Untangling word webs: graph theory and the notion of density in second language word association networks
Clarissa Wilks and Paul Meara
Second Language Research 2002; 18; 303
DOI: 10.1191/0267658302sr203oa

The online version of this article can be found at:
http://slr.sagepub.com/cgi/content/abstract/18/4/303

Published by:
http://www.sagepublications.com

Additional services and information for Second Language Research can be found at:

Email Alerts: http://slr.sagepub.com/cgi/alerts
Subscriptions: http://slr.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.co.uk/journalsPermissions.nav
Citations http://slr.sagepub.com/cgi/content/refs/18/4/303
Untangling word webs: graph theory and the notion of density in second language word association networks

Clarissa Wilks Kingston University and Paul Meara University of Wales Swansea

This article examines the implications of the metaphor of the ‘vocabulary network’. It takes a formal approach to the exploration of this metaphor by applying the principles of Graph Theory to word association data in order to compare the relative densities of first language (L1) and second language (L2) lexical networks. Earlier graph theoretical research into L2 word associations is reviewed and methodological flaws in this work discussed. It describes the development of a new elicitation tool which is able to provide a means of quantifying lexical density levels. Levels of linkage in the L1 and L2 lexical networks are shown to be higher than previously assumed in the literature. It is argued that it will be helpful to develop a more complex interpretation of the notion of lexical density.

I Introduction

The work reported in this article was motivated by an underlying desire to explore some of the implications of one particularly dominant metaphor in word association research in both first language (L1) and second language (L2): that of the ‘lexical network’. It builds on earlier research which has applied the principles of the branch of mathematics known as Graph Theory to word association data in order to compare L1 and L2 vocabularies. The notion of the mental lexicon as resembling a network or, more familiarly, a ‘web of words’ is very common currency in lexical research. Researchers have used it as a convenient shorthand to convey the idea that speakers store words in some kind of interconnected system. More than this, they have been attracted to the intuitive implications of the network/web analogy, which suggest a kind of multidimensional complexity. This has seemed appropriate in trying to capture the multiple levels of interconnections between items in the mental lexicon, in which different types of semantic link are supplemented by phonological, orthographic, syntactic and
encyclopaedic connections between words. Indeed, as with many successful metaphors, the network image in vocabulary research seems almost to have developed a life of its own. We can illustrate this by looking at the way in which the conception of the network is progressively extended and built-upon by different researchers. Aitchison, for example, in seeking to bring to life the idea of the semantic network, invites us to imagine ‘a gigantic multi-dimensional cobweb’ (Aitchison, 1987: 72). One step further on, Bogaards, summarizing Aitchison’s comments, picks up and develops the connotations of this image:

Et en effet, le lexique évoque l’image des toiles d’araignée qui flottent au vent: les matériaux lexicaux se présentent dans des structures ultra-légères qui s’adaptent avec une souplesse et une flexibilité incroyables aux nécessités contextuelles du discours. [The lexicon does indeed conjure up the image of cobwebs floating in the wind: lexical material is built into ultra-light structures which adapt with incredible suppleness and flexibility to the needs of particular linguistic contexts; our translation.] (Bogaards, 1994: 71–72)

The result of this process of assimilation and development of the network metaphor may be to make us lose sight of the fact that we are still dealing with an analogy. Our conventionally accepted terminology does not in fact translate any empirically verifiable reality. When we return from the poetic language of the floating cobweb, to the apparently more prosaic terminology of the ‘network’, we may forget that the ‘network’ itself is also a metaphor. Ironically, this intuitive acceptance of the fitness of the network metaphor may ultimately tend to encourage imprecision regarding the nature of the lexicon. The belief it fosters that the lexicon ‘is’ a network may obscure a whole range of complex issues concerning the hierarchy and structure (or structures) of the lexicon and the exact nature of the types of linkages it contains.

Any metaphor as thoroughly conventionalized as this thus deserves to be fully investigated. The use of metaphor in any kind of scientific enquiry is not a neutral matter: metaphors can and do change our insights. The constructivist view of metaphor would argue that the use of metaphor in scientific theory helps not only to explain ideas, but also to shape them: ‘metaphors are constitutive of the theories they express, rather than merely exegetical’ (Boyd, 1993: 486). Changing our conception of the metaphors that we use to describe phenomena may, then, help to further our understanding of those phenomena themselves. It is, therefore, worthwhile to try to follow up the full implications of the analogies we are using to describe the mental lexicon. Metaphors are
important tools in furthering the development of theory in any field; we need to be sure that we have fully exploited their power, and fully appreciated their potential shortcomings. As Boyd points out: ‘It is part of the task of scientific theory construction involving metaphors (or any other sort of theoretical terminology) to offer the best possible explication of the terms employed’ (Boyd, 1993: 488).

