

Methodological Considerations of Mental Lexicon Research: Focusing on Representational Links among Lexical Entries

Simon FRASER

Institute for Foreign Language Research and Education

Hiroshima University

Naoki SUGINO

College of Information Science and Engineering

Ritsumeikan University

Yuya KOGA

Department of Human Communication

Shukugawa Gakuin College

Hirokatsu KAWASHIMA

Department of English

Nagasaki College of Foreign Languages

Chika IKEDA

Faculty of Letters

Aichi Prefectural University

Naoki SUGIMORI

College of Information Science and Engineering

Ritsumeikan University

It is widely acknowledged that vocabulary items in the mental lexicon are interrelated and form a network (see e.g., McCarthy, 1990; Aitchison, 2003). This network consists of numerous sub-networks, with multiple layers corresponding to the various linguistic levels (i.e., semantic, phonological/orthographical, and syntactic). Such a view of the lexicon obviously has significant implications for foreign language teaching and testing.

However, the network conception of the mental lexicon has to this date remained largely theoretical. To our knowledge, vocabulary tests that directly assess the qualitative state of lexical knowledge are scarce. While there has been no shortage of practical suggestions on how best to present and teach vocabulary, these are generally much less systematic than desired; the selection of vocabulary included in word mapping, for example, depends largely on the teacher's subjective judgment and/or textbook materials. From the perspective of the learners, very few learning tools have been developed that will enable them to visualize the structure of their mental lexicon (that is, how a word is known in relation to other words), and thus help them to monitor and assess their own progress in vocabulary learning.

The objective of our project is to address this need by developing a computer-based vocabulary-learning tool to aid learners. To do this, we need to find ways to picture how words are interconnected at the different linguistic levels in the mental lexicons of individual

learners. The important questions to ask are:

- How are words grouped together, and how is a new candidate incorporated into a network?
- How can the “distance”, or “relatedness” between the lexical items in a network be determined and measured?

To be able to answer these questions, it is first necessary to establish what we already know from L1, L2, and bilingual research. To that end the present paper reviews some of the leading network models of the mental lexicon; of particular interest to us are studies that attempt to show how sub-networks are formed at the different levels, and how representational links might best be defined and measured empirically.

NETWORK CONCEPTIONS OF THE MENTAL LEXICON

Since Oldfield (1966) first proposed the idea of a “mental dictionary”, the storage system in the mind has been likened variously to a thesaurus, an encyclopedia, a library, and a computer. All of these share the idea of input, storage, and retrieval. Nowadays, though, the metaphor of choice is that of a network, or “giant multidimensional cobweb”, as Aitchison (2003) puts it. McCarthy (1990:43) envisages a three-dimensional model of the lexicon, with “phonological nets crossing orthographic ones and criss-crossing semantic and encyclopedic nets”. The lexical network is in a perpetual state of flux, constantly receiving input to be incorporated into the existing word store. Evidence that the mental lexicon is a web of words connected to each other comes from word-association experiments and native-speaker “slips of the tongue” observations. The findings of such studies point towards both L1 and L2 lexicons being phonologically arranged. Access appears to be primarily via phonological networks for low proficiency learners and via semantic networks for high proficiency learners.

Network theories draw on connectionist models of cognition (see, e.g., Bechtel and Abrahamsen, 2001). These take inspiration from the structure of the brain, with the terms “nodes” and “links” being used to represent nerve cells (neurons) and synapses. How such a network operates depends on which cells are wired together, and the strength or efficiency of the connections between them.

In a lexical network, vocabulary size can then be seen as corresponding to the number of nodes in the network; adding a new node (increasing “breadth”) and adding new links (increasing “depth”) both have implications for the rest of the network (Meara and Wolter, 2004). It seems likely that both semantic and phonological/orthographical networks are organized this way, with the representation of a word involving multifarious interconnections. Evidence that words in the network are related meaningfully comes from priming experiments, which have found that it takes less time to identify a target word when it is immediately preceded by a semantically related word. Priming effects can also be found using stimuli that are phonologically, orthographically, or morphologically related (see, e.g., Taft, 1991).

COMPARING L1 AND L2 LEXICONS

Past studies have generally supported the idea of fundamentally different L1 and L2 lexicons (e.g., Meara, 1984; Channell, 1990). The word association task is a frequently employed tool for investigating the organization of the mental lexicon of L2 learners. Three types of word association have been postulated: *paradigmatic*, *syntagmatic*, and *clang associations*. Paradigmatic associations are formed with words from the same word class as the stimulus and which fulfill the same grammatical function (*dog* – *cat*). This type of association includes coordinates, subordinates, and synonyms. Syntagmatic associations bear sequential relationships to the stimulus, which are generally from a different word class (*hunt* – *dog*). Clang associations are responses which are phonologically similar to the stimulus, and bear no semantic connection to the stimulus (*dog* – *log*).

It has been shown that the type of word association changes with the development of a learner's level of English proficiency. Meara (1980, 1982) found that L2 learners' association responses are far less regular than those of native speakers, and often different in type. However, as the learners' proficiency increases, their responses become increasingly similar to those of native speakers. Söderman (1993) found that the number of clang and syntagmatic associations tended to decrease, while that of paradigmatic associations increased the more her students were exposed to the second language.

Can we be sure, though, that the deviations between L1 and L2 lexicons are the result of structural differences, or of an L2 lexicon that is similar in structure, but smaller? There is some evidence for structurally similar L1 and L2 lexicons (e.g., Postman, 1970; Stolz and Tiffany, 1972), and Wolter (2001) believes that the L2 lexicon may in fact be much more structured than past research has tended to indicate. Wolter compared NS and NNS patterns of response in light of depth of word knowledge scores, and concluded that individual words progress from a state in which phonological and other nonsemantic connections are dominant to one where syntagmatic or paradigmatic connections take precedence. This challenges the belief that lexical development is indicated by a shift from predominantly syntagmatic to predominantly paradigmatic responses.

