Entangling words and meaning

P.D. Bruza & K. Kitto
Queensland University of Technology
Australia
p.bruza@qut.edu.au

D. Nelson & K. McEvoy
University of South Florida
USA

Abstract

Human memory experiments appear to be generating "non-local" effects (Nelson & McEvoy 2007). In this paper the possibility that words might be entangled in human semantic space is seriously entertained. This approach leads to a very natural picture of the way in which context might affect word association via the standard interpretation of quantum measurement. Two possible scenarios for testing such a hypothesis are suggested, both based upon potential violations of the CHSH inequality.

Introduction

Human memory experiments study how words are recalled in a variety of experimental settings. The basic elements are a cue q and a word w, with the result being represented as the probability $\Pr(w|q)$ - the probability word w was recalled in relation to the cue q. In the light of this conditional probability, q can be viewed as the context in relation to which the probability of w is given.

These conditional probabilities allow words to be modelled as a network of associations. Consider figure 1 which has been adapted from (Nelson & McEvoy 2007). Nodes in this figure correspond to words, where "Target" denotes a specific word of interest and "Associate 1" and "Associate 2" are two different associations of "Target". The figure also reflects "Associate 1" being an associate of word "Associate 2". Edges reflect conditional probabilities collected from free association experiments. For example, the probability that subjects produce "Associate 2" when cued with the word "Target" is x. See (Nelson & McEvoy 2007) for a more detailed exposition of such word association networks. In this work, the authors argued convincingly that spreading a activation model underestimates the probability of recall of a target word. The crux of their argument can be understood with reference to figure 1. Firstly, we observe that there are no arrows (associate-to-target links) going back to the target which would allow probability to flow back to contribute to the target's recall. Despite this lack of direct links, the dynamics of the recall of the target can change. It turns out that it is the number of associate-associate links which plays a fundamental role in establishing an accurate estimation of

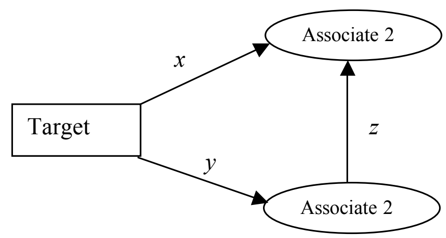

Figure 1: Word association network

a given targets probability of recall. For this reason, Nelson, McEvoy and their colleagues have proposed a formula for calculating a given target's recall based on the intuition of "spooky activation at a distance". In this equation the target and the associates in the network are deemed to activate in synchrony. As a consequence, the equation assumes that each link in the network contributes additively to the net strength irrespective of whether they are associate-to-target links, or not.

Nelson and McEvoy have recently begun to consider the spooky-activation-at-a-distance formula in terms of quantum entanglement, "The activation-at-a-distance rule assumes that the target is, in quantum terms, entangled with its associates because of learning and practicing language in the world. Associative entanglement causes the studies target word to simultaneously activate its associate structure" (Nelson & McEvoy 2007, p3). This is speculation the present authors seriously entertain. The aim of this article is to investigate quantum-like entanglement of words in human memory and to propose experiments which may be performed to provide evidence for its existence. We shall start by considering the contextuality of word meanings and associations, as this contextuality will be used in the construction of the experiments.

Words, Context and Hilbert Space

Often, the meanings attributed to a word depend strongly upon the context in which that word appears. For example, the word bat has a number of possible meanings, or senses, depending upon the context; it could mean the small furry creature generally found in caves (e.g. a vampire bat); it could be a sporting implement (e.g. a cricket bat); it could also mean take the more colloquial usage of a strange old lady (e.g. an old bat). A number of verb meanings are also possible: an idea might be batted around; someone might bat at a ball; they may even bat their eyelashes in order to attract attention and admiration. Clearly, we can only distinguish between this range of different senses by looking at the context in which the word occurs. There are very few formalisms capable of modelling such contextual dependencies, quantum theory is one (Kitto 2006). It seems possible that words are somehow entangled, with the different senses of words all being active in some way until we actively create meaning by collapsing what is potentially a massively entangled state. This would select out one possible sense from the large array of possibilities with reference to the way in which the word is being used in the current context. If this was the case then it should be possible to generate violations of Bell-type inequalities which would indicate that the processing involved in determining the meaning of words is not in fact separable from either the context in which they are used, or even from other meanings. This paper shall discuss some of the issues that arise with the attempt to apply a quantum formalism to human memory experiments.