The network metaphor is by no means exclusive to vocabulary research. The analogy of the network is prevalent in a wide range of sciences and social sciences. Many of these disciplines, however, have exploited the formal properties of the network analogy to a far greater extent than has been common in lexical research. Most work on the mental lexicon, as we have indicated, has tended to focus on the qualitative aspects of the network metaphor: the images of intricacy and multi-layeredness that it evokes. A notable exception to this general trend, are the few studies of the mental lexicon that have combined the use of word association data with the resources of Graph Theory to formalize some of the implications of the network metaphor (e.g., for L1, see Rapaport et al., 1966; Kiss, 1968; for L2, see Welsh, 1988; Meara, 1992; Wilks, 1999). The work to be reported here follows up the approach adopted by these studies in an attempt to address some of the unanswered questions that they have raised and, thus, to take our understanding of the power and appropriacy of the network metaphor a step further. We are particularly concerned with the use of the network metaphor as a way of understanding lexical development in L2 speakers. The approach we adopt here in our attempt to explore formally some of the implications of the network metaphor does not set out to model every dimension of the hugely complex human vocabulary storage system. The objective of the ‘graph theory model’ we propose (or indeed of any experimental model of this type) is not to describe the nature and composition of the phenomenon. Rather its purpose is to offer a simple framework for investigating how one aspect of the network metaphor – in this case the relative densities of the L1 and L2 networks – may be formalized and explored. As we shall see, formalizing a metaphor in this way can sometimes lead to unexpected results, suggesting that the metaphor itself may need to be treated with considerable caution.
II Background

1 Word associations and Graph Theory

Graph Theory is concerned with the presentation and exploration of ‘relational data’, that is, the contacts or connections that relate one entity or agent to another. Hence its appeal to those interested in word associations. Some lexical researchers in L1 and L2 (Pollio, 1963; Kiss, 1967) have turned to Graph Theory as a way of avoiding the basic design flaw that they perceived in much conventional word association research. Mainstream word association work in L1 (and almost all work on L2 associations) has looked principally at the classification and development of different types of word association response (e.g., Politzer, 1978; Söderman, 1992). This work can be criticized for attempting to generalize the features and properties of a complex large-scale phenomenon – the mental lexicon – on the basis of small-scale snapshots of that phenomenon: experiments which test only small numbers of lexical items often selected from words with similar characteristics (Meara, 1992). By applying the formal framework that Graph Theory offers, it may be possible to clarify and make explicit some of the unspoken assumptions that inform conventional research into lexical networks and to find ways of predicting how such networks will behave. It may also be possible to explore questions of the ‘texture’ of the mental lexicon and to examine the relative densities and structures of L1 and L2 vocabularies. In this section, therefore, we briefly sketch some key concepts of Graph Theory, and then go on to look at how this framework has so far been used to explore L1 and L2 vocabularies.

The graphs of Graph Theory are not those familiar from, for example, the financial pages of a daily newspaper, where statistical information is plotted on x–y axes. Rather, in Graph Theory a graph is simply a set of lines which connect a given number of points. The number of points that a graph contains is referred to as the ‘order’ of the graph. The ‘size’ of the graph is the number of ‘lines’ (i.e., connections) that join the points. The series of lines that join any two points is referred to as the ‘path’ between those points. The number of connections that any one point has is known as the ‘degree’ of that point. This information may be presented in diagrammatic form, or as a matrix (Figure 1). The only concern of Graph Theory is the pattern of relations between those points. Taking this simple definition of a graph – as a structure made up of points joined by lines – we can then say that Graph Theory consists of ‘a body of mathematical axioms and formulae which describe the properties of the pattern formed by the lines’ (Scott, 1991:13).
Lexical researchers, as we have said, have been particularly interested in the concept of ‘network density’, that is, in the extent to which items in the L1 and L2 vocabulary are interconnected. This notion derives from research in a variety of real world applications that suggests that very few networks (even very small ones) are ‘complete’, i.e., have every two points connected by a line as illustrated in Figure 2. The concept of network density, therefore, attempts to measure how far the graph is from a state of completion by looking at the overall distribution of lines within it. Most research into L1 and L2 vocabulary networks has relied on the basic premise that the more points that are connected to one another in a network, the denser that network will be. This is an entirely reasonable assumption, and one consistent with the principles of Graph Theory. However, the evidence of the research reported in this article suggests that this approach to network density is not...
unproblematic, and that earlier work may have relied on too simplistic a conception of the phenomenon.