Graph Theory is a branch of mathematics concerned with the contacts or connections that relate one entity to another. Its principles have been applied to word association data in order to compare the relative densities of L1 and L2 lexical networks. Wilks and Meara (2002) estimated the mean number of associational links between small sets of words, and found clear differences between native and L2 speakers. They concluded that levels of linkage in the L1 and L2 lexical networks are higher than previously assumed, but do not come close to their potential for connection. Wilks and Meara's model is very simple, though, and does not distinguish between paradigmatic and syntagmatic relations, for example.

It is important to remember that although studies often assess the L2 lexicon as an isolated entity, it is highly likely that strong links exist between words in a given speaker's L1 and L2 mental lexicons. A number of computational models have been proposed with the aim of shedding light on the way in which a bilingual's lexicons interact, and answering questions

such as whether bilinguals have single, separate, or overlapping language processing systems for their two languages.

MODELS OF THE BILINGUAL LEXICON

The two main approaches favored by bilingual researchers in the connectionist tradition involve the use of *localist* and *distributed* models. In a localist model, each entity being represented (e.g., letter features, letters, or words) is assigned a single unit, and connection strengths are generally “hard-wired”. Distributed representations, on the other hand, involve identification not of a single unit but of a code over several units. In this type of network, connection strengths are initially randomized, and the system follows a learning rule which allows it eventually to relate each word to its meaning. The leading localist models are the Bilingual Interactive Activation (BIA) model, the Semantic, Orthographic, and Phonological Interactive Activation Model (SOPHIA), and the Bilingual Interactive Model of Lexical Access (BIMOLA).

The Bilingual Interactive Activation Model

The BIA model can be seen as an extension of McClelland & Rumelhart’s Interactive Activation Model (1981). It features three levels: letter features, letters, and words (only four-letter words can be represented), with a fourth level added: language nodes. Activating a particular language node allows selection of words in that language while inhibiting words in other languages. Words from different languages are represented in an integrated lexicon at the word level. BIA+ is an extension of this model, and includes phonological and semantic representations. SOPHIA is the implementation of the BIA+ model, with layers for letters and phonemes, orthography and phonology, and semantics and language nodes. This enables it to account for considerably more data than the BIA model.

The BIMOLA Model

The BIMOLA model is an extension of the TRACE (auditory word recognition) model (McClelland & Elman, 1986). It focuses on spoken word recognition, and has distinct modules at phoneme and word levels. The lexicons are separate, and there is only within-language competition at the word level. Unlike the BIA, there are no language nodes, and emergent top-down language effects arise from inhibition rather than activation.

These models attempt to explain the processing structures in the adult bilingual, and can be applied to cross-modal priming, recognition of cognates and homographs, and semantic effects. Distributed models, on the other hand, are geared more towards addressing issues of language acquisition and loss. Examples include the Bilingual Single Network Model (BSN), the Bilingual Simple Recurrent Network Model (BSRN), and the Self-Organizing Model of Bilingual Processing (SOMBIP).

Bilingual Single Network and Bilingual Simple Recurrent Network Models

The BSN model (Thomas, 1997a, 1997b) is an extended word recognition/reading model

with layers for orthographic input, semantic input, and language context. It captures evidence both of the independence of lexical representations and of between-language effects (interlingual homographs). With the BSRN, French (1998) made use of a simple recurrent network in which input to two artificial languages (Alpha and Beta) consists of nouns and verbs in SVO sentences. Each word can be processed in the context of the word preceding it; the network's task, given these simple sentences, is to predict the next word. French found that differences in word order were sufficient to lead to differentiated representations of each language.

The Self-Organizing Model of Bilingual Processing

The architecture of the SOMBIP model (Li and Farkas, 2002) is based on two self-organizing maps – one representing the meaning of words (English/Chinese), the other phonology (English/Chinese). Associative links are learned between these two maps. In this model, related words have more similar representations than unrelated words. Both the semantics and the phonology of the two different languages self-organize separately, which results in language-specific lexicons in a single integrated network. Unlike the BSN model, which maps word orthography to word meaning, input (conversations from the bilingual CHILDES corpus) is realistic.

Random Autonomous Boolean Network Models

Meara (1999, 2004, 2006) has adopted a rather different approach in his modelling of the bilingual lexicon. His models are based on Random Autonomous Boolean Networks (Kauffman and Glass, 1973; Kauffman, 1993), and aim to show how emergent network properties of the lexicon might explain aspects of language shifting between L1 and L2. In this type of model of the mental lexicon, L1 and L2 each form an *attractor state*, in which some words are permanently activated, and others permanently deactivated. The network will be in either of the two states. When the network is in L1, for example, increasing exposure to L2 words may

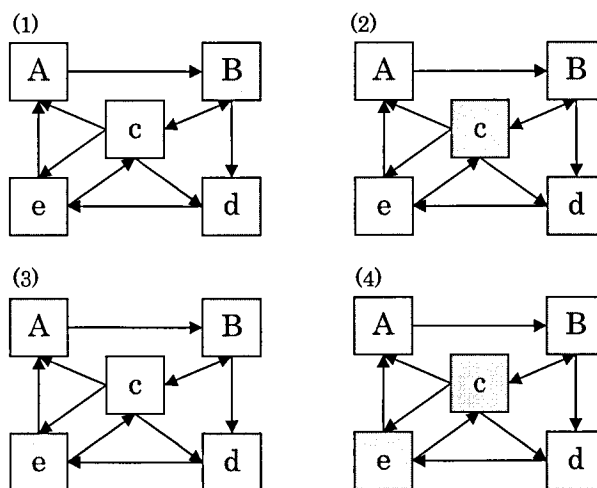


Fig. 1: A Random Autonomous Boolean Network Model of the Lexicon

cause the attractor to rapidly shift to L2, which will become the dominant lexical process.

Figure 1 (adapted from Meara, 2006:627) shows how a random autonomous Boolean network responds to an external stimulus. The arrows show the direction of activation. Words A and B are only activated if both of their inputs are already activated; the other words (labelled with lower case letters) will become activated if either one of their inputs is already activated. Activated units are shaded dark gray. In (1), Word B has been activated by an external stimulus, (2) which causes Word C and Word D to become activated; (3) this activates both Word E and Word D, (4) resulting in the reactivation of Word C. The drawback of this approach, which Meara readily admits, is that because the models are so general, and the words in the network have neither semantic nor phonological entries, their applicability to real lexical networks is limited. Nevertheless, there is undoubtedly much that we can learn about the behavior of real lexicons from an examination of the properties of this type of simple network structure.

It will be apparent by now that there are two opposing views of how the bilingual's language system relates to that of the monolingual. The first sees a separate language system for the bilingual's L2, with independent representations for a word and its translation at the lexical level, as in the BIMOLA model. The alternative is that the two languages serve as subdivisions within a single system, as is the case in the BIA model. Strong support for the single network view comes from the BSN model, which captures evidence of between-language effects (interlingual homographs). However, there are limitations with the model; training input is unrealistic (mapping word orthography to word meaning), for example, and other information sources may be used, such as the familiarity of pronunciation that the input string produces.

APPLICATIONS: DICTIONARIES AND THESAURUSES

Applying this network conception of the mental lexicon, various types of electronic dictionaries (e-dictionaries) have been proposed. One such application is *Visual Thesaurus* by Thinkmap, Inc. (see Figure 2). In general, e-dictionaries have a number of advantages over their paper counterparts, including easier cross reference, faster search, and less interference with the understanding process in reading. There is also the benefit of a "history/log" of the words that have been looked up. In addition, such network visualizations of related words can help learners grasp not only which words are connected with the target word but also *how* they are connected. However, being a thesaurus, any networks that are presented will be normative in nature; they do not reflect how a learner has learned the words in relation to other words at a particular stage in the learning process, and they may contain words that are still unfamiliar to the learner. This may therefore be a less effective way of aiding the learner's network construction process and the development of meta-strategies in vocabulary building.

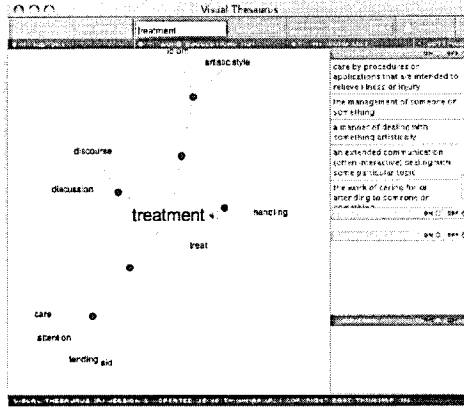


Fig. 2: *The Visual Thesaurus Electronic Dictionary*

MENTAL LEXICON SUB-NETWORKS

The Semantic Level

The Hierarchical Network Model

In the Hierarchical Network Model (Collins and Quillian, 1969), nodes representing semantic concepts of lexical entries are interconnected and organized into a hierarchy (see Figure 3). In this model, broad conceptual categories are divided into narrower categories, which in turn are subdivided into still narrower groupings.

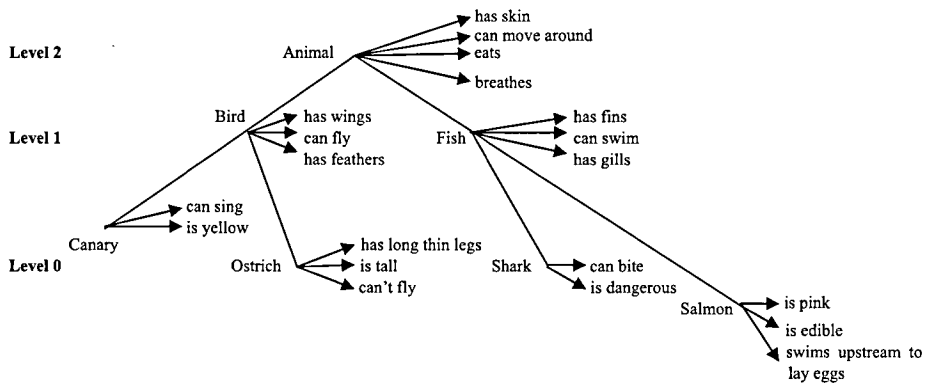


Fig. 3: *The Hierarchical Network Model*

The interrelatedness of lexical items is determined by the number of nodes between them. Lexical access will therefore take longer if there is a greater difference between the node levels of items (the semantic distance effect), or if the number of items in a category is greater (the category size effect). The model can explain the following observations ($<$ = less time to judge in semantic verification task):

Category level effect: *A canary is a bird.* < *A canary is an animal.*

Cognitive economy: *A bird has feathers.* < *A bird has skin.*

Animals can breathe. < *Birds can breathe.* < *Canaries can breathe.*

Critically, however, by assuming that all the lexical items in one category belong to the same node level, it fails to explain the typicality effect; also, the number of node levels does not necessarily reflect the difficulty of semantic verification of lexical items.

Typicality effect: *A robin is a bird.* < *An ostrich is a bird.*

Reversals of category level effect:

A dog is an animal. < *A dog is a mammal.*

A chimpanzee is an animal. < *A chimpanzee is a primate.*

The Spreading Activation Model

The Spreading Activation Model (Collins and Loftus, 1975) was an attempt to address these theoretical flaws. Instead of having a hierarchical structure, this model consists of a web of interconnected nodes. Each node represents a word, and the length of the pathways (i.e., the distance between nodes) reflects the degree of conceptual relevance (semantic similarity) between each word. In the Hierarchical Network Model, the assumption is that connections between concepts are based on logic. However, in the Spreading Activation Model, the connections are not necessarily logical; they are, rather, based on personal experience.

