

[**To appear in:** Andor, J. & Pelyvas, P. (Eds.). *Empirical, Cognitive-Based Studies in the Semantics-Pragmatics Interface*. Berlin: Elsevier Science.]

Frame-shifting and Frame Semantics:

Joke Comprehension on the Space Structuring Model

Seana Coulson

Surface Mail:

Seana Coulson

Cognitive Science 0515

9500 Gilman Drive

La Jolla, CA 92093-0515

USA

Email: coulson@cogsci.ucsd.edu

Phone: 1-858-822-4037

Fax: 1-858-534-1128

Frame-shifting and Frame Semantics:

Joke Comprehension on the Space Structuring Model

Consider how one might understand the following ad, "For Sale, Parachute. Used once; never opened; small stain." Although there are no explicit cues to do so, the reader naturally recruits knowledge about how and why parachutes are normally used in order to construe the listed features of this chute as being causally related to one another. Empirical support for the importance of such causal and relational information in language comprehension comes from the consistent finding that memory for short texts includes inferred material (Sanford, 1981). Further, computational considerations imply that language comprehension involves a capacity for dynamic inferencing based on general knowledge represented as frames (Lange, 1989).

Invented somewhat independently by researchers in the field of linguistics (Fillmore, 1968), psycholinguistics (Sanford, 1981), artificial intelligence (Minsky, 1975), and natural language processing (Schank, 1977), frame-type data structures are often invoked to account for inferential aspects of language comprehension (Fillmore, 1982). As used here, frames are representations with slot/filler structure, default values, and weak constraints on the type of fillers for a given slot (reviewed in Barsalou, 1992). Frames contain causal and relational information, are organized hierarchically so as to allow recursive embedding of frames within frames, and can be used to represent knowledge about a wide variety of objects, actions, and events.

In linguistics, frame semantics is a research program in which a word's semantic properties are described with respect to the way that they highlight aspects of an associated frame. For example, *buy* and *sell* both evoke the Commercial Transaction frame. But *buy* highlights the buyer and the goods, while *sell* highlights the seller and the money. Background knowledge thus figures prominently in the establishment of meaning, as language functions against the backdrop of conceptual structure. Below we examine the utility of frame semantics for models of language comprehension, and focus on its capacity to explain the sorts of inference needed to understand the parachute joke.

In section 1, we consider the role that pragmatics plays in on-line meaning construction, contrasting the view that context-independent meaning is *supplemented* with background and contextual knowledge with the view that background and contextual

knowledge is *essential* for the construction of meaning. Although most people working in pragmatics subscribe to varying versions of the latter, these portraits of the semantics/pragmatics interface have had almost no influence on language researchers in psychology, computer science, and cognitive science. One exception to this trend is the space structuring model (Coulson, 2000), a model of language comprehension inspired by ideas in cognitive linguistics. Section 2 outlines some assumptions of the model, and section 3 describes experimental work on joke comprehension designed to test some of those assumptions. Section 4 reviews the adequacy of some frame-based models in cognitive science, and section 5 offers our speculations about the future of this approach.

1. Pragmatics: What's that?

On the classical approach to language, pragmatics is defined largely by what semantics leaves out (Gazdar, 1979). This definition involves positing an implicit model of comprehension in which purely linguistic knowledge of meaning is used to determine sentence meaning. To determine *utterance* meaning, all that remains is to bring background knowledge and contextual information to bear on sentence meaning. Katz & Fodor (Katz, 1963), for example, argued that a viable theory of semantics would require a theory of pragmatics that would disambiguate sentences vis a vis their contexts of utterance. Indeed the core phenomena in pragmatics are deixis, speech acts, and implicature – all topics in which the primary concern is how contextual factors modulate interpretation.

One problematic aspect of the classical view of meaning is that it relegates pragmatics to the role of disambiguator. For example, in a sentence about a pawn in a chess game, pragmatics is supposed to answer the question "Which pawn?" Similarly, a prominent function of pragmatics on the traditional view is to give content to indexicals. Thus, on such an account, pragmatics is what answers questions such as "Who are *you*?" and "Where is *here*?"

Although few people working in pragmatics are likely to agree to exactly this characterization, unfortunately, it still captures the way that many linguists and psycholinguists think about language in context. In psycholinguistics, in particular, the role of context is almost exclusively in the resolution of ambiguities in lexical meaning and syntactic structure. A key issue in research on word recognition has concerned whether

context affects the initial choice of a word's meaning (see (Simpson, 1994) for review), while a major focus of sentence processing research is whether contextual factors influence the initial assignment of a syntactic representation, and how context affects syntactic reanalysis (e.g. (Frazier, 1995); (Spivey-Knowlton, 1994)). To the psycholinguist, then, pragmatics is what answers the question, "Is that *funny* 'haha' or *funny* 'weird'?" Similarly, when confronted with the statement "Tonight we'll discuss sex with the President," pragmatics is what tells us whether the President will be the *topic* of discussion, or a *participant* in it.

