Meaning in context

Kirsty Kitto and Peter Bruza

Queensland University of Technology
Brisbane, Australia*
kirsty.kitto@qut.edu.au, p.bruza@qut.edu.au

Abstract. Language exhibits a number of contextuality and non-separability effects. This paper reviews a new set of models showing promise for capturing this complexity which are based upon a quantum-like approach.

1 Meaning and Context

How do humans understand language? Every day, we are confronted with novel words and combinations, often in completely new contexts, and yet we are usually able to extract meanings from these. People will often even agree with one another, interpreting the same novel scenarios in a similar manner.

Language is inherently contextual. Consider for example the word 'bat'. This word has at least two senses in its noun form; it might refer to a flying mammal that lives in caves, or alternatively it might refer to a sporting implement (and a variety of these are possible). Generally we can tell the sense that another speaker intends through a consideration of the context in which the word appears. Thus, if I were to claim that "the bat flew over the horizon" then it is highly unlikely that you would think I was talking about a sporting implement.

These different *senses* of a word can be explored via word association experiments. In free association, words are presented to large samples of participants who produce the first associated word to come to mind. The probability or strength of a pre-existing link between words is computed by dividing the production frequency of a response word by its sample size. For example, the University of South Florida free association norms (1) give a set of free association probabilities for a set of 5,019 cue words. Thus, used as a cue word, 'bat' produces 'ball' 25% of the time, 'cave' 13% of the time etc. (the total set of free association probabilities are shown in figure 1).

We can also find out which words are likely to produce the word 'bat' (now called a target). One way of achieving this involves a process known as *extra-list cuing*. Here, subjects typically study a list of to-be-recalled target words shown on a monitor for 3 seconds each (e.g. 'bat'). The study instructions ask them to read each word aloud when shown and to remember as many as possible, but participants are not told how they will be tested until the last word is shown. The test instructions indicate that new words, the test cues, will be shown and

* This work was supported in part by the Australian Research Council's Discovery Projects Scheme (numbers DP1094974 and DP0773341).

that each test cue (e.g., 'ball') is related to one of the target words just studied. These cues are not present during study (hence, the name extralist cuing). As each cue is shown, participants attempt to recall its associatively related word from the study list.

However, these associates of 'bat' are themselves capable of generating their own associations and these too can be probed experimentally. The full associative network for 'bat' is depicted in figure 1. Interestingly, word association

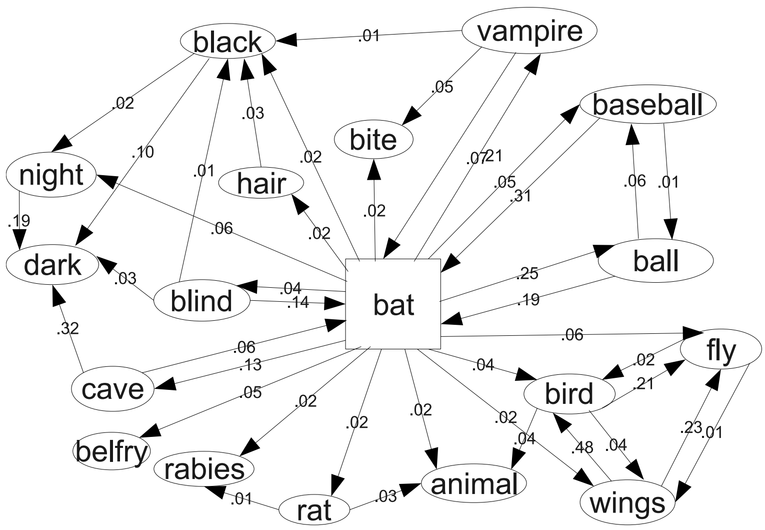


Fig. 1. The associative network of 'bat' (1). See http://web.usf.edu/FreeAssociation/ for the full data set.

experiments have shown that such semantic networks are small-world networks (2) as an average of only three links is needed to move from one word to any other in the system.

