008. Frames and semagrams; meaning description in the general Dutch dictionary. In: *Proceedings of the Thirteenth Euralex International*

Congress, EURALEX, pp. 561-569, Barcelona, Spain.

Morais, A. S., Olsson, H. and Schooler, L. J. 2013. Mapping the Structure of Semantic Memory. Cognitive Science, 37: 125–145.

Moss, H., Older, L., & Older, L. J. 1996. *Birkbeck word association norms.* Hove, UK: Psychology Press.

Motter, A. E., de Moura, A. P. S., Lai, Y.-C. and Dasgupta, P. 2002. Topology of the conceptual network of language. Physical Review E, 65.

Nadeau, D. and Sekine, S. 2007. A survey of named entity recognition and classification. *Lingvisticae Investigationes*, 30(1), 3-26.

Navigli, R. and Ponzetto, S. 2012. BabelNet: The Automatic Construction, Evaluation and Application of a Wide-Coverage Multilingual Semantic Network. *Artificial Intelligence* 193:217-250.

Nelson, D. L., C.L. McEvoy and T.A. Schreiber. 1998. The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/.

Nelson, D., McEvoy, C. and Schreiber, T. 2004. The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods* 36(3):402–407.

Nickels, L. A. 2002. When the words won't come: Relating impairments and models of speech production. In Wheeldon, L. (Ed.). *Aspects of language production.* (pp. 115-142). Psychology Press.

Oldfield R. C. 1963 Individual vocabulary and semantic currency: A preliminary study, British Journal of Social and Clinical Psychology, 211 122-130.

Ostermann C. 2015 . Cognitive Lexicography: A New Approach to Lexicography Making Use of Cognitive Semantics. De Gruyter

Palmer, M., Gildea, D., and Kingsbury, P. 2005. The proposition bank: An annotated corpus of semantic roles. Computational Linguistics 31, 71–106.

Pilehvar, M. T. 2019. On the Importance of Distinguishing Word Meaning Representations: A Case Study on Reverse Dictionary Mapping. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1* (pp. 2151-2156).

Polguère, A. 2014. From writing dictionaries to weaving lexical networks. *International Journal of Lexicography*, 27(4):396–418.

Posner, M., and Snyder, C. 1975. *Facilitation and inhibition in the processing of signal*. In S. Rabbitt& S. Dornic (Eds.), Attention and performance (pp. 669–682). New York: Academic Press.

Postman, L & G. Keppel (Eds.). 1970. Norms of Word Association. NY: Academic Press.

Pulvermüller, F. 1999. Words in the brain's language. In: *Behavioral and Brain Sciences 22*, 253–336.

Pulvermüller, F., 2002. *The neuroscience of language: On brain circuits of words and serial order*. Cambridge University Press.

Quasthoff, U., Richter, M. and Biemann, C. 2006. Corpus Portal for Search in Monolingual Corpora. *In: Proceedings of LREC-06*, pp. 1799-1802, Genoa, Italy.

Rapp, B. and Goldrick, M. 2006. Speaking words: Contributions of cognitive neuropsychological research. *Cognitive Neuropsychology*, 23 (1):39-73.

Rapp, B., & Goldrick, M. 2000. Discreteness and interactivity in spoken word production. *Psychological review, 107*(3), 460

Rapp, R. 2017. The Reverse Association Task. In Sharp, B., Sedes, F., & Lubaszewski, W. (Eds.). *Cognitive Approach to Natural Language Processing* (pp. 63-89). Elsevier.

Rastle, K. and D. Burke.1996. Priming the tip of the tongue: Effects of prior processing on word retrieval in young and older adults. Journal of Memory & Language, 35, 586-605.

Reiter, E. and Dale, R. 2000. *Building natural language generation systems* (Vol. 33). Cambridge: Cambridge University Press.

Reyes-Magaña, J., Bel-Enguix, G., Sierra, G., & Gómez-Adorno, H. 2019. Designing an electronic reverse dictionary based on two word association norms of English language. In *Proceedings of Electronic Lexicography in the 21st Century Conference* (Vol. 2019, pp. 865-880).

Richardson S, W., B. Dolan and L. Vanderwende. 1998. MindNet: Acquiring and Structuring Semantic Information from Text. In: *Proceedings of ACL-COLING'98*, pp. 1098-1102, Montreal, Canada.

Robert, P., Rey, A. & Rey-Debove, J. (1993) Dictionnaire alphabétique et analogique de la Langue Française. Paris : Le Robert

Robin, J. 1990. A Survey of Lexical Choice in Natural Language Generation,

Technical Report CUCS 040-90, Dept. of Computer Science, University of Columbia

Roelofs, A. 2003. Modeling the relation between the production and recognition of spoken word forms. *Phonetics and phonology in language comprehension and production: Differences and similarities*, 115-158.