While these are all valuable contributions, and disambiguator is a perfectly respectable occupation, it is not (I believe) the role that background and contextual knowledge in fact plays in meaning construction. Though presumably motivated by the observation that many words have many meanings, the pragmatics as disambiguator model is ultimately undermined by careful consideration of the way in which word meaning changes from context to context. Although *tomato* is not ambiguous in the examples below, Johnson-Laird has pointed out that different features of tomatoes are salient in (1) than in (2) (Johnson-Laird, 1993) .

(1) The tomato rolled across the floor

(2) He accidentally sat on a tomato.

Examples such as this would seem to argue against a view in which the speaker's knowledge of the context of utterance helps her to adjudicate between a finite set of determinate meanings. Rather, background knowledge about tomatoes is used to enrich the cognitive models constructed to represent the referential situation depicted in the sentences.

Moreover, various pragmatic functions are available that allow any noun to refer to an infinite set of related phenomena. For example, *BMW* can be used to refer to all sorts of BMW-related things, including stock in the company, as in (3), the building that houses the company, as in (4), or shattered BMW car parts as in (5).

(3) He bought 6,000 shares of BMW.

(4) He works next door at BMW.

(5) My ex- planted a bomb in my car and I've been cleaning up BMW ever since.

Similarly, Clark has pointed out that these sorts of mechanisms allow speakers to improvise word meanings *ad infinitum* (Clark, 1983) . For example, in an article about choosing a college roommate based on the appliances she owns, "computer" can mean roommate who owns a computer, as in (6).

(6) Our computer fell in love and left school.

The problem with the pragmatics as disambiguator model is that while it presumes that computing utterance meaning is difficult because language is semantically *ambiguous*, the difficulty really lies in the fact that language is semantically *indeterminate*. Rather than helping to adjudicate between pre-specified meanings, it would seem that context plays a constitutive role in the construction of meaning. Moreover, any mechanism that can account for the way that people can deal with semantic indeterminacy, ought to account quite trivially for semantic ambiguity as well. Rather than attempting to explain how pragmatic factors disambiguate a fully specified meaning constructed by semantics, then, it makes more sense to explore how pragmatic mechanisms contribute to the specification of meaning in the first place.

2. Space Structuring Model

On the space structuring model (Coulson, 2000), a model of language comprehension motivated by ideas in mental space theory (Fauconnier, 1994), conceptual blending theory (Fauconnier, 1998; Fauconnier, 2002), and cognitive grammar (Langacker, 1987), perceptual input, language input, social context, and the speaker's current cognitive state all contribute to the construction of cognitive models of the discourse situation. This can include models of the referential aspects of sentences as well as models relevant to the agent's social and material goals. Both linguistic and non-linguistic cues prompt the retrieval of frames from long-term memory, and these frames are exploited in the construction of cognitive models of the message-level representation.

Three assumptions of the model include:

(1) *the embodiment assumption*, that the structure of language at least partially reflects bodily constraints on perception and action;

(2) *the immediacy assumption*, that the integration of linguistic and non-linguistic information occurs rapidly, and does not (necessarily) require the prior construction of a propositional representation of sentence meaning;

(3) *the elaboration assumption*, that language comprehension involves animating the cognitive models constructed by the listener.

In traditional frame-based approaches (such as Schank, 1977), comprehension requires frames to be bound to contextually available elements. However, in the space structuring model, this need not be the case. At times comprehension can proceed by binding slots or attributes in the activated frame. However, often the frame serves only to constrain the construction of a cognitive model that is particularized to the discourse situation. Like the frames that inform them, these cognitive models are hierarchically organized, have attribute/value structure, and a mechanism that assigns default values for unspecified attributes. Though schematic and partial, these models are detailed enough to enable small-scale simulations of the scenarios they represent (as in the mental models described by (Norman, 1974).

For example, if a listener heard the sentence in (7), she might, at one level at least, respond by constructing a model of the referential situation described by the speaker.

(7) When I asked the bartender for something cold and full of rum, he recommended his daquiri.
Of course, at other levels the listener might be building models related to why the speaker might make this particular statement, what the speaker's attitude is toward her, or many other things. Nonetheless, at the referential level, the listener combines linguistic information with background knowledge to build a cognitive model of an interaction between a customer and a bartender.