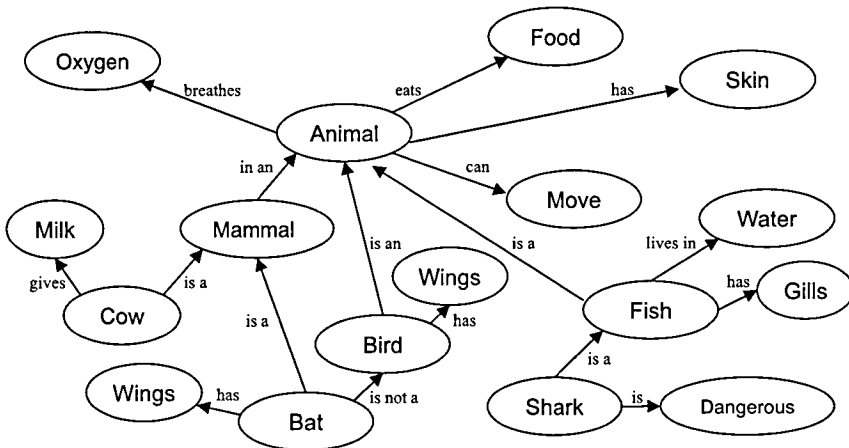


Fig. 4: *The Spreading Activation Model*

Much of the research into the organization of semantic lexical models has made use of the semantic priming effect. When subjects are asked to read two words consecutively and perform a lexical decision task on the second word, they usually process the second word more quickly if the first word is semantically related (Meyer and Schvaneveldt, 1971). However,

the results obtained from such tasks are controversial because semantic relationships between words are not as clearly defined as associative ones which reflect word use rather than word meaning (McNamara, 1992; Thompson-Schill et al., 1998). Examples of semantic relatedness, reflecting similarity in meaning or overlap in “features”, would be *whale* – *dolphin* (same category), *hammer* – *nail* (functional), or *stem* – *flower* (part-whole). Associative relatedness, on the other hand, might be represented by pairs such as *needle* – *thread* and *spider* – *web*. An important consideration is that the degree of association between different pairs of words is highly variable. For instance, with the pair *coat* – *rack* the words are strongly associated, yet semantically dissimilar; with *radish* – *beets*, however, the converse is true – the words are weakly associated, yet semantically similar.

In order to determine the validity of semantic lexical models, it is also necessary to consider the possible influence of the processing strategies used by experimental subjects, such as *retrospective processing* (Neely et al., 1989) and *prospective expectancy generation* (Becker, 1980). Retrospective processing is when the participants evaluate the relationship between the prime and the target after the target is presented; any relationship that the participants detect between the words can be used. Prospective expectancy generation refers to participants noticing that prime-target pairs are related and so generating possible related targets in every trial. On some of the related prime-target trials (but on none of the unrelated pairs), participants can eventually generate the correct targets, and as a consequence, priming effects might be overestimated.

In a self-organizing map (Kohonen, 1997), the nodes are arranged spatially, typically in a two-dimensional rectangle (see Figure 5). Semantic relatedness is represented by statistical regularities in input patterns. When a pattern is presented, one of the nodes will be activated most – the “winner”. The weights to the winner, and of the nodes in the neighborhood, are moved in the direction of the input pattern.

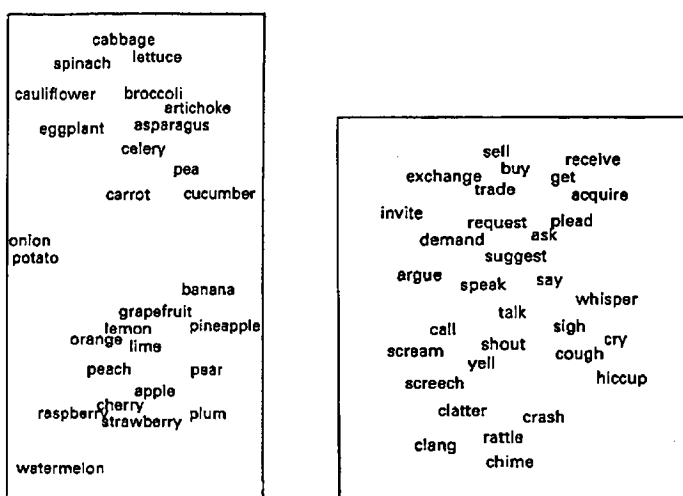


Fig. 5: *Self-organizing Maps*

Computer-generated semantic networks are able to provide a good graphic conceptualization of the structure of the lexicon. Software such as SemNet™ (available at www.biologylessons.sdsu.edu/about/semnetdown.html) enables learners to construct their own semantic networks of concepts and relationships by using the software's graphic interface tools, and has the potential to help them to organize their knowledge for better comprehension and retention.

The Pathfinder procedure is an algorithm that was developed in cognitive science to determine the most important links in a network (Schvaneveldt, 1990). It can generate networks empirically from estimates of psychological distance, and these networks have much in common with the above-mentioned semantic networks. In a Pathfinder Network, the relatedness between nodes is represented by how closely they are linked. Pathfinder does this by removing duplicate lines (“pruning”) and computing the simplest path between any two nodes in the system. The resulting proximity matrix data is processed to produce a graphic network representation, as shown in the following hypothetical example showing the links between five words.

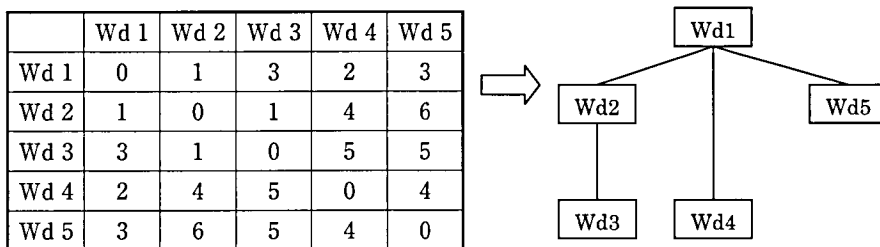


Fig. 6: A Simple Pathfinder Network Generated from Proximity Data

The Phonological/Orthographical Level

There is considerable interest in how words in the mental lexicon are related to each other phonologically and orthographically. However, the state of the art consensus (see, e.g., Read, 2000; Schmitt, 2000; Nation, 2001; Daller et al., 2007) is that there is little systematic understanding of phonological/orthographical distances, with methods of measuring such distances systematically being particularly limited. Three different approaches to the problem of determining phonological/orthographical distance can be discerned in the research. These look into possible applications of the Neighborhood Size Effect (Coltheart et al., 1977), the concept of synforms (e.g., Laufer, 1988; Laufer-Dvorkin, 1991) and usage-based phonological and morphological models (Bybee, 2001).

The Neighborhood Size Effect, or neighborhood density, is measured by the number of words which differ by a single letter in the same position. Thus, *date*, *gate*, *hate*, *rate* and other words with *-ate* are neighbors, for example, and this kind of group of words would be regarded as having higher density than, say, that consisting of the phonological neighbors *gate* and *wait*. In psycholinguistic studies, it has been shown that neighborhood density affects the

efficiency of lexical access (Goldinger, Luce and Pisoni, 1989, Cluff and Luce, 1990). In the lexical decision task used to test the neighborhood density effect, the degree of closeness among neighbors is not directly measured; however, it could be argued that the degree of closeness can be seen in terms of differential response latencies and that the larger the neighborhood size, the longer it takes to recognize the word.

The Neighborhood Size Effect suggests one possible direction we can take in order to develop ways of determining the phonological/orthographical distance between words. It must be noted, however, that there are a great number of combinations of words whose phonological/orthographical distance cannot be measured by neighborhood density alone, such as *tool* and *cook*. Also, it is important to remember that the Neighborhood Size Effect is applicable only when language users have a certain level of proficiency.

Similarity of lexical forms is known as *synformy*. Synforms share some general characteristics such as belonging to the same word class, or being identical in all but one of their phonemes. Laufer-Dvorkin (1991) lists ten categories of synforms; Category 1, for example, comprises synforms which have the same root, productive in present-day English but different in suffix (e.g., *considerable/considerate*). In Cvikić (2007), too, three types of synforms are dealt with: *synphones*, *syngraphs*, and *synmorphs*. Synphones are words which are similar in their sound, such as *live* and *leave*; syngraphs are words which are similar in their script, such as *excerpt* and *expert*, and synmorphs are words which are similar in their forms, as with *comprehensible* and *comprehensive*.

These two studies both report that, unsurprisingly, synformy is often a cause of confusion and therefore a factor of difficulty in vocabulary learning. In order to understand the nature of vocabulary learning, though, attention must also be paid to how similar or different a certain two or more synformic words such as *comprehensive* and *apprehensive* may be, and how we might go about examining the precise degree of difference. Still, the concept of synforms could be the basis of a measurement tool which can determine the phonological/orthographical distance between words.

Bybee's (2001) Usage-based Model for Phonology and Morphology holds that linguistic items are not stored in a long, unstructured list but that, instead, storage is structured by the regularities and similarities observable in linguistic items. In this model, similar or identical properties of meaning and form are associated with one another across items, generating an associative network. It is shown, for example, that lexical connections are established by the phoneme [b] among the words: *bee*, *bet*, *bed*, *ban*, and *bin*. It is likewise claimed that morphological relations (e.g., *play* + the suffix *-ing* = *playing*) may form an internal structure, in which words such as *playing*, *banning*, *ramming*, and *spoiling* are all connected by the same suffix with the same sound at the end.

This model holds that relations are formed between objectively different sounds, and that there is a high degree of similarity between some of these sounds at some cognitive level. It must be pointed out, however, that little attention seems to be directed at how to systematically or objectively determine phonologically different and similar sounds.

Measuring Phonological Distance

How, then, might we measure the phonological distance between lexical items? One possible way is to take phonemes as basic “distance units” and then follow this procedure:

1. Compare the total numbers of phonemes in target words e.g., SCHOOL [sku:l] (consisting of 4 phonemes) vs. TOOL [tu:l] (consisting of 3 phonemes).
2. Equalize the total number of phonemes by adding one or more quasi-phonemes (+) to the word with fewer phonemes, e.g., SCHOOL [sku:l] (4 phonemes) vs. TOOL [+tu:l] (1 + 3 = 4 phonemes).
3. To determine the distance, calculate the ratio of the number of shared phonemes to the total number of phonemes. In this example: 2 (the number of the same phonemes: /u:/and /l/) ÷ 4 (the total number of phonemes) = 0.5.

There are, however, problems with calculating distance based on the number of different or identical phonemes alone. Consider the following example comparing the words BAD, SAD, and MAD, where the number of phonemes in target words (D=different, S=same, T=total) may show:

BAD [bæd] (D1+S2=T3) vs. SAD [sæd] (D1+S2=T3)

→ Distance : 2/3=0.67

BAD [bæd] (D1+S2=T3) vs. MAD [mæd] (D1+S2=T3)

→ Distance : 2/3=0.67

Although ostensibly, the distances obtained are identical, intuitively, the two pairs do not have the same distance.

Phonemes and the Sonority Scale

Speech sounds can be ranked in terms of their intrinsic “sonority”, or “relative loudness compared to other sounds” (Giegerich, 1992:132). Lowest on the sonority hierarchy are plosives, moving up through fricatives, nasals, approximants, to vowels, with open vowels being the most sonorous:

| | | | | | | |
|-----|------------|---|------------|---|--------|------|
| LOW | Oral Stops | → | Fricatives | → | Nasals | HIGH |
| | 1 | | 2 | | 3 | |
| | /b/ | | /s/ | | /m/ | |

The possibility exists of using the number of phonemes and weighting the differences according to the Sonority Scale at the same time. Look at the following two examples, in which the distance of 1) is smaller than that of 2) because of the differences in the sonority rank (represented by the number below each of the consonants):

| | | | | | | |
|--------|-------|-----|-----|-------|---|------------------|
| 1) BAD | [bæd] | vs. | SAD | [sæd] | → | smaller distance |
| | 1 | | | 2 | | |
| 2) BAD | [bæd] | vs. | MAD | [mæd] | → | greater distance |
| | 1 | | | 3 | | |

The Syntactic Level

Another layer in the mental lexicon is that representing the syntactic level of organization, where words are grouped together on the basis of shared syntactic behaviors. In this respect, previous studies on the lexico-semantic features of verbs and their projections onto surface structures are of particular importance. The correspondence between syntax and lexico-semantics has for some time been one of the liveliest arenas for discussion in both first and second language acquisition and in language processing (e.g., Pinker, 1989; Jackendoff, 1990; Goldberg, 1995; Juffs, 1996). Investigations into the subcategorization properties of verbs, and their surface-structure projections, have led to researchers classifying verbs into several subclasses (such as *verbs of creation*, *verbs of future having*, *verbs of permission*, etc.) whose members share not only common semantic features but also co-occurring argument constructions.

However, it is the case that even though a pair of verbs may be semantically similar, it is not always possible to use both of them in the same construction (e.g., Joe told/*whispered Mary a story). In order to account for such lexical idiosyncrasies, Goldberg (1995, Ch. 5) circumscribes the verbs in a subclass into implicitly represented verb clusters. The members of each cluster are interrelated by means of the argument structure construction with which the verbs can co-occur.

Of particular relevance to the mental lexicon is that, in Goldberg's constructionist account, the verbs are grouped into subclasses or clusters, and that these subclasses are interrelated at the level of argument structure construction. It should also be noted that type frequencies of verbs and of constructions are conceived of as the determinants of the "closeness" among the members of the relevant networks.

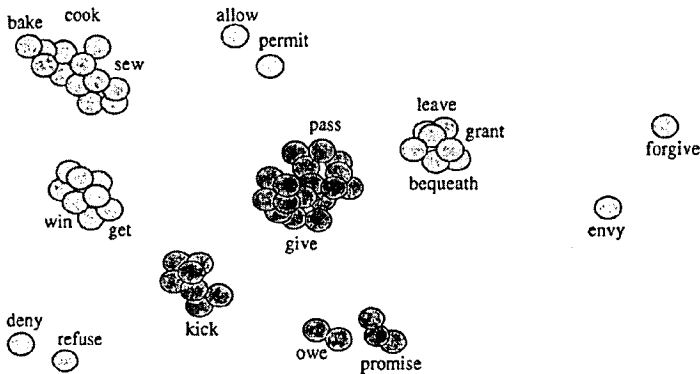


Fig. 7: A Cluster of Verbs

FUTURE DIRECTIONS

What has emerged from our review of the literature is that, even though there are models that account for the composition of the various sub networks, there is little understanding of how to measure the degrees of closeness of, or distances between, the constituents. Our first task will therefore be to consider how we can determine the length, and the salience, of the links between the lexical items in the mental lexicon. In this respect, it should be kept in mind that at any time the links in the network will be dynamic rather than stable. In one sense, they are dynamic because an increase in frequency of co-occurrence will mean that a link between a given pair of words will be shorter and more salient. Furthermore, as a consequence of the unlearning process or of noticing, a link connecting two words may later be revised, resulting in different pairs being linked.

Once the various issues are resolved and the vocabulary-learning tool developed, careful tracking and monitoring of a learner's construction of her/his mental lexicon using the tool will bring about significant impact on, and strengthen the ties with, related disciplines. Take, for example, second language acquisition research. A language acquisition model that can be applied to all of the three layers discussed so far is the Usage-based Model (Langacker, 1987, 1991; Barlow and Kemmer, 2000; Tomasello, 2003). One of the basic tenets of the Usage-based Model is that language acquisition is an exemplar-based generalizing process, which implies that it is statistic (based on the interpretation of numerical data), probabilistic, and stochastic (subject to behavior based on the likelihood of occurrence) in nature. The construction of the mental lexicon by an individual learner is not rule-based, and therefore it is impossible to be absolutely certain about the state of a learner's lexicon at any given stage in language acquisition, or to predict how the network construction process might proceed. However, after accumulating sufficient data from the use of the vocabulary-learning tool, and with the aid of corpus linguistics, "probabilistic forecasting" may become possible.

Another practical application of the vocabulary-learning tool may be to employ it in scrutinizing the relationship between success in inferring the meaning of unfamiliar words and a learner's existing lexical knowledge. So far, this question has been raised based only on the "list" notion of vocabulary organization.

Below are two possible ways of representing the interconnectedness between lexical items based on the above discussion. In the layer model (see Figure 8), the target word (in this case, *acquire*) functions as a kind of "pivot" that connects the networks at the different levels. In this model, more words can be represented, but the salience of each connection is difficult to represent. The axis model shown in Figure 9, on the other hand, is better suited to showing the distances between the interrelated words by placing the target word at the origin of the three axes. In either case, if we can determine the distances a learner may attribute to the relations between given pairs of words, it should be possible to display the state of that individual's mental lexicon, or, at least, the part of it which is "activated" by the target word.

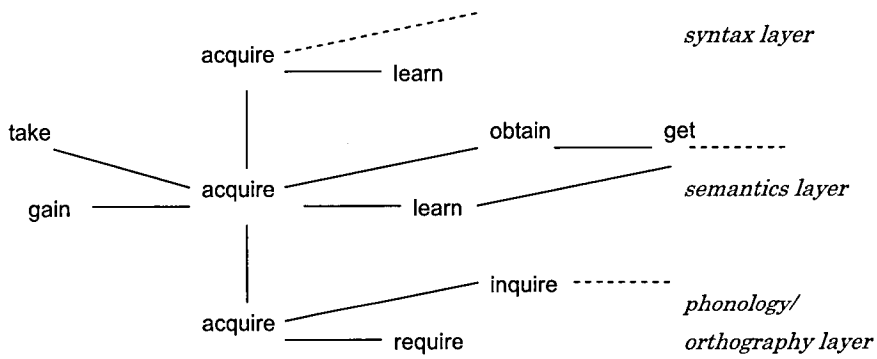


Fig. 8: A Hypothetical 3D Representation: Layer-Model

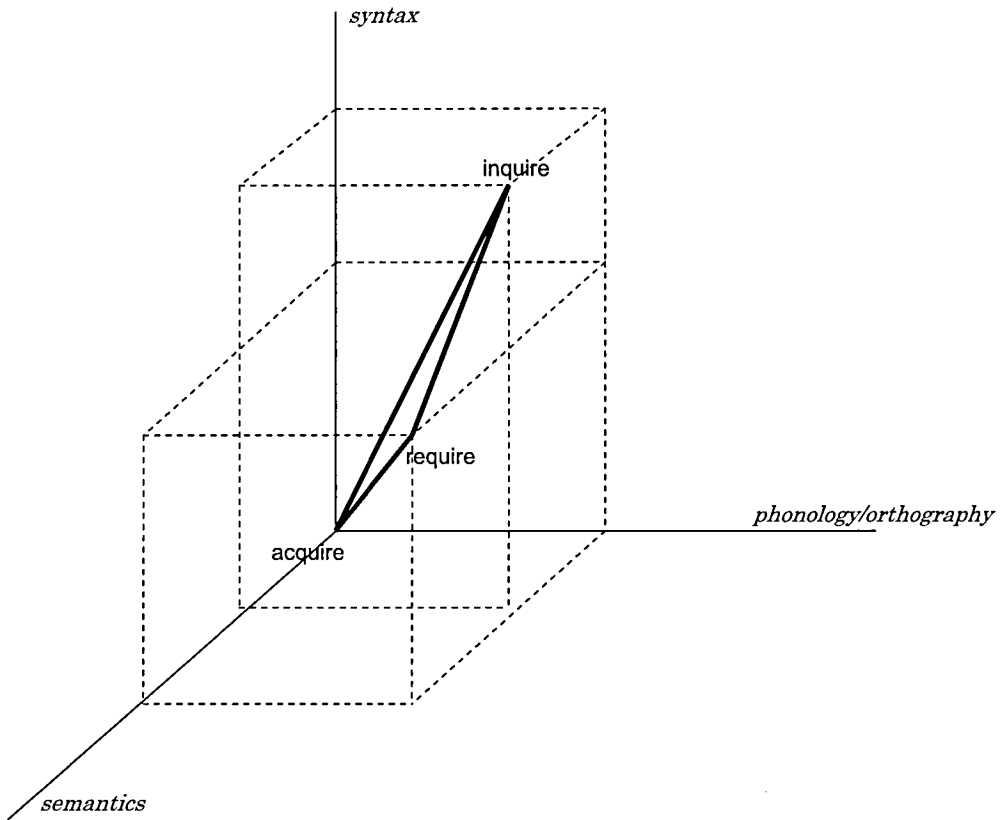


Fig. 9: A Hypothetical 3D Representation: Axis-Model

REFERENCES

- Aitchison, J. (2003). *Words in the mind: An introduction to the mental lexicon*. (3rd ed.). Oxford, UK: Blackwell.
- Anderson, J. R. (1995). *Cognitive psychology and its implications*. (4th ed.) New York: W. H. Freeman.

- Barlow, M. & Kemmer, S. (Eds.) (2000). *Usage-based models of language*. Stanford, CA: CSLI Publications.
- Bechtel, W. & Abrahamsen, A. (2001). *Parallel processing, dynamics, and evolution in networks*. (2nd ed.) Oxford, UK: Blackwell.
- Becker, C. A. (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory and Cognition*, 8, 493-512.
- Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.
- Channell, J. (1990). Vocabulary acquisition and the mental lexicon. In J. Tomasczyk, & B. Lewandowska-Tomasczyk (Eds.) *Meaning and lexicography*, 21-31. Amsterdam: Benjamins.
- Cluff, M. & Luce, P. (1990). Similarity neighborhoods of spoken two-syllable words: Retroactive effects on multiple activation. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 551-563.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407-428.
- Collins, A. & Quillian, M. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 22, 88-104.
- Coltheart, M., Davelaar, E., Jonasson, J.T. & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.) *Attention and Performance*, VI. Lawrence Erlbaum Associates Inc.
- Cvikić, L. (2007). The importance of Language Specific Features for Vocabulary Acquisition: An Example of Croatian. In Z. Lengyel & J. Navracics (Eds.) *Second language lexical processes: Applied linguistic and psycholinguistic perspectives*, 146-165. Multilingual Matters Ltd.
- Daller, H., Milton, J., & Treffers-Daller, J. (Eds.) (2007). *Modelling and assessing vocabulary knowledge*. Cambridge: Cambridge University Press.
- French, R. M. (1998). A simple recurrent network model of bilingual memory. In M. A. Gernsbacher & S. J. Derry (Eds.) *Proceedings of the 20th Annual Conference of the Cognitive Science Society*, 368-373. Mahwah, NJ: Erlbaum.
- Giegerich, H. J. (1992). *English phonology: An introduction*. Cambridge: Cambridge University Press.
- Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: The University of Chicago Press.
- Goldinger, S., Luce, P., & Pisoni, D. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language*, 28, 501-518.
- Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: The MIT Press.
- Juffs, A. (1996). *Learnability and the lexicon: Theories and second language acquisition research*. Amsterdam: John Benjamins.
- Kauffman, S. A. (1993). *The origins of order: Self-organization and selection in evolution*. New York: Oxford University Press.
- Kauffman, S. A., & Glass, K. (1973). The logical analysis of continuous, nonlinear biochemical control networks. *Journal of Theoretical Biology*, 39, 103-129.