Roelofs, A., Meyer, A. S. and Levelt, W. J. 1998. A case for the lemma/lexeme distinction in models of speaking: Comment on Caramazza and Miozzo (1997). *Cognition*, 69(2), 219-230.

Roget, P. 1852. *Thesaurus of English Words and Phrases.* Longman, London.

Rundell, M. 2007. The dictionary of the future. Optimizing the role of language in Technology-Enhanced Learning, 49.

Rundell, M., and Fox, G. (Eds.). 2002. *Macmillan English Dictionary for Advanced Learners.* Macmillan, Oxford.

Sahlgren, M. 2006. The Word-Space Model: Using distributional analysis to represent syntagmatic and paradigmatic relations between words in high-dimensional vector spaces. PhD thesis, Institutionen för lingvistik, Stockholm University.

Schriefers, H., Meyer, A. S. and Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, 29, 86–102.

Schulte im Walde. S. and Melinger, A. 2008. An in-depth look into the co-occurrence distribution of semantic associates. *Rivista di Linguistica, Special Issue on From Context to Meaning: Distributional Models of the Lexicon in Linguistics and Cognitive Science*, 20(1):89-128.

Schvaneveldt, R. (Ed.). 1989 Pathfinder Associative Networks : studies in knowledge organization. Ablex, Norwood, New Jersey

Schwartz, B. L. 2002. *Tip-of-the-tongue states: Phenomenology, mechanism, and lexical retrieval.* Mahwah. New Jersey: Lawrence Erlbaum Associates.

Schwartz, B. L. 2006. Tip-of-the-tongue states as metacognition. *Metacognition and Learning, 1*(2), 149-158.

Sekine, S. and Nobata, C. 2004. Definition, Dictionaries and Tagger for Extended Named Entity Hierarchy. In *Proceedings of Language Resources and Evaluation.*

Sérasset, G. 2012. DBnary: Wiktionary as a Lemon-based multilingual lexical

resource in RDF. Semantic Web Journal-Special issue on Multilingual Linked Open Data.

Shaw, R., Datta, A., VanderMeer, D. & Dutta, K. 2011. Building a scalable database-driven reverse dictionary. *IEEE Transactions on Knowledge and Data Engineering*, 25(3), 528-540.

Shen, D. & Lapata, M. 2007. Using semantic roles to improve question answering. In Proceedings of the Joint conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)(pp. 12-21).

Siew, C. S., Wulff, D. U., Beckage, N. M., & Kenett, Y. N. (2019). Cognitive Network Science: A review of research on cognition through the lens of network representations, processes, and dynamics. *Complexity, 2019*.

Sinclair, J. 1987. The dictionary of the future. Library Review 36, 4, 268–278.

Sinopalnikova, A. and P. Smrz. 2006. Knowing a word vs. accessing a word: WordNet and word association norms as interfaces to electronic dictionaries. Proceedings of the Third International WordNet Conference. Korea, 265–272

Speer, R., Chin, J. and Havasi, C. 2017. Concept net 5.5: An open multilingual graph of general knowledge. In *Thirty-First AAAI Conference on Artificial Intelligence.*

Spitzer, M. 1999. The mind within the net : models of learning, thinking and acting. A Bradford book. MIT Press, Cambridge

Sporns, O. 2011. Networks of the Brain. MIT Press, Cambridge.

Stede, M. 1995. Lexicalization in Natural Language Generation: a survey. Artificial Intelligence Review 8. pp. 309-336

Stemberger, J. 1985. The lexicon in a model of language production. New York, Garland Publishing

Steyvers, M. and J. Tenenbaum. 2005. Graph theoretic analyses of semantic networks: Small worlds in semantic networks. Cognitive Science, 29:41-78

Steyvers, M., Shiffrin, R., Nelson, D. 2002. Semantic spaces based on free association that predict memory performance. Festschrift honoring Lyle Bourne, Walter Kintsch, and Tom Landauer, 24.

Steyvers, M., Tenenbaum, J. 2010. The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. Cognitive science 29(1), 41–78

Storjohann, P. (Ed.). 2010. *Lexical-semantic relations: theoretical and practical perspectives* (Vol. 28). John Benjamins Publishing.

Strube, G. 1984. *Assoziation.* Berlin: Springer.

Suchanek, F.M., G. Kasneci, G. and Weikum, G. 2007. YAGO: A Core of Semantic Knowledge Unifying WordNet and Wikipedia. In: *WWW*. Banff, Alberta, Canada: ACM, pp. 697–706.

Svensén, B 2009. A Handbook of Lexicography: the theory and practice of dictionary-making, Cambridge: Cambridge University Press.

Thorat, S. and Choudhari, V. 2016. Implementing a Reverse Dictionary, based on word definitions, using a Node-Graph Architecture. In: *Proceedings of COLING 2016*, Osaka, Japan.

Trappes-Lomax, H. 1997. *Oxford Learner's Wordfinder Dictionary.* Oxford: Oxford University Press.

Tulving E. 1983. *Elements of Episodic Memory.* Oxford: Clarendon.

Turney, P., and Pantel, P. 2010. From frequency to meaning: Vector space models of semantics. Journal of artificial intelligence research 37, 1, 141–188.

Van Sterkenburg, P. (ed.) 2003. A Practical Guide to Lexicography. John Benjamins Publishing

van Sterkenburg, P. (ed) 2003. A Practical Guide to Lexicography. John Benjamins, Amsterdam.

van Turennout, M., Hagoort, P., & Brown, C. M. 1998. Brain activity during speaking: From syntax to phonology in 40 milliseconds. Science, 280(5363), 572-574.

Vigliocco, G.; Antonini, T. and Garrett, M. F. 1997. Grammatical gender is on the tip of Italian tongues. Psychological Science, 8, 314-317.

Vitevitch, M. S. 2008. What can graph theory tell us about word learning and lexical retrieval? Journal of Speech, Language, and Hearing Research, 51, 408–422.

Vitevitch, M., Goldstein, R., Siew, C. & N. Castro 2014. Using complex networks to understand the mental lexicon. Yearbook of the Poznań Linguistic Meeting, pp. 119–138

Wanner, L. 1996. Lexical Choice in Text Generation and Machine Translation. Special Issue on Lexical Choice. Machine Translation. L. W. (ed.). Dordrecht,

Kluwer Academic Publishers. 11: 3-35.

Wettler, M. and Rapp, R. 1993. Computation of word associations based on the co-occurrences of words in large corpora. In: *Proceedings of the 1st Workshop on Very Large Corpora*, pp. 84-93, Beijing, China.

Wettler, M., Rapp, R. and Sedlmeier, P. 2005. Free word associations correspond to contiguities between words in texts. *Journal of Quantitative Linguistics, 12*(2-3).

Wheeldon, L. (Ed.). 2002. Aspects of language production. Psychology Press.

Widdows, D. 2004. Geometry of Meaning. University of Chicago Press.

Wilks, Y., Slator, B. and Guthrie, L. 1996. Electric Words : Dictionaries, Computers, and Meanings. Cambridge : The MIT Press.

Yang, B. and Mitchell, T. 2017. A joint sequential and relational model for frame-semantic parsing. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. (pp. 1247-1256).

Zhang, L., Qi, F., Liu, Z., Wang, Y., Liu, Q. & Sun, M. 2019. Multi-channel Reverse Dictionary Model. *arXiv preprint arXiv:1912.08441.*

Zock, M. 2002. Sorry, what was your name again, or how to overcome the tip-of-the tongue problem with the help of a computer? In: Proc. of the *SemaNet workshop*, COLING-2002.

Zock, M. 2019. *Eureka! A Simple Solution the Complex 'Tip-of-the-Tongue'-Problem.* In Bastardas-Boada, A; Massip Bonet, A & Bel-Enguix, G. (eds.). Complexity Applications in Language and Communication, Springer. Berlin, pp. 251-272

Zock, M. 2015. *Introduction* to the special issue of 'Cognitive Aspects of Natural Language Processing", title: Words in Books, Computers and the Human Mind. Journal Journal of Cognitive Science. 16-4: 355-378. Institute for Cognitive Science, Seoul National University (http://j-cs.org/gnuboard/bbs/board.php?bo_table=__vol016i4)

Zock, M. and D. Schwab. 2011. Storage does not guarantee access. The problem of organizing and accessing words in a speaker's lexicon. *Journal of Cognitive Science* 12(3):233-258.

Zock, M. and Biemann, C. 2016. Towards a resource based on users' knowledge to overcome the Tip-of-the-Tongue problem. Proceedings of the COLING Workshop Cognitive Aspects of the Lexicon (CogALex-V), Osaka, Japan, pp. 57-68

Zock, M. and Schwab, D. 2016. *WordNet and beyond*. In Fellbaum, C., Vosse, P., Mititelu, V. and D. Forăscu (Eds.). 8th International Global WordNet conference. Bucharest (http://gwc2016.racai.ro), pp. 436-444

Zock, M. and Tesfaye, D. 2015. *Automatic creation of a semantic network encoding part_of relations*. In Journal of Cognitive Science 16, 4, pp. 431-491, Institute for Cognitive Science, Seoul National University

Zock, M., Ferret, O. and D. Schwab. 2010. *Deliberate word access : an intuition, a roadmap and some preliminary empirical results*. In A. Neustein (Ed.) 'International Journal of Speech Technology', 13(4), pp. 201-218, 2010. Springer Verlag

Zortea, M., Menegola, B., Villavicencio, A., & Salles, J. F. D. (2014). Graph analysis of semantic word association among children, adults, and the elderly. *Psicol. Reflex. Crit. 27, 27*(1), 90-99.