2 Earlier graph theory models of the mental lexicon

It was this very straightforward understanding of the idea of network density that informed a series of experiments and replications conducted in the late 1980s and the 1990s by Welsh (1988), Meara (1992) and Wilks (1999). All of these studies applied graph theoretical principles to empirical word association data in order to investigate the relative densities of L1 and L2 networks. This work assumed that L2 lexicons would be less dense than the lexicons of L1 speakers. It was thus able to exploit the mathematical relationships between the number of points in a graph, the degree of those points (the number of connections each point has) and the path length between points (i.e., the number of steps it takes to get from one point to another). Meara and Wilks drew on two related properties of graphs: first, the fact that paths between points get longer as the number of points in the graph increases; and, secondly, that paths get shorter as the degree of the points increases. The model of the mental lexicon that underpins these experiments therefore suggested that as the number of interword connections in the lexicon increases (i.e., the denser it becomes), more and shorter routes should become available between a given lexical item (a ‘start word’) and any other item (a ‘target word’). We might expect this development to manifest itself as shorter paths of association between start and target words.

In a series of experiments Wilks collected data from learners and native speakers of French, German, Spanish and English using a new word-association technique devised by Welsh and Meara. Informants were asked to link together two ostensibly unrelated words by a chain of associations. Thus, for example, the start word *cold* and the target word *desire* might be linked by the chain: *cold>*hot>*passion>*desire. Working on the ‘common-sense’ assumption that L1 word association networks would be more densely interconnected than those in L2, with shorter paths between items, the model predicted that association chains produced by native speakers would be shorter than those produced by learners. In the earliest of these studies, steps taken to move between start and target word were simply counted and L1 and L2 path lengths compared. Later, in response to experimental findings, Wilks developed a somewhat more sophisticated weighted network model that attempted to account not only for the length of the association chains produced by informants, but also for the strength
of their constitutive links. (For a detailed account of all these studies, see Wilks, 1999.)

Results from these experiments were intriguing but unexpected. Whilst a number of very interesting and suggestive patterns emerged, it was not – contrary to the predictions of the lexical model upon which these studies were premised – possible to find consistent and replicable differences in length between L1 and L2 word association chains. In attempting to evaluate the significance of her own and earlier findings from these ‘graph theory studies’, Wilks scrutinized the effectiveness of the word association chain-building methodology that they had used (Wilks, 1999). She found that it was prone to encourage idiosyncratic and unpredictable associative behaviour even among native speakers. The methodology produced unforeseen artefacts that undermined its usefulness as an instrument for exploring lexical networks. Nor was it possible to adapt the chain-building methodology in any straightforward way so as to counterbalance these artefacts. Her attempt to refine the original model by replacing it with a weighted network model that would take account of the strength of association chains as well as their length did not prove successful. Again, one of the major difficulties encountered lay in the practical limitations of the methodology. Even with small groups of participants and small numbers of words, the methodology generates huge amounts of data, and this made it unworkable to run large-scale experiments.

Wilks suggested that incorrect assumptions in Meara’s (1992) model – which provided the basis for these experiments – may explain the lack of effectiveness of the chain-building methodology. Recall that this model exploited the mathematical relationships between the degree of points in a graph and the average path length between them. In lexical terms this was interpreted as meaning that denser vocabulary networks, with more interconnections between items, would generate shorter chains of association between randomly chosen words. However, the mathematics of Graph Theory indicate that very small differences in average path length between the points in any two graphs may correspond to very considerable differences in the degree of those graphs. If we look at this the other way round, it is clear that graphs with a marked difference in degree would be expected to generate paths between points that were only slightly different in average length. For example, the average path length for a 50-point graph of degree 3 is only a very little longer than what one would expect from a 50-point graph with degree 4. In terms of vocabulary networks, then, the model would predict that even where two networks differed
markedly in terms of their density (as we would have expected when comparing L1 and L2), they would nevertheless generate association chains that were only slightly different in average length. This feature of the model may then mean that such small differences in chain length may be difficult to detect reliably unless we use very large participant samples and very long lists of start and target items. Such a constraint would have obvious practical drawbacks.

It was against this background that the present research study was designed. It was conceived in an attempt to explore ways in which the methodological difficulties outlined above might be overcome and an alternative methodology developed that would allow us to tap into the information provided by word association data to explore the large-scale properties and/or the density of L2 lexical networks.

III Preliminary work

1 Exploratory model

In view of the methodological pitfalls of the earlier graph theory studies identified above, our first task was to design a sensitive, reliable and practical elicitation tool that could be used to compare the density of lexical networks in L1 and L2. Our aim was to make this instrument as simple and easy to use as possible. We hoped in this way to reduce the productive burden imposed on informants and to attenuate the risk of generating unspontaneous and idiosyncratic associative behaviour. We therefore took as a hypothetical starting point a simple questionnaire design in which comparable groups of learners and native speakers might be asked to indicate perceived association pairs in randomly chosen strings of words. The ‘common-sense’ model of the mental lexicon that we referred to earlier would lead us to predict that the number of associations perceived by native speakers in such a set of random words would be greater than the number picked out by learners. In other words, given a set of random words, and asked to find any associations among them, L1 speakers would have a higher ‘hit rate’ than we would expect from L2 speakers. This basic idea was further developed through a series of computer simulations.

