The concept of Lexical Priming in the context of language use[1]

Michael T.L. Pace-Sigge, University of East Finland

Abstract

Corpus Linguistics is becoming an increasingly important part of language research; also interest in incorporating Corpus Linguistics into language teaching has become pronounced (cf. O'Keefe et al. 2007). Hoey (2005) has presented a theory that provides a set of rules that are possibly underpinning why Corpus Linguistics works as a way of analysing language: Lexical Priming. This theory can be seen as an explanation of why collocations exist. A listener will recognise a word more quickly when a related word is given (i.e. body-heart). However, the theory is dependent on evidence of psycholinguistic and cognitive science claims. We cannot modify an approach to language description unless we are confident that we have sound evidence. Since Lexical Priming will have implications for how we deal with the results of computer-based language analysis, this article will show a range of arguments to support Lexical Priming as a linguistic theory and provide, to an extent, material to supplement Hoey's 2005 book.

1 Introduction

Corpus Linguistics has become increasingly relevant for the teaching of English language (see, amongst others, O'Keefe *et al*. (2007)). It also becomes an increasingly important tool for translation methods. In order to help make the use of Corpus Linguistics more widely applicable, we see that sections of the research community are now moving towards developing helpful theoretical models.

One such model is the theory of *Lexical Priming* which has been developed over the last decade. The aim is to connect Corpus Linguistics (the *lexical* aspect) with Psycholinguistics (the *priming* aspect). These aspects will be described in more detail below.

The book *Lexical Priming* (Hoey 2005) looks at Corpus Linguistic evidence, but the psychological (and psycholinguistic) background is only thinly represented and can be seen as insufficient to protect the theory submitted insufficiently from charges of circularity in its argumentation.

This paper will attempt to describe the possibility that there is a far stronger link between corpus linguistics and psycholinguistics. While *Lexical priming* (Hoey 2005) has focussed on the Corpus Linguistics aspect, this paper aims to supply the other half, looking at research into *Artificial Intelligence*, *Cognitive Science* and experiments by psychologists to show the link between the psycholinguistic aspects brought to light by corpus linguistic research.

2 What Lexical Priming is and where it came from

Some ideas need incubation time and new people, new techniques, and new technologies to finally make an impact. Even as early as in the 1920s, Palmer started what would become a cornerstone of British Applied Linguistics: Palmer devised lists of the most frequently used words and phrases, constructed what he later termed Pattern Grammar (which was then refined by A. S. Hornby in 1954[2] and taken up by Hunston and Francis in 1999) and gave a detailed study of collocations to the Carnegie Conference in 1934.[3] A decade after Palmer, Firth said, "A language is not merely a community of sounds or even of grammar and dictionary. It is also a community of usage and idiom..." (Firth 1957: 155). Firth appears to have sown a seed for John Sinclair's corpus work and a great deal of empirical research into language use that is based on corpora of naturally occurring language.

Subsequently, Corpus Linguistics appears to be built on three pillars[4] (cf. Nelson 2000: 122):

- *Collocation* – words that are found in a close environment to each other, like *naked* **with** *eye,* or *you* **with** *know.*
- *Colligation* – words that are a direct grammatical construction. For example, while *that winter* is always in the past tense, *in the winter* is used only half the time in the past tense.
- *Semantic Association*[5] *– that winter* refers to a 'specific event' while *in winter* mostly refers to a 'timeless truth'.

All three elements indicate that it can be observed that certain words or word forms either attract or avoid each other. While we can find *beer* + *glass* as well as *beer* + *mug*, a quick check with a search engine will show that there is clearly greater attraction of *beer* to *glass* than *beer* to *mug*. Similar experiments can be

undertaken to observe the occurrence of not just word collocations, but also their colligations and semantic associations.

Yet, while Corpus Linguistics has been able to show us that these relations between words will provide no more than a representation of what can be observed: They simply reflect occurrence patterns of the language investigated. It does not, however, explain why it occurs in the first place. There must be a reason (or reasons) why words (or clusters of words) collocate, group into certain constructions but avoid others, and have certain semantic associations.

At this point, Lexical Priming gives us an idea why it is likely (or unlikely) that certain words stand in such relations to each other. Hoey (2005: 7) indicates that *collocation* is a concept that is pervasive. Yet it is more than that: it is also subversive.

The subversion can be found in a subliminal, sub-conscious psychological concept: namely that repeat-use reinforces an idea of an occurrence pattern that seems natural. This 'repeat occurrence' acts to prime one's mind to make automatic connections. Priming can provide for sets of actions, or, in the lexical field, sets of words. So, for example, a listener, hearing the word *bread* will recognise words like *baker, butter, knife* more quickly than unrelated words like *doctor, mortar, radiator*.

Wondering what implications his theory of Lexical Priming has for language learners, Hoey states:

> (...) Priming is the result of a speaker encountering evidence and generalising from it. [Primings come] from single focussed and generalising encounters. Language teaching materials and language teachers can provide essential shortcuts to primings. (Hoey 2005: 185f.)

Yet to accept the theory of Lexical Priming as valid and useful in second language learning, we have to establish that it is compatible with the known evidence of how the mind works with regard to language. What is at issue is the point that we could teach an L2 learner a list of colours – *yellow, green, orange, blue, purple etc.* We are aware, however, that each word is usually found to be in a rather different "list" (set of collocations) – the ones native speakers are primed for (expect to) hear: *yellow with envy, green tea, blue skies* etc.

Hoey points out that

> Priming need not be a permanent feature of the word or word sequence; in principle, indeed, it never is. Every time we use a word, and every time we encounter it anew, the experience either reinforces

> the priming by confirming an existing association between the word and its co-texts and contexts, or it weakens the priming, if the encounter introduces the word in an unfamiliar context or co-text or if we have chosen in our own use of it to override its current priming.
>
> (Hoey 2005: 9)

This, in turn, can be seen as both a plus and a difficulty when teaching a language. The plus is that even the native speaker will change his or her understands of a word. The negative side is that a greater awareness of the use over time (for example the term 'gay') needs to be taught. When this awareness exists, however, deeper understanding should be the logical consequence. For example, the outmoded term for a 'radio' – 'the wireless' – has now been superseded by the use of wireless as in 'wireless network' or 'wireless headphones'. Here we can see that the collocations, colligations and semantic associations disambiguate which type of 'wireless' is referred to.

This paper aims to present a clearer picture of what priming in the context of language occurrence and language use means and why it is relevant to language instructors.

3.1 Priming

Hoey's theory of Lexical Priming is firmly grounded in corpus linguistic work carried prior to his development of the theory. He outlines the usefulness of the theory for teaching purposes (Hoey 2003; 2005: 183ff.). This article goes back to the roots and looks at what we understand as 'priming'.

The *Sage handbook of social psychology* provides the following:

> A (…) factor that influences the accessibility of information in memory is priming. The activation of stored knowledge through experiences in the immediate context can make prime-relevant information more accessible in memory, and such recent construct activation can influence inferences, evaluations, and decisions on subsequent tasks (…)
>
> A second factor that influences the accessibility of information in memory is the frequency with which a construct has been primed (…). Traits, attitudes, or stereotypes that have been frequently activated in past experience are more available in memory than those that have been less frequently primed. Such frequency of activation, if it occurs on a regular and continuing basis, can result in certain constructs becoming chronically accessible, such that no external priming in the immediate context is necessary to make them highly accessible (Hig-

> gins *et al*., 1982). Moreover, because people differ in the kinds of experiences they have that would generate such routine construct activation, individuals quite naturally differ in the particular constructs that are chronically accessible (…). (Sherman *et al*. 2003: 55)

This highlights all the relevant aspects of the notion of priming. We see how the human brain does not access memory in a random way, since information can be accessed all the easier when it can be linked to other known information. This link is strengthened the more (often) a person absorbs the same (or slight variations of) connected information.

Priming as such is not a linguistic but a psychological concept. Though it appears as if most research focuses on lexical priming (where tests under laboratory conditions are undertaken with words) the wider application of priming is widely acknowledged. (See, for example Habib: 2001).[6] The term does not, however, appear until the later 20th century.

The early literature in which the term appears seems to be mostly concerned with the priming of language – words read and heard. According to Collins and Loftus (1975) it was Ross M. Quillian who first used the term: "Quillian's theory of semantic memory search and semantic preparation, or **priming**" (my highlighting). As can be seen, Quillian (1961, 1962, 1966, 1967, 1969; see also Collins and Quillian 1969) laid the groundwork for all the research to come in the field of priming since the early 1960s. Papers written by Quillian and Collins (1969) and Collins (1969; 1970; 1972a/b; 1975) where these two authors looked at the process they name "retrieval from the semantic memory", will be discussed in some detail below. All the seminal works that past and current research are based on go back to this early research, such as the investigations by Meyer and Schvanefeldt (1971), who Posner and Snyder (1975a/b) later based their research on. James H. Neely's (1976 and 1977)[7, 8] papers are entitled *Semantic priming and retrieval from lexical memory.* Neely very clearly refers to the work of these researchers as his main influence.[9] 'Priming' together with Lexical appears, however, to be first mentioned by Neely.

3.2 M. Ross Quillian and the language learning machine

A researcher in Artificial Intelligence, M.R. Quillian (1962, 1969) describes, in theory, how to construct an *Understanding machine* (1962), a *teachable language comprehender* (1969). Talking about language translation, he states: "… human translators do not translate 'directly', and … really good mechanical ones cannot hope to either" (Quillian 1962: 17).

In providing the theoretical blueprint for a mechanical translator, he tries to simulate how the human mind learns language.[10] While the term priming is not yet introduced, Quillian deals with a number of issues that will resurface, over forty years later, in Hoey's Lexical Priming.

On of Quillian's initial concern of was how to deal with polysemy.

> The resolution of a polysemantic ambiguity, by whatever method of translation, ultimately consists of exploiting clues in the words, sentences or paragraphs of text that surround the polysemantic word, clues which make certain of its alternate meanings impossible, and, generally, leave only one of its meanings appropriate for that particular context. The location and arrangement in which we find such clues is itself a clue, or rather a set of clues, which we may call syntactic clues. (Quillian 1962: 17)

His theoretical outline foreshadows Hoey's work. The problem of polysemy exists in an ambiguous sentence like 'He reached the bank' but not in 'He got a loan from the bank'. Polysemy and disambiguity being key issues in language teaching, it is helpful to know that, in the latter example, Quillian sees the clues as sufficient:

> Thus, in our example, a reference to money is one such semantic clue, and one which, should it appear in the sentence, could be exploited no matter what word it occurred in, whether one of those on our list or not. (…) Learning to understand a language would consist of learning which readings on which scales should be activated in response to each word of that language. (Quillian 1962: 18)

This is the part of Lexical Priming referred to by Hoey as semantic association.

Quillian actively spurns transformational linguistics.[11] In line with Brazil (1995) he seems to prefer the concept of linear grammar when he says

> This seems to me a crucial advantage over those other approaches to mechanical translation which, lacking any manageable representation of meaning, have to proceed as though the only clues that are useful in resolving polysemantic ambiguities are those in grammatical features and their locations, or else in established idiomatic phrases. *That human beings do not so limit themselves, but also utilize semantic clues extensively, would appear obvious from the fact that people are able to understand language that is full of grammatical and syntactical errors*.[12] (My italics) (Quillian 1962: 18)

In fact, by the time Quillian (1969) discusses his *Teachable Language Comprehender* (TLC), he speaks of a machine that resembles in some ways the computational, algorithm-based electronic calculation machine of the most successful company of the day (2012) Google.[13] In refining their search algorithms, Google engineers take refinement of queries and whether or not people stay at a proposed link (the so-called *long click*) as feedback to teach the system:

> "We discovered a very early nifty thing", says search engineer (...), who worked hard on synonyms. "People change words in their queries. So someone would say, 'Pictures of dogs', and then they'll say 'Pictures of puppies.' That said that maybe dogs and puppies were interchangeable. We also learned that when you boil water it's hot water. We were learning semantics from humans, and that was a great advance." (Levy 2011: 48)

This resembles the concept described by Quillian, an early pioneer of *Artificial Intelligence (AI)* of a machine reader that has built up a semantic web in its memory:

> This memory is a "semantic network" representing factual assertions about the world. The program also creates copies of the parts of its memory which have been found to relate to the new text, adapting and combining these copies to represent the meaning of the new text. By this means, the meaning of all text the program successfully comprehends is encoded into the same format as that of the memory. In this form it can be added into the memory. (Quillian 1969: 459)

Though the wording is different, it does not sound unlike Hoey's everything heard or read, everything said or written that primes a person to use words in one way and not another. In his paper on the TLC, Quillian gives the example of a text that is easily comprehended because it is natural.

> What the reader must have, then, as he reads the text[14] (…), is an extremely versatile ability to recognize the appropriate chunk of memory information from among literally thousands of others he may since have learned about 'Presidents', about 'fruit trees', and about 'fathers'. (…) we assume that there is a common core process that underlies the reading of all text – newspapers, children's fiction, or whatever – and it is this core process that TLC attempts to model. (Quillian 1969: 461)

This, I would claim, is the first step Quillian takes towards identifying lexical priming as a psychological process. In fact, Quillian proposes to prime the machine in a way similar to how a young person would be primed to figure out words in contexts. He proposes to give

> (...) twenty different children's books dealing with firemen and have TLC read all of these [and reckons that the machine] will require less and less input as it accumulates knowledge. (Quillian 1969: 464)

In his references to natural language, he goes well beyond that:

> Natural language text communicates by causing a reader to recall mental concepts that he already has. It refers him to such already known concepts either with isolated words or with short phrases, and then specifies or implies particular relations between these.
>
> (Quillian 1969: 474)

This appears to be very close to Sinclair's Idiom Principle. In other words, in natural language the mind is primed to connect concepts on hearing or reading words and short phrases.

It might be argued, however, that Quillian simply philosophises over the problem. He does not quote other research, and he makes only a few references to other works. Neither are his descriptions backed up by successful experiments at this stage. However, he makes clear that he is providing a theoretical basis for building an actual machine. Most importantly, his ideas have stood the test of time and provided a theory that is still quoted by Artificial Intelligence (AI) researchers in the 21st century. In fact, the following:

> Essentially, it asserts that to read text a comprehender searches his (her, its) memory, looking for properties which can be considered related to that text. (Quillian 1969: 474)

sounds remarkably familiar to those who have read these lines from Hoey (2005):

> I have talked of the language user as having a mental concordance and of the possibility that they process this concordance in ways not unrelated to those used in Corpus Linguistics. (Hoey 2005: 14)

Quillian reckons that his TLC is fully teachable – not by working on big structures but by learning piece by piece. The structure would thereby develop through what is feasible and what is not. Once we substitute Speaker/Writer for

the term machine, it becomes clear that Quillian gives a good grounding for the priming research to come:

> Overall, the most distinctive features of this theory, as compared with other models and theories of language of which we are aware, are its explicitness and detail and its reliance on "knowledge of the world".
> (Quillian 1969: 475)

3.3 Facilitating access to the semantic memory

Moving on from the original theory of *priming*, Quillian and Collins (1969, 1970 and 1972a), discussing retrieval from the semantic memory, publish the results of a series of experiments. The last of these makes use of the term priming. The research involved checking the reaction times of volunteers to find out that true sentences (*tennis is a game*) have a shorter reaction time than false[15] ones (*football is a lottery*). They link these findings to what is termed semantic memory:

> Priming is understood to be a process by which concepts and their meanings in semantic memory are activated, regardless of the origin of that activation.
> (Collins and Quillian 1972, quoted in Ashcraft 1976: 490)

This work in turn started a whole lot of experiments by psycholinguists like Loftus (1973), Posner and Snyder (1975a), Collins and Quillian (1975), Ashcraft (1976) and, significantly, led to the seminal paper by Meyer and Schvanefeldt, *Facilitation in recognizing pairs of words. Evidence of a dependence between retrieval operations* (1971).

The important phrase in this title is 'pairs of words', which links the study to J. R. Firth's notion of collocation, the importance of which has been highlighted also by Halliday (1959, etc.) Sinclair (1991) and Hoey (2005 etc.) amongst many others. Meyer and Schvanefeldt's paper links an insight derived from psycholinguistic experimental evidence with a theoretical concept that has acquired significance in corpus linguistics:

> We showed that such decisions are faster when one word (e.g., 'nurse') is preceded by another semantically related word (e.g., 'doctor') [than linked with a unassociated word, e.g. *bread*].
> [Positive] responses averaged 85 ± 19 msec. faster for pairs of associated words than for pairs of unassociated words.
> (Meyer and Schvanefeldt [1971] 1984: 20)

The response time for collocates, therefore, was shown to be decisively quicker than the one for unrelated terms. This indicated that the mind of the reader/listener has a mental, subconsciously-made connection between these two nodes. Meyer and Schvanefeldt point out that "the results of [their experiment] suggest that degree of association is a powerful factor affecting lexical decisions in the (…) task." (Meyer and Schvanefeldt 1971: 229)

Sinclair's (1991) view that collocations mainly occur within five steps on either side of a word is an observation of how words appear in texts. That there is a possible link to how words are linked in one's memory finds support in the following results described by Meyer and Schvanefeldt:

> (…) responses to pairs of associated words would be faster than those to pairs of unassociated words. This follows because the proximity of associated words in the memory structure permits faster accessing of information for the second decision. The argument holds even if the accessed information is (*a*) sufficient *only* to determine whether a string is a word and (*b*) does not include aspects of its meaning.
>
> (Meyer and Schvanefeldt 1971: 232)

The key here is the proximity of associated words – one word acts as prime and the mind is already set to expect a limited set of options to follow. Meyer and Schvanefeldt go on to claim that this is a mental process that does not only reside in the short-term memory:

> (…) any retrieval operation R_2 that is required sufficiently soon after another operation R_1 will generally depend on R_1. This would mean that human long-term memory, like many bulk-storage devices, lacks the property known in the computer literature as *random access* (cf. McCormick, 1959, p. 103).
>
> (Meyer and Schvanefeldt 1971: 232)

This would explain why computer users, understandably, feel that their machine cannot think or is illogical. The fact is that the logic of a RAM (Random Access Memory) has little in common with the network that binds information together in the human memory.

Finally, Meyer and Schvanefeldt refine Quillian's concept of linking words as nested strings.[16] They note:

> We previously have argued that processing normally begins with a decision about the top string and then proceeds to a decision about the bottom one. Let us now assume that memory is organized by familiar-

> ity as well as by meaning, with frequently examined locations in one "sector" and infrequently examined locations in another sector.
> (Meyer and Schvanefeldt 1971: 232)

This means that *priming* of words (that is, high exposure) is not necessarily linked with meaning but, can to an extent, be described as 'familiarity'.

Meyer and Schvanefeldt (1976) claim, in their paper with the unambiguous sub-title "People's rapid reactions to words help reveal how stored semantic information is retrieved", that their set-up differs from most other experiments in the field, in that they do not seek to measure speakers' mistakes but the reaction time people take making lexical choices. Interestingly, the rate of error is remarkably low, indicating how sure-footed language users are in their native language:

> But the reaction times depended significantly on the set relations between the categories. When the meanings of the category names were closely related to each other, reaction times tended to be shorter.
>
> (…)
> People were about 55 ± 7 milliseconds faster on the average at recognizing a word like BUTTER if it followed the related word BREAD than if it followed the unrelated word NURSE (20).
> (Meyer and Schvanefeldt 1976: 30)

The difference in milliseconds becomes significantly large when compared at this level. Meyer and Schvanefeldt do not use the term lexical priming, but it is clear to readers familiar with concordances that BREAD and BUTTER are likely to be in each other's company, while BREAD and NURSE are not. This, then, would experimentally confirm that the notion of lexical priming, though not named as "lexical priming", is supported by another set of experiments described by the authors. Once words are made harder to decipher, the semantic memory assists recognition:

> Degrading the legibility with [a] pattern of dots increased reaction times by more than 100 Milliseconds. The harmful effect of degradation was significantly less, however, for related words than for unrelated words, suggesting that semantic relatedness helped to overcome the visual distortions produced by the degradation.
> (Meyer and Schvanefeldt 1976: 30)

Hoey (2005) notes that lexical priming does not simply mean connecting lexically/semantically related words. In fact, some primes (e.g. VERY) have little

lexical content. For language teaching, however, it is highly relevant to have these words in particular connected with their correct collocates. That these words play an important part of the semantic memory is pointed out by Quillian (1969). Meyer and Schvanefeldt highlight that it is not necessarily the "meaning" of a word that makes it act as a prime and, consequently, ask for further investigation:[17]

> It is not true, however, that close relations of meaning always facilitate mental processing of words. Some processes are actually inhibited when they must deal with two words that have related meanings. (…) The apparent inhibition raises more questions about what semantic information is stored in human memory and how the information is used. (Meyer and Schvanefeldt 1976: 31)

This could be seen as an explanation of why synonyms, though clearly related, are not fully interchangeable in all contexts. As spoken language production is not pre-planned and aims to be fluent with as little hesitation as possible, the words (chunks of words) that have least inhibition will tend to be the preferred choice.

3.4 "Semantic Priming of the Lexical Memory"

J. H. Neely's papers (1976, 1977) built on Meyer and Schvanefeldt's work and are cited in Hoey (2005). Neely's 'Semantic Priming of Lexical Memory,' for the first time, connects the words 'priming' and 'lexical'. In his 1976 experiment, volunteers see a Related (R), Unrelated (U) or Neutral (Nx)[18] semantic term as a prime before a target word. Exposure to these primes varies between extremely short (360 msec), medium (600 msec) and very long times (2,000 msec). Whatever the exposure, the R prime provoked a shorter reaction time. During short exposure, the difference between R and U is 40 msec (Nx lies in between). However the gap becomes marked for 600msec and longer exposure times. As with the Meyer and Schvanefeldt experiment, Neely's informants' error rate was remarkably low. He relates in his discussion that

> Activation spreads from the [target word] for the priming word to the [word forms] for semantically related words, and (2) the subject uses the priming word to direct his (…) attention for words that are semantically related to the priming word. (Neely 1976: 652)

Neely appears to say that the threshold of perception of the target word is directed by the level of semantic relatedness. His final conclusions point in the direction of lexical priming:

> (…) In comparison to a noninformative and semantically neutral warning-signal prime, a word prime (1) facilitates lexical decisions about a subsequently presented semantically related word, (2) inhibits lexical decisions about a subsequently unrelated word, and (3) facilitates decisions about a subsequently presented nonword. (Neely 1976: 654)

With this, Neely underlined the importance of Meyer and Schvanefeldt's findings, while at the same time rebutting a theory of Posner and Snyder, who had postulated that priming was expectancy-based and under the subject's control.

At this stage, experimental linguists had opened a gate to connect lexical decisions with concepts formed in the mind. That grammatical choices and lexical choices are entwined was under serious discussion. Zimmermann (1972), discussing automated text lemmatisation, comments:

> The conception of a dictionary comprises almost totally the conception of a grammar: lexicon and rule system are one. (…) It is syntax analysis (or, in a wider sense, context analysis) that creates the basis for lemmatising texts. The structures that are ambiguous on the surface (of words or sentences) are to be disambiguated with the information gathered from its context and to be integrated into the framework of the text (or the sentences). (My translation). (Zimmermann 1972: 3)

Nevertheless, despite the occasional paper linking into this type of research in the eighties – notably by Neely (1989) himself, the citation index of the papers published shows that the notion of priming, in the context of lexical memory and sentence grammar (semantic or syntactical) has not become prominent in linguistic discussion until quite recently.[19]

3.5 Collocates and semantic memory

The foundations for research on dependent clusters can be found in Murphy (1988). Here, the complex concept is defined as lying between the simple – that "can be represented as a single lexical item", and the "lexicalized (i.e. idiomatic) expression". In his paper, he quotes the example of 'corporate lawyer' which is a fixed, complex adjective-noun expression. Murphy notes that the noun-noun expression '*corporation lawyer*' is not available for use, and, furthermore, that expressions like 'corporate stationery' mean something very different from the term 'corporate'. Murphy hints at the fact that the listener would have to know which of the specific meanings a non-predicating term like "corporate" has, and his paper can be seen as another stepping-stone towards acceptance of fixed collocations as a psycholinguistic notion.

Ratcliff and McKoon (1988) go much further in their research. The hypothesis they outline is that of compound cue priming. In terms of retrieval from memory, they advance the theory that it is not concept trees (bird – animal – flight) but words that go together that make it possible to associate. Referring to their earlier (1981) work, Ratcliff and McKoon (1988: 389) point out that "they have shown that priming can be obtained between concepts that are much more than four words apart." They (and others) therefore raise an important issue about *collocation*, since it appears to contradict Sinclair's (1991) claim that there are no valid collocations beyond the five-word mark on either side.

The concept of lexical access appears to be very close to lexical priming. De Mornay Davies is more explicit when he states:

> Even if two words are not 'semantically related' in the strictest sense (i.e. they do not come from the same superordinate category), their frequent association produces a relationship at the "meaning" level.
>
> (de Mornay Davies 1998: 394)

This foreshadows Hoey's assertion that each term is primed to mean something as a result of frequent association.

De Mornay Davies finds that there is still a strong drive by researchers to try and find a meaning-driven correlation of words. However, this would neither explain idiomatic use, nor his findings with brain-damaged patients. There is, however, a lexical and semantic automatism:

> .. activation in the lexical network could be controlled by co-occurrence frequency, such that words which often co-occur in speech or text ('collocates') would be more strongly linked in a phonological or orthographic lexical network. Lexical co-occurrence, therefore, has no connection with meaning-level representations, and many researchers argue that associative priming results from lexical-level co-occurrence.
>
> (de Mornay Davies 1998: 402)

Being more specific than Ratcliff and McKoon, he anticipates Hoey's later claim that it is the property of each word to be primed to either prefer or avoid the company of other specific words, noting that this is the case because the mind co-associates these words, rather than because it links each individual word to concepts or meanings. This is backed up by Trofimovich (2005), who looks at word priming in a spoken context, comparing learners both in L1 and L2 contexts. He finds that there is intrinsic value in repeated exposure and use of words in their contexts for the learners:

> In contrast to the facilitative effects of a repeated phonological context or of a semantically related word which rarely last more than a second, auditory word-priming-effects are long lasting. For example, reliable processing benefits for repeated spoken words are maintained over delays of 8 s[econds] (Cole, Coltheart, and Allard, 1974), minutes (Church and Schachter, 1994), days, and even weeks (Goldinger, 1996). These findings suggest that auditory word-priming effects have a long-term memory component. (Trofimovich 2005: 481)

In support, Trofimovich quotes Church and Fisher (1998) who say that:

> (we have) recently identified auditory word priming as a likely mechanism supporting spoken-word processing and learning. (…) **because auditory word priming does not require access to word meaning**, it may reflect the process whereby listeners build and use presemantic *auditory* representations. (My highlights). (Trofimovich 2005: 482)

Furthermore, Hernandez *et al*. (2001) show that there is enough evidence to conclude that priming is an automatic process, a single process not split into stages.

3.6 The value of context

In a continuum from collocation to colligation we can find the propensity already discussed by Quillian in 1962 for word meaning to be disambiguated by the context it is found in. A considerable number of words have little concrete meaning by themselves, either because of the level of de-lexicalisation they have undergone or because of their role as function words.

Novick *et al*.'s 2003 paper on spoken word recognition reads like a blueprint for the theory that Michael Hoey started to outline during conferences from the same year onwards. Several conclusions about the nature of sentence comprehension arise from these results:

1. Lexical knowledge encodes detailed information about the syntactic possibilities for words, directly influencing the manner in which words are combined to form sentence-level representations. (…)
2. Those lexical-combinatory representations are encoded in a distributed manner and shared between words in a way that crosses grammatical class boundaries.
3. The lexical representations that guide sentence processing include combinatory information of a sort that may go beyond classical syntactic notions. This information may include event-structural information, including infor-

mation about which specific classes of arguments a particular word tends to associate with.

4. The findings in general align well with constraint-based lexicalist theories of parsing. Word recognition appears to play an important role in the grammatical analyses of sentences. (Novick *et al.* 2003: 72)

I have quoted the conclusions in full to highlight the parallel conclusions drawn between Novick *et al*. and Hoey. Though lexical priming is not mentioned as such, constraint-based *lexicalist* theories of parsing would certainly include it. Points one and two mirror the concept of colligation, while points two and three also encompass semantic association. Point four, to conclude, highlights that grammar is lexically-driven and that lexical occurrence and position determine the grammatical structure, not vice versa. This also links to research into a subcategory of priming named *anchoring:*

> ... researchers found that 68^0F made it easier to recognize summer words (like *sun* and *beach*) and 40^0F facilitated winter words (like *frost* and *ski*). The selective activation of compatible memories explains anchoring: the high and low numbers activated different sets of ideas in memory.
>
> (Mussweiler, 2010 quoted in Kahneman, 2011: 123)

3.7 Priming and the corpus

Up to this point, all the evidence for the existence of Lexical Priming and its workings has been based on experimental evidence by researchers into artificial intelligence (AI), cognitive linguists and psycholinguists. Only a few people have tried to find proof for this notion in the real-occurring texts produced by writers and speakers – the corpus.

Looking at the work by Gries (2005) and Nick Ellis *et al*. (2006a, 2006b), it becomes apparent that the two strands of empirical research – experiment based and corpus based – are finally brought together. Ellis *et al*. quote Meyer and Schvaneveldt (1971), while Gries highlights the fact that:

> … although it has sometimes been argued that only experimental data can contribute to studies of priming, the analysis shows that ... the corpus based results for datives are very similar to the experimental ones." (Gries 2005: 365)

Gries introduces his study with a brief overview, stating that

> … syntactic priming: (…) Levelt and Kelter (1982) and Branigan *et al.* (1999) report that priming (in spoken and written production respectively) is fairly short-lived. (Gries 2005: 368)

That priming is a short-lived and short-term memory issue, however, is only discussed in earlier syntactic priming discussions. Later research has accommodated the notion that there is also the long-term, more fixed priming.

Gries notes that his colleagues appear to be locked into their traditional methods; he quotes Branigan *et al.* (1995: 492) and asks the readers also to consult Pickering and Branigan (1999: 136):

> Corpora have proved useful as a means of hypothesis generation, but unequivocal demonstrations of syntactic priming effects can only come from controlled experiments. (Quoted in Gries 2005: 369)

It appears from this that neither Branigan nor Pickering did any work with corpora at all. Branigan and Pickering seem unwilling to consider looking beyond the scope of 'controlled experiments'; yet Gries notes that corpora have been used for psycho-linguistic research since 1997 (cf. Gries 2009: 222).

While using data from the ICE-GB corpus, Gries analyzes two different pairs of syntactic patterns, the so-called 'dative alternation' and "particle placement of transitive phrasal verbs" Gries himself seems to be taken aback by how well the data from his corpus match experimental results:

> In the present data, the ratios of the primed structure vs. the non-primed structure are 1.5 and 1.9 for prepositional datives and ditransitives respectively. By comparison, in her classic study, Bock (1986: 364) reports (…) the corresponding ratios are 1.5 and 2.1 for prepositional datives and ditransitives respectively; the differences between her ratios and mine are obviously negligible. (…)
> In sum, not only has the corpus-based analysis of syntactic priming revealed significant priming effects for ditransitives and prepositional datives, the results are also strikingly similar to those of previous experimental studies in terms of strength of effects, (…) (i.e. the time course of priming). (Gries 2005: 373f.)

Gries' results are remarkable. All hypotheses were matched, with a very small reported rate of error. It is remarkable how well theory and results match. Throughout a great number of experiments discussed, Gries is able to find significant priming effects.

The results presented by Gries make a good case for corpus linguistics working in tune with psycho-linguistic methods:

> While I do not rule out discourse-motivated factors of priming at all, it is hard to explain all the similarities between the different kinds of results and still simply uphold the claim that all this is epiphenomenal. Without doubt, (…) it seems as if the utility of corpus-based, explorative results should not be underestimated prematurely.
> (…) **the fact that lexical activation decays too fast makes it unlikely that the long duration of priming effects observed here and in other (experimental studies) is just a lexical memory effect.** (My highlights). (Gries 2005: 387)

This appears to indicate to Gries that priming effects go beyond syntactic priming found in exchanges. Priming effects work on a far deeper and more profound level.

Ellis *et al*. (2006b) seems to mirror and expand the experiments discussed above, where native speakers are compared with non-native ESL speakers.[20] In experiments that, similarly to Gries', compared volunteers' reaction times with (BNC) corpus evidence, Gries' results are replicated. Having Sinclair's (1991) idiom principle in mind, however, Ellis *et al*. outline that primings work in different ways for the two groups:

> Fluent Native speakers much more affected by MI (**M**utual **I**nformation)
> Non Native ESL speakers more affected by Frequency
>
> (Ellis *et al.* 2006b[21])

This is based on the following definitions:

> Frequency: need to have come upon the string before (strong effects of frequency in vocabulary acquisition and processing).
> MI: the bindings of words within a formula which make the formula distinctive and functional as a whole.
>
> (Ellis *et al.* 2006b)

There is logic to this. All listeners / readers can be sure that "high frequency patterns are processed more fluently" (Ellis *et al*. 2006b). For all that, a learner of a new language will merely recognise strings he or she has been exposed to frequently before. A native speaker, however, is not just more likely to have heard / read the formula before: they will also be more open to a looser form of repetition – as long as the bindings of the words remain consistent.

In other work, Ellis *et al*. (2006a) look at collocations and semantic prosody.[22] Ellis describes the set up of their tests as straightforward:

> We investigated the frequency and strength of these collocations in the BNC then looked for processing effects using the lexical decision paradigm. (Ellis *et al.* 2006a)

This means that the researchers extracted frequently occurring collocates (clusters) (for example: lose weight – frequent; receive virginity – infrequent) and then measured the reaction time (RT) it took to make a lexical decision. As a result, the team found that "Language processing (as indexed by this lexical decision task) is intimately sensitive to patterns of collocations in usage." (Ellis *et al*. 2006a). The graphs of the corpus-occurrence patterns and the reaction times run in close correlation to each other for all the above-mentioned patterns. The picture is not as clear-cut, however, when it comes to semantic prosody. This may be due to the fact that semantic prosody is a vulnerable concept, as it is not easily replicable,[23] and has been disputed.

The researchers conclude that –

- Written language processing is intimately tuned to frequencies of actual usage
- We process frequent collocations faster than infrequent ones
- We do *not* see ready evidence of *semantic* generalization here
- It appears the fluent processing associated with spread of activation in 'semantic priming effects' a due to memory for particular word associations.
- There is little by way of semantic generalization at this level of processing at least. (Quoted from Ellis *et al*. 2006a)

The first point is in total agreement with what Gries found in his experiments and what Hoey (2005) claims. The second and the last points also confirm what de Mornay Davies and others have claimed – that priming is not down to something that is based on semantic generalisations but more likely due to automatic decisions made because of word associations in the memory.

4 Conclusion

O'Keefe *et al.* point out that many text books are still over-reliant on intuition when technological and scientific advances already bring profound changes in, amongst other things, teaching methods (cf. O'Keefe *et al.* 2007: 21). With this article I hope to have shown that the notion of priming is a well-established and

widely-tested phenomenon of the human mind. It has been predicted theoretically and, by Google, mechanically replicated through the use of Artificial Intelligence; it has been, since the early 1970s, confirmed to exist through numerous experiments under laboratory conditions. The results have, in the 2000s, been replicated by corpus-based trials. All in all, the theory of Lexical Priming appears to give a sound explanation of why collocations, colligations and semantic associations exist, and this can be seen as one of the 'scientific advances' O'Keefe *et al*. (2007) speak about.

Notes

1. Parts of this article use the same material as M. Pace-Sigge, *Lexical priming in spoken English*. (forthcoming). Houndmills, Basingstoke: Palgrave Macmillan. Reproduced with permission of Palgrave Macmillan.
2. Hunston and Francis (3: 1999). See Hornby's 1954 book *A guide to pattern and usage of English* as their forerunner.
3. Cf. Richard Smith 1999.
4. Cf. Sinclair 1996, Stubbs 1996 and Hoey 2005.
5. *Semantic Association* is Hoey's term. He tries to combine what has previously been referred to as *semantic prosody* or *semantic preference.* See Hoey (2005: 22ff.).
6. "Priming" designates hypothetical processes that underlie the priming effect, the empirical finding that identification of objects is facilitated by the individual's previous encounter with the same or similar *objects*. (My italics). (Habib 188: 2001).
7. Which Hoey refers to in *Lexical priming* (8: 2005).
8. Hoey (2005) names the same paper as published in 1977. JHN published two parts of the same paper (with different subtitles) in two different publications in two consecutive years.
9. Neely (1991) describes how Posner and Snyder's work was his main influence and how Posner and Snyder were influenced by Meyer and Schvanefeld. All four appear in the bibliography of Neely (1977).
10. "The program's strategy is presented as a general theory of language comprehension." Quillian (459: 1969).
11. The relation between Teachable Language Comprehender, a semantic performance model, and the syntactic 'competence' models of transformational linguistics (Chomsky 1965) is not clear. The efforts that have been made so far to attach 'semantics' to transformational models seem, to this writer at least, to have achieved little success (Quillian, 461: 1969).

12. Meyer and Schvanefeldt (1976) suggest that Quillian is right in an experiment where words are made harder to read.
13. The man who heads Google's Research Division is the co-author of *Artificial Intelligence: A modern approach* (1995, 2003, 2010), Peter Norvig, who, in his work, quotes Quillian extensively.
14. Here Quillian refers to a story about (President) George Washington who felled his father's cherry tree. This is apparently a commonly known story in the USA.
15. "True" and "false" sentences – their terminology.
16. See Quillian (472: 1969): "[this] does not output as a parsing a tree structure, but rather a set of nested strings. However, in building these strings it succeeds in 'undoing' a number of syntactic transformations, replacing deleted elements and rearranging others".
17. I discuss the link of "meaning" and "priming" in Section 3.4.1.
18. Described by Neely (649: 1976) as follows: "a semantically neutral warning prime consisting of a series of Xs".
19. To find this, I have made use of the University of Liverpool's Summon, Scopus and Discover systems. A further search was made on Google Scholar.
20. The research is based on the most frequent phrases found in spoken and written academic texts in the BNC.
21. Ellis *et al.* 2006a and Ellis *et al.* 2006 are PowerPoint presentations; hence no page numbers are given.
22. Semantic prosody, based on the definitions by Louw and Sinclair, is here described as the consistent aura of meaning with which a form is imbued by its collocates and the general tendency of certain words to co-occur with either negative or positive expressions.
23. This fact has been highlighted by John Sinclair.

References

Ashcraft, Mark H. 1976. Priming and property dominance effects in semantic memory. *Memory and Cognition* 4 (5): 490–500.

Brazil, David. 1995. *A grammar of speech*. Oxford: Oxford University Press.

Collins, Alan M. and R. Michael Quillian. 1969. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behaviour* 8: 240–248.

Collins, Alan M. and R. Michael Quillian. 1970. Facilitating retrieval from semantic memory: The effect of repeating part of an inference. *Acta Psychologica* 33: 304–314.

Collins, Alan M. and R. Michael Quillian. 1972a. Experiments on semantic memory and language comprehension. In L. W. Gregg (ed.). *Cognition in learning and memory,* 117–137. New York: Wiley.

Collins, Alan M. and R. Michael Quillian. 1972b. How to make a language user. In E. Tulving and W. Donaldson (eds.). *Organisation of memory,* 319–322. New York: Academic Press.

Collins, Alan M. and Elizabeth F. Loftus. 1975. A Spreading-Activation Theory of semantic processing. *Psychological Review* 82 (6): 407–428.

Ellis, Nick E., Eric Frey and Isaac Jalkanen. 2006a. The psycholinguistic reality of collocation and semantic prosody. Presentation. Exploring the Lexis–Grammar Interface. Hanover, Germany 5–7th October.

Ellis, Nick E., Rita Simpson-Vlach and Carson Maynard. 2006b. The processing of formulas in native and L2 speakers psycholinguistic and corpus determinants. Presentation. Exploring the Lexis-Grammar Interface. Hanover, Germany 5–7th October.

Firth, John R. 1957. *Papers in linguistics 1934–1951*. London: Oxford University Press.

Gries, Stefan Th. 2005. Syntactic priming. A corpus-based approach. *Journal of Psycholinguistic Research* 34 (4): 365–399.

Gries, Stefan Th. 2009. *Quantitative corpus linguistics with R*. New York and London: Routledge.

Habib, Reza. 2001. On the relation between conceptual priming, neural priming and novelty assessment. *Scandinavian Journal of Psychology* 42: 187–195.

Halliday, M.A.K. 1959. *Language of the Chinese "Secret history of the Mongols"* London: Longman.

Hernandez, Arturo E., Christine Fennema-Notestine, Care Udell and Elizabeth Bates. 2001. Lexical and sentential priming in competition: Implications for two-stage theories of lexical access. *Applied Psycholinguistics* 22: 191–215.

Hoey, Michael. 2003. Why grammar is beyond belief. *Belgian Essays on Language and Literature* 2003: 183–196. Liège: University of Liège.

Hoey, Michael. 2005. *Lexical Priming. A new theory of words and language.* London: Routledge.

Hogg, Michael and Joel M. Cooper (eds.) 2003. S*AGE handbook of sosial psychology.* London: SAGE.

Hornby, Albert S. 1954. *A guide to patterns and usage in English.* London: OUP.

Hunston, Susan and Gill Francis. 1999. *Pattern Grammar. A corpus-driven approach to the lexical grammar of English.* Amsterdam: John Benjamins.

Kahneman, David. 2011. *Thinking fast and slow*. New York: Farrar, Straus and Giroux.

Levy, Steven. 2011. *In the Plex. How Google thinks, works and shapes our lives*. New York: Simon and Schuster.

Loftus, Elizabeth F. 1973. Activation of semantic memory. *American Journal of Psychology* 86: 331–337.

Meyer, David E. and Roger W. Schvanefeldt. 1971. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology* 90 (2): 227–234.

Meyer, David E. and Roger W. Schvanefeldt. 1976. Meaning, memory structure and mental processes. *Science* 192 (2): 27–33.

Meyer, David E. and Roger W. Schvanefeldt. 1984. Review of David E. Meyer and Roger W. Schvanefeldt, "Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations" (1971). A "Citation classic". *Current Contents: Social and Behavioral Sciences* (1984) 16 (No. 47): 20.

de Mornay Davies, Paul. 1998. Automatic semantic priming: The contribution of lexical- and semantic-level processes. *European Journal of Cognitive Psychology* 10 (4): 389–412.

Mussweiler, Thomas. 2000. The use of category and exemplar knowledge in the solution of anchoring tasks. *Journal of Personality and Social Psychology* 78: 1038–1052.

Neely, James H. 1976. Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory and Cognition* 4 (5): 648–654.

Neely, James H. 1977. Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology* 106: 226–254.

Neely, James H. 1989. Experimental dissociations and the episodic/semantic memory distinction. In H. L. Roediger and F. I. M. Craik (eds.). *Varieties of memory and consciousness: Essays in honour of Endel Tulving***,** 229–241. Hillsdale, NJ: Erlbaum.

Neely, James H. 1991. Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner and G. W. Humphreys (eds.). *Basic processes in reading: Visual word recognition*, 264–336. Hillsdale, NJ: Erlbaum.

Nelson, Michael. 2000. *The methodological background: British traditions of text analysis, correlative register analysis and corpus linguistics.* http://users.utu.fi/micnel/thesis/Thesis1.doc. Last accessed 11.02.2013.

Novick, Jared M., A. Kim and J. Trueswell. 2003. Studying the grammatical aspects of word recognition: Lexical priming, parsing and syntactic ambiguity resolution. *Journal of Psycholinguistic Research* 32 (1): 57–75.

O'Keefe, Anne, Michael McCarthy and Ronald Carter. 2007. *From corpus to classroom*. *Language use and language teaching*. Cambridge: Cambridge University Press.

Pace-Sigge, Michael T. L. Forthcoming. *Lexical priming in spoken English*. Houndmills, Basingstoke: Palgrave Macmillan.

Posner, Michael I. and Charles Richard R. Snyder. 1975a. Facilitation and inhibition in the processing of signals. In P. M. A. Rabbitt and S. Domic (eds.). *Attention and performance* V: 669–681. New York: Academic Press.

Posner, Michael I. and Charles Richard R. Snyder. 1975b. Attention and cognitive control. In: L. Solso (ed.). *Information processing and cognition: The Loyola symposium,* 55–85. Hillsdale, NJ: Erlbaum.

Quillian, R. Michael. 1961. The elements of human meaning: A design for and understanding machine. *Communications of the ACM* 4 (9): 406–408.

Quillian, R. Michael. 1962. A revised design for an understanding machine. *Mechanical translation* 7: 17–29.

Quillian, R. Michael. 1966. Semantic memory. Unpublished doctoral dissertation, Carnegie Institute of Technology. Reprinted in part in M. Minsky (ed.). *Semantic information processing*. (no pages available). Cambridge, Mass.: M.I.T. Press, 1968.

Quillian, R. Michael. 1967. Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral Science* 12: 410–430.

Quillian, R. Michael. 1969. The Teachable Language Comprehender: A simulation program and theory of language. *Computational Linguistics* 12 (8): 459–476.

Sherman, Steven J., Matthew T. Crawford, David L. Hamilton and Leonel Garcia-Marques. Social inference and social memory: The interplay between systems. In M. Hogg and J. Cooper (eds.) 2003. *The SAGE handbook of social psychology*, 45–68. London: SAGE Publications Ltd.

Sinclair, John M. 1991. *Corpus, concordance, collocation.* Oxford: Oxford University Press.

Smith, Richard E. 1999. *The writings of Harold E. Palmer: An overview.* Tokyo: Hon-no-Tomosha.
http://www.warwick.ac.uk/~elsdr/WritingsofH.E.Palmer.pdf.
Last accessed 12.02.2013.

Stubbs, Michael. 1996. *Text and corpus analysis. Computer-assisted analysis of language and culture*. Oxford: Basil Blackwell.

Trofimovich, Pavel. 2005. Spoken-word processing in native and second languages: An investigation of auditory word priming. *Applied Psycholinguistics* 26: 479–504.

Zimmermann, Harald H. 1972. Zur Konzeption der automatischen Lemmatisierung von Texten SFB 100 "Elektronische Sprachforschung": Aspekte d. automatischen Lemmatisierung. Bericht 10–72. *Linguistische Arbeiten* 12: 4–10.