In quantum theory, the context of a system is represented via the choice of measurement settings, which is implemented with a choice of the basis used to describe the system. Many different choices of experimental settings are often possible, and these are represented by different choices of basis, which then affects the results obtained via the measurement postulate (Kitto 2008b).

When applying the quantum formalism to memory experiments, a specific cue q is taken to specify a basis in a Hilbert space representing that particular meaning of the of the word w, which is represented in this picture as a state vector. Thus, the basis takes the form $\{|0_q\rangle, |1_q\rangle\}$, where the basis vector $|0_q\rangle$ represents the basis state "not recalled" and $|1_q\rangle$ represents the basis state "recalled" in relation to the cue q. A word w is assumed to be in a state of superposition reflecting its potential to be recalled, or not, in relation to the given cue q:

$$|w\rangle \quad = \quad b_0|0_q\rangle + b_1|1_q\rangle$$

However, as we saw above, a word w can have a number of different meanings, each of which can be represented through a different choice of basis. Thus, the same word w's state in relation to another cue p is accordingly:

$$|w\rangle \quad = \quad a_0|0_p\rangle + a_1|1_p\rangle$$

This idea of a word in the context of two different cues, represented via two different bases is depicted in figure 2. Word w is represented as a unit vector and its representation is expressed with respect to the bases $\mathcal{P} = \{|0_p\rangle, |0_p\rangle\}$ and $\mathcal{Q} = \{|0_q\rangle, |0_q\rangle\}$ as described above.

In a quantum formalism, superposed states such as w are not measured, rather, a word is either recalled $|1_q\rangle$ or it is not recalled $|0_q\rangle$, with a certain probability. This probability is obtained via one of the core postulates of quantum theory,

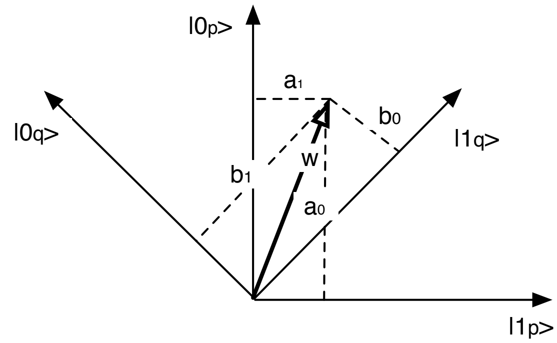

Figure 2: Word vector w with respect to two different bases

the projection postulate, which considers some observable quantity A, represented by a self-adjoint operator, $\hat{A}$, and a (normalised) state $w \in \mathcal{P}$ and returns the expected value:

$$\langle A\rangle_w = \langle w, \hat{A}w\rangle = \sum w^*\hat{A}w \tag{1}$$

where we have made use of the discrete structure of the Hilbert spaces as they were defined above. At this point it is necessary to elaborate upon the nature of the operator $\hat{A}$, and the way in which they interact with the state w. Self-adjoint operators are defined as those for which $\hat{A} = \hat{A}^\dagger$, which implies that when the expected value is calculated for the vector $w \in \mathcal{H}$ and another vector $v \in \mathcal{H}$ (also defined in the same subspace), that

$$\langle v, \hat{A}w\rangle = \langle \hat{A}v, w\rangle \tag{2}$$

$$= \langle v, \hat{A}w\rangle^* \tag{3}$$

hence self-adjoint operators are *probability conserving*, a characteristic which is essential to the interpretation of the projection postulate (discussed in more detail in (Kitto 2008b), this volume). Self-adjoint operators have a number of useful properties which relate to this property of probability conservation; their eigenvalues are real numbers, and the eigenvectors that correspond to two different eigenvalues are orthogonal. Clearly both of these characteristics aid in the interpretation of measurement in quantum theory.

$\hat{A}$ is often taken to be a projection operator P. These are operators that project a vector onto some subspace of Hilbert space, which by definition have eigenvalues of 0 and 1 alone, making calculations much more simple than in the general self-adjoint case; projection operators correspond to the binary case of respond or not. We shall make use of projection operators in what follows for the sake of simplicity alone.