The simulator allowed us to specify a range of simple models of lexical networks by manipulating two basic network parameters: the size of the lexicon and the number of connections that each word in the lexicon has to other words. For example, suppose that we...
want to simulate a lexicon of 1000 words\(^1\) with 3 connections per word. For this specification, the simulator generates a set of strings, that might look like this:

```
word0001 123 145 160
word0002 99 182 279
word0003 10 382 761
...
word0999 135 856 687
word1000 72 65 321
```

Each ‘word’ in the lexicon is connected to 3 other words. For this particular simulation run, ‘word0001’ is linked with words 123, 145 and 160; word 2 is connected to words 99, 182 and 279 and so on. Other runs, making different random choices, would produce different numbers. With a basic network of this type we can investigate the following questions:

- If word \(x\) and word \(y\) are chosen at random from the lexicon, what are the chances that word \(x\) will appear in the association list for word \(y\).
- By extension, if we choose a set of \(n\) words from the lexicon, what are the chances that at least one of these words appears in the association lists of the other words in the set.

The simulator allows us to run very large numbers of ‘pseudo-experiments’ in which we set up a network with specific characteristics, randomly select sets of words from the network and look for associations in these sets. The simulator calculates the number of associations that we find in these randomly chosen sets of words. Running a series of simulations of this sort gives us a good idea of the general properties of a lexicon with the specified characteristics.

The next step is to use the simulator to explore how a given lexicon would perform in a range of ‘typical’ experimental designs. For example, in a real experiment, we might ask 40 Ss to look for associations in 20 sets of 10 randomly chosen target words. We can simulate this task by asking the simulator to perform the following tasks:

---

\(^1\) In discussing the size of the lexicon, the simplification ‘words’ has been used in this section in place of the preferred ‘items’ in order to avoid terminological confusion between ‘lexical items’ and ‘questionnaire items’. 

Downloaded from http://slr.sagepub.com at Shanghai Jiaotong University on March 7, 2009
Do this 40 times { (number of pseudo-subjects)
  Do this 20 times { (number of trials per pseudo-subject)
    Generate a set of 10 target words (number of items per trial)
    Calculate the number of hits in this set
    Report the hit rate
  }
  Report the mean hit rate for these 20 sets of words
}
Report the mean hit rates for the 40 pseudo-subjects

To recap, the simulator allows us to specify simple network models of different sizes and densities (numbers of connections between words) and to compute the number of associations that we can expect to find in a randomly chosen set of words from the lexicon. It also allows us to vary the experimental conditions. Once these network parameters and test conditions are set we can explore the effect on hit rate of altering any one of the elements. The network simulator thus serves two purposes. First, it can inform experimental design by allowing us to test the sensitivity of a range of different instruments that use different numbers of trials and different numbers of words per set. This has clear practical attractions since it obviates the necessity of running time-consuming pilot versions of a range of experiments. Secondly, the network simulator can help us to establish a valuable point of comparison against which we can set actual word association behaviour. Knowing that networks of certain densities generate word association ‘hit rates’ of a certain magnitude, we can subsequently project back from actual hit rate to an estimation of the density of the network producing it.

Using the simulator suggested that the simplest questionnaire design likely to yield consistent results with a group of 30 informants would use 40 items each consisting of a set of 5 randomly chosen words.

2 Preliminary simulations

Following this preparatory work, we ran a further set of large-scale simulations. These simulations modelled an experiment with 30 participants, each performing 40 trials with a set of 5 words in each trial. Various parameter settings for a simulated mental lexicon were tested in these conditions. In each case, the size of the lexicon was fixed at 1000 words in order to simulate an experiment using a core vocabulary that would be common to both native speakers and relatively low level learners. Although the definition of what constitutes a ‘core’ vocabulary for learners remains very
controversial, some recent researchers report that learners may know around 2000 words after 5 or 6 years of language learning at school level (Bogaards, 1994). We may thus fairly safely assume that a vocabulary of 1000 words will be core for our purposes. The number of links between words was progressively increased from 4 to 10, in order to reflect a range of different densities in the networks. The range 4 to 10 was suggested by the literature on L2 word associations, which suggests that learners have very few links in their networks. Woolfenden (1983), for example, found that even native speaker networks were of low density, with few connections between words, and that learner networks were extremely sparse. Beck (1981) also found that density was low in nonnative speaker networks.

By progressively increasing the number of connections for each word in the lexicon we can simulate vocabulary networks of increasing density. Assuming that L1 networks are denser than L2 networks, we can then make theoretical predictions about the likely behaviour of native speakers and learners of varying degrees of competence under these test conditions.

Each group of parameter settings was run 15 times. On each of the 15 occasions, the mean hit rate per participant group was calculated and, subsequently, an average hit rate worked out over the 15 simulations. This procedure allowed a very large-scale survey to be simulated which would provide a robust and generalizable estimation of expected performance. The results of these simulations are given in Table 1.

We can interpret these findings as follows: in a series of repeated studies each using 30 participants, each completing a 40-item questionnaire, with 5 words per item, we would expect the mean hit rates as shown in Table 1. For example, we would predict that a group of participants with a relatively sparse lexicon where each word has only 4 connections in a network of 1000 words would have an average hit rate of just over 3 associations per questionnaire. On the other hand, a group of participants whose lexicon was denser at – for example, 10 connections per word – would be expected to

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Mean hit rates over 15 simulations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4 links</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.325</td>
</tr>
</tbody>
</table>

Notes: Lexicon = 1000 words; pseudo-subjects = 30; trials = 40; words per set = 5
score a hit rate of over 7 associations per questionnaire. Of course at this stage the results of the simulations merely establish a basis for speculation. Clearly no real experimental study is going to be dealing with entirely homogeneous groups of participants with lexicons of uniform density. We may therefore expect that the real data will show a greater degree of variation within participant groups. Moreover, the network models we have been working with are for the time being simply ‘best guesses’, and estimates of relative densities in native speaker and learner vocabularies may well need to be revised. However, having said this, the results of these simulations do provide us with some general guidelines against which we may set actual results from real data in an attempt to gauge the general level of linkage within learner and native speaker vocabularies.

IV Main study

1 Method

For the real study 60 informants were used: half were learners of French whose L1 was English and half were native speakers of French. The 30 learners were all students at Kingston University and had completed one year of study in French at undergraduate level. Of the 30 native speakers of French, six were members of the Kingston University teaching staff and the rest were students on a range of undergraduate and post-graduate programmes at the university. All but six subjects (three learners and three native speakers) were aged between 19 and 25.

Participants completed a 40-item questionnaire whose structure was identical to the one we used in the simulation experiments. Each item consisted of a set of 5 words randomly chosen from the Français Fondamental list: approximately the first 1000 most frequent words in French excluding grammatical items (Gougenheim et al., 1956). Written instructions asked participants to read each set of words and to circle any two words in a set that they considered to be associated, for example:

blouse cheminée coûter feu tort

The instructions stipulated that where participants perceived no links between any of the words in the set, they should write nothing. In addition, instructions stated that where they considered more than two of the words in the set to be associated they should circle the two with the strongest link.
Participants were instructed to work instinctively and to spend no more than 20 minutes on the exercise. In practice, the learners took an average of 10.38 minutes to complete the questionnaire and the native speakers an average of 11.25 minutes. (The fact that learners did not work more slowly than native speakers as we might have expected is taken up in Section V.) The questionnaires completed by native speakers and learners were identical except for the instructions, which were given in the L1 for each group. After completing the exercise, participants were given the opportunity to discuss their experience of the exercise and some of the learners did further work based on the questionnaire as part of their teaching programme. Their comments are reported in Section V.

2 Results

The completed questionnaires were scanned to determine the number of ‘hits’ (i.e., pairs of words circled) for each participant. Results are shown in Table 2. Statistical tests show that native speakers do perceive significantly more associations (i.e., they have a higher hit rate) than do their nonnative counterparts ($t = 6.47, p < .001$). Note that the mean hit rate is very much higher for both native speakers and for learners than our simulations had predicted. Every item of the questionnaire elicited some responses from both the learners and the native speakers. We address this point in some detail in Section V.

V Discussion

In terms of the model of the mental lexicon that we have proposed, these findings show clearly that, as we would expect, native speaker lexical networks have higher levels of linkage between items than those of learners. In other words, their lexicons are denser, at least in our straightforward understanding of the term. However, the implications of our findings go beyond this simple observation.

One of the most striking things about the results we have reported is the discrepancy between the actual findings and the predictions of our large-scale simulations. Recall that in these

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Mean hit rate per group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nonnative speakers</td>
</tr>
<tr>
<td>Mean</td>
<td>19.00</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>7.56</td>
</tr>
<tr>
<td>n</td>
<td>30</td>
</tr>
</tbody>
</table>
repeated simulations we examined the expected hit rates for a
groups of pseudo-subjects with 4, 5, 6 or 7 links per word. Our
simulations suggested that we would expect to find about 3 hits per
questionnaire for the first group, and just over 4 hits per
questionnaire for the last group. Compare this with the actual
findings of the current study which uses the same basic test
conditions as the simulations, that is, a 40-item questionnaire with
5 words per set and a population of 30 participants. The actual data
show a mean hit rate of 19.6 associations for learners and 30.9
associations for native speakers.

How can we account for this marked difference between
simulated and real hit rate? There are two obvious explanations
that we need to consider. One explanation is that the simulations
have underestimated the average number of connections between
items in the lexicon of both learners and native speakers. With this
in mind, the simulator was used again to calculate the approximate
numbers of links per word of the lexicon that would be needed to
generate the kinds of hit rates that emerged in the current study.
From this it emerges that we would need to raise the links
parameter to approximately 36 links per word in order to produce
mean hit rates of 19–20 associations per 40-item questionnaire. To
achieve mean hit rates of some 30 per questionnaire, as found for
the native speaker group, implies more than 45 links per word in a
1000 word lexicon. In Figure 3 we illustrate the simulator’s
predictions of how levels of linkage in a 1000 word lexicon must
rise in order to generate increasing hit rates.

The implications of this explanation are that, in a restricted
lexicon, interconnections between items are a much more
significant feature of the network than we had originally assumed.
Recall that our ‘best guess’ regarding the L1 and L2 lexicons, which
was based on the findings of earlier research (e.g., Beck, 1981;
Woolfenden, 1983), predicted that even in native speaker networks
each word would have relatively few associations to other words.
Our simulations in fact initially imagined that a notional ceiling of
10 links per item would easily account for the densest lexicon. Our
current findings, however, suggest that this assumption was wrong.
Our calculations imply that even learners must have far higher
numbers of links between words in their core vocabulary in order
to be able to generate the ‘hit rates’ recorded on our test.

How, then, should we interpret this departure from the findings
of low network density in learners as reported in earlier studies? It
is important to note that almost all the earlier studies of L2 word
association networks we have referred to have relied on
methodologies that require participants to produce associations

prompted by given stimulus words in some form or another. The current experiment, however, relies on the recognition of associations. Whilst the notion of separate ‘passive’ and ‘active’ lexicons is now largely dismissed in the literature (McCarthy, 1990), it nevertheless seems plausible to assume that the distinction between a receptive and productive activity is an important one. Where no productive task is involved, attentional control is freed to some extent for greater concentration on higher order processing, e.g., at a syntactic or semantic level (Garman, 1990). This may allow learners (and even native speakers) to tap into a denser web of associations. Quite simply, then, associations may be more readily recognized than they are produced. This impression is reinforced by the comments of the informants collected after they had completed the questionnaire. Both learners and native speakers were enthusiastic about the exercise and said that they found it easy to complete. They reported very few difficulties in understanding the words used in the questionnaire. Of course, it is possible that a few participants were either under-reporting the gaps in their knowledge or had misunderstood some vocabulary, but this seems

![Figure 3](https://slr.sagepub.com)  
**Figure 3** Expected number of hits generated by varying the degree of each word in a simulated lexicon
unlikely. It is worth recalling that the learners completed the test, on average, slightly faster than the native speakers. This would seem to confirm that they were not struggling with unknown words and that the task of identifying associations between words was a relatively spontaneous one.

The second possible explanation of the discrepancy between the hit rate we had predicted and our findings may lie in the approach we have been taking to the calculation of network density. The design of both the earlier graph theory experiments (Welsh, 1988; Meara, 1992; Wilks, 1999) and of the current study were built around the notion that the density of a network is dependent upon the degree of points within that graph. This notion presupposes, in other words, that the higher the number of connections between points in the network, the denser that network will be. As we have seen from the results of these various studies, the assessment of density by any of these methods does not appear to be a straightforward matter. As we have noted, the earlier chain-building and weighted network experiments found few systematic differences in either path length or path strength for learners and native speakers. The alternative methodology reported here, on the other hand, appeared to suggest that the number of connections per word in both L1 and L2 networks was far greater than other work in the field had led us to expect.

These combined findings invite the view that our original conception of network density may have been too simplistic and may warrant a more detailed exploration of the concept. As we have indicated, the fundamental notion of density as dependent on the degree of points in the graph is widely accepted in graph theoretical literature. There are, however, a number of somewhat less obviously intuitive factors underlying this basic assumption which the Welsh/Meara model may have left unexplored. One important factor is that the density of networks depends on two parameters of network structure: ‘inclusiveness’ and the ‘sum of degrees’ of points (Scott, 1991). Inclusiveness refers to the number of points that are included within the various connected parts of the graph. Thus, the inclusiveness of a graph is evaluated by looking at the total number of points in the graph minus the number of isolated points. The more inclusive a graph is, the denser it will be.

When we consider the second parameter – the sum of degrees of points in a network – we should note an important consideration: that is, that although we may refer for convenience to average numbers of links per point, it is very likely that in most natural networks points will vary in their degree of connection. Thus, some
points will be connected to many others, and some to only a few. We might certainly expect this to be the case when dealing with a network of word associations.

Looking at these two parameters, which must be considered when assessing network density, we can see that the density of a graph and its structure are not the same thing. The notion of the density of a graph does not convey its structural properties (Hage and Harary, 1983). Given this, perhaps the attempt simply to compare average numbers of connections in the L1 and L2 lexical networks may be more problematic and more misleading than we had assumed. It is true that we have found clear indications of differences in the number of connections in learner and native speaker networks. What we cannot claim to have identified, however, is how the structure of these L1 and L2 networks might differ. Indeed, the more sophisticated reading of the concept of density that we have outlined shows us that two networks with the same density could in fact be quite differently arranged in terms of how the connections between points are disposed and how many isolated points are present in the network. Thus, average degree may not be a good predictor of the way the association network will behave.

Another issue that we should consider when looking at our findings in the light of a re-evaluated conception of network density is the question of the actual measurement of this dimension. Earlier studies have approached the question of lexical network density only in terms of a comparison between learner and native speaker networks. It may however prove revealing to attempt to quantify L1 and L2 network densities in concrete terms. The issue of quantifying network density is particularly pertinent here given the surprising finding that the ‘hit rates’ of association pairs identified were far higher than both the literature and the simulation studies had led us to expect.

One accepted formula for calculating network density (Hage and Harary, 1983; Scott, 1991), which incorporates the two parameters of inclusiveness and sum of degrees as outlined above, involves comparing the actual number of lines (connections between points) within a graph with the number of lines that would be present if the graph were complete (i.e., if every two of its points are connected by a line). If we look at the basic construction of any graph, it is easy to see that the actual number of lines in the graph is equal to half the sum of the degrees of that graph. The maximum number of lines in a graph, of \( n \) points, on the other hand, can be calculated by the formula: \( n(n - 1)/2 \), since each point can be connected to all others except itself, and each line connects two
points. Bearing this in mind, we can see that the following formula gives us a measurement of network density:

\[
\text{network density} = \frac{l}{n(n - 1)/2}
\]

where \( l \) is the actual number of lines, and \( n \) is the number of points in the graph. Density measurements calculated using this formula range from 0 (a completely unconnected set of points) to 1 (a complete graph with every pair of points connected).

Let us now apply this method of density measurement to the figures for the experiment here. For the native speaker group the hit rate recorded in the study led us to assume an average of 45 links per word. In a network of 1000 items, this would imply a total of 45 000 links. In graph theoretical terms, then, the sum of the degrees of the graph is 45 000, and the actual number of lines (\( l \)) in that graph is half that figure: 22 500. Thus, the density of the network is calculated as follows:

\[
\frac{22 500}{1000 (1000 - 1)/2} = 0.045
\]

For the learner group, the average number of links per item in the 1000 word lexicon was found to be 36, making the sum of degrees in the L2 graph 36 000 and the actual number of lines in that graph 18 000. The parallel density calculation for learner network density can thus be expressed as follows:

\[
\frac{18 000}{1000 (1000 - 1)/2} = 0.036
\]

What are the implications of these figures? At first sight it might seem that these concrete measurements of L1 and L2 network density are somewhat difficult to interpret since they are in some sense decontextualized measures. We do not, after all, at this stage have well-established norms for L1 and L2 lexical network density against which we can set these results. Despite this, however, the figures are revealing for two reasons. First, they indicate that even if the number of links is greater than we had expected, this does not necessarily mean that networks are extremely dense in absolute terms. It would seem that even core lexical networks are a very long way from the state of ‘completion’ that they could theoretically
reach. In other words, the potential for interconnections between words in a core lexicon is much greater than appears to exist in reality even for native speakers. This is worth bearing in mind given the tendency in the literature on vocabulary to refer, in metaphorical terms at least, to the great density or richness of lexical networks. Aitchison (1987: 72), for example, invites us to imagine the mental lexicon as a web in which ‘every item is attached to scores of others’. In a similar vein, McCarthy (1990: 42) refers to the ‘labyrinthine connections between words’ of the native speaker’s mental lexicon. These images are, of course, in part a reference to the complexity of the mental lexicon in terms of the multiple types of links that exist between items (on the orthographic, phonological, semantic, syntactic, encyclopaedic levels, etc.) Nevertheless, they do also allude to complexity in terms of numbers of links between items. Perhaps, then, the concrete figures presented here might act as a small corrective to intuitive, but not necessarily accurate, assumptions about the intense nature of network density. Whilst we would not in any way want to argue that lexicons are not complex webs, it is worth nevertheless pointing out that in terms of overall density they are not nearly as complex as they could be.

In the light of this observation it is interesting to note that one contention found in the network theory literature is that in many cases there may be limits to the number of relations that any one point in the network can sustain (Scott, 1991: 77). Where this is the case, the actual number of lines possible in any graph will be limited by the number of points in the graph and therefore, all other things being equal, larger graphs will have lower densities. How applicable is this idea to word association structures? It is not unreasonable to speculate that there may in fact be constraints on the number of associations that can be supported by any word in the lexicon. An unlimited capacity to form an ever increasing number of associations could have implications for processing loads within the mental lexicon. Such constraints would not of course imply that word association networks must remain static; indeed, it would be difficult to argue against the conception of a dynamic lexical network in which ‘webs of meanings and associations constantly shift and re-adjust’ (McCarthy, 1990: 42). However, what the notion of upper limits on the numbers of links sustained by any one item might imply is that there is not a constant build-up of links for any one item, but rather that some associations may atrophy or become dormant as others are formed. The shift in associations would thus be more a matter of rotation and/or replacement rather than a purely incremental process. We could further speculate that not all
associations to an item would have the same status in terms of their permanence or impermanence. There might be some types of links—such as the most conventionalized and stereotypical associations of relatively frequent words—that would be less subject to shift and replacement.

The second reason why the concrete density measurements that we have produced may be revealing is that they may hint at some interesting structural properties even within small core networks. One possible interpretation of the low overall density scores that we have recorded for both the L1 and L2 networks could be that, even in a 1000-word core lexicon, structure is not uniform. Some parts of the core network, even in native speakers, may have much higher levels of linkage, whilst other parts are relatively sparsely connected or contain isolated items that would tend to reduce overall density scores. In associative terms, then, we might imagine that some lexical items are ‘more important’ than others, either in that they have more connections than other points, or in that they hold positions of strategic significance in the overall structure of the association network by acting as link points between different clusters of associations within the network. Graph Theory addresses questions of the relative ‘importance’ of points in this sense in terms of their ‘centrality’ within the graph. The concept of centrality has been variously defined in the graph theory literature, but three definitions in particular of this notion seem potentially pertinent to the discussion of lexical network structures. The first of these focuses on the degree of a point within a graph and argues that this can be interpreted as an index of its ‘communication activity’. Thus, points with the highest degree are considered to be most central in the network. A second definition of centrality revolves around the notion of the ‘communication efficiency’ of a particular point. This is assessed by calculating the distance or closeness of that point from the sum of the lengths of the shortest paths between that point and all other points in the graph (Hage and Harary, 1983). The third definition of centrality focuses on the ‘between-ness’ of a point, that is, the extent to which it lies between other points in the graph. According to this measure, even a point with a relatively low degree (few direct connections) may play an important intermediary role and so be very central to the overall network (Scott, 1991). An important issue for future research will therefore be to investigate ways of determining whether certain lexical items, or types of lexical items, play ‘central’ roles in the vocabularies of learners and native speakers. Watts’ (1999) recent ideas about ‘small world’ graphs may be relevant here.
VI Conclusions

In this article we have shown how we used simple Monte Carlo methods to run simulations of some simple network models of L2 mental lexicons. This work has allowed us to develop an elicitation tool capable of detecting clearly quantifiable differences between L1 and L2 association networks. This tool is more robust than some of the earlier methodologies we used in that it does not encourage idiosyncratic associative behaviour. Using this tool and the predictions of the simulator it has been possible to collect word association data that can be used to explore various graph theoretical approaches to the question of network density in the vocabularies of learners and native speakers. We have demonstrated that the notion of ‘density’ as understood by some earlier research may have been too simple and that the parameters of ‘inclusiveness’ and ‘sum of degrees’ should also be taken into account. Our discussion led us to propose that the simple measurement of density of the lexical networks in L1 and L2 might not be sufficient to convey important differences in the structural properties of those networks.

The elicitation tool that we have developed has also allowed us to quantify differences in L1 and L2 network density. This measurement has shown that, whilst levels of linkage in both L1 and L2 networks may be far higher than suggested in some of the word association literature, they nevertheless do not come close to their potential for interconnection. These suggestions show that the notion of ‘word web’ may not always lead us in the right direction and should alert us to the dangers of accepting too superficial an analysis of the popular metaphor of the lexical network rather than opting for a more formal approach. Clearly, the model we have used is a very simple one. It has, for example, treated all association types as similar and does not distinguish between, say, paradigmatic vs. syntagmatic relations or whole–part vs. exemplar–category associations. The implications of expanding the model to include these types of dimensions are as yet unclear. However, it seems highly likely to us that such an expansion of the model would complicate the network metaphor even further than our current study has suggested. This reinforces our belief that the network metaphor should be applied with greater caution than is apparent in much of the current literature.
VII References