The fact that speakers build models like this, and the extent to which those developing models guide their expectations is perhaps best appreciated by examining examples in which those expectations are violated. For example, (8) is very similar to (7) but prompts the construction of a very different cognitive model.

(8) When I asked the bartender for something cold and full of rum, he recommended his wife.

Rather than recommending a drink, the bartender in (8) has just insulted his wife! The semantic and pragmatic reanalysis that reorganizes existing elements in the message-level representation is known as *frame-shifting* (Coulson, 2000). With the activation of background knowledge and the establishment of mappings between

counterpart structure in the old frame and in the new one, the bartender's wife is accused of being a frigid lush.

In fact, jokes are deliberately constructed to suggest one frame while evoking elements consistent with another. While frame-shifting is not unique to jokes, jokes differ from more 'everyday' examples in the extent to which the need to shift is clearly demarcated. For example, in (9) the reader begins by evoking a frame in which a busy professional pays an accountant to do his taxes.

(9) I let my accountant do my taxes because it saves time: last spring it saved me ten years.

At the disjunct *years* however, the reader is forced to go back and reinterpret *time* to evoke a frame where a crooked businessman pays an accountant to conceal his illegal business dealings. The word *time* is called a *connector* because it serves as a bridge between the two frames. Merely knowing that *time* refers to time in prison does not in and of itself explain why the accountant is doing the man's taxes, or how doing so will prevent a prison sentence. A full understanding of (9) requires recruitment of background knowledge about the particular sorts of relationships that can obtain between business people and their accountants so that the initial busy professional interpretation can be mapped into the crooked-businessman frame.

Frame-shifting involves a dramatic reorganization of the message-level aspects of the utterance, most of which can't be attributed to compositional mechanisms of reanalysis. For example, most examples of frame-shifts in jokes don't require the listener to instantiate a new structural analysis of the sentence. Though the listener is led down a pragmatic garden path, it is often the case that pragmatic reanalysis proceeds without syntactic reanalysis. In (8), for example, *wife* is the object of *recommended* just as *daquiri* is in (7), the straight version of (8). In (9), the joke interpretation actually requires the reader to abandon the fully grammatical reading of "saved me ten years" for something akin to "saved me jail time," which is questionable at best.

Further, such cases of frame-shifting frequently require the creation of nonce senses. For example, in the accountant joke in (9) "saves time" is re-interpreted as meaning "prevents me from having to do time." In the case of the bartender's wife, "full of rum" comes to mean "alcoholic." However, the construction of these somewhat novel

phrasal meanings is as much the *effect* of frame-shifting as the *cause*. That is, it seems likely that understanding the novel reading of “full of rum” at least partially depends on the construal of the bartender’s speech act as an insult. Moreover, the adaptation of the idiomatic meaning of *time* in (9) as in (“do time”) is only congruous because of a stereotyped scenario that involves accountants obscuring illegal business dealings.

3. The Psychological Reality of Frame-Shifting

Demonstration of the psychological reality of frame-shifting, then, would suggest a role for pragmatics that goes beyond pragmatics the disambiguator. In particular, frame-shifting suggests that lexical processing does not simply benefit from context, but actively contributes to it. Moreover, background and contextual knowledge do not merely help the listener to specify the meaning of indexicals and disambiguate the meaning of lexical items, but, rather, are crucial for the construction of the message-level representation. Below we discuss the results of studies using three different techniques that establish the psychological reality of frame-shifting: self-paced reading times, eye movement registration, and event-related brain potentials.

3.1 Self-Paced Reading Times

To demonstrate the psychological reality of frame-shifting, Coulson & Kutas (Coulson, 1998) conducted a variety of experiments using the self-paced reading time technique. In this experimental paradigm, the task is to read sentences one word at a time, pressing a button to advance to the next word. As each word appears, the preceding word disappears, so that the experimenter gets a record of how long the participant spent reading each word in the sentence.

Stimuli for this experiment were comprised of one-line jokes that required frame-shifting for their comprehension, and straight versions of the same sentences that did not require a frame-shift. Moreover, because we wanted to be able to detect the effect of frame-shifting on the processing of a single word, the *disjunct*, or frame-shifting trigger, was always a sentence-final noun. In order to find out what sort of non-joke frames people constructed for these sentences, we performed a norming task in which people were given the jokes minus the last word and asked to complete the sentence with the first word or phrase that came to mind. This is known as a *cloze task*, and the percentage of

people who offer a given word in a given sentence context is known as the *cloze probability* of that word in that particular sentence context.