Consider once again word w in relation to the cue q, with $|w\rangle = b_0|0_q\rangle + b_1|1_q\rangle$. When a subject is presented with q, w will, or will not, be recalled. This is akin to a quantum measurement which "collapses" the superposition onto the corresponding basis state yielding an outcome. For a projector, the probability of obtaining a measurement result $\lambda_i, i \in \{0, 1\}$ corresponding to basis state $|i\rangle$ is given by a simplified version of the projection postulate (1):

$$\Pr(\lambda_i) = \langle i|w\rangle^2 = b_i^2 \tag{4}$$

Referring to figure 2, one see that the probabilities are of a geometric nature, and due to Pythagoras' theorem $b_0^2 + b_1^2 = 1$ (recall w is represented by a vector of unit length) which is a unique characteristics of quantum probability (Isham 1995). When a subject is presented with a cue q, w collapses onto the basis state $|0_q\rangle$ ("spin down— w not recalled") or $|1_q\rangle$ ("spin up — w recalled") with probabilities b_0^2 or b_1^2 respectively.

For this simplified projection picture, the spectral decomposition of Z relates the outcomes to the projection operators P_i, one projection operator per outcome (due to the fact that eigenvalues are unique and correspond to a unique eigenfunction) to the expected value:

$$\begin{aligned} Z &= \lambda_0 P_0 + \lambda_1 P_1 \\ &= \lambda_0 |0\rangle\langle 0| + \lambda_1 |1\rangle\langle 1| \end{aligned}$$

The projector $P_0 = |0\rangle\langle 0|$ is associated with the outcome of not recalling w and $P_1 = |1\rangle\langle 1|$ is associated with w being recalled. In quantum mechanics (QM), measuring a quantum system as described above unavoidably disturbs it leaving it in a state $|0\rangle$ or $|1\rangle$ determined by the outcome. We feel this phenomenon carries across to memory experiments in the sense that recalling a word, or not, unavoidably disturbs the state of the memory of subject in question.

It is important to realise that during the process of measurement, a word is actively entangled with the measuring apparatus, in this case the cue word. This is the very source of the measurement problem in quantum theory, and was discussed very early in the development of QM by (von Neumann 1955), who showed that an infinite regress looms when quantum theory is assumed to describe all of reality. This is not such a problem in the quantum interaction picture where a quantum—classical divide is not so problematic. With reference to the interpretation of quantum theory discussed in (Kitto 2008b) it would merely indicate an interaction between a very complex system and another that was more amenable to reductive analysis; in other words a controllable interaction. We can sketch out a simple model of this form of entanglement using an argument similar to Von Neumann's. First let us return to the superposed word $w = b_0|0_q\rangle + b_1|1_q\rangle \in \mathcal{H}_1$ representing a combined state of non-recall and recall of the word given the cue q. Towards the end of eventually constructing a full model of the collapse of meaning onto one specific outcome, we can construct a toy model by considering this word in the fuller context of a persons mind, say in an initially pure state $|M_i\rangle \in \mathcal{H}_2$ for the sake of simplicity. We have explicitly specified two different Hilbert spaces for the word and the mind states, there is no reason to suppose that the two spaces are equivalent, indeed it is far more likely that $\mathcal{H}_1 \neq \mathcal{H}_2$ since words generally have contexts beyond the representations that an individual might have. When presenting a person with a cue, it is expected that this will cause the target word w to become entangled with the cognitive state $|M_i\rangle$, with the associated enlargement of the Hilbert space via the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$. While the details of the actual process of time evolution are not known, something similar to the time evolution that occurs in the Schrödinger equation is postulated to occur:

$$|M_i\rangle \left(b_0|0_q\rangle + b_1|1_q\rangle\right) \rightarrow b_0|M_0\rangle|0_q\rangle + b_1|M_1\rangle|1_q\rangle \quad (5)$$

thus a combined state will form that cannot be written as a simple product of the two initial states; it is a superposition of product vectors (similar to the Schrödinger cat type states that form in standard quantum theory). Such states do not occur in classical formalisms, and can be considered as exemplars of the strange effects of which the quantum formalism is capable. However, in a quantum model there is a second dynamical step that occurs, where the state represented by (5) collapses onto either:

$$\begin{cases} |M_0\rangle|0_q\rangle \text{ with probability } |b_0|^2, \text{ i.e. word is not recalled, or} \\ |M_1\rangle|1_q\rangle \text{ with probability } |b_1|^2, \text{ i.e. word is recalled.} \end{cases} \quad (6)$$

Thus, even at this stage, this model is markedly different from the standard more 'classical' accounts, however, another possibility presents itself, namely that words themselves might become entangled, leading to states similar to the Einstein–Podolsky–Rosen (EPR) entangled states of standard quantum theory.