Attempts to map the associative lexicon of English soon made it clear that some words produce more associates than others. This feature is called *set size* and it indexes a word's associative dimensionality (3). Mapping the lexicon also revealed that the associates of some words are more interconnected than others. Some words have many such connections, whereas some have none, and this feature is called *connectivity* (4). Experiments have shown that link strengths between words, the set size and connectivity of individual words have powerful effects on recall which existing theories cannot satisfactorily explain. A seminal model is based on the idea that activation spreads through a fixed associative network, weakening with conceptual distance (e.g., (5)). This model only allows activation of a target if there are direct links between it and its associates, but there are cases where targets are activated via indirect links (6). Furthermore, the contextuality of recall is not represented in this model, and yet cuing a word differently changes the relative probabilities of recall. In what follows, we shall explore a series of models that work on this principle. Starting with a general model based upon the notion of *superposition* as it arises in Quantum Theory (QT), we shall create a model of the observed behaviour of associative networks (in section 2.1), and then develop more sophisticated models of how concepts combine (7; 8). These models treat context by representing it as cue words, or co-appearing words, and experiments are currently underway to test their validity (see sections 2.2 and 2.3).

2 Ambiguous Words and their Representation

The fundamental units of the mental lexicon show a kind of duality similar to a quantum particle. They can appear point-like or wave-like depending on how they are measured. For example, many words can take different senses depending upon the context in which they occur. When shown out of context, 'bat' reminds people of 'cave,' and 'vampire', but also of 'baseball', 'glove', and so on. Associates related to both meanings are likely to be activated when the word 'bat' is read in isolation (i.e. with minimal contextual cues), and such activation can be understood as wave-like, or distributed. Conversely, what is the the probability of recalling 'bat' when some context is present?

The recall (or not) of a word can be represented using a *superposition* state, such as the one appearing in figure 2(a).

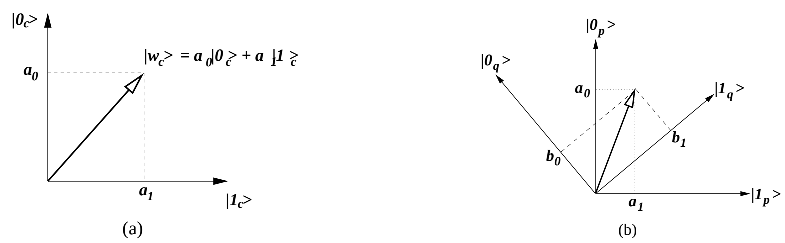


Fig. 2. A word w, for example bat, is represented in some context c which takes the form of a basis. (a) The word is recalled $|1\rangle$, or not, $|0\rangle$, in some context. Thus, if the context is the extra list cue 'cave', then the subject might recall 'bat' from a prior target list with a probability a_1^2, or they might fail to recall bat, with the probability a_0^2. Here, as in all quantum superpositions, $a_0^2 + a_1^2 = 1$. (b) Changing the cue to 'ball' might significantly change the chances of recall.

Here, we have the word w, represented in some context c, as a superposition of recalled, $|1\rangle$ and not recalled $|0\rangle$. Thus, the word 'bat' might be a target word, expected to be recalled in an extra-list cueing experiment upon presentation of the cue word 'cave' which in this case acts as the context c. The probability of 'bat' being recalled in this context is represented by a_1^2, as per the measurement postulate of quantum theory (9), but can be easily related to the Pythagorean theorem for the above diagram (which explains its origins). Thus, with reference to the University of South Florida word association data (1), we could represent 'bat' as the superposition: $\sqrt{0.94}|0\rangle + \sqrt{0.06}|1\rangle$ (see figure 1 for this data).

This model is made more interesting in figure 2(b), where we have represented the fact that a different context might result in a different set of recall probabilities. Thus, when given the cue word 'ball' we could represent 'bat' as the new superposition $\sqrt{0.81}|0\rangle + \sqrt{0.19}|1\rangle$. In this case we see that 'bat' is more likely to be retrieved from memory when a subject is presented with the cue 'ball' than the cue word 'cave'. A finding suggesting that this formalism provides a very natural representation of contextual effects as they actually occur in language.

Experimental findings show that when context is absent an extralist cue recalls all senses relating to a cue word (10). Indeed, a cue that activates one meaning of a polysemous word (i.e. 'bat') can often retrieve both its animal and sport senses in the extralist task. Such wave-like behavior, however, changes dramatically when the target appears in a context. For example, given 'Baseball Bat' as a cue, a subject is likely to recall only meanings related to the compound's sport senses, and the same principle applies equally well to all words in the human mental lexicon, e.g., 'Guitar Piano', 'Music Piano', and 'Keyboard Piano' generate somewhat different senses of 'Piano'. Thus, context can modify the meaning of a target word even as it is recalled. It can serve to localise the meaning of a word. In QT, such behaviour is termed wave-particle duality. A particle changes its behaviour from a wave-like to point-like when measured, and analogously, in semantic processing, a word changes from an extended to a local meaning state when encoded in a context, and this results in different outcomes when a subject must recall a word.

So far, we have represented context in our model through reference to the possible multiple senses of a polysemous word. Here, the context of a target word can be understood as the cue word provided in an extra-list cueing experiment, but it is clear that this idea only scratches the surface of contextuality in semantics. For example, we are not as yet even considering the interactions between multiple words.

Our current example can be made more sophisticated with reference to this interaction. We might choose to combine our representation of 'bat' in the context of the cue word 'cave' with another word, representing a more detailed context. What if our subject was, for a reason to become apparent shortly, concurrently thinking of a 'boxer'? Both 'boxer' and 'bat' have animal senses, and sporting senses, and can thus be sensibly represented as recalled, or not, with respect to these contexts. So, if we continue with the context that is suggested by the animal sense of bat, then at least four possibilities arise when a subject is asked to consider the combined system 'boxer bat'. Firstly, a subject might take a 'boxer' to be a dog, hence recalling the animal sense of 'boxer', and a 'bat' to be an animal. This could be represented as $|11\rangle$. Similarly, a subject might not recall either of these words in the animal sense and this would be represented as $|00\rangle$. However, they might also recall one of the words in an animal sense and the other in the sporting sense, and we could represent these two possibilities as $|01\rangle$ and $|10\rangle$, depending upon which word was recalled in which sense. This list of all four possibilities could be represented as the following state, obtained through use of the tensor product:

$$|boxer\rangle \otimes |bat\rangle = (a_0|0\rangle + a_1|1\rangle) \otimes (b_0|0\rangle + b_1|1\rangle) \tag{1}$$
$$= a_0b_0|00\rangle + a_1b_0|10\rangle + a_0b_1|01\rangle + a_1b_1|11\rangle, \tag{2}$$

where $|a_0b_0|^2 + |a_1b_0|^2 + |a_0b_1|^2 + |a_1b_1|^2 = 1$. However, it may not always be the case that all possibilities are available (6). Perhaps a subject can be biased in such a way that the two words influence one another non-separably. We could represent this scenario in quantum theory through use of the notion of *entanglement*. Here, we would represent the combined state 'boxer bat' using the same formalism, but it might be the case that either both animal senses are recalled, or neither are recalled, and we would represent this as

$$\psi_t = x|00\rangle + y|11\rangle, \text{ where } x^2 + y^2 = 1. \tag{3}$$

This state cannot be described as the product of its component states. That is, it cannot be represented in the form of equation (1). It is a complex superposition state, consisting of two components. When this state is "measured" during testing, the superposition is said to "collapse" to one state or the other, and it would be expected to remain in this state if tested again after a short period.

What could such entangled states signify for the human mental lexicon? Essentially, they would account for an "all or nothing" recall (6), where, if one word is recalled then its entire associative network related to that word is also recalled, rather than the more common spreading activation models (4).

What evidence exists to suggest that words might behave in this manner?

2.1 Word Association Networks and Recall

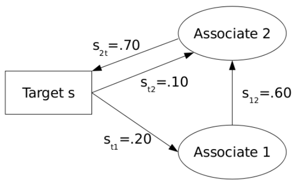
A comprehensive set of experiments have accumulated data over decades that suggests a number of key findings from word association experiments (4; 6). Here, we shall see that the basic model introduced above can be extended, leaving us with a model of word associations encompassing an entire association network for a word.

Figure 3 shows a hypothetical target having two target-to-associate links, it also contains a table listing the association probabilities depicted in this figure, and a set of superposition states that must be somehow combined in a model.

Making use of the "all or nothing" assumption discussed above, we shall choose to model this network as an entangled state,

$$\psi_t' = \sqrt{p_0}|000\rangle + \sqrt{p_1}|111\rangle. \tag{4}$$

This formula expresses a superposed state in which the entire associative structure is activated ($|111\rangle$) or not at all ($|000\rangle$). Choosing the values of the probabilities p_0 and p_1 is problematic, since there is no model of the time evolution



	t	a_1	a_2
t		0.2	0.1
a_1			0.6
a_2	0.7		
	$p_t = 0.7$	$p_{a_1} = 0.2$	$p_{a_2} = 0.7$

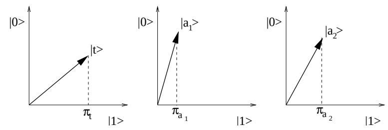


Fig. 3. A hypothetical target with two associates and single associate-to-target and associate-to-associate links. To the left, is a matrix corresponding to hypothetical association network. Free associations probabilities are obtained by finding the row of interest (the cue) and running across to the associate word obtained. The corresponding three bodied quantum system of words is underneath. The projection of the qubit onto the $|1\rangle$ basis relates to the probabilities in the table.

for semantic spaces, we are forced to speculate. However, again working with the "all or nothing" assumption, we could reasonably surmise that the lack of activation of the target is determined solely in terms of lack of recall of any of the associates. That is, $p_0 = \bar{p}_t \bar{p}_{a_1} \bar{p}_{a_2}$. Consequently, the remaining probability mass contributes to the activation of the associative structure as a whole.

$$p_1 = 1 - \bar{p}_t \bar{p}_{a_1} \bar{p}_{a_2} \tag{5}$$
$$= 1 - (1 - p_t)(1 - p_{a_1})(1 - p_{a_2}). \tag{6}$$

In a more detailed paper (6), we discuss the relationship between this simple prediction and the more standard recall models (such as the Spreading Activation model) and more recent work (in preparation) analyses the relative merits of the different models. The above model can be readily extended to more complicated networks, such as that depicted in figure 1, but there is every chance that the "all or nothing" assumption is somewhat simplistic. Could a more accurate model be obtained by considering activation across one specific sense? This is an area for future investigation.

We shall now move to a discussion of two experiments that are being performed with the intention of further investigating the effects of context upon recall. Is it possible to somehow prove that word senses and meaning should be considered non-separable in a way that can be well defined?

2.2 Shifting the Sense of a Bi-ambiguous Word

The first experiment (reported in more detail in (11)) looks at the manner in which the interpretation attributed to a novel word pairing can be influenced by the context in which the combination appears. Each word in this experiment was chosen carefully for its bi-ambiguous nature; they have two regular senses, one dominant and one subordinate in that the dominant one is more likely to be recalled (1). Participants completed a web-based task in which they provided an interpretation for twelve bi-ambiguous compounds, e.g., ‘boxer bat’. (See table 1), and context was modified through a careful choice of cue words.

The compounds were only seen once by each participant. Participants were assigned to one of ten groups based on the order in which they logged into the experiment. For eight groups the compound was preceded by a priming word (e.g., ‘vampire’) and for two groups (baseline) the compound was preceded by a neutral prime (e.g., 36). For the priming groups, participants classified the priming word as “natural” or “non-natural”. The goal of the classification task was to activate the prime in memory. The baseline groups decided whether a number was odd or even. This classification task was neutral and not expected to interfere with the interpretation of senses, although it was chosen in such a manner that it would balance the amount of cognitive processing across all subjects. The priming groups received one of four potential primes for each compound: Prime 1: Word 1 dominant sense (e.g., ‘fighter’), Prime 2: Word 1 subordinate sense (e.g., ‘dog’), Prime 3: Word 2 dominant sense (e.g., ‘vampire’), Prime 4: Word 2 subordinate sense (e.g., ‘ball’). The purpose of the primes was to influence the sense attributed to either word one or word two. Participants received different primes according to their group, with the constraint that over the twelve compounds they all received three each of the four potential priming words. After interpreting the compound participants were asked to clarify which sense they chose for each word. Thus subjects could claim that they had interpreted a ‘bat’ as one of: (A) An animal (B) A piece of sporting equipment, or (C) Other (which they were asked to specify).

	Word 1		Word 2	
Compound	Prime 1 (dom)	Prime 2 (sub)	Prime 3 (dom)	Prime 4 (sub)
boxer bat	fighter	dog	ball	vampire
bank log	money	river	cabin	journal
star charge	moon	movie	account	volt
apple suit	banana	computer	vest	slander
stock tick	shares	cow	flea	mark
fan post	ceiling	football	lamp	web
ring pen	diamond	oval	ink	pig
seal pack	walrus	envelop	suitcase	leader
spring plant	summer	coil	seed	factory
racket pitch	tennis	noise	tone	throw
toast gag	jam	speech	choke	joke
poker strike	cards	fire	lightning	union

Table 1. The compounds that were used, along with their dominant and subdominant prime cues.

The data obtained from this experiment has been analysed through use of what amounts to a probabilistic hidden variables model (11), although more data is required for a definitive result. Thus, the combination of two words is modelled by two random variables A and B, where A corresponds to the first word in the combination and B corresponds to the second word.

The variable A ranges over $\{a_1, a_2\}$ corresponding to its two underlying senses, whereby a_1 is used to refer to the dominant sense of first word in the combination and a_2 refers to its subordinate sense. Similarly B ranges over $\{b_1, b_2\}$.

Primes are designed to span four mutually exclusive cases. By way of illustration, the primes used for "boxer bat" are {fighter, dog, vampire, ball}. The primes are modelled as a random variable λ ranging over $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$. In a probabilistic setting, separability is formalized by assuming the joint probability is factorizable:

$$\Pr(A, B|\lambda) = \Pr(A|\lambda)\Pr(B|\lambda) \tag{7}$$

Using Bayes' rule, this can be rewritten as:

$$\Pr(A, B, \lambda) = \Pr(A|\lambda)\Pr(B|\lambda)\Pr(\lambda) \tag{8}$$

Assuming the law of total probability:

$$\Pr(A, B) = \sum_{1 \leq i \leq 4} \Pr(A|\lambda_i)\Pr(B|\lambda_i)\Pr(\lambda_i). \tag{9}$$

This final equation opens the door to test the separability assumption. From the baseline group, the joint probability function is given empirically: $\Pr(A, B) = \{\Pr(a_1, b_1), \Pr(a_1, b_2), \Pr(a_2, b_1), \Pr(a_1, b_2)\}$, where $\Pr(a_i, b_j)$ is shorthand notation for $\Pr(A = a_i, B = b_j)$.

Each priming condition expresses a distribution $\Pr(A, B|\lambda_i), 1 \leq i \leq 4$. For "boxer bat" the four distributions were empirically determined (see table 2):

	b_1	b_2
a_1	0.2	0.8
a_2	0	0

boxer bat:λ_1 = fighter, $n = 10$

	b_1	b_2
a_1	0.71	0
a_2	0.29	0

boxer bat:λ_3 = vampire, $n = 7$

	b_1	b_2
a_1	0.2	0.8
a_2	0	0

boxer bat:λ_2 = dog, $n = 11$

	b_1	b_2
a_1	0.5	0.5
a_2	0	0

boxer bat:λ_4 = ball, $n = 6$

Table 2. The data obtained from a pilot study investigating the separability of word compounds. Each table lists the probability of obtaining the respective sense of the compound 'boxer bat' for each of the different cue words listed (λ_i), n is the sample size for each set of data reported.

These data allow the joint probability distribution $\Pr(A, B)$ to be computed assuming only separability (represented by equation (9)) and a assuming uniform prior probability of the primes. The baseline joint distribution is depicted in table 3(a), while the corresponding joint distribution calculated from this data is alongside in table 3(b). A chi-square goodness-of-fit test at the 95%

	b_1	b_2
a_1	0.3	0.3
a_2	0.1	0.3

(a)

	b_1	b_2
a_1	0.42	0.41
a_2	0.16	0

(b)

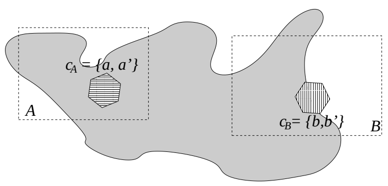
Table 3. A comparison between (a) the base joint distribution of 'boxer bat' for a sample size of ($n = 10$), and (b) that derived through the application of equation (9) to the data recorded in table 2.

confidence level shows that these two distributions are significantly different $\mathbf{p} = \mathbf{0.00014}$, ($n = 10, n_p = 34$). Thus, there is reason to believe that this analysis shows some promise in providing a test of non-separability in semantic spaces. However, more data, and some more stringent testing is required to fully establish these results and this work is currently underway.

This analysis was been motivated by similar considerations as they arose in the construction of the Bell-type inequalities of QT (9), however, it does not directly correlate with that analysis. A more direct experiment suggests itself, based more fully upon a direct experimental test of the separability assumption used in equation (9), the Clauser–Horne–Shimony–Holt (CHSH) inequality. This is discussed in the next section.

2.3 A Direct Non-separability Test

The CHSH inequality of quantum theory (9), provides an experimental test for distinguishing between local hidden variables theories and entangled quantum systems. For a system that is in some sense spatially distant, it makes sense to assume that actions performed upon one region will not effect the results that are obtained in the other region. A general scenario representing this case is illustrated in figure 4. Here, a system consisting of two components is frequently



1. (a, b)=(fighter,ball) $\sim$ $(sport, sport)$ senses
2. (a, b')=(fighter,vampire) $\sim$ $(sport, animal)$ senses
3. (a', b)=(dog,ball) $\sim$ $(animal, sport)$ senses
4. (a', b')=(dog,vampire) $\sim$ $(animal, animal)$ senses

Fig. 4. A general scenario illustrating the separability of a system consisting of two distant components. Alongside are a set of experimental scenarios that can be applied to word combination experiments testing the CHSH inequality in semantic spaces for the novel word pair 'boxer bat'.

considered separable due to the fact that the components are in some sense distant. Thus, making a choice in region A to measure some characteristic c_A of the system is deemed not to effect the results that will be obtained when the characteristic c_B is measured in region B.

In this experiment we calculate expectation values for the four available combinations of two different experimental scenarios, a, a', b, b':

$$E(i,j) = \frac{N_{11} + N_{00} - N_{10} - N_{01}}{N_{11} + N_{00} + N_{10} + N_{01}} \text{ where } i \in \{a, a'\}, j \in \{b, b'\}. \quad (10)$$

If the two different sides of this experiment can be considered separately, then the expectation values for this experimental scenario will satisfy the CHSH inequality:

$$-2 \leq E(a,b) - E(a,b') + E(a',b) + E(a',b') \leq 2 \quad (11)$$

which provides us with a numerical test for the separability (or not) of a quantum system. If the system can be considered separable then the CHSH inequality will be satisfied. Applying this inequality to semantic structures is non-trivial, two cues must be used, 'entangling' the concept combinations and 'polarising' the sense in which they will be interpreted. Essential to this experimental protocol was a need to specify compounds which had 'overlapping' polariser settings, as this makes it possible to explore the concept of coincidence between word pairs. So, returning to the example of the compound 'boxer bat', the two senses of these two words overlap, in that boxer has a sport sense and an animal sense, as does bat. This makes it possible to define the polariser settings (primes) listed above to generate the four necessary experiments.

The results of this experiment are reported in detail in (12).

3 Context and Non-separability in the Human Mental Lexicon

The forms of contextuality and non-separability presented in this paper have been applied to modelling words in human semantic space. We have primarily modelled words with two senses, concentrating in particular upon the manner in which the dominant and subordinate senses can be considered in a vector space approach (13; 14). This approach becomes less straightforward if our words have more than two senses, as the identification of which sense is being recalled (or not) is made more difficult. However, it is possible to represent higher-dimensional quantum systems using a tensor combination of q-bits, each representing in this case the recall (or not) of one particular word sense so extension of this simple model is feasible.

We have seen that QT also lends itself to a very natural model of the interactions between word senses, and we have seen this done both for separable senses (via the tensor product) and for non-separable senses (via entanglement). Further research is required. Without a proper model for the time evolution of the states of words in human semantic space, it will be very difficult to justify the choice of one or the other model. Indeed, in section 2.1 we made a somewhat arbitrary decision about the cognitive state of the subject as recalling "all or nothing" of a given associative network. While this decision is empirically justified, a more complete model would provide constructive arguments for the choices made, and work is progressing on this.

However, it is important to recognise that context is created by more than just semantics. Historical contingencies, social interactions and conditioning all play an important role in the choices that we make when interpreting the language we see around us every day. For example, students performing free association tests in South Florida have a high likelihood of choosing the fighter sense of *boxer*[1], but Australians appear to be following a trend of preferentially listing the dog sense when confronted with the same word.[2] It is important to understand these nuances. Diplomacy, trade, social exclusion etc. are all affected by these different interpretations of the same sentences. However, there is very little data available at present which can highlight these differences. We think it likely that this lack of data is due to the lack of models capable of dealing with context. With a new context-rich model of semantics, we see a way in which to progress in our models of complex social interactions, and this will be an area of future investigation.

We shall conclude with some thoughts regarding the generality of this model, and its future applicability.

QT provides a very natural model of most of the key effects displayed by human subjects in experiments surrounding word association and recall. However, this finding raises a particularly interesting question; why is it so? This question is made more compelling when we realise that QT has been used in the description of a wide range of systems not normally considered 'quantum' (see e.g. (15; 13; 14; 6; 16; 17)) which itself raises a new question; is there a unifying feature to these systems that makes QT a good descriptor of them?

One possible answer to this question lies in complex systems science. Every one of the systems for which a quantum-like theory has been posited are indisputably complex, indeed, they are likely to lie on the high end of a scale of complexity (18). This has led to the suggestion that QT is not just a theory of the microscopic, rather that it is a theory of contextually dependent complex systems. If this is indeed the case then a new possibility for interpreting QT presents itself, and the domain of application for QT becomes specifiable. The tests for non-separable behaviour that are presented in this paper appear to be quite generalisable, and it is likely that they will prove very useful for the future modelling of non-separable complex systems.

Thus, starting with words and meaning, we have finished with a new potential avenue for the understanding QT itself. It is clear that simply changing the frame of complex systems science, and asking questions about the most complex systems we can find, can lead to new important ideas and genuine progress.

---

[1] See the data at http://web.usf.edu/FreeAssociation/AppendixC/Matrices.A-B .

[2] Although not enough data has been gathered at this stage to be sure.

Bibliography

[1] Nelson, D., McEvoy, C., Schreiber, T.: The University of South Florida, word association, rhyme and word fragment norms. Behavior Research Methods, Instruments & Computers **36** (2004) 408–420

[2] Steyvers, M., Tennenbaum, J.: The large scale structure of semantic networks: statistical analyses and a model of semantic growth. Cognitive Science **21** (2005) 41–78

[3] Nelson, D.L., Schreiber, T.A., McEvoy, C.L.: Processing implicit and explicit representations. Psychological Review **99**(2) (1992) p322 – 348

[4] Nelson, D., McEvoy, C., Pointer, L.: Spreading activation or spooky action at a distance? Journal of Experimental Psychology: Learning, Memory and Cognition **29**(1) (2003) 42–52

[5] Collins, A.M., Loftus, E.F.: A spreading-activation theory of semantic processing. Psychological Review **82**(6) (1975) p407 – 428

[6] Bruza, P., Kitto, K., Nelson, D., McEvoy, C.: Is there something quantum-like about the human mental lexicon? Journal of Mathematical Psychology **53** (2009) 362–377.

[7] Wisniewski, E.J.: When concepts combine. Psychonomic Bulletin & Review **4**(2) (1997) 167–183

[8] Wilkenfeld, M.J., Ward, T.B.: Similarity and Emergence in Conceptual Combination. Journal of Memory and Language **45** (2001) 21–38

[9] Laloë, F.: Do we really understand quantum mechanics? Strange correlations, paradoxes, and theorems. American Journal of Physics **69**(6) (2001) 655–701

[10] Gee, N.R.: Implicit memory and word ambiguity. Journal of Memory and Language **36** (1997) 253–275

[11] Bruza, P.D., Kitto, K., Ramm, B., Sitbon, L., Blomberg, S., Song, D.: Quantum-like non-separability of concept combinations, emergent associates and abduction. Logic Journal of the IGPL (2010) In Press.

[12] Kitto, K., Ramm, B., Bruza, P.D., Sitbon, L.: Testing for the non-separability of bi-ambiguous words. Submitted to QI 2010. (2010)

[13] Widdows, D.: Geometry and Meaning. CSLI Publications (2004)

[14] Van Rijsbergen, C.: The Geometry of Information Retrieval. CUP (2004)

[15] Bruza, P., Sofge, D., Lawless, W., van Rijsbergen, K., Klusch, M., eds.: Proceedings of the third conference on Quantum Interaction. Volume 5494 of Lecture Notes in Artificial Intelligence., Springer (2009)

[16] Baaquie, B.E.: Quantum Finance: Path Integrals and Hamiltonians for Options and Interest Rates. Cambridge University Press (2004)

[17] Pothos, E.M., Busemeyer, J.R.: A quantum probability explanation for violations of 'rational' decision theory. Proceedings of the Royal Society B **276** (2009) 2171–2178

[18] Kitto, K.: High End Complexity. International Journal of General Systems **37**(6) (2008) 689–714