

# Wordfinding Problems and How to Overcome them Ultimately With the Help of a Computer

Michael Zock

LIF-CNRS / TALEP

163, Avenue de Luminy

13288 Marseille / France

michael.zock@lif.univ-mrs.fr

## Abstract

Our ultimate goal is to help authors to find an elusive word. Whenever we need a word, we look it up in the place where it is stored, the dictionary or the mental lexicon. The question is how do we manage to find the word, and how do we succeed to do this so quickly? While these are difficult questions, I believe to have some practical answers for them. Since it is unreasonable to perform search in the entire lexicon, I suggest to start by reducing this space (step-1) and to present then the remaining candidates in a clustered and labeled form, i.e. categorial tree (step-2). The goal of this second step is to support navigation.

Search space is determined by considering words directly related to the input, i.e. direct neighbors (associations/co-occurrences). To this end many resources could be used. For example, one may consider an associative network like the Edinburgh Association Thesaurus (E.A.T.). As this will still yield too many hits, I suggest to cluster and label the outputs. This labeling is crucial for navigation, as we want users to find the target quickly, rather than drown them under a huge, unstructured list of words. Note, that in order to determine properly the initial search space (step-1), we must have already well understood the input [*mouse*<sub>1</sub> / *mouse*<sub>2</sub> (rodent/device)], as otherwise our list will contain a lot of noise, presenting 'cat, cheese' together with 'computer, mouse pad', which is not quite what we want, since some of these candidates are irrelevant, i.e. beyond the scope of the user's goal.

## 1 Introduction

Whenever we read a book, write a letter, or launch a query on Google, we always use words, the short-hand labels for more or less well specified thoughts. No doubt, words are important, a fact nicely expressed by Wilkins (1972) when he writes: *without grammar very little can be conveyed, without vocabulary, nothing can be conveyed*. Still, ubiquitous as they may be, words have to be learned, that is, they have to be stored, remembered, and retrieved. Given the role words play in our daily lives, it is surprising to see how little we have to offer so far to help humans to memorize, find or retrieve them. Hoping to contribute to a change for this, I have started to work on one of these tasks: word access, also called *retrieval* or *wordfinding*.

Imagine the following situation: your goal is to express the following ideas: *superior dark coffee made of beans from Arabia* by a single word, but you cannot access the corresponding form *mocha*, even though you know it, since you've used it not so long ago. This kind of problem, known as the *tip-of-the-tongue* (TOT)-problem, has received a lot of attention from psychologists (Schwartz, 2002; Brown, 1991). It has always been pointed out that people being in this state know quite a bit concerning the elusive word (Brown and McNeill, 1996). Hence, using it should allow us to reduce the search space. Put differently, it would be nice to have a system capable to use whatever you have, incomplete as it may be, to help you find what you cannot recall. For example, for the case at hand, one might think of *dark, coffee, beans, and Arabia*, to expect from system a set of reasonable candidates, like *arabica, espresso, or mocha*. In the remainder of this paper I will try to show how this might be achieved, but before doing

---

This work is licensed under a Creative Commons Attribution 4.0 International Licence. Page numbers and proceedings footer are added by the organisers. Licence details: <http://creativecommons.org/licenses/by/4.0/>

so, I would like to clarify what I mean by *computer-aided lexical access*, what characterizes the problem of word production, i.e. the process.

## 2 Computer-aided lexical access

Under normal circumstances, words are accessed on the fly, that is, the lexical access is immediate, involuntary and autonomous. Also, it takes place without any external help. As we all know, things do not always work that smoothly, which is why we may ask for help. In this latter case, lexical access is deliberate, incremental (i.e., distributed over time), and may be mediated via some external resource (another person or a dictionary). This situation may well arise in writing, where we are much more demanding and where we have much more time. Hence words are chosen with much more care than during speaking, i.e., spontaneous discourse.

I view *computer-aided lexical access* as an interactive, cognitive process. It is *interactive* as it involves two cooperative agents, the user and the computer, and it is *cognitive* as it is largely knowledge-driven. The knowledge concerns words, i.e. meanings and forms, as well as their relations to other words. Since the knowledge of both agents is incomplete, they cooperate: neither of them alone can point to the target word ( $t_w$ ), but by working together they can. It is as if one had the (semantic) *map* and the other the *compass*, i.e., the knowledge to decide where to go. Since both types of knowledge are necessary, they complete each other, helping ultimately the user to find the elusive word, which is the goal.