Entanglement of Words

Particularly interesting effects can be achieved with the direct entanglement of words, which occurs when words are combined in certain ways. By way of illustration, consider the words u and v in relation to the cue q. Following from the preceding section this means they are represented by the kets $|u\rangle, |v\rangle$ in the basis $\mathcal{Q} = \{|0_q\rangle, |1_q\rangle\}$. Further, assume $|u\rangle = a_0|0\rangle + a_1|1\rangle$ and $|v\rangle = b_0|0\rangle + b_1|1\rangle$ whereby the subscript q on the basis vectors has been momentarily dropped for reasons of notational convenience.

One possible state of the combined system would be a vector denoted $|u\rangle \otimes |v\rangle$ in the tensor space $\mathcal{Q} \otimes \mathcal{Q}$ whereby:

$$\begin{aligned} |u\rangle \otimes |v\rangle =& (a_0|0\rangle + a_1|1\rangle)(b_0|0\rangle + b_1|1\rangle) && (7) \\ =& a_0b_0|0\rangle|0\rangle + a_0b_1|0\rangle|1\rangle + a_1b_0|1\rangle|0\rangle \\ &+ a_1b_1|1\rangle|1\rangle && (8) \end{aligned}$$

The latter is often written as:

$$|u\rangle \otimes |v\rangle = a_0b_0|00\rangle + a_1b_1|10\rangle + a_0b_1|01\rangle + a_1b_1|11\rangle \quad (9)$$

as the basis of the tensor space $\mathcal{Q} \otimes \mathcal{Q}$ is denoted by $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. These basis vectors are formed by taking the tensor product of the original basis vectors:

$$|00\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad (10)$$

Similarly for the other basis vectors in $\mathcal{Q} \otimes \mathcal{Q}$. (See (Lomonaco 2002, p 35.))

In the light of the running example, the basis vector $|00\rangle$ corresponds to the outcome where neither u nor v are recalled in response to cue q; $|01\rangle$ corresponds to the outcome u is *not recalled* and v *is* recalled with respect to q etc.

Equation (9) represents a superposition is 4 dimensional space. The probability of each outcome is once again given by Born's rule. So, for example, the probability that both u and v are recalled is $|a_1 b_1|^2$. It is important to realise however that (9) is *not* an entangled state. This is due to the way in which it was obtained, we simply took the product of $u \otimes v$, and by definition an entangled state is one that it is impossible to write as a product.

To obtain an entangled state, we might consider the the state ψ when words u and v are either *both* recalled, or both *not* recalled in relation to q:

$$\psi \quad = \quad \alpha|00\rangle + \beta|11\rangle \tag{11}$$

where $\alpha^2 + \beta^2 = 1$. This state is impossible to write as a product state, thus it differs markedly from (9). This seemingly innocuous state is one of the so-called *Bell states* in QM. The fact that entangled systems cannot be expressed as a product of the component states makes them non-separable. More specifically, there are no coefficients a_0, a_1, b_0, b_1 which can decompose equation (11) into a product state exemplified by equation (7). For this reason ψ is not written as $u \otimes v$ as it can't be represented in terms of the component states $|u\rangle$ and $|v\rangle$. The physicist Erwin Schrödinger once said, "I would not call [entanglement] one but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought" (Schrödinger 1935, p 807). He was right to emphasise this uniqueness. We have now seen two examples of entangled states, quantum measurement, and the Bell-type entanglement of two states, and both are characteristic of the weird results of quantum mechanics; such behaviour is not exhibited by classical formalisms.

However, as was discussed in (Kitto 2008b), and in more detail in (Kitto 2008a) such behaviour is often exhibited by systems exhibiting very complex behaviour, or high-end complexity. Thus there is a need for us to develop our formalisms to the extent that they are capable of modelling such nonseparable behaviour. One such attempt falls very much under the umbrella of quantum interaction, in modelling systems not generally deemed as quantum systems using the quantum formalism much can be learnt about the nonseparable, and contextual aspects of those systems behaviour, as well as about the nature of the quantum formalism itself.

The remainder of this discussion will focus upon particular experimental scenarios that can be investigated which will test the assumption that human memory experiments can be modelled using the quantum formalism.

Bell inequalities for general systems

Following (Aerts *et al.* 2000) we shall construct Bell inequalities by considering some entity S and four experiments e_1, e_2, e_3, e_4 that can be performed upon it. Each experiment has two possible outcomes: up and down, which we shall denote $|1\rangle$ and $|0\rangle$ respectively in keeping with the above formalism. Some of the different experiments can be combined, which leads to a number of *coincidence experiments* $e_{ij}, i, j \in \{1, 2, 3, 4\}, i \neq j$. Combinatorically, there are four outcomes possible from performing such a coincidence experiment, $|11\rangle, |10\rangle, |01\rangle$, and $|00\rangle$. Note that there is no *a priori* reason to expect the results of the coincidence experiments to be compatible with an appropriate combination of the results of the two *separate* experiments, although the utility of such reductive assumptions in science has generally been considerable.

Expectation values[1] can be introduced which sum the probabilities for the different coincidence experiments, weighted by the outcome itself; it is the sum of all the values that the two experiments can yield, weighted by the (experimentally obtained) probabilities of obtaining them. For the current setup, if an $|11\rangle$ is obtained the value of the experiment is +1 (denoting a correlation), and the result is the same if a $|00\rangle$ results, whereas if either $|10\rangle$ or $|01\rangle$ is obtained then a -1 is recorded. Thus, for the experimental combination e_{ij}:

$$\begin{aligned} E_{ij} = +1.P(|11\rangle) + 1.P(|00\rangle) \\ - 1.P(|10\rangle) - 1.P(|01\rangle). \end{aligned} \tag{12}$$

These expectation values allow us to test the appropriateness of the reductive separability assumption, via a Bell-type inequality:

$$|E_{13}E_{14}| + |E_{23} + E_{24}| \leq 2 \tag{13}$$

which assumes in its derivation that the outcome of one sub-experiment in a coincidence pair does not affect the other and vice versa. If (13) is violated then this implies that this assumption is invalid, and that the coincidence experiment is fundamentally different from the combination of the two separate sub-experiments. This particular inequality is due to (Clauser *et al.* 1969) and is known as the CHSH inequality.

If we take the three noun-type senses of bat mentioned earlier, then it should be possible to generate an entangled state in the memory of a subject through an appropriate priming procedure. Such a state would take the form of a complex activation of all possible meanings in the mind of the subject, which could then be collapsed onto one specific meaning when they were subjected to cues that provide a limiting context in which to understand the words. If entanglement has indeed occurred, then the associations that a subject produces after this collapse should be changed by the prior entangled state, and a count of the associations produced should yield different results from if the counts of the associations were combined posthoc via a separability assumption. This would suggest that a violation of (13) should be exhibited.

The preparation of an appropriate entangled state is key to all of the experimental procedures that follow in this section. To this end, subjects would take part in an extralist cuing procedure whereby a subject is first asked to study a list of words with the intention of the experiment not being

[1] In probability theory the expectation of a discrete random variable is the sum of the probability of each possible outcome of the experiment multiplied by the outcome value, this translates in quantum theory into the predicted mean value of the result of an experiment. The general expression for a quantum mechanical expectation value is $\langle \hat{A} \rangle_\psi = \sum_j a_j |\langle \psi, \phi_j \rangle|^2$, where a_j are the eigenvalues, and ϕ_j are the eigenvectors associated with the operator $\hat{A}$.

disclosed. Each word is studied in isolation for a couple of seconds before the next study word is produced. After all words are studied, the subject is presented with a cue word and asked to recall a single word, or words, from the list just studied. It is supposed that by studying the words an entangled state is created it can be probed in a number of different ways. This section will now turn to a discussion of two different experimental procedures that might be conducted to investigate such an entangled state.