Results of the cloze task enabled Coulson & Kutas to ascertain readers' default (non-joke) interpretation for the sentences. However, it also revealed a disparity in the cloze probability of the most popular response for the items, suggesting that some of the sentence fragments provided a more constraining context than others. For example, (10) elicited a similar response from 81% of the participants, while (11) elicited many different responses, albeit mostly from the gambling frame.

(10) I asked the woman at the party if she remembered me from last year and she said she never forgets a (face 81%).

(11) My husband took the money we were saving to buy a new car and blew it all at the (casino 18%).

As a result, two types of jokes were tested: high constraint jokes like (10) which elicited at least one response with a cloze probability of greater than 40%, and low constraint jokes like (11) which elicited responses with cloze probabilities of less than 40%. To control for the fact that the joke endings are (by definition) unexpected, the straight controls were chosen so that they matched the joke endings for cloze probability, but were consistent with the frame evoked by the context. For example, the straight ending for (10) was *name* (the joke ending was *dress*); while the straight ending for (11) was *tables* (the joke ending was *movies*). The cloze probability of all four ending types (high and low constraint joke and straight endings) was equal, and ranged from 0% to 5%.

Given the impact of frame-shifting on the interpretation of one-line jokes, one might expect the underlying processes to take time, and, consequently result in increased reading times for jokes that require frame-shifting than "straight" versions of the same sentences. Coulson & Kutas (1998) found that readers spent longer on the joke than the straight endings, and that this difference in reading times was larger and more robust in the high constraint sentences. This finding suggests there was a processing cost associated with frame-shifting reflected in increased reading times for the joke endings, especially in high constraint sentences that allow readers to commit to a particular interpretation of the sentence.

3.2 Eye Movement Registration

The self-paced reading paradigm is a good technique for establishing that one kind of sentence requires more processing time than another (presumably very comparable) type of sentence. However, one drawback to this technique is its lack of ecological validity. In contrast to normal reading, participants in a self-paced reading task are permitted to see only one word at a time, and moreover, are not permitted to look back at earlier regions of the sentence. In contrast, in free reading, people frequently move their eyes leftward (or *regress*) to re-examine earlier parts of the text. Another deficit of the self-paced reading paradigm is that it provides little information about the *nature* of the processing difficulty that readers encounter. For example, longer reading times for sentences that ended as jokes than straight controls in Coulson & Kutas (1998) suggests the jokes engendered more processing difficulty. However, reading time data do not indicate whether this difficulty occurs in the initial stages of word processing, or later as the reader moves on to inferential aspects of processing.

To address these questions, Coulson & Urbach (Coulson, In Preparation) conducted an eye movement study comparing reading times for sentences that ended as jokes to reading times for the same sentences with the unexpected straight endings (as in Coulson & Kutas, 1998). They found that just as in the reading time study, readers spent reliably longer on the joke endings, and that this difference was far more pronounced for the high constraint sentences. Moreover, this effect arose in the later stages of processing associated with subsequent fixations of the sentence final word (rather than in the initial phase of processing). Further, in both high and low constraint sentences, participants were more likely to regress when they encountered joke endings than straight ones. This finding is consistent with the psychological reality of frame-shifting, suggesting readers literally revisit aspects of the preceding context in order to get the jokes.

3.3 Event-Related Brain Potentials

Another way of assessing readers' on-line comprehension of language materials is to use event-related brain potentials (ERPs). ERPs provide an on-going record of brain activity related to various kinds of sensory, motor, and cognitive processing events. The physical basis of the ERP signal is the fact that when large groups of neurons (on the order of tens of thousands) fire simultaneously, they create an electrical field in the brain

that can be detected with electrodes at the scalp via the electroencephalogram (EEG). The ERP is obtained by applying electrodes to the scalp, recording participants' EEG, and averaging across events within experimental categories. Because the averaging process presumably cancels out the EEG that is not related to the experimenter's categories, the remaining signal represents the brain activity related to the processing of the experimental stimuli. By comparing the ERPs to different sorts of stimuli, the researcher can assess how changing the nature of the cognitive task modulates the brain response.

Because eye movements necessary for normal reading produce artifacts in the EEG, ERP reading experiments typically involve presenting sentences one word at a time in the center of a computer monitor. EEG can thus be time-locked to the onset of each word on the monitor, and the resultant ERP represents brain activity associated with reading a particular category of words (i.e., the last word of an incongruous sentence). The ERP is a waveform with a series of positive and negative peaks (often called *components*) