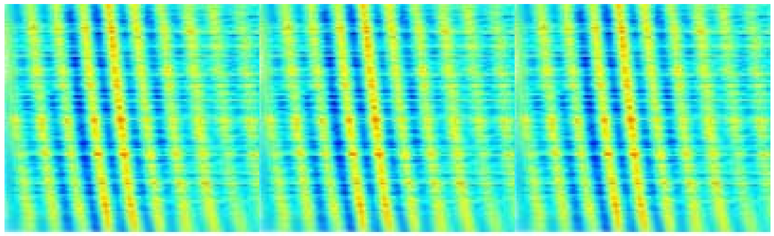

Cognitive Science Online

Vol 3, Issue 1, Winter 2005
http://cogsci-online.ucsd.edu

Letters

Letter from the editors: A new kind of brain drain
Arielle Borovsky i

Articles

Language origins without the semantic urge
Martin I. Sereno 1

Conceptual blending theory and psychiatry
Michael Kiang 13

Information

Cognitive Science Online is an online journal published by the UCSD Cognitive Science Department and seeks to provide a medium for the cognitive science community in which to exchange ideas, theories, information, advice and current research findings. This online publication is a peer-reviewed and highly interdisciplinary academic journal seeking contributions from all disciplines and methodologies investigating the mind, cognition and their manifestation in living, and possibly artificial, systems. For more information about this journal, submissions, back issues, please visit our website at http://cogsci-online.ucsd.edu

Contact Information

Department of Cognitive Science
University of California San Diego
9500 Gilman Drive
La Jolla, CA 92093-0515
cogsci-online@cogsci.ucsd.edu

Editors

Advisory Editorial Board

A new kind of brain drain

Wars in the course of history were largely territorial – to defend territories or destroy them. In today's age, the object of wars, as perceived by terrorist organizations, is to prevent science from changing archaic agendas. -- March 6, 2003 Jerusalem Post, Shimon Perez

Recently, there has been a rash of alarmed articles envisaging the end of the United States brain drain. A recent survey from the Council of Graduate Schools reveal that foreign student applications to American universities dropped by 28 percent this year from the previous year, and overall foreign student enrollments were down as well. At the same time, universities in other parts of the world are also enrolling more international students. Analysts have begun to voice concerns that America might begin losing its lead in research and have squarely laid much of the blame upon a number of post 9/11 political factors that discourage foreign visitors, including more stringent security measures and visa application procedures.

Even the UCSD Cognitive Science department seems to be experiencing its own brain drain -- but interestingly -- not directly related to security woes. In the fall of 2004, the department added seven new graduate students to its ranks, which is around its average enrollment for the past few years. Yet, each and every one of these students was an American citizen. This contrasts sharply with historic enrollment patterns where 33% of incoming students have been non-citizens. What happened? Did we have fewer foreign applicants? No – according to the head of this year's graduate admissions committee; although the overall number of applications did drop slightly, there was still a comparable ratio of foreign students in the mix. Instead, sheer economics shaped the 2004 class – budget cuts in California have reduced the amount of money available to the department, which at the same time is also under pressure to increase the number of graduate students enrolled to meet increased projections for undergraduate enrollment in the next few years. Under the UC system, non-citizen students are much more expensive to admit because they never qualify for in-state tuition, whereas out-of-state American graduate students qualify after their first year. Simply put, the CogSci department was faced with pressure to maintain previous levels of admission with less money – and chose students that its budget could afford.

However, while this year's graduate student enrollment figures might only be related to temporary economic pressures, there are other more troubling signs within the faculty ranks. This quarter, the department saw two of its faculty members make plans to move to one permanent and one temporary position at foreign institutions because their (foreign) partners faced difficulties staying and working in this country. In one case, a professor's partner was unable to practice medicine in this country with a European medical degree, and in another case, a highly-educated partner (with 2 PhDs) and a job offer outstanding was forced to return to her home country for at least two years to fulfill a Fulbright fellowship requirement.

On the face of it, these cases seem to have little to do with post 9/11 security changes – and more to do with fallout from budget cuts, medical licensing issues and fellowship obligations. Yet, even here, one might be tempted to become suspicious. Is it now more difficult to find jobs for bright foreign workers in this country because of our security policies? Is there less funding for Cognitive Science research and students because more research money is being spent upon defense and security related science?

Our field relies upon the work of talented minds from around the world, and upon the ability for Cognitive Scientists to collaborate and communicate openly and freely. However, it is unsettling that there is a potential for American students to become isolated from their foreign counterparts when their graduate education is 'cleansed' of interaction with non-citizens due to shortsighted budgetary and political decisions. The impact of prolonged nationalist isolation would clearly be detrimental to all of Cognitive Science.

With the Bush administration already hinting that there could be drastic declines in science spending in the 2006 budget, and our ballooning defense expenditures showing no sign of falling off, there seems little hope that interest in science is increasing or the administration's war on terror is winding down. At this point, we can only hope that our country might heed Perez's words of warning and invest again in research – with the hope of reaping rewards of economic prosperity from scientific progress, and international respect from increased scientific interaction.

Arielle Borovsky

Department of Cognitive, UCSD

Cognitive Science Online, Vol 3.1, pp.1-12, 2005 http://cogsci-online.ucsd.edu

Language Origins Without the Semantic Urge

Martin I. Sereno
Department of Cognitive Science
University of California San Diego
sereno@cogsci.ucsd.edu

Abstract

Despite the paucity of direct evidence about the origin of human language, the great intrinsic interest in this question has made it difficult for writers to resist speculating about it (Harnad et al., eds., 1976; Merlin, 1991; Deacon, 1997; Jablonski & Aiello, eds., 1998; King, ed., 1999; Knight et al., eds., 2000). The following attempts to bring a fresh perspective on this old question, using an analogy with the origin of cellular coding systems and applying it to what we know about the evolution of vocal behavior in animals. In other places (Sereno, 1991b), I have argued that DNA and protein based life and language based human thought may have enough in common as the only two naturally occurring examples of a code-using system to make it useful to take an analogical look at one system in order to make predictions about the other. Rather than rehearsing those arguments, I will only visit two jumping off points reached while developing that analogy: the difference between origin and evolution, and the foundational role of an intermediate string of "symbol representation" segments with properties partway between symbol and meaning.

Origins versus evolution

Discussions of how life came into existence (Wills & Bada, 2000) often distinguish *origin* of life from the *Darwinian evolution* of life. The core of every living cell is a system for converting genes into proteins; that is, DNA sequences into amino acid chains that spontaneously fold into the 3D molecular machinery of the cell including enzymes, receptors, force producing strands, and so on.

The problem of the *origin* of this system is not really an evolutionary problem in the usual Darwinian sense of the word. If we shrink ourselves down to molecular size and look at what cells are doing, it becomes clearer that cells have invented a new kind of molecular level intentionality as a way to partly overcome the deterministic thermodynamic buffetings to which all matter is subject that went far beyond the chemical dynamics of the landscape before there was life. This does not imply that cells create mysterious, irreducible holistic forces; in fact, we know quite a lot about how they work. But it is a natural way of characterizing what goes on in cells that distinguishes them from the prebiotic chemical cycles in clouds, rock piles, streams, beaches, and ocean floors. The prebiotic soup was already a complex, energy-dissipating system containing many different types of dynamically stable subunits. Cells, however, invented a way to encode, use, and reproduce information about how to cause thousands of different chemical reactions in this soup to happen. The tricky part is that the information, as well as all of the interpreting apparatus has to be in the soup where everything is still subject to the soup's deterministic buffetings. The cellular system speeds many chemical reactions, slows or prevents others, invents many new ones that never used to happen at all before, and above all, orders and organizes the reactions. In short, code-using cells have taken over forceful control of chemical phenomena in local regions of the otherwise still prebiotic soup. But the 'evolution' from prebiotic to biotic systems was not *Darwinian* evolution. Until the DNA code-using system was in place, Darwinian evolution as it is usually defined, "heritable variations in fitness" (Lewontin, 1970), was not possible. The central problem of the origin of the coding system in life is to try to imagine how such an intentional system could have arisen out of prebiotic situations lacking intentionality.

In thinking about the origin of language and cultural evolution, the situation is more complex since human language was built upon a pre-existing genetic system already capable of Darwinian evolution. There has been a recent revival of interest in 'evolutionary psychology' or human sociobiology, which attempts to come up with plausible scenarios for how biological evolution might have directly driven the origin of many human behaviors. One difficulty with this approach is the huge increase in the pace of behavioral change supported by the origin of language, which is what most distinguishes humans from other animals. Cultural evolution is so fast that it makes biological evolution look effectively stationary; this great difference in velocity makes it difficult for culturally transmitted memes to be fixed in much more slowly evolving genes.