To be more concrete, consider some user input (one or several words), the system reacts by providing all directly associated words. Since all words are linked, they form a graph, which has two major consequences : the system knows *everyone*, the immediate neighbors, the neighbors' neighbors, etc. and the user can initiate search from anywhere, to continue it until he has reached the target word,  $t_w$ . Everything being connected, everything is reachable, at least in principle. Search may require several steps, but in most cases the number of steps is surprisingly small.

As mentioned already, the user definitely has some knowledge concerning words, their components and their organisation in the mental lexicon, but this knowledge is by no means complete. The user also has some knowledge (or, more precisely, meta-knowledge) concerning the topology of the graph,<sup>1</sup> but he certainly does not know as much as the system. The fact that an author does have this kind of knowledge is revealed via word associations (Cramer, 1968; Deese, 1965; Nelson et al., 1998; Kiss et al., 1972) and via the observed average path length (Vitevitch, 2008) needed in order to get from some starting point ( $s_w$ ) to the goal ( $t_w$ ). This path is generally quite short. It hardly ever exceeds three steps, and in many cases even less: search is launched via an item directly related to the  $t_w$  (direct neighbor).

If the user does not know too much concerning the topology of the network, he does know quite a bit concerning the  $t_w$ ,<sup>2</sup> information the system has no clue of at this point. Eventhough it knows 'everyone' in the network, it cannot do mind-reading, i.e. guess the precise word a user has in mind ( $t_w$ ) when providing a specific input ( $s_w$ ). Yet the user can. Even if he cannot access the word at a given moment, he can recognize it when seeing it (alone or in a list). This fact is well established in the literature on the 'tip-of-the-tongue problem' (Aitchison, 2003).

## 3 From mind to mouth, or what characterizes the process of word production?

According to the father of modern linguistics (de Saussure, 1916), word forms (signifier) and their associated meaning (signified) are but one, called the sign. They are said to be an inseparable unit. This is in sharp contrast to what psychologists tell us about words synthesis. For example, one of the leading specialists of language production (Levelt, 1989; Levelt, 1999) has convincingly shown that, when speaking

<sup>1</sup>For example, he knows that for a given word form there are similar forms in terms of sound or meaning. There are also words that are more general/specific, or others meaning exactly the opposite than a given input. This kind of knowledge is so obvious and so frequent that it is encoded in many resources like WordNet, Roget's thesaurus or more traditional dictionaries (including synonym and rhyming dictionaries).

<sup>2</sup>For example, parts of the *form* (rhymes with x: health/wealth) or *meaning*, like the 'type' (animal), the 'function' (used for eating) or the 'relationship' (synonym, antonym, ...) with respect the source word ( $s_w$ ). He may even be able to provide parts of the definition (say, 'very small' for 'liliput'). His main problem resides in the fact that he cannot access at this very moment the exact word form (he experiences the so called 'tip-of-the-tongue problem, TOT), which is why he tries to find it in a lexical resource (dictionary).

we go, step by step, from *meanings* (concepts), to the *lexical concept* (also called lemma) to the *sound* (written or spoken form). Depending on the theory, there may be retroaction or not, a lower level, say, phonology, influencing a higher level, the lexical concept.

Note that the notion of *lemma* has completely different meanings in psychology and in lexicography. While for linguists it is roughly speaking the word's base-form or dictionary-form, for psycholinguists it is a schema, i.e. an abstract form representing a specific meaning (a lexicalized concept) and a syntactic category (part of speech), but it lacks entirely specific values concerning the form (sounds/graphemes). This is being taken care of at the next step (sound form encoding). In short, in contrast to Saussure's view, the information contributing to what we commonly call words (lemma or word forms) is distributed. This is a well established empirical fact observed by psychologists working on the time course of word production (Stemberger, 1985; Levelt and Schriefers, 1987; Dell et al., 1999), as by those who analyze and interpret speech errors (Fromkin, 1973; Fromkin, 1980; Fromkin, 1993).

Yet, what concerns us here in particular is the following: as noted, speakers go from meanings to sounds via lexical concepts (abstract word forms). More importantly, the conceptual input may lack information to determine a precise lexical form. Put differently, rather than starting from a full fledged definition or complete meaning representation, authors may well start from an underspecified input ('small bird' rather than 'sparrow'). Note that the specific requirements of a culture may help us to clarify our thoughts, as well as induce biases or imprecisions because of lexical gaps. Hence we end up using an existing word (eventhough it does not express exactly what we had in mind) rather than coining a new one fitting better our purpose (expressibility problem). For a psycholinguistic explanation concerning gradual refinement, see (Zock, 1996).

Let me briefly illustrate this here via an example, and comment then on the way how specific knowledge states may ask for different kind of information from the lexicon. Suppose you wanted to talk about a given reptile having certain features (dangerous, size, living space, ...). If you cannot come up immediately with the intended word, any of the following could be candidates: alligator, crocodile, cayman. At some point you need to make up your mind though, as the form synthesizer needs to know what items to activate so that it can produce the corresponding form (graphemes, sounds).

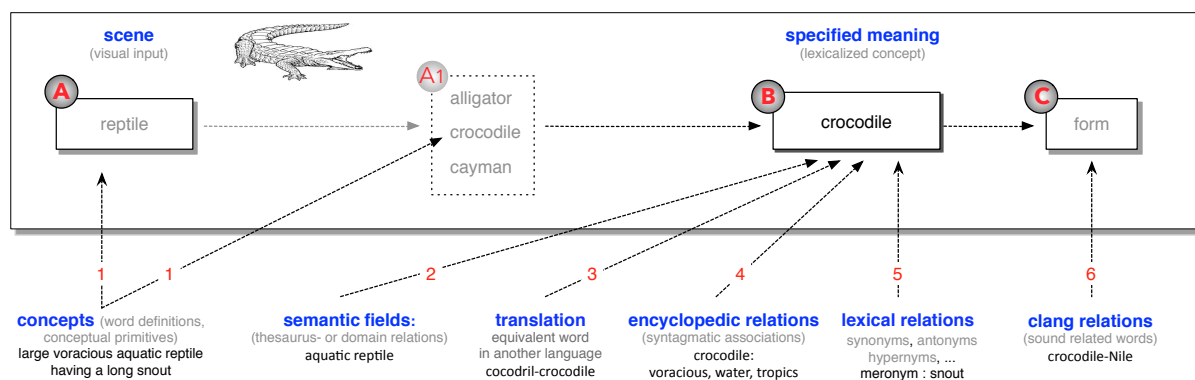


Figure 1: Underspecified input and progressive refinement

As we can see in the figure above, there are two critical moments in word production: meaning *specification* (A-B) and *sound-form encoding* (B-C). It is generally this latter part that poses problems. How to resolve it has been nicely illustrated in an experiment done by (James and Burke, 2000). They showed that phonologically similar words of the target could resolve the TOT state. To show this they put participants into the TOT state by presenting them low-frequency words: *abdicate*, *amnesty*, *anagram*,.... Those who failed were used for the experiment. Next the experimenters read a list of words containing parts of the syllables of the TOT word. For example, if the definition 'to renounce a throne' put a participant into a TOT state, he was asked to read aloud a list of ten words, like *abstract*, *indigent*, *truncate*, each of which contains a syllable of the target. For the other half, participants were given a list of 10 phonologically unrelated words. After that participants were primed again to produce the elusive word (*abdicate*). As the results clearly showed those who were asked to read phonologically related words resolved more

TOT states than those who were presented with unrelated words.

This is a nice example. Alas, we cannot make use of it, as, not knowing the target word we cannot increase (directly) the activation level of the phonological form. Hence we have to resort to another method, namely, association networks (see section 3). Let us see how search strategies may depend on cognitive states.

#### 4 Search strategies function of variable cognitive states

Search is always based on knowledge. Depending on the knowledge available at the onset one will perform a specific kind of search. Put differently, there are different information needs as there are different search strategies.

There are at least three things that authors typically know when looking for a specific word: its *meaning* (definition) or at least part of it (this is the most frequent situation), its *lexical relations* (hyponymy, synonymy, antonymy, etc.), and the collocational or encyclopedic relations it entertains with other words (Paris-city, Paris-French capital, etc.). Hence there are several ways to access a word (see Figure 1): via its meaning (concepts, meaning fragments), via syntagmatic links (thesaurus- or encyclopedic relations), via its form (rhymes), via lexical relations, via syntactic patterns (search in a corpus), and, of course, via another language (translation). Note that access by meaning is the golden route, i.e. the most normal way. We tend to use other means only if we fail to access straight away the desired word.

I will consider here only one of them, word associations (mostly, encyclopaedic relations). Note that, people being in the TOT-state clearly know more than that. Psychologists who have studied this phenomenon (Brown and McNeill, 1996; Vigliocco et al., 1997) have found that their subjects had access not only to meanings (the words definition), but also to information concerning grammar (gender) and lexical form: sound, morphology and part of speech. While all this information could be used to constrain the search space, the ideal dictionary being multiply indexed, I will deal here only with semantically related words (associations, collocations in the large sense of the word). Before discussing how such a dictionary could be built and used, let us consider a possible search scenario.

I start from the assumption that in our mind, all words are connected, the mental lexicon (brain) being a network. This being so, anything can be reached from anywhere. The user enters the graph by providing whatever comes to his mind (source-word), following the links until he has reached the target. As has been shown (Motter et al., 2002), our mental lexicon has small-world properties: very few steps are needed to get from the source-word to the target word. Another assumption I make is the following: when looking for a word, people tend to start from a close neighbour, which implies that users have some meta-knowledge containing the topology of the network (or the structure of their mental lexicon): what are the nodes, how are they linked to their neighbours, and what are more or less direct neighbours? For example, we know that black is related to white, and that both words are fairly close, at least a lot closer than, say, black and flower.

Search can be viewed as a dialogue. The user provides as input the words that a concept he wishes to express evokes, and the system displays then all (directly) connected words. If this list contains the target search stops, otherwise it will continue. The user chooses a word of the list, or keys in an entirely different word. The first part described is the simplest case: the target is a direct neighbour. The second addresses the problem of indirect associations, the distance being bigger than 1.

Before presenting our method in section 3, let us say a few words about existing resources. Since the conversion of meaning to sounds is mediated via a lexicon, one may wonder to what extent existing resources can be of help.

#### 5 Related work

While there are many kinds of dictionaries or lexical resources, very few of them can be said to meet truly the authors' needs. To be fair though, one must admit that great efforts have been made to improve the situation both with respect to lexical resources and electronic dictionaries. In fact, there are quite a few *onomasiological dictionaries* (van Sterkenburg, 2003). For example, Roget's Thesaurus (Roget, 1852), analogical dictionaries (Boissière, 1862; Robert et al., 1993), Longman's Language Activator

(Summers, 1993), various network-based dictionaries: WordNet (Fellbaum, 1998; Miller et al., 1990), MindNet (Richardson et al., 1998), HowNet (Dong and Dong, 2006), Pathfinder (Schvaneveldt, 1989), 'The active vocabulary for French' (Mel'čuk and Polguère, 2007) and Fontenelle (Fontenelle, 1997). Other proposals have been made by Sierra (Sierra, 2000) and Moerdijk (2008). There are also various collocation dictionaries (Benson et al., 2010), reverse dictionaries (Bernstein, 1975; Kahn, 1989; Edmonds, 1999) and OneLook,<sup>3</sup> which combines a dictionary (WordNet) and an encyclopedia (Wikipedia). Finally, there is MEDAL (Rundell and Fox, 2002), a thesaurus produced with the help of Kilgarriff's Sketch Engine (Kilgarriff et al., 2004). There has also been quite a lot of work on the time-course of word production, i.e. the way how one gets progressively from a more or less precise idea to its expression, a word expressed in written or spoken form. See for example (Levelt et al., 1999; Dell et al., 1999). Clearly, a lot of progress has been made during the last two decades, yet more can be done especially with respect to indexing (the organization of the data) and navigation.

Two key ideas underlying modern lexical resources are the notions of 'graphs' and 'association'. For a useful introduction to graph-based natural language processing, see (Mihalcea and Radev, 2011). Associations have a long history. The idea according to which the mental lexicon (or encyclopedia) is basically an associative network, composed of nodes (words or concepts) and links (associations) is not new at all. Actually the very notion of association goes back at least to Aristotle (350BC), but it is also inherent in work done by philosophers (Locke, Hume), physiologists (James & Stuart Mills), psychologists (Galton, 1880; Freud, 1901; Jung and Riklin, 1906) and psycholinguists (Deese, 1965; Jenkins, 1970; Schvaneveldt, 1989). For good introductions see (Hörmann, 1972; Cramer, 1968) and more recently (Spitzer, 1999). The notion of association is also implicit in work on semantic networks (Quillian, 1968), hypertext (Bush, 1945), the web (Nelson, 1967), connectionism (Dell et al., 1999) and, of course, in WordNet (Miller, 1990; Fellbaum, 1998).

## 6 The framework for building and using our resource

To understand the problems at stake, I describe the communicative setting (system, user), the existing and necessary components, as well as the information flow (see figure 2).

Imagine an author wishing to convey the name of a special beverage ('mocha') commonly found in coffee shops. Failing to do so, he tries to find it in a lexicon. Since dictionaries are too huge to be scanned from beginning to the end, I suggest another solution: reduce the search space based on some input (step-1) and presentation of the results (all directly related words) in a clustered form (step-2). More concretely speaking, I suggest to have a system that accepts whatever comes to an author's mind, say 'coffee' in our 'mocha' case, to present then all directly associated words. Put differently, given some cue, we want the system to guess the user's goal (the elusive word). If this list contains the target, search stops, otherwise the user will pick one of the associated terms or provide an entirely new word and the whole process is repeated again, that is, the system will come up with a new set of proposals.

What I've just described here corresponds to step-1 in figure 2 (see next page). While there are a number of resources that one could use to allow for this transition, I rely here on the E.A.T., i.e. the 'Edinburgh Association Thesaurus'. Note that the output produced by this resource is still too big to be really useful. Suppose that each input word yielded 50 outputs (the EAT often presents 100, and one could think of a lot more). Having provided three words the system will return 150 outputs. Actually, it will take an intersection of the associated words to avoid redundancies. Since this list is still too big to be scanned linearly (one by one), I suggest to structure it, by clustering words into categories (step-2).

This yields a tree whose leaves are words (our potential targets) and whose nodes are categories, that is, also words, but with a completely different status, namely to group words. Category names function like signposts, signalling the user the direction to go. Note that it is not the system that decides on the direction, but the user. Seeing the names of the categories he can make reasonable guesses concerning their content. Categories act somehow like signposts signaling the user the kind of words he is likely to find if he goes one way or another. Indeed, knowing the name of a category (fruit, animal), the user can guess the kind of words contained in each bag, a prediction which is all the more likely as each

---

<sup>3</sup><http://onelook.com/reverse-dictionary.shtml>

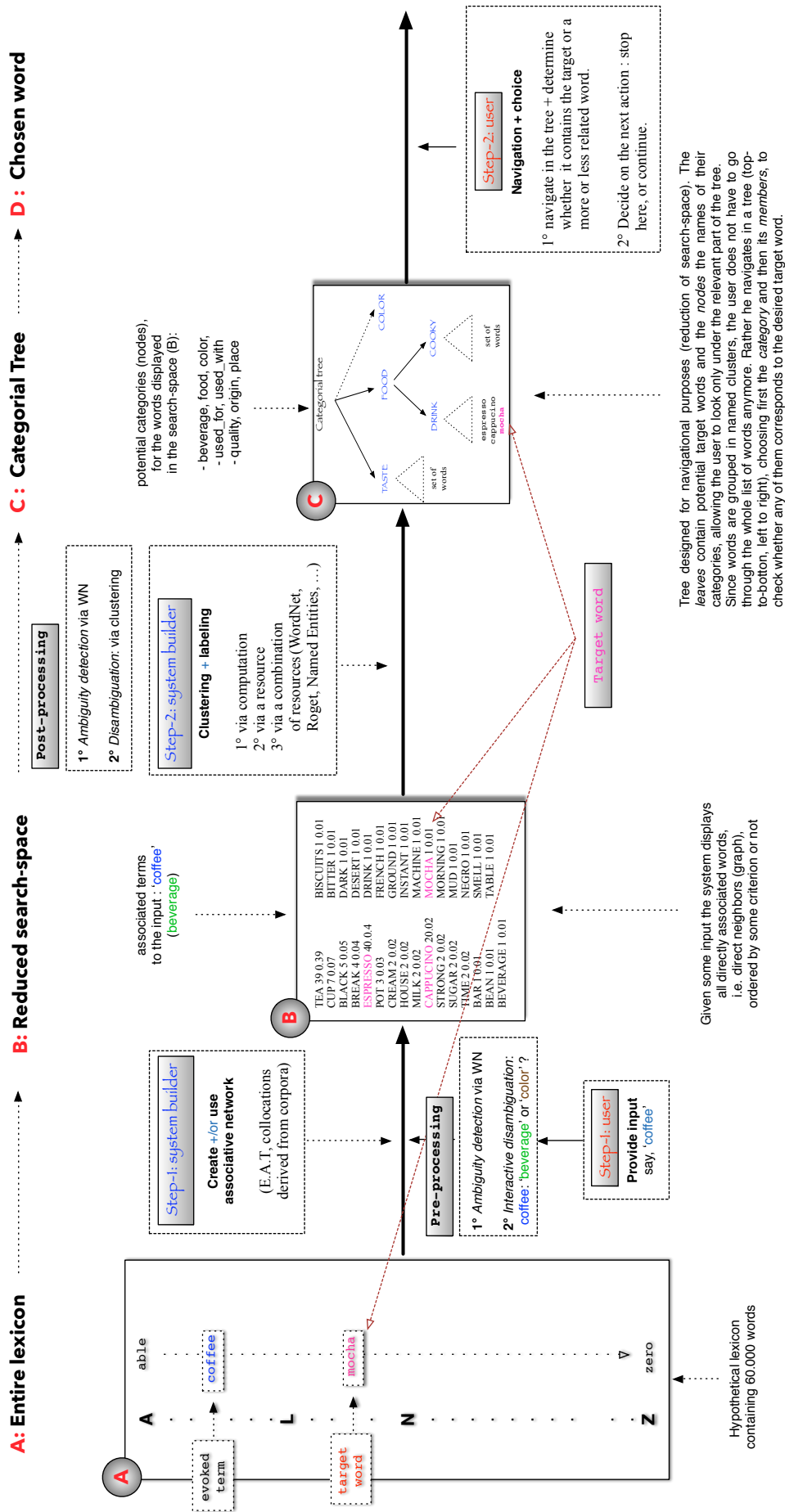


Figure 2: Architecture of the components and information flow

category contains only terms directly associated with the source word. Assuming that the user knows the category of the searched word,<sup>4</sup> he should be able to look in the right bag and take the best turn. Navigating in a categorial tree, the user can search at a fairly high level (class) rather than at the level of words (instances). This reduces not only the cognitive load, but it increases also chances of finding the target, while speeding up search, i.e. the time needed to find a word.

Remains the question of how to build this resource and how to accomplish these two steps. I have explained already the first transition going from A-B. The system enriches the input by taking all associated words, words he will find in the EAT. Obviously, other strategies are possible, and this is precisely one of the points I would like to experiment with in the future : check which knowledge source (corpus, association thesaurus, lexical resource) produces the best set of candidates, i.e. the best search space and the best structure in order to navigate. The solution of the second step is quite a bit more complicated, as putting words into clusters is one thing, naming them is another. Yet, arguably this is a crucial step, as it allows the user to navigate on this basis. Of course, one could question the very need of labels, and perhaps this is not too much of an issue if we have only say, 3-4 categories. I am nevertheless strongly convinced that the problem is real, as soon as the number of categories (hence the words to be classified) grows. To conclude, I think it is fair to say that the first stage is clearly within reach, while the automatic construction of the categorial tree remains a true challenge despite the vast literature devoted to this topic or to strongly related problems (Zhang et al., 2012; Biemann, 2012; Everitt et al., 2011).

## 7 Outlook and conclusion

I have started from the observation that words are important and that their accessibility can be a problem. In order to help a dictionary user to overcome it I have presented a method showing promise. In particular, I have shown how to reduce the search space, how to present a set of plausible candidates and what needs to be done next (clustering and naming them) to reduce the search space and to support navigation. In particular, I have proposed the creation of a categorial tree whose leaves contain the (potential target) words and the nodes the names of their categories. The role of the latter is to avoid the user to search in non relevant parts of the tree. Since words are grouped in named clusters, the user does not have to go through the whole list of words anymore. Rather he navigates in a tree (top-to-bottom, left to right), choosing first the category and then its members, to check whether any of them corresponds to the desired target word.

Even if the details of this work turn out to be wrong (this is just preliminary work), I believe and hope that the overall framework is of the right sort, allowing for a rich set of experimentation in particular with respect to determining the search space and the clustering. Concerning evaluation, the ultimate judge will be, of course, the user, as only s/he can tell us whether our resource fits his/her needs or goals.

## References

- Jean Aitchison. 2003. *Words in the Mind: an Introduction to the Mental Lexicon (3d edition)*. Blackwell, Oxford.
- Morton Benson, Evelyn Benson, and Robert A Ilson. 2010. *The BBI Combinatory Dictionary of English*. John Benjamins, Philadelphia.
- Theodore Bernstein. 1975. *Bernstein's Reverse Dictionary*. Crown, New York.
- Chris Biemann. 2012. *Structure Discovery in Natural Language*. Springer.
- Jean Baptiste Prudence Boissière. 1862. *Dictionnaire analogique de la langue française : répertoire complet des mots par les idées et des idées par les mots*. Larousse et A. Boyer, Paris.
- Roger Brown and David McNeill. 1996. The tip of the tongue phenomenon. *Journal of Verbal Learning and Verbal Behaviour*, 5:325–337.
- Allan S. Brown. 1991. The tip of the tongue experience a review and evaluation. *Psychological Bulletin*, 10:204–223.

---

<sup>4</sup>A fact which has been systematically observed for people being in the 'tip of the tongue state' who may tell the listener that they are looking for the name of "a fruit typically found in PLACE", in order to get 'kiwi'.

- Vannevar Bush. 1945. As we may think. *The Atlantic Monthly*, 176:101–108.
- Phebe Cramer. 1968. *Word association*. Academic Press, New York.
- Ferdinand de Saussure. 1916. *Cours de linguistique générale*. Payot, Paris.
- James Deese. 1965. *The structure of associations in language and thought*. Johns Hopkins Press.
- Gary Dell, Franklin Chang, and Zenzi M. Griffin. 1999. Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, 23:517–542.
- Zhendong Dong and Qiang Dong. 2006. *HOWNET and the computation of meaning*. World Scientific, London.
- David Edmonds, editor. 1999. *The Oxford Reverse Dictionary*. Oxford University Press, Oxford, Oxford.
- Brian S. Everitt, Sabine Landau, Morven Leese, and Daniel Stahl. 2011. *Cluster Analysis*. John Wiley and Sons.
- Christiane Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database and some of its Applications*. MIT Press.
- Thierry Fontenelle. 1997. *Turning a Bilingual Dictionary into a Lexical-Semantic Database*. Max Niemeyer, Tübingen.
- Sigmund Freud. 1901. *Psychopathology of everyday life*. Payot, Paris, 1997 edition.
- Victoria Fromkin, editor. 1973. *Speech errors as linguistic evidence*. Mouton, The Hague.
- Victoria Fromkin. 1980. Errors in linguistic performance: Slips of the tongue, ear, pen and hand.
- Victoria Fromkin. 1993. Speech production. In J. Berko-Gleason and N. Bernstein Ratner, editors, *Psycholinguistics*. Harcourt, Brace, Jovanovich, Fort Worth, TX.
- Francis Galton. 1880. Psychometric experiments. *Brain*, 2:149–162.
- Hans Hörmann. 1972. *Introduction à la psycholinguistique*. Larousse, Paris, France.
- Lori James and Deborah Burke. 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*, 6(26):1378–1391.
- James Jenkins. 1970. The 1952 Minnesota word association norms. In L. Postman and G. Kepper, editors, *Norms of Word Association*, pages 1–38. Academic Press, New York, NY.
- Carl Jung and Franz Riklin. 1906. Experimentelle Untersuchungen über Assoziationen Gesunder. In C. G. Jung, editor, *Diagnostische Assoziationsstudien*, pages 7–145. Barth, Leipzig, Germany.
- John Kahn. 1989. *Reader's Digest Reverse Dictionary*. Reader's Digest, London.
- Adam Kilgarriff, Pavel Rychlý, Pavel Smrž, and David Tugwell. 2004. The Sketch Engine. In *Proceedings of the Eleventh EURALEX International Congress*, pages 105–116, Lorient, France.
- George Kiss, Christine Armstrong, and Robert Milroy. 1972. *The associative thesaurus of English*. Edinburgh University Press, Edinburgh.
- William Levelt and Herbert Schriefers. 1987. Stages of lexical access. In G. Kempen, editor, *Natural Language Generation: New Results in Artificial Intelligence, Psychology, and Linguistics*, pages 395–404. Nijhoff, Dordrecht.
- William Levelt, A. Roelofs, and A. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1):1–75.
- William Levelt. 1989. *Speaking : From Intention to Articulation*. MIT Press, Cambridge, MA.
- William Levelt. 1999. Language production: a blueprint of the speaker. In C. Brown and P. Hagoort, editors, *Neurocognition of Language*, pages 83–122. Oxford University Press.
- Igor Aleksandrovič Mel'čuk and Alain Polguère. 2007. *Lexique actif du français : l'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. Champs linguistiques. De Boeck, Bruxelles.



- Rada Mihalcea and Dragomir Radev. 2011. *Graph-Based Natural Language Processing and Information Retrieval*. Cambridge University Press, Cambridge, UK.
- George A. Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine Miller. 1990. Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography*, 3(4), pages 235–244.
- George Armitage Miller. 1990. Wordnet: An on-line lexical database. *International Journal of Lexicography*, 3(4).
- Adilson E. Motter, Alessandro P. S. de Moura, Ying-Cheng Lai, and Partha Dasgupta. 2002. Topology of the conceptual network of language. *Physical Review E*, 65(6).
- Douglas Nelson, Cathy McEvoy, and Thomas Schreiber. 1998. The university of South Florida word association, rhyme, and word fragment norms.
- Ted Nelson. 1967. Xanadu projet hypertextuel.
- Ross Quillian. 1968. Semantic memory. In M. Minsky, editor, *Semantic Information Processing*, pages 216–270. MIT Press, Cambridge, MA.
- Stephen Richardson, William Dolan, and Lucy Vanderwende. 1998. Mindnet: Acquiring and structuring semantic information from text. In *ACL-COLING'98*, pages 1098–1102.
- Paul Robert, Alain Rey, and J. Rey-Debove. 1993. *Dictionnaire alphabetique et analogique de la Langue Française*. Le Robert, Paris.
- Peter Roget. 1852. *Thesaurus of English Words and Phrases*. Longman, London.
- Michael Rundell and G. (Eds.) Fox. 2002. *Macmillan English Dictionary for Advanced Learners*. Macmillan, Oxford.
- Roger Schvaneveldt, editor. 1989. *Pathfinder Associative Networks: studies in knowledge organization*. Ablex, Norwood, New Jersey, US.
- Bennett Schwartz. 2002. *Tip-of-the-tongue states: Phenomenology, mechanism, and lexical retrieval*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Gerardo Sierra. 2000. The onomasiological dictionary: a gap in lexicography. In *Proceedings of the Ninth Euralex International Congress*, pages 223–235, IMS, Universität Stuttgart.
- Manfred Spitzer. 1999. *The mind within the net: models of learning, thinking and acting*. MIT Press, Cambridge, MA.
- Joseph Paul Stemberger. 1985. An interactive activation model of language production. In A. W. Ellis, editor, *Progress in the Psychology of Language*, volume 1, pages 143–186. Erlbaum.
- Della Summers. 1993. *Language Activator: the world's first production dictionary*. Longman, London.
- Piet van Sterkenburg. 2003. Onomasiological specifications and a concise history of onomasiological dictionaries. In *A Practical Guide to Lexicography*, volume A Practical Guide to Lexicography, pages 127–143. John Benjamins, Amsterdam.
- Gabriella Vigliocco, Tiziana Antonini, and Garrett Merrill. 1997. Grammatical gender is on the tip of italian tongues. *Psychological Science*, 4(8):314–317.
- Michael Vitevitch. 2008. What can graph theory tell us about word learning and lexical retrieval? *Journal of Speech, Language, and Hearing Research*, 51:408–422.
- David Wilkins. 1972. *Linguistics and Language Teaching*. Edward Arnold, London.
- Ziqi Zhang, Anna Lisa Gentile, and Fabio Ciravegna. 2012. Recent advances in methods of lexical semantic relatedness – a survey. *Journal of Natural Language Engineering, Cambridge University Press*, 19(4):411–479.
- Michael Zock. 1996. The power of words in message planning. In *Proceedings of the 16th conference on Computational linguistics*, pages 990–995, Morristown, NJ, USA. Association for Computational Linguistics.