Non-separability in semantic space

The first experimental procedure is intended to probe the nonseparable nature of word meanings via a direct violation of a Bell-type CHSH inequality. We shall choose four sub-experiments, each of which arise from exposing a subject to a different cue (q):

$$e_1 : q = \text{bat},\ e_2 : q = \text{stick},\ e_3 : q = \text{lady},\ e_4 : q = \text{animal} \tag{14}$$

and then asking them to give a set of associations from the list of priming words that they have seen during the preparation phase. Rather than testing these sub-experiments, four coincidence experiments will be run, based upon the following four coincidence possibilities:

$$e_{13} : (\text{bat}, \text{lady}) \tag{15}$$
$$e_{14} : (\text{bat}, \text{animal}) \tag{16}$$
$$e_{23} : (\text{stick}, \text{lady}) \tag{17}$$
$$e_{24} : (\text{stick}, \text{animal}) \tag{18}$$

In the preparation phase for this experimental procedure a subject might see the following list of words to study: *furry, cricket, skirt, sport, grey hair, gloves, ball, fly, female, night, strange, rodent, pitch, blind*. Notice that some of these words can actually be associates of more than one word meaning. For example, the animal bat is blind, and an old lady might also be blind, while a blindside can occur in sport (although not generally legally). This should assist in the formation of an entangled mind state.

The subject will then be instructed they will see two cue words *simultaneously* in the same way as they were exposed to the priming words and asked to recall as many words as they can from the priming list. At this point one of the four e_{ij} experiments will be run. After the experiment has concluded each association word that was recalled will be examined for its agreement with the meaning of the experimental setup. An association word that correlates with both of the cuing words with be recorded as returning an $|11\rangle$, a word that correlates with neither of the cuing words will be considered as returning as $|00\rangle$, an association appropriate to the first cue but not the second will be considered as a $|10\rangle$, and the converse case will lead to the recording of a $|01\rangle$.

It is expected that when these results are counted and the appropriate expectation values obtained, that equation (13) will be violated.

A direct entanglement effect

Illustrating the non-separable nature of words is one thing, but it may be possible to directly illustrate entanglement of the entanglement of two words. Consider once again the word *bat*. Assume, for simplicity, it has two senses, the animal sense, and a sporting sense. Several authors have put forward the proposition that within cognition *bat* can be modelled as $|b\rangle$ a superposition of its senses (Widdows 2004; Bruza & Cole 2005; Aerts, Broekaert, & Gabora 2005; Bruza, Widdows, & Woods 2007). As indicated earlier, *bat* can be represented as the superposition:

$$|b\rangle \quad = \quad a_0|0\rangle + a_1|1\rangle$$

where $|0_q\rangle$ now corresponds to the sport sense being recalled and $|1_q\rangle$ to the animal sense being recalled in response to a given cue q. When *bat* in seen in context, for example, *vampire bat* the superposition collapses onto the animal sense. The coefficients a_0 and a_1 relate to the probabilities of collapse as described earlier. Quite remarkably, some knowledge of the *bat* superposition, can be recovered by studying free association norms using the word *bat* as cue. It turns out the the majority of the free associates are related to the animal sense of *bat*. However, the most probable associate *ball* ($p(ball|bat) = 0.25$) relates to the sport sense (Nelson, McEvoy, & Schreiber 1998).

The word *boxer* also has a sport and animal sense, i.e., the breed of dog. Examination of the free association norms reveals the sport sense is heavily favoured via associates such as *fighter, gloves, fight, shorts, punch*. The sole free associate related to the animal sense is *dog* ($p(dog|boxer) = 0.08$) (Nelson, McEvoy, & Schreiber 1998).

Consider the possibility both *bat* and *boxer* are superpositions in a particular entangled cognitive state. Just like photons, senses of a word can be thought of as polarizing, that is, the sense can be "up" (animal sense) or "down" (sport sense). The direct entanglement of words in memory relies on the hypothesis that the collapse of one word onto a sense, influences the collapse of the other word. In accord with the above background whereby particles are prepared in a certain way, and measurement of one yielding $|1\rangle$ (spin "up") also leads to $|1\rangle$ (spin "up") in the other sub-experiment, the senses of a word will be viewed as polarizing either "up" or "down" with respect to a given cue.

The simplest type of entangled state is a Bell state as exemplified in equation (11). In experiments involving the entanglement of photons, a Bell state must first be prepared, which is the subjected to subsequent experiment using spatially separated detectors. This raises the thorny question of whether it is possible to actually create a Bell-type state for words in human memory. In other words, a state ψ must be prepared of the form depicted in equation (11) whereby the state is a superposition of both senses being "up" and both "down". One possibility for potentially creating such a state is to alter the preparation phase of extra-list cuing experimental framework. Recall in this framework the subject is first asked to study a list of words with the intention of the experiment not being disclosed. Each word is studied in isolation for a couple of seconds before the next study word is produced. It was shown above in relation to *bat* and *boxer* how the senses of a word are reflected in the free associates of the word. It is possible to turn this around and use the free associates to prepare a Bell-type entangled word state as reflected in equation (13)? By way of illustration, consider once again the words *bat* and *boxer*. The study list would comprise two parts - each part relating to one of the senses. For example, the first part would contain those associates of *boxer* and *bat* related to the animal sense such as *dog, cave, vampire, night, blind*. The intention here is to correlate *boxer* and *bat* only within the animal sense, represented by the basis vector $|11\rangle$ ("up","up") in equation (13). The second part of the study list is made up of associates related to the sport sense such as *gloves, punch, fight, ball, baseball*. The intention of the second part of the list is to correlate the words only within the sport sense, represented by the basis vector $|00\rangle$ ("down", "down"). By studying the associates of both words *within* a given sense, the hope is that the anti-correlations $|10\rangle$ ("up","down") and $|01\rangle$ ("down","up") will be negligible.

Assuming that something akin to a Bell state can be prepared in human memory, the next step is to devise an experiment so the CHSH inequality of equation (13) can be applied. Unlike the experiment of the previous section, the coincidence experiments are realised by cuing a subject *twice* after the study period. Example cues are as follows:

$$e_1 : q = \text{black, bat},\ e_2 : q = \text{bat},$$
$$e_3 : q = \text{boxer},\ e_4 : q = \text{black, boxer} \quad (19)$$

The subjects are asked to recall the first word from the study list that comes to mind in response to a cue. The four coincidence experiments are set out as follows:

$$e_{13} : (\text{black bat}), (\text{boxer}) \quad (20)$$
$$e_{14} : (\text{black bat}), (\text{black boxer}) \quad (21)$$
$$e_{23} : (\text{bat}), (\text{boxer}) \quad (22)$$
$$e_{24} : (\text{bat}), (\text{black boxer}) \quad (23)$$

A given subject would take part in one experiment and be "measured" by two cues. For example, in a coincidence experiment e_{13}, a given subject would be first cued with *black bat* and a word recalled, say "vampire". In this case observe how the recalled word reflects the animal sense of "bat", i.e., spin "up". This is not surprising as the context word *black* could easily promote the animal sense, though note this is not a certainty as baseball bats are often black. This is a deliberate design which is akin to setting a polarizer in a certain direction. What happens next is potentially more interesting. The same subject would then be presented with the cue *boxer* and a word recalled, say "dog" (spin "up"). In this example, the recalled word reflects the animal sense of *boxer*. Recall the sport sense of *boxer* is heavily dominant. The supposition here is that the collapse of *bat* onto the animal sense has influenced the collapse of *boxer* onto the animal sense, and this influence has occurred in the face of an *a priori* strong tendency towards the sport sense of *boxer*. It is this influence which is central to entanglement and is the essence of why it is considered so weird and surprising. In this connection, it is important that the context word associated with a second cue be chosen more or less neutrally so the influence of the first cue is not washed out by a local context effect when the second cue is activated. Note that the second cue *black boxer* is ambiguous and leaves room for both the animal and sport senses to manifest.

The above experiment would be run using n subjects. The subjects would be divided into four groups G1, G2, G3, G4 of equal size. All subjects would be prepared in the same fashion. Subjects in G1 would be given experiment e_{13}, subjects in G2 given e_{14}, subjects in G3 given e_{23} and subjects in G4 given e_{24}. The expectations of the experiments can be calculated using equation (12). For example, in order to calculate E_{13} requires $P(|11\rangle)$ to be estimated. This can be computed by counting the number of subjects in G1 which collapsed the state onto the animal sense in response to the cues *black bat* and *boxer*. That is, both outcomes are spin "up". This value is then divided by $\frac{n}{4}$, the size of G1. Similarly the for the other three cases: $|00\rangle$ (both outcomes spin "down"), $|10\rangle$ (first outcome spin "up", the second spin "down"), and $|01\rangle$ (first outcome spin "down", the second spin "up").

What would it mean if the inequality of equation (13) is violated? As described in the previous section, it suggests that *bat* and *boxer* are inseparable in the cognitive state resulting from the preparation. Granted the preparation proposed above is highly artificial, however, one can speculate whether in certain circumstances context acts like the preparation procedure above, yielding something like a Bell state in memory. For example, in general the words *Reagan* and *North* would be distant in human semantic space, however, according to the intuition above, seeing *Reagan* in the context of *Iran* leads the collapse of *Reagan* onto a basis state (sense) of President Reagan dealing with the Iran-Contra scandal which in turn may influence the collapse of *North* onto the Iran-Contra basis state, i.e., *Oliver North* who was a central figure in the scandal.

Summary and Outlook

This paper has entertained the speculation that quantum-like entanglement of words may be at play in human memory. Two experimental frameworks are put forward as a means of potentially testing for the existence of such entanglement. These experiments both rely on the CHSH inequality, a variant of Bell's theorem. Should strong evidence of the entanglement of words appear, it is hard to predict what the consequences of such a discovery might be. From a philosophical point of view, current reductionist models of human memory would be seriously undermined; words just cannot be considered as isolated entities in memory. Granted, at first sight, the connected topology of models underpinning spreading activation seem not to treat words as isolated entities, yet Nelson & McEvoy's recourse to "spooky-action-at-a-distance" models seems to suggest that spreading activation may fundamentally be a reductionist model. Moreover, the theory presented here may lead to a more fully fledged quantum like model of word association as advocated by (Nelson & McEvoy 2007). More practically, one wonders whether entanglement of words may be somehow leveraged, just as in quantum computing where entanglement is seen as a resource to be exploited. Some preliminary thoughts in this direction centre around whether entanglement of words may be leveraged for knowledge discovery (Widdows & Bruza 2007).

Acknowledgements

This project was supported in part by the Australian Research Council Discovery project DP073341.

References

Aerts, D.; Aerts, S.; Broekaert, J.; and Gabora, L. 2000. The Violation of Bell Inequalities in the Macroworld. *Foundations of Physics* 30:1387–1414.

Aerts, D.; Broekaert, J.; and Gabora, L. 2005. A Case for Applying an Abstracted Quantum Formalism to Cognition. In Bickhard, M., and Campbell, R., eds., *Mind in Interaction*. John Benjamins: Amsterdam.

Bruza, P., and Cole, R. 2005. Quantum Logic of Semantic Space: An Exploratory Investigation of Context Effects in Practical Reasoning . In Artemov, S.; Barringer, H.; d'Avila Garcez, A. S.; Lamb, L. C.; and Woods, J., eds., *We Will Show Them: Essays in Honour of Dov Gabbay*, volume 1. London: College Publications. 339–361.

Bruza, P.; Widdows, D.; and Woods, J. A. 2007. Quantum Logic of Down Below. In Engesser, K.; Gabbay, D.; and Lehmann, D., eds., *Handbook of Quantum Logic and Quantum Structures*, volume 2. Elsevier.

Clauser, J.; Horne, M.; Shimony, A.; and Holt, R. 1969. Proposed experiment to test local hidden-variable theories. *Physical Review Letters* 23:880–884.

Isham, C. J. 1995. *Lectures on Quantum Theory*. London: Imperial College Press.

Kitto, K. 2006. *Modelling and Generating Complex Emergent Behaviour*. Ph.D. Dissertation, School of Chemistry Physics and Earth Sciences, The Flinders University of South Australia.

Kitto, K. 2008a. High End Complexity. *International Journal of General Systems*. In Press, available online at http://www.informaworld.com/smpp/content?content=10.1080/03081070701524232.

Kitto, K. 2008b. Why quantum theory?

Lomonaco, S. J., ed. 2002. *Quantum Computation (Proceedings of Symposia in Applied Mathematics)*. American Mathematical Society.

Nelson, D., and McEvoy, C. 2007. Entangled associative structures and context. In Bruza, P.; Lawless, W.; Rijsbergen, C. v.; and Sofge, D., eds., *Proceedings of the AAAI Spring Symposium on Quantum Interaction*. AAAI Press.

Nelson, D.; McEvoy, C.; and Schreiber, T. 1998. The university of south florida, word association, rhyme and word fragment norms.

Schrödinger, E. 1935. Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften* 23:807.

von Neumann, J. 1955. *Mathematical Foundations of Quantum Mechanics*. Princeton: Princeton University Press.

Widdows, D., and Bruza, P. 2007. Quantum information dynamics and open world science. In Bruza, P.; Lawless, W.; Rijsbergen, C. v.; and Sofge, D., eds., *Quantum Interaction*, AAAI Spring Symposium Series. AAAI Press.

Widdows, D. 2004. *Geometry and Meaning*. CSLI Publications.