It is certainly true that there is one great point of interaction between the DNA based genetic system and the language based human cultural system, which is the genetic basis of the peculiar human ability to readily learn a language. However, I think we may be able to make more progress by considering the *origin* of language as essentially a *pre*-evolutionary problem -- that is, as the second origin of a symbol using evolution supporting system, one that partly relies on DNA based symbols for its persistence, but that is largely decoupled from biological evolution in its content. In fact, human language might best be thought of as a brain operating system that allowed us to partly overcome the constraints on the biological evolution of behavior much in the way that cells have partly overcome the deterministic constraints on the 'evolution', now in the physicist's sense of 'the evolution of a dynamical system', of prebiotic soups.

The semantic urge and the 'RNA world'

I want to take issue with an assumption that lies behind almost every language origin scenario, something that could be called the 'semantic urge'. This widespread intuition grows out of the fact, just mentioned above, that the human linguistic coding system was built on top of a lower level biological coding system already capable of constructing sophisticated, nonlinguistic cognitive systems such as those in parrots and primates. The sustained goal directedness of animals makes it very hard to avoid the notion that human language must have somehow grown out of an insistent craving of inarticulate hominins to communicate complex meanings to each other ("hominins" because of a demotion of one level in the hierarchy from hominids, since chimpanzees are now thought to be more closely related to humans than to gorillas).

This is at heart a Baldwinian picture, where behavior provides a selective context that drives standard Darwinian evolution (Baldwin, 1902). As noted above, there is little tendency to fall back on intuitions like this in thinking about the origin of cellular life because cellular life had no analogous pre-existing code using system capable of goal directed behavior beneath it. Thus, one influential picture about the origin of life is that protosymbol chains emerged first without standing for anything and then only later were taken over as a code for other chains that could fold up and control chemical reactions; there is no Baldwinian urge of prebiotic soups to control their surroundings that drives the emergence of cellular symbol chains. The corresponding picture of protolanguage that we arrive at by analogy is somewhat peculiar, but it fits better with what we know about the evolution of vocal behavior in other animals. Let's first review some ideas about protosymbols at the cellular level.

The idea of an ‘RNA world’ (Gilbert, 1986) as a predecessor of modern DNA and protein based life was put forward soon after demonstrations in the early 1980's that RNA could act as a bona fide, enzyme-like catalyst (specifically, a spliced out segment of RNA in the single-celled organism *Tetrahymena* was discovered to fold up and catalyze RNA splicing). Three main arguments for the foundational role of RNA come from observing its current position in cells:

- RNA can act either as a 1D symbol string (mRNA) *or* a 3D controller of chemical reactions (structural RNA's, based on the protein-like ability of RNA to form precisely shaped surface cavities with high specificity for particular substrates)
- Modern cells mostly use proteins, not RNA’s, to control and catalyze reactions
- The instances where RNA *is* used as a protein-like structure stand at the very center of code use in cells – splicing of code-like RNA (nucleolus, spliceosome), recognizing words in code-like mRNA (tRNA), and assembling amino acids into proteins (rRNA, SRP RNA)

This idea that the dual roles of RNA as code and catalyst might have bootstrapped life have gained support in recent years as additional catalytic RNA's were discovered, as large scale RNA's (ribosomes) were finally crystallized showing that RNA itself, not proteins, catalyzed the attachment of each amino acid onto the growing protein chain (Ban et al., 2000; Nissen et al., 2000; Yusupov et al., 2001), and most recently, when a small RNA was created (in a laboratory selection experiment) that catalyzed the attachment of an amino acid onto the small RNA itself. This last observation was particularly evocative since RNA-amino acid bonds are made and broken during each chain lengthening step in modern cellular protein synthesis (Illangasekare et al., 1997; Zhang and Cech, 1997; Jenne and Famulok, 1998).

Despite the intuitive attractiveness of the ‘RNA world’, however, it has turned out to extremely difficult to find plausible prebiotic synthesis pathways for nucleotides, the subunits of RNA, which stands in sharp contrast to the easy prebiotic availability of amino acids (Miller and Orgel, 1974; Schwartz, 1998). This led many origin of life researchers to search for prebiotic precursors of RNA constructed from other more easily obtainable subunits (Joyce et al., 1987). A key feature of this search, so obvious to those within the field that it is rarely explicitly stated, is to find reasons *other* than the ability to code for proteins (or to catalyze chemical reactions in a protein-like way) as to why a pre-RNA-like molecule might have come into existence. Once RNA or something like it existed, its dual role as a catalyst and a code chain could then be discovered, as it were, leading perhaps to something like an RNA world, and then finally DNA/RNA/protein life.

In turning back to language, many language origins scenarios start with a repertoire of already meaningful vocalizations like those used by many different animal species (Hauser, 1995) and then attempt to come up with a reason, typically, the semantic urge – for why they might have multiplied. Several people (Zahavi, 1993; Knight, 2002) have pointed out that there is a major problem with this standard scenario. Animal calls, such as the well studied set of vervet monkey alarm calls, are laden with emotional meaning (Cheney and Seyfarth, 1990). The usual explanation for this is that alarm calls are emitted in life and death situations, which generates strong selection pressure to maintain call reliability, and that this reliability seems to have been ensured across many species by tightly tying calls to the emotional state of the sender and the receiver. This linkage, however, presents a problem for the scenario of calls as a proto-language; the number of different emotional states is rather small, and emotional states don't follow each other in a quick, regular succession like words do.

The origin of language required the development of a large inventory of thousands of words whose meanings are made more specific by assembling them into strings that are deployed in a rapid, regular sequence at a rate of several words per second. Individual words, especially the high frequency polysemous words that are central to every language such as “over”, “put”, “give”, “line”, “big”, or “hand”, are freed from emotion and bleached of emotion when compared to animal calls. Certainly, some single words such as epithets can be intrinsically emotive, but these are in a small minority. Perhaps the difficulty of imagining a path from a handful of emotive calls to the five thousand word core of emotionally neutral words in human language stems from the fact that the two are phylogenetically unrelated. The analogy with the RNA world and the pre-RNA world suggests that perhaps we should instead try to find a way by which a large pool of pre-RNA-like prewords might have been generated as units that are like words in some sensory and motor respects, but that *don't* yet stand for anything.

Birdsong and language preadaptations

The relation between birdsong and speech was noted early on. Darwin, who in *Origin of Species* (1859) often discussed the relation between biological and linguistic evolution (interestingly, to argue that biological evolution might be like language evolution, not *vice versa*!), turned briefly, in the *Descent of Man* (1871), to language origins. Darwin was especially fond of scenarios in which a structure had initially evolved for one purpose only to become a ‘preadaptation’ for another. Breaking with a common view that language arose from a gestural substrate, he suggested that language developed out of a form of “rudimentary song,” a kind of purely prosodic pre-language that conveyed emotions and other broad, unitary meanings in much the same way that pitch modulation and emphasis are used in modern speech. Darwin mentions the flashy hooting vocalizations of gibbons, which are generated during territorial and courtship displays, as something like what he had in mind, but pointed out that birdsong provided “in several respects the nearest analogy to language” (p. 55), citing the work of Daines Barrington (a century before) on the extended learning period for birdsong, the initial “babbling” stage, and the development of birdsong dialects. I think Darwin's idea should be revived but also revised in light of newer work on birdsong and primate vocalizations.

Modern research on birdsong has provided a neurobiological foundation for these earlier hunches, but has also revealed a system that looks a good deal more like human-style, 'left-hemisphere' speech than like the call systems of other animals, including New and Old World monkeys and gibbons (Konishi, 1985; Nelson and Marler, 1989; Doupe & Kuhl, 1999), but also the call systems of songbirds themselves, who have retained their limited set of emotional calls alongside song. There is a powerful perennial tendency outside fields explicitly focussed on evolutionary processes to think of evolution in terms of a "Great Chain of Being" and to ignore the mosaic nature of evolution. Thus, birdsong has often been dismissed as a model of human language for the reason that monkeys seem much smarter than some birds, or that monkey calls seem to have more semantic content than birdsong. In fact, the importance of birdsong in the present context (only dimly glimpsed by Darwin) is exactly the fact that a set of language-like features have evolved *in the absence of* a semantic function.

Birdsong requires a significant learning period, during the early parts of which the bird is silent. If a bird is not exposed to a tutor song within a certain early "critical period", it will produce only a crude version of its species' song. Normally exposed young birds initially produce sounds called "subsong" which resemble the progression of types of "babbling" in baby humans: Initially a broad range of sounds are produced, followed by an unorganized recombination of species specific song fragments, and then finally, adult song. Within a species, there are regional "dialects" which are learned from a bird's regional peers; artificial rearing experiments show that birds learn the dialect of their tutors, regardless of their genetic background. Adult song repertoires can be quite considerable; some wrens produce hundreds of distinct songs each containing 5-20 'syllables', while mockingbirds produce virtually endless sequences of different syllables in variable orders. Good singers may have a thousand or more distinct 'syllables' (a syllable consists of a particular figure sometimes repeated once or twice; in this respect, it is unlike a phonetic syllable which consists of one or more consonants and a vowel). If a songbird is deafened before learning to sing, it will fail to produce songlike sounds as an adult. By contrast, nonsong birds and many other animals including nonhuman primates (including gibbons) that do not learn complex serial vocal patterns from their peers, still come to produce their species specific sound repertoire when deafened at birth (Merker, 2000). In many respects, it might be more accurate to call it 'birdspeech', since birdsong differs from human singing and musical performance in many ways, birdsong lacks a regular meter, musical tonality, and harmony.

The parallel evolution of fine grained vocal control in singing birds affords a crucial comparative perspective on the anatomical and neural constraints on auditory-motor learning and performance. Birdsong is initiated in a structure called the syrinx, which is evolutionarily related to (and controlled by the same nerve as) the tongue, but which corresponds functionally to the human larynx. Human speech sounds are generated by filtering and modulating the higher harmonics of the fundamental frequency of the vocal cords (by controlling the position of the tongue in the pharyngeal and oral cavities), making the higher frequency parts of sounds independent of fundamental frequency (voice pitch). Birdsong, by contract, is generated primarily by directly controlling of the fundamental frequency produced by the syrinx. Nonetheless, in other respects, birdsong is much more like human speech than are the vocalizations of other animals, some of which can (like monkeys and male deer) modulate laryngeal harmonics (Owren et al., 1997; Fitch & Reby, 2001) in a human speech-like fashion.

There are intriguing clues about the evolution of fine vocal control from neuroanatomy of the song system. For example, motor output neurons in the forebrain (in nucleus RA) of songbirds have gained direct access to motoneurons controlling the syrinx vocalization musculature. Projections from RA also bypass the brainstem pattern generator circuitry for calls through which all forebrain outputs must pass in nonsong birds like ducks (Nottebohm et al., 1976; Arends & Dubbeldam, 1982) but also in squirrel monkeys (Ploog, 1981; Kirzinger & Jurgens, 1991) and macaque monkeys (Simonyan & Jurgens, 2003). There is a striking parallel here to the evolution of fine finger control in primates (but also finger control in raccoons, as a yet another reminder that evolution is a bush, not a linear Great Chain of Being), where motor cortex neurons have also come to contact finger motoneurons directly, bypassing pattern generators for coordinated limb movement in the spinal cord; hand motor cortex in cats, by contrast, contacts primarily the spinal pattern generators, which then have the only private access to motoneurons. The more direct access afforded the forebrain in the case of the songbird syrinx and the primate and raccoon hand presumably underlies more complex, differentiated, learned control observed in these systems. Note that this means that the relevant forebrain output areas have essentially come to assume a *lower* level in the motor control hierarchy, allowing the development of higher forebrain pattern generating centers that can operate somewhat independently of the brainstem and spinal pattern generating circuitry that is still needed for locomotion and nonsong vocalization.

Speech-like birdsong carries *less* meaning than vocal call systems do

The most striking characteristic of birdsong, however, in light of its prodigious complexity, is its essential lack of semantic content. Individual syllables or song fragments do not seem to have any specific meaning outside of being part of a particular song; and particular songs do not seem to convey specific content. Nor do birds appear to produce anything like ‘words’ by recombining their ‘syllables’ in order to signify concepts. Despite having vocal and auditory equipment ideally suited to support the recombinable speech symbol half of a language-like meaning conveying system, birdsong seems to communicate only very general meanings. Songs serve to mark territories, identify the singer's species, attract mates, often all at once. The messages communicated by birdsong are, in fact, *less* content filled than the messages communicated by, say, vervet monkey calls, which have been shown to signify rather elaborate distinctions among predators and conspecifics (Cheney & Seyfarth, 1990), despite the comparative simplicity of those unlearned monkey calls. This difference in referential content is particularly obvious when we consider the ‘meaning’ of a single syllable of a songbird's song; though emotion is keenly involved in motivating the bird to begin singing, the identity and order of syllables carry no specific emotional baggage.

One plausible theory about birdsong is that it was a product of runaway sexual selection, like the male peacock tail or outsize antlers in males, or huge sexual swellings in female baboons and chimpanzees. Elaborate singing abilities seem to have been preferred by mates, despite making little contribution to fitness beyond the fact that they were preferred. Sexual selection stands in contrast to natural selection, which rewards improved function like a stronger beak or more efficient wings. Certainly, a complex song could initially have been a sign of a mate fitter in other nonsong respects. But it is harder to explain the maintenance of extreme examples this way, especially when sexually selected features impede other functions (huge antlers) or attract predators (elaborate vocal displays). Sexual selection is not confined to female choice affecting male characters. And in some songbird species, both the male and female sing. Bay wren male-female monogamous pairs, for example, execute precisely coordinated unison “duets”, and the song control nuclei are large and hormone sensitive in females as well as males (Brenowitz & Arnold, 1985). The generally accepted explanation for this behavior is that the attractiveness of the male's song to listening females is reduced when a duetting female is singing along.

Several whale species have evolved a vocal learning system that resembles birdsong in a large number of respects, and provides a key additional example of how a speech-like vocal learning system can evolve without a ‘semantic urge’ (Tyack & Sayigh, 1997). Humpback whales learn to precisely reproduce long sequences of sounds and culturally transmit them to animals that are genetically unrelated. The main difference is that whale songs are lower in pitch, and individual songs unfold over several minutes instead of several seconds. The underwater acoustic environment of the ocean is quite reverberant, due to the faster and more efficient propagation of sound as well as the air-water boundary, which may be one reason for whale's more leisurely tempi. As with birdsong, whale song has social and sexual functions.

A birdsong-like ‘RNA world’ for pre-language

With the context provided above, we can see our way to a surprising extension of Darwin's language origins theory. On the evidence of the avian case, it seems possible that early hominids might have initially evolved an elaborate system of essentially phonetic vocalizations -- a kind of “talking song” with no component semantics -- as a result of sexual selection. In this view, a number of the specializations for auditory-vocal control evolved for entirely nonsemantic reasons. Perhaps early hominid pairs initially duetted like bay wrens, innocent of reference for a million years. Turning standard language origins scenarios on their heads, the preadapted ‘symbols’-without-meaning system might have only later been taken over for use as a semantic vehicle, that is, speech-sound ‘song’ before referential speech. An odd scenario, but then birdsong is quite odd itself.

This scenario contrasts with Fitch's idea that laryngeal descent in hominids (which occurs early in development) might have occurred as strategy to indicate large size in males, by analogy with the realtime laryngeal descent that occurs during calls made by rutting male deer (Fitch & Reby, 2001). A functionally similar sort of call, though using air sacs instead of laryngeal descent, is already well known in gibbons and orangs, and it closely resembles emotional meaning laden signals in standard animal call systems; and unlike birdsong, these primate calls develop even in deafened animals, indicating that learning is not required. The birdsong model suggests instead that there was runaway selection for complex *sequences* of essentially meaningless segments -- each untied from particular emotions -- as opposed to selection for a large sounding roar; it's the elaborateness of the sequences that the mates found attractive, not their throatiness.

As mentioned above, RNA serves both as a code (mRNA), but also as a noncode-like, folded word recognition device (tRNA) and chain assembly device (rRNA). By analogy, the internal representations of speech sound sequences that a primate neurobiologist would expect to find in the human lateral temporal cortex may have some other function besides merely serving as internal copies of the speech stream;

these uninterpreted speech sound representations could also be involved in word recognition and assembly of primarily visual meaning units into coherent discourse structures. By this account, what distinguishes humans is the ability to use a sequence of symbol patterns from another modality to cause the assembly of meaning patterns in higher visual cortex. But the product of that assembly may be very similar to patterns assembled from direct visual inputs arriving from earlier visual areas during scene comprehension, which also involves rapid serial assembly (of successive glances). The implication is that the trick of language was not to have invented the basic meaningful units but to have found a way of making standardized connections between them (see Sereno, 1984; 1991b).

One longstanding problem in the origin of human language is its sudden appearance. Most commentators agree that modern-style human language is less than 100,000 years old and probably less than 50,000 years old, based on the appearance in the cultural record of modern appearing artifacts together with *Homo sapiens*. Set against this is the much longer time that it must have taken for the anatomical and neural structures that control human language production to have evolved from their primitive condition in all other anthropoid primates. The birdsong/RNA world picture presented above provides one way out of this problem. Perhaps there was a long period of a million years or more in which the neural and anatomical basis of language production evolved for essentially *nonsemantic* reasons in early *Homo* or even *Australopithecus* species. This would have set the stage for the emergence of a linguistic 'RNA world', where the word recognition and chain assembly properties of meaningless speech-sound representations could be discovered, and then eventually grafted on to a productive meaning construction system that began to use visual representations to do most of the work (Sereno, 1991a), the analogue of the mostly protein based world at the cellular level.

Homo floresiensis and the ebu gogo

After this paper was finished, a new dwarf hominin species, *Homo floresiensis*, was discovered on the island of Flores in Liang Bua cave in eastern Indonesia (Brown et al., 2004). Dating by radiocarbon, luminescence, uranium series, and electron spin resonance indicates that the species existed from before 38,000 years ago until at least 18,000 years ago. It was associated with a large number of small stone tools (Morwood et al., 2004). Dwarfing and giantism is a common evolutionary response in species that are confined to islands, and this dwarf species of *Homo* coexisted on Flores island with dwarf *Stegodon* elephants (elephants are good long distance swimmers), giant rats, giant tortoises, Komodo dragons (Komodo is a small island just west of the much larger Flores island), and an even larger extinct varanid lizard. *Homo floresiensis* is thought to be most closely related to *Homo erectus* on the basis of the dentition (lightly built jaw containing small canines and small premolars and molars), the skeleton (indicating it was an obligate biped, though with somewhat long arms and fingers), and the extremely small brain case (slightly smaller than a chimpanzee). Stone tools on the island previously attributed to *Homo erectus* date back to 800,000 years ago. There is archaelogical evidence that anatomically modern *Homo sapiens* had already arrived in the area (East Timor, Australia) by 40,000 years ago, but no *Homo sapiens* fossils have yet been found on Flores. No specimens of the dwarfed *Stegodon* were found above a 12,000 year old tuffaceous deposit that resulted from a large eruption of the Flores volcano. Given the small number of hominin fossils found, however, it is less clear that they went extinct, too.

Austronesian speaking immigrants arrived on the island about 2,000 years ago to find indigenous Melanesians (the current inhabitants speak Austronesian languages). There was contact with India and China; and then the Portugese first arrived in 1520, and a century later, the Dutch.

A decade before the *Homo floresiensis* find, Gerd van den Bergh, a paleontologist working on the faunal remains who speaks Indonesian, had heard stories from villagers living in several different towns near the foot of the volcano about a race of hairy, three foot tall people, the “ebu gogo” (literally, ‘the grandmother who eats anything’). The ebu gogo were long-haired, potbellied cave dwellers with protruding ears, and long arms and fingers, and they walked with a slightly awkward gait and would climb small trees. The villagers said that the last ebu gogo was seen in the 19th century, when the Dutch settled in central Flores (Roberts, 2004). Although the folklore of many groups around the world mention small people (leprechauns in Ireland, menehune on Hawai'i), the ebu gogo stories are unique among them in matching several specific physical aspects of local subfossil remains.

In the context of the present paper, the most poignant aspect of these stories concerns the putative vocal abilities of the ebu gogo (and by implication, of *Homo erectus*!) that were observed by the villagers as they tolerated the ebu gogo raiding their crops, and during closer encounters when the villagers provisioned them with grains, vegetables, fruits, and meat, all of which the ebu gogo ate raw. The ebu gogo “murmured at each other and could repeat words verbatim” in a parrot-like fashion; for example, “in response to ‘here's some food’ [in Indonesian], they would respond ‘here's some food’” (Roberts, 2004). Although this evidence is incomplete and indirect, there is an uncanny fit to the scenario introduced above in which modern *Homo sapiens* style language emerges ‘at the last minute’ from an initial set of auditory and motor system modifications of much greater antiquity that had originally evolved to support *nonsemantic*, birdsong-like vocalizations. The remote but exciting possibility that the ebu gogo still exist might someday make it possible to test these ideas directly.

References

Arends, J.J. & Dubbeldam, J.L. (1982). Exteroreceptive and proprioceptive afferents of the trigeminal and facial motor nuclei in the mallard (Anas platyrhynchos L.). *Journal of Comparative Neurology, 209*, 313-329.

Baldwin, J.M. (1902). *Development and Evolution*. New York: Macmillan.

Ban, N., Nissen, P., Hansen, J., Moore, P.B. & Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science, 289*, 905-920.

Brenowitz, E.A. & Arnold, A.P. (1985). Lack of sexual dimorphism in steroid accumulation in vocal control brain regions of duetting song birds. *Brain Research 344,* 172-175.

Brown, P., Sutikna, T., Morwood, M.J., Soejono, R.P., Jatmiko, Saptomo, I.E. & Due, R.A (2004). A new small-bodied hominin from the Late Pleistocene of Flores, Indonesia. *Nature, 431*, 1055-1061.

Cheney, D.L. & Seyfarth, R.M. (1990). *How Monkeys See the World*. Chicago:

University of Chicago Press.

Darwin, C. (1859). *On the Origin of Species.* Cambridge, MA: Harvard University Press.

Darwin, C. (1971). *The Descent of Man, and Selection in Relation to Sex* (2nd ed. Rev). New York: Appleton.

Doupe, A.J. & Kuhl, P.K. (1999) Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience, 22*, 567-631.

Deacon, T.W. (1997). *The Symbolic Species*. New York: Norton.

Fitch, W.T. & Reby, D. (2001). The descended larynx is not uniquely human. *Proceedings Royal Society of London B, 268*, 1669-1675

Gilbert, W. (1986). The RNA world. *Nature, 319*, 618.

Harnad, S.R., Steklis, H.D. & Lancaster, J. eds. (1976) *Origins and Evolution of Language and Speech*. Annals of the New York Academy of Sciences, v. 280: New York Academy of Sciences.

Hauser, M. (1996). *The Evolution of Communication*. Cambridge, MA: MIT Press.

Illangasekare, M., Kovalchuke, O. & Yarus, M. (1997). Essential structure of a self-aminoacylating RNA. *Journal of Molecular Biology 274*, 519-529.

Jablonski, N.G., & Aiello, L.C. eds. (1998). *The Origin and Diversification of Language.* San Francisco: University of California Press.

Joyce, G.F., Schwartz, A.W., Miller, S.L. & Orgel, L.E. (1987). The case for an ancestral genetic system involving simple analogues of the nucleotides. *Proceedings of the National Academy of Sciences 84*, 4398-4402.

Jenne, A. & Famulok, M. (1998). A novel ribozyme with ester transferase activity. *Chemical Biology, 5*, 23-24.

King, B.J., ed. (1999). T*he Origins of Language: What Nonhuman Primates Can Tell Us*. Santa Fe, N.M.: School of American Research Press.

Kirzinger, A. & Jurgens, U. (1991). Vocalization-correlated single-unit activity in the brain stem of the squirrel monkey. *Experimental Brain Research, 84*, 545-560.

Knight, C., Studdert-Kennedy, M. & Hurford, J.R. eds. (2000). *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form.* New York: Cambridge University Press.

Knight, C. (2002). Language and revolutionary consciousness. In A. Wray (ed.), *The Transition to Language* (pp. 138-160). Oxford: Oxford University Press.

Konishi, M. (1985). Birdsong: From behavior to neuron. *Annual Review of Neuroscience, 8*, 125-170.

Lewontin, R.C. (1970). The units of selection. *Annual Review of Ecology and Systematics, 1*, 1-18.

Merker, B. (2000). Gibbon songs and human music from an evolutionary perspective, In N.L. Wallin, B. Merker, and S. Brown, eds., *The origins of music* (pp. 103-124). MIT Press.

Merlin, D. (1991). *Origins of the Modern Mind*. Harvard University Press.

Miller, S.L. & Orgel, L.E. (1974). *The Origins of Life on Earth*. Englewood Cliffs, NJ: Prentice-Hall.

Morwood, M.J., Soejono, R.P., Roberts, R.G., Sutikna, T., Turney, C.S.M., Westaway, K.E., Rink, W.J., Zhao, J.-X., van den Bergh, G.D., Due, R.A., Hobbs, D.R., Moore, M.W., Bird, M.I. & Fifield, L.K. (2004). Archaeology and age of a new hominin from Flores in eastern Indonesia. *Nature*, *431*, 1087-1091.

Ploog, D. (1981). Neurobiology of primate audio-vocal behavior. *Brain Research, 228,* 35-61.

Nelson, D.A. & Marler, P. (1989). Categorical perception of a natural stimulus continuum: birdsong. *Science, 244*, 976-978.

Nissen, P., Hansen, J., Ban, N., Moore, P.B. & Steitz, T.A. (2000). The structural basis of ribosome activity in peptide bond synthesis. *Science, 289*, 920-930.

Nottebohm, F., Stokes, T.M. & Leonard, C.M. (1976). Central control of song in the canary, Serinus canarius. *Journal of Comparative Neurology, 165*, 457-486.

Owren, M.J., Seyfarth, R.M. & Cheney, D.L. (1997). The acoustic features of vowel-like grunt calls in chacma baboons (Papio cynocephalus ursinus): implications for production processes and functions. *Journal of the Acoustic Society of America, 101*, 2951-2963.

Roberts, R. (2004, October 28). Villagers speak of the small, hairy Ebu Gogo. *Daily Telegraph (UK)*.

Schwartz, A.W. (1998). Origins of the RNA world. In A. Brack (ed.), *The Molecular Origins of Life* (pp. 237-254). Cambridge: Cambridge University Press.

Sereno, M.I. (1984). *'DNA' and Language: the Nature of the Symbolic-representational System in Cellular Protein Synthesis and Human Language Comprehension*. Unpublished doctoral dissertation, University of Chicago, Illinois.

Sereno, M.I. (1991a). Language and the primate brain. *Proceedings, Thirteenth Annual Conference of the Cognitive Science Society* (pp. 79-84). Lawrence Erlbaum Associates.

Sereno, M.I. (1991b). Four analogies between biological and cultural/linguistic evolution. *Journal of Theoretical Biology, 151*, 467-507.

Simonyan, K. & Jurgens, U. (2003). Efferent subcortical projections of the laryngeal motorcortex in the rhesus monkey. B*rain Research, 974*, 43-59.

Tyack, P.L. & Sayigh, L.S. (1997) Vocal learning in cetaceans. In C.T. Snowdon, M. Hausberger et al. (eds.), *Social Influences on Vocal Development* (pp. 208-233). Cambridge, England: Cambridge University Press.

Wills, C. & Bada, J. (2000). *The spark of life: Darwin and the primeval soup*.

Cambridge, MA: Perseus Publishing.

Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate, J.H.D. & Noller, H.F. (2001). Crystal structure of the ribosome at 5.5 A resolution. *Science, 292*, 868-299.

Zahavi, A. (1993). The fallacy of conventional signaling. *Philosophical Transactions of the Royal Society of London B, 340*, 227-230.

Zhang, B. & Cech, T.R. (1997). Peptide bond formation by in vitro selected ribozymes. *Nature, 390*, 96-100.

Cognitive Science Online, Vol 3.1, pp.13-24, 2005 http://cogsci-online.ucsd.edu

Conceptual blending theory and psychiatry

Michael Kiang

Cognitive Science Department
University of California, San Diego
mkiang@cogsci.ucsd.edu

Abstract

Conceptual blending has been proposed as a model for a variety of cognitive phenomena, and has the potential to contribute to our theoretical understanding of psychopathology. For example, from the perspective of conceptual blending theory, delusions are the product of conceptual integration processes similar to those present in many normal cognitive functions; however, there is aberrant selection of an input space, resulting in blends with relations that would normally be regarded as unfounded. Similarly, chronic interpersonal difficulties can be viewed as the repetitive application of a particular input-space frame to the construction of blends. Future research could examine whether this framework can be applied more generally to other psychiatric disorders, and whether the explicit reconstruction and examination of conceptual integration networks can be useful as a psychotherapeutic technique.

Introduction

Since its inception, conceptual blending theory has been proposed as a model for a variety of cognitive activities. In conceptual blending, elements from different mental spaces are aligned with one another as well as selectively projected to a blended space, yielding emergent meaning (Fauconnier & Turner, 1998). Conceptual blending has been postulated as a general principle underlying the construction of meaning in a range of mental phenomena of different complexity. As described comprehensively by Gilles Fauconnier and Mark Turner (2002) in The Way We Think, such phenomena include relatively simple linguistic constructions like the sentence Tom is my father; as well as more complex operations used in metaphor, analogical problem-solving, scientific creativity, literature, ritual, and humor.

Conceptual blending theory is predicated on the existence of different mental spaces, each comprising a set of interrelated elements that can be activated as a unit. An example of an extremely simple mental space with minimal interrelation between the elements would be one consisting of two people (e.g. Jane and Robert) with no specified relation between them. An example of a mental space whose elements are related in a specific organization or frame would be one in which Julie buys coffee at Peet's Coffee Shop (from Fauconnier & Turner, 2002). Its elements include Julie, the coffee shop, a cup of coffee, a salesperson, money, and so on. The frame specifies interrelations among these elements, e.g. that the salesperson gives the cup of coffee to Julie, that Julie gives money to the salesperson, and so on.

The process of conceptual blending involves the construction of relations between several different mental spaces in a conceptual integration network. These mental spaces include two or more input spaces, a generic space, and a blended space (the blended space can also be referred to simply as the blend). Elements in the generic space each map to an element in each of the input spaces, and represent what the input spaces have in common. In addition, selective mapping between elements across input spaces occurs, along with selective projection of elements from input spaces to the blended space. By establishing relationships between elements in different spaces, these processes create emergent meaning.

So far, there has been relatively little examination of how psychiatric problems might be viewed within the framework of conceptual blending theory. As an example of blending, Fauconnier and Turner (2002) cite the phenomenon of "lottery depression," in which holders of lottery tickets become depressed after losing, even though they acknowledge that the odds of winning were minute. This exemplifies the construction of a blend representing an imagined or counterfactual situation, in which the individual wins the lottery (with input spaces corresponding to the individual's actual situation, and to that of a prototypical lottery winner), with further elaboration of the blend by addition of imagined objects that the individual would acquire. Apparently, some individuals invested so much emotional energy in this counterfactual blend that they became depressed when they learned that it would not come true, feeling as if they had "lost" the fantasy objects they had acquired. Fauconnier and Turner (2002) also cite the example of anosognosia, a syndrome in which individuals acquire a neurological deficit but have no conscious awareness of it, as when a stroke patient with a paralyzed left arm is unaware of the paralysis, and gives other explanations for why she is not moving it which are false but which she consciously believes. This can be seen as another example of a counterfactual blend, but one that has taken on a character that the patient experiences as fully real.

Further examination of psychiatric problems from the perspective of conceptual blending theory could potentially contribute to our theoretical understanding of the cognitive abnormalities that lead to particular symptoms, and inform our efforts to treat them. In addition, if psychiatric disorders can be described in terms of abnormalities of conceptual blending, this would help support the ecological validity of this theoretical model, and could add to our knowledge of the empirical principles that govern which particular blends (out of all possible ones) people do or do not construct in different situations. The purpose of this article is to apply conceptual blending theory in an exploratory manner toward the analysis of two selected examples of psychiatric symptoms: delusions and chronic interpersonal difficulties.

Delusions as Conceptual Blends

One common psychiatric symptom, seen in a variety of disorders, is the delusion. Delusions have been defined as beliefs that are false, not accounted for by the patient's cultural background, and fixed (i.e. the patient does not consider any alternative interpretations as possible) (Kaplan & Sadock, 1998). Delusions can develop in different disorders, including primary psychotic disorders like schizophrenia or delusional disorder, mood syndromes like depression or mania, and delirium due to medical causes or substance intoxication. Thus, analogous to a physical symptom like fever, delusions are regarded not as having a single possible cause, but rather as being the final common pathway of a range of underlying pathophysiologies. The content of delusions is also to some extent characteristic of the particular disorder causing them - for example, a belief that one is Jesus Christ would be most typical of mania, a delusion that all one's money has been taken away as punishment for a bad deed in the past would be characteristic of depression, and a delusion that one's thoughts or actions are being controlled by an external force (e.g. an implanted microchip receiving signals from the CIA) would be characteristic of schizophrenia.

To look more closely at delusional thinking from the perspective of conceptual blending, let us examine an example of such a case, one of delusional disorder of the paranoid type (described in detail in Spitzer, Gibbon, Skodol, Williams, and First (1994)). In this case, a 42-year-old married postal worker and father of two, with no previous history of any psychiatric problems, is brought to the emergency room by his wife because he has been insisting "there is a contract out on my life." According to the patient, four months previously, his supervisor wrongly accused him of tampering with a package. The patient appealed and was exonerated in a hearing. Two weeks later, the patient began to notice that when he approached co-workers, "they'd just turn away like they didn't want to see me." He then became sure that they were talking about him behind his back, and that all this was because his supervisor had taken out a contract on the patient's life, because the supervisor was "furious" that he had been "publicly humiliated." The patient then began noticing unfamiliar "large white cars" repeatedly driving up and down his street, became convinced that these cars contained "hit men" seeking to kill him, and became terrified.

The paranoid ideation in this case can be seen as involving the construction of new conceptual blends. These are double-scope blends (Fauconnier & Turner, 2002), containing input spaces with different organizing frames, and a blended space with its own organizing frame that draws structure from each of the input frames, yielding emergent structure. One input space (which we will term the thematic input space) has a frame incorporating the theme of the delusion, and another input space (the situational input space) includes elements of the patient's actual experience. Figure 1 schematically illustrates the conceptual integration network for the case presented above. The thematic space on the left has a frame in which hit men drive cars along the street where their quarry lives, intending to kill him for their boss. The situational input space on the right contains elements from the patient's experience - himself, his supervisor, and white cars which he has apparently seen driven repeatedly up and down the street where he lives. The generic space includes elements common to both input spaces - in this case, the presence of cars driven on a street on which a person lives. In constructing the network, the patient has made selective alignments between elements in the two input spaces - for example, mapping the apparently observed drivers and their cars to hit men and their cars, his supervisor to the hit men's boss, and himself to the hit men's target. Both members of each mapped pair are then projected to the same element in the blended space, so that in the blended space, the drivers whom the patient has seen become hit men, the patient's supervisor becomes the hit men's boss, and the patient becomes the intended victim. In addition, each input space has some elements or relations that have no counterpart in the other space, but are projected into the blended space, yielding emergent meaning in the overall blend. For example, the relation between the hit men and their victim in the thematic space is that they intend to kill him. In the situational space, there is no such relation between the drivers and the patient - the white cars could conceivably be limousines that take this street to get to the airport, or belong to neighbors living on the street. Similarly, the relation between the hit men and their boss in the thematic space is that the boss has hired the hit men to kill the target, whereas in the real situation no such relation exists between the patient's supervisor and the drivers of the cars. Thus, relations from the thematic space are projected to, and lend their structure to, the blended space, adding organizing (but erroneous) meaning to elements projected from the situational space. Meanwhile, in the situational space, the cars' characteristics of being white and of being seen to drive up and down the street repeatedly, as well as the workplace relationship between the patient and his supervisor, are properties not specified in the thematic space. Thus, in the blend, the hit men have a predilection for white cars, and drive them up and down the street repeatedly without actually killing the patient, and the hit men's boss is targeting his subordinate in the workplace – emergent characteristics not present in a stereotypical "hit man" frame.

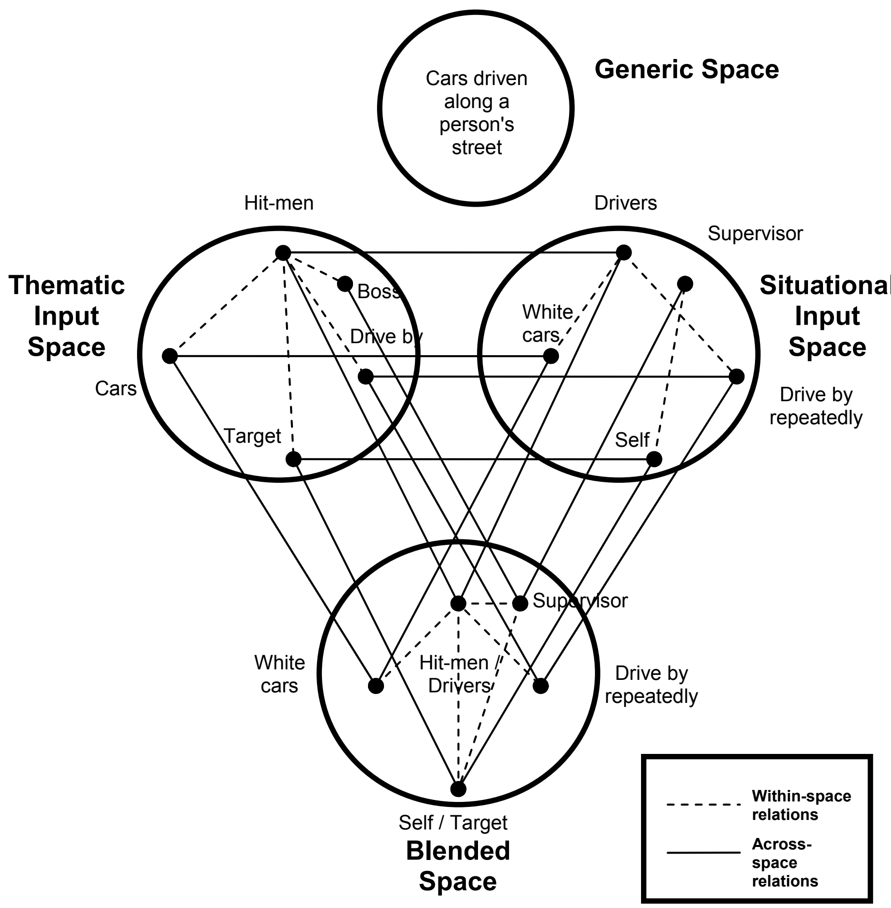

Figure 1. Blending diagram of hit man scene

In general, then, delusions can be seen as the product of the same kind of conceptual integration processes present in many normal cognitive capacities like analogy, creativity, and metaphor; it is the selection of input spaces that is aberrant, resulting in blends with relations that would normally be regarded as unfounded. In the above example, the blend depends on the establishment of mappings between elements in the "hit man" thematic input space and the situational input space containing objects or events seen by the patient. To a normal individual, however, there is no evidence to support these alignments as appropriate. According to Fauconnier and Turner (2002), two of the principles governing the construction of blends are the need to maximize topological connections between mental spaces, and the need to maintain appropriateness of these connections. These two principles can oppose each other in a given situation, requiring the individual to find a blend that compromises by satisfying both to some degree. In the anosognosia example, the cross-space mapping between the patient and a person who can move her arm is seen by others as pathological because they see evidence to contradict it; thus, in this case topology is maximized at the expense of appropriateness. In contrast, Fauconnier and Turner's (2002) "Hide the Penny" example, in which a woman draws a parallel between her brother's present relationship difficulties and his childhood behavior, the mapping between the childhood and adult actions is seen as appropriate because of our belief that people's behavior patterns persist over time; however, in a hypothetical culture where there was no such belief, the mapping would be seen as unfounded. Thus, delusions are characterized not by an inability to construct conceptual blends, but rather by the selection of inappropriate combinations of inputs for these blends.

Although conceptual blending theory offers a characterization of the cognitive abnormality in delusions - namely, the inappropriate selection of input spaces for conceptual integration, it can not explain why a specific type of frame tends to be consistently overselected as thematic space in a particular illness. For example, it would be characteristic of a manic patient with grandiose delusions to believe that white cars repeatedly driving by his home contained spies seeking to kidnap him to find out the secret behind his special powers, while a woman with an erotomanic delusional disorder would be apt to believe that such cars contained a prominent person who was secretly in love with her and wanted to catch a glimpse of her. Conversely, although an unlimited number of thematic frames are possible, many of these are not seen empirically in delusions. For example, it would be unusual in any psychiatric disorder for a patient to develop a belief that he is a non-human animal or an inanimate object (such that, upon encountering a patient who said this, one would likely question whether he was feigning illness). As stated by Fauconnier and Turner (2002):

> "What counts as a 'natural' match will depend absolutely on what is currently activated in the brain. Some of these activations come from real-world forces that impinge on us…others from bodily states…and many others from internal configurations of our brains acquired through personal biography, culture, and, ultimately from biological evolution."

A framework for further research is to explore the factors - whether biological, psychological, environmental, or a combination thereof – that determine why delusions in different disorders are characterized by the activation of particular thematic-space frames.

Conceptual blending theory does postulate, however, that, for blends in general, once their particular input frames are activated, the binding of elements in the subsequent conceptual integration network is often unconscious and can be very entrenched. The processes involved in constructing the blend "operate for the most part automatically and below the horizon of conscious observation" which "makes the detection of biases difficult" (Fauconnier & Turner, 1998). This is consistent with the observation that delusions are relatively impervious to conscious questioning, whether by the individual herself, or in challenges from others. One factor contributing to this robustness of blends, including delusions, appears to be the rewarding sense of insight that they appear to engender. Achieving the topological alignments in a blend seems to "give us the feeling that 'one thing' is giving us insight into 'another thing'…strong emotions emergent in the blend can induce the feeling of global insight, because the highly compressed blend remains actively connected to the entire network" (Fauconnier & Turner, 2002). It also appears that maximizing the number of alignments can increase this feeling of insight, as long as the individual considers them appropriate. This process can be seen in the elaboration of blends, in which additional aligned elements are recruited, with this pattern completion contributing to the feeling of understanding. Thus, in the Hit Men example, the patient might next see the driver of a car talking on a cell phone, and then believe in addition that this "hit man" must be communicating with an associate in order to co-ordinate the hit. From an evolutionary point of view, a feeling of insight, security or certainty triggered by the elaboration of blends might be adaptive, since our evolutionary ancestors encountered situations where they needed to be able to recognize and complete patterns quickly in order to survive. For instance, if they perceived a large object moving quickly toward them in the dark, the most adaptive response would be to assume that this was a predator and to flee, even if this assumption might not be correct.

The role of this sense of insight in the formation and maintenance of conceptual blends is consistent with some existing theories of delusion formation. According to one theory of delusion formation in schizophrenia, the primary abnormality is one of dysregulated dopaminergic transmission which causes environmental stimuli to develop an abnormally salient character (Kapur, 2003). This theory is based on phenomenological observations in which schizophrenia patients report that, immediately before becoming delusional, they experienced a state in which normally unremarkable stimuli took on an unusually salient character, seeming to hold some strange but unknown significance. Thus, Sullivan (1994/1927) describes how "all sorts of trifling and wholly unrelated occurrences seemed fraught with great personal import, to bear in some signal but incomprehensible way upon him." This abnormal salience is also reflected in patients' own words: "It was as if parts of my brain awoke which had been dormant...I became interested in a wide assortment of people, events, places, and ideas which normally would make no impression on me" (Bowers & Freedman, 1966); or "my senses seemed alive, colors were very bright, they hit me harder...I noticed things I had never noticed before" (Bowers, 1968). Schneider (1959) described a patient with schizophrenia who attributed strange significance to a dog he had seen:

A dog lay in wait for me as he sat on the steps of a Catholic convent. He got up on his hindlegs and looked at me seriously. He then saluted with his front paw as I approached him. Another man was a little way in front of me. I caught up to him hurriedly and asked if the dog had saluted him too. An astonished 'no' told me I had to deal with a revelation addressed to me.

The strange sense of significance leaves patients perplexed and seeking some kind of explanation for it - e.g. “I was really groping to understand what was going on” (Bowers, 1968). Kapur (2003) proposes that delusions are a “top-down” cognitive explanation that the individual then imposes upon these experiences in an effort to make sense of them. Delusions are thus to some extent imbued with the individual’s particular personal history, psychodynamic themes and cultural context, so that within schizophrenia, one patient might attribute experiences caused by aberrant salience to the evil ministrations of a shaman, and another to technological machinations by the CIA. Once the patient arrives at such an explanation, though erroneous, it provides a “psychotic insight” which relieves her previous sense of uncertainty (Bowers & Freedman, 1966; Roberts, 1992). In terms of conceptual blending theory, in this model of delusion formation, the patient constructs new conceptual blends to explain the strange aberrant salience of stimuli. The patient’s individual psychodynamics constrain the elements selected for the thematic input space, and hence the specific subject matter of her delusions. The reinforcing sense of insight gained by completing the blend might explain why delusions are so resistant to psychotherapy or rational challenge by others. An open question is whether this model of delusion formation - in which aberrant salience of stimuli is the primary cause, and the construction of explanatory blends is secondary – is generalizable to other disorders involving delusions, or whether some of these disorders may involve an overactive input frame as the primary pathology.

Conceptual blending and interpersonal difficulties

A different type of psychiatric problem to which conceptual blending theory might be fruitfully applied is that of interpersonal difficulties customarily treated with psychotherapy. Unlike delusions, which typically arise acutely in a previously normal individual, are believed to result from an illness state, and are treated most effectively by pharmacotherapy, these interpersonal difficulties tend to be chronic and are thought to result from the patient’s personality development. Mitchell and Black (1995) describe a typical case, that of Emily, a professionally successful young woman.

> Emily came to treatment because of her difficulties in establishing satisfying relationship with other people, both men and women. She tended to put people off for reasons she did not understand; she found herself irritated and impatient with others. She tended to feel that in most of her activities - in her work, around her house, sexually in her bed - she could do it better by herself than relying on others. She was so good at what she did that the wisdom of this approach seemed to be substantiated again and again.

Over the course of psychotherapy, the therapist began to consistently feel that Emily was subtly dismissive of his efforts to help her.

> He became increasingly aware that he always felt he was interrupting something he perhaps should be staying out of…Emily at first…rejected this observation, as if to reassure the analyst that his efforts were appreciated. Eventually Emily was able to reflect more fully…On the one hand, she came for treatment because she knew she needed help and felt a great regard for the analyst’s professional abilities. On the other hand…she did feel that she had been thrown off the track whenever the analyst spoke…it meant she had to deal with him, his thoughts, which could only be a distraction from…her own sense of where she needed to go.

It then came to light that during her childhood Emily perceived her father, a businessman, as self-absorbed and having little time for his children. He treated her mother as “imbecilic and incompetent,” and her mother in turn “continually turned to Emily tearfully for reassurance.” Thus, “neither parent seemed capable of attending to Emily’s needs; they broke into her world only when they needed something from her.”

A psychodynamic formulation of this case might draw parallels between the interpersonal dynamics of Emily’s childhood world and her current interactions with others, including with her therapist. Thus, her preconception of others as being unreliable and self-absorbed, learned from her childhood experiences, have led her to approach current situations with the priority of maintaining independence or control, at the expense of strengthening interpersonal relationships which would benefit her in the long run.

In terms of conceptual blending theory, the way Emily approaches her current relationships can be seen as the repetitive creation of a particular kind of double-scope blend, containing a recurring thematic input space whose elements she has learned over the course of her personal developmental history, and a situational input space whose elements vary depending on the current interpersonal situation in question. As shown in Figure 2, the thematic input space includes herself as a child, and her parent when she was a child. The elements in the situational input space include Emily and another person with whom she interacts in the present (in this example, the therapist), and these are both projected into the blended space. The mapping between herself in the present and herself as a child involves an across-space identity relation between two different entities - herself at different times. The frame in the thematic space also contains another type of element – aspects of her parent, such as unreliability or self-absorption. These are projected to the blended space, so that these aspects are connected to the therapist.

This type of blend shares similarities with the blend that we proposed as a model for delusions. In both cases, the frame in the thematic space varies from one patient to another. Its qualitative character cannot be predicted by conceptual blending theory, but depends to varying degrees on the individual’s personal history, biology or temperament. Nevertheless, it can be seen empirically that once a particular type of thematic-space frame becomes activated in an individual, it tends to be applied recurrently to different focuses, and the resulting integration networks are usually quite entrenched and not immediately conscious to introspection.

One difference between this blend and the delusion blend is that the latter contains a greater number of inappropriate projections. In the Hit Men example, the hit men and the drivers from different input spaces are projected to the same element in the blended space, leading to an identity compression, so that the patient believes that the drivers are actually hit men. An analogous process occurs for the patient and the intended victim, so that the patient believes he is actually a target of hit men. On the other hand, in the case of Emily, although there is an analogy mapping across input spaces between present-day Emily and Emily as a child, or between the parent and the therapist, only one of each pair is projected to the blended space. In other words, Emily does not actually believe that the therapist is really her parent, or that she is simultaneously both herself and herself as a child. Thus we can see that the delusional individual, as compared to the neurotic individual, is increasing topology to a greater degree, at the expense of appropriateness.

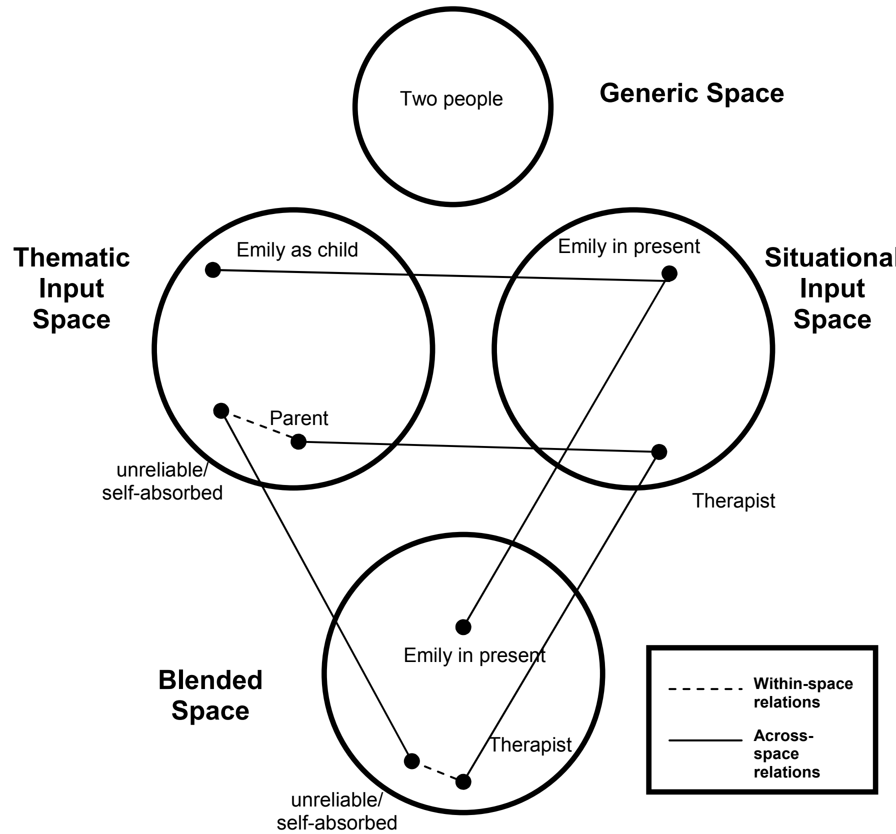

Figure 2. Blending diagram of Emily

Another contrast between the two blends is that in the chronic interpersonal problem the thematic input space is comprised of elements from the patient's childhood, whereas in the delusion, the thematic input space has elements generally characteristic of the illness rather than reflective of the patient's specific past experiences. This reflects the view that Emily's dysfunctional cognitive schema recapitulates conflicts from her early life, whereas that of the postal worker is a more generic one determined primarily by his biological illness. This comparison also highlights how the elements we, as theoreticians, choose to incorporate in the thematic input space when modeling psychopathology with a blend depend on our formulation of the etiology of the symptoms. Thus, in the case of the postal worker, an author who believed that the structure of the delusion was determined primarily by childhood interpersonal experiences might instead construct the thematic input space with elements from the patient's early life. Alternatively, since a conceptual blend can have more than two input spaces (Fauconnier & Turner, 2002), it could also reflect a formulation in which developmental experiences and a biologically-driven schema both contribute to the delusion, by representing each of these two frames with a separate thematic input space.

Perspectives on treatment

Although it appears that conceptual blending theory may help us better characterize the cognitive processes in psychiatric problems, to what extent can it inform our approach to treating them? In the case of delusions, blending theory may not be directly applicable to how one interacts with a patient, since - despite some reports of the efficacy of cognitive-behavioural therapy for delusions in a subset of patients (Rector & Beck, 2001) - psychotherapy has not in general been found to be effective in changing delusions, and pharmacotherapy is the treatment of choice. From the perspective of blending theory, some aspect of the pathogenesis of delusions limits the patient's conscious apprehension of the blended space, so that the patient while ill can only, in the words of Fauconnier & Turner (2000), "live in the blend" and is incapable of living "in the full integration network." It is unlikely that attempts to help the delusional patient achieve insight into the existence of the full integration network would be successful. In the case of interpersonal problems, however, psychotherapy is considered potentially effective, and thus conceptual blending theory could perhaps be applied toward refining the current repertoire of psychotherapeutic techniques.

Two different psychotherapeutic approaches in common use include psychodynamic therapy, and cognitive-behavioural therapy. One goal of psychodynamic therapy - used for example in the case of Emily described earlier (Mitchell & Black, 1995) - is for the patient to achieve understanding into how she manifests similar patterns of cognitions and behaviours across a range of interpersonal situations, including with the therapist, and to relate these patterns to historical situations in the patient's development that are believed to have contributed to these patterns (e.g. Book, 1998). Somewhat differently, cognitive-behavioural therapy focuses on identifying the "automatic" or barely conscious thoughts that the patient experiences in present-day situations, and the distorted assumptions that underlie these thoughts (Beck, 1995). The therapist aims to help the patient identify empirical evidence contradicting these assumptions, sometimes through assigned experiments designed to test them. The patient is also asked to come up with alternative thoughts or assumptions and to test their validity. Unlike in psychodynamic therapy, the genesis of distorted assumptions in childhood experiences is not an essential focus in cognitive-behavioural therapy, and is generally explored relatively late in the therapy, if at all, depending on the patient's apparent degree of access to these experiences and acceptance of their relevance.

It appears that implicit in the process and goals of psychodynamic therapy is insight into the "full integration network" which has previously not been fully conscious. From the perspective of conceptual blending theory, the essential goal of understanding how certain properties of childhood figures are repeatedly projected onto present-day ones can be viewed as the achievement of insight into how a particular thematic input space is repeatedly used in constructing blends. The patient comes to understand how persons from a situational space and properties from a thematic space are projected into the blended space, to construct an emergent stereotyped meaning that may lead the patient to respond in a self-defeating way.

In contrast, insight into the full integration network is not a crucial part of cognitive-behavioural therapy, where the emphasis is limited mainly to identifying the elements of the blended space. For example, if used with Emily, cognitive-behavioural therapy would primarily focus on how, across different situations, the blended spaces she constructs consistently involve the compression of a person with the properties "self-absorbed" or "unreliable." The therapy would then examine evidence for or against the appropriateness of this compression. In addition, the development and testing of alternative thoughts and assumptions would involve the conscious construction of alternative blended spaces - ones in which different properties like "caring" or "considerate" are attached to the other person instead. The therapy, however, would generally focus on the blended spaces, as opposed to examining elements and their relations in the rest of the conceptual integration networks to which the blended spaces belong. Whether or not examining the network outside the blended space has any differential effect on the outcomes of cognitive-behavioral therapy is a potential question for empirical clinical study.

In psychotherapy, conceptual blending theory could potentially be used even more explicitly than in current methods of psychodynamic therapy. The therapy might then involve reconstructing, from the clinical data, the characteristic type of conceptual integration network that the patient appears to be activating that is responsible for their presenting difficulties. This would involve identification of the elements of the input and blended spaces, and the interrelations between them. The therapy would aim at achieving insight into these components of the network, and how they are activated repeatedly across different situations. Clinical studies could then test the effectiveness of this approach in achieving subjective and behavioural change.

Conclusions

The examples presented here suggest that conceptual blending theory can provide a useful framework for understanding the cognitive processes involved in different classes of psychopathological symptoms. Both delusions and chronic interpersonal difficulties involve the application of a characteristic type of input-space frame to the construction of conceptual blends, while the actual processes involved in constructing the blend are similar to those used in many normal cognitive activities. Patients often appear to have little or no insight into the processes involved in constructing these pathological blends, which is consistent with the resistance to conscious apprehension that has been noted for many other types of conceptual blends. A framework for further research is to examine whether conceptual blending theory can provide a general framework for understanding symptoms across the spectrum of psychiatric disorders. In terms of this model, future work can elucidate how biological, psychological and cultural factors determine the particular thematic-space frames that tend to be activated in specific disorders or particular individuals. Future research could also evaluate whether, and in what circumstances, the explicit reconstruction and examination of conceptual integration networks might be useful as a psychotherapeutic technique.

Acknowledgements

I am grateful to Gilles Fauconnier, Rafael Nuñez and an anonymous reviewer for their helpful comments.

References

Beck, J. S. (1995). *Cognitive therapy: Basics and beyond.* New York: Guilford Press.

Book, H. E. (1998). *How to practice brief psychodynamic psychotherapy.* Washington, DC: American Psychological Association.

Bowers, M. B. (1968). Pathogenesis of acute schizophrenic psychosis: An experimental approach. *Archives of General Psychiatry, 19*(3), 348-355.

Bowers, M. B., & Freedman, D. X. (1966). "Psychedelic" experiences in acute psychoses. *Archives of General Psychiatry , 15*(3), 240-248.

Fauconnier, G., & Turner, M. (1998). Conceptual integration networks. *Cognitive Science, 22*(2), 133-187.

Fauconnier, G., & Turner, M. (2000). Compression and global insight. *Cognitive Linguistics, 11*(3/4), 283-304.

Fauconnier, G., & Turner, M. (2002). *The way we think: Conceptual blending and the mind's hidden complexities.* New York: Basic Books.

Kaplan, H. I., & Sadock, B. J. (1998). *Kaplan and Sadock's synopsis of psychiatry: Behavioural sciences, clinical psychiatry.* Philadelphia: Lippincott Williams & Wilkins.

Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry, 160*, 13-23.

Mitchell, S. A., & Black, M. J. (1995). *Freud and beyond.* New York: Basic Books.

Rector, N. A., & Beck, A. T. (2001). Cognitive-behavioural therapy for schizophrenia: An empirical review. *Journal of Nervous and Mental Disease, 189*(5), 278-287.

Roberts, G. (1992). The origins of delusion. *British Journal of Psychiatry, 161*, 298-308.

Schneider, K. (1959). *Clinical psychopathology.* New York: Grune and Stratton.

Spitzer, R. L., Gibbon, M., Skodol, A. E., Williams, J. B. W., & First, M. B. (Eds.). (1994). *DSM-IV casebook*. Washington, DC: American Psychiatric Press.

Sullivan, H. S. (1994/1927). The onset of schizophrenia. *American Journal of Psychiatry, 151*(6S), 134-139.