- Kohonen, T. (1997). *Self-organizing maps*, 2nd edition. Heidelberg: Springer-Verlag.
- Langacker, R. W. (1987). *Foundations of cognitive grammar*, Vol. 1: *Theoretical prerequisites*. Stanford, CA: Stanford University Press.
- Langacker, R. W. (1991). *Foundations of cognitive grammar*, Vol. 2: *Descriptive application*. Stanford, CA: Stanford University Press.
- Laufer, B. (1988). The concept of 'synforms' (similar lexical forms) in L2 learning. *Language and Education*, 2, 113-132.
- Laufer-Dvorkin, B. (1991). *Similar lexical forms in interlanguage*. Tuebingen: Gunter Narr Verlag.
- Li, P., & Farkas, I. (2002). A self-organizing connectionist model of bilingual processing. In R. Heredia & J. Altarriba (Eds.) *Bilingual sentence processing*, 59-85. North-Holland, The Netherlands: Elsevier Science.
- Meara, P. M. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching and Linguistics: Abstracts*, 13(4), 221-246.
- Meara, P. M. (1982). Word associations in a foreign language: A report on the Birkbeck Vocabulary Project. *Nottingham Linguistic Circular*, 11(2), 29-37.
- Meara, P. M. (1984). The study of lexis in interlanguage. In A. Davies, A. Howatt, & C. Criper (Eds.) *Inerlanguage*, 225-235. Edinburgh: Edinburgh University Press.
- Meara, P. M. (1999). Simulating recovery from bilingual aphasia. *International Journal of Bilingualism*, 3(1), 45-54.
- Meara, P. M. (2006). Emergent properties of multilingual lexicons. *Applied Linguistics*, 27(4), 620-644.
- Meara, P. M. & Wolter, B. (2004). V-links: Beyond vocabulary depth. In *Angles on the English-speaking World*, 4, 85-96.
- McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception, Part I: An account of basic findings. *Psychological Review*, 88, 375-405.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1-86.
- McNamara, T. P. (1992). Priming and constraints it places on theories of memory and retrieval. *Psychological Review*, 99, 650-662.
- McRae, K., Seidenberg, M. S., & de Sa, V. R. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, 126 (2), 99-130.
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90, 227-234.
- Moss, H. E., Ostrin, R. K., Tyler, L. K., & Marslen-Wilson, W. D. (1995). Accessing different types of lexical semantic information: Evidence from priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 863-883.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University

Press.

- Neely, J. H., Keefe, D. E., & Ross, K. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 1003-1019.
- Oldfield, R. C. (1966). Things, words and the brain. *Quarterly Journal of Experimental Psychology*, *18*, 340-253.
- Perea, M. & Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychological Research*, *66*, 180-194.
- Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: The MIT Press.
- Postman, L. (1970). The California norms: Association as a function of word frequency. In L. Postman & G. Keppel (Eds.) *Norms of word association*, 241-320. New York: Academic Press.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schvaneveldt, R. W. (1990). *Pathfinder associative networks: Studies in knowledge organization*. Norwood, NJ: Ablex.
- Söderman, T. (1993). Word associations of foreign language learners and native speakers - A shift in response type and its relevance for a theory of lexical development. In H. Ringbom (Ed.) *Near-native proficiency in English*. Åbo: Åbo Akademi.
- Stoltz, W. S., & Tiffany, J. (1972). The production of "child-like" word association by adults to unfamiliar adjectives. *Journal of Verbal Learning and Verbal Behavior*, *11*, 38-46.
- Taft, M. (1991). *Reading and the mental lexicon*. Hove, UK: Lawrence Erlbaum Associates.
- Thomas, M. S. C. (1997a). Connectionist networks and knowledge representation: The case of bilingual lexical processing. Unpublished doctoral thesis, Oxford University, England.
- Thomas, M. S. C. (1997b). Distributed representations and the bilingual lexicon: One store or two? In J. Bullinaria, D. Glasspool, & G. Houghton (Eds.) *Proceedings of the Fourth Annual Neural Computation and Psychology Workshop*, 240-253. London: Springer.
- Thompson-Schill, S. L., Kurtz, K. J., & Gabrieli, J. D. E. (1998). Effects of semantic and associative relatedness on automatic priming. *Journal of Memory and Language*, *38*, 440-458.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, Mass.: Harvard University Press.
- Vinson, D. P. & Viglicco, G. (2002). A semantic analysis of grammatical class impairments: Semantic representations of object nouns, action nouns and action verbs. *Journal of Neuro-linguistics*, *15*, 317-351.
- Wilks, C. & Meara, P. M. (2002). Untangling word webs: Graph theory and the notion of density in second language word association networks. *Second Language Research*, *18*(4), 303-324.
- Wolter, B. (2001). Comparing the L1 and L2 mental lexicon. *Studies in Second Language Acquisition*, *23*, 41-69.

要 約

心内辞書研究における語彙間の結びつき

フレイザー, サイモン

広島大学外国語教育研究センター

杉野直樹

立命館大学情報理工学部

古賀友也

夙川学院短期大学人間コミュニケーション学科

川島浩勝

長崎外国語短期大学英語学科

池田周

愛知県立大学文学部

杉森直樹

立命館大学情報理工学部

本論文は、英語学習者個々人が持つ目標言語の心内辞書における語彙ネットワーク構造を視覚的に提示することにより語彙学習を支援するツール開発を目的としたプロジェクトのための理論的基礎研究である。まず、心内辞書の構造についての主要なモデルを概説した後、意味・音韻／書記素・統語の各層において、(1) ある(サブ)ネットワークがどのようにして構成され、新規メンバーとなる語彙がどのようにそのネットワークに取り込まれるのか、という点、及び、(2) メンバーとなる個々の語彙間にある(心理的な)結びつきの強さがどのように決定され、それがどのように測定可能であるのか、という点に関して考察を行う。更に、得られた知見より本プロジェクト全体の方向性を示し、語彙ネットワーク構造の三次元モデル化のための検討課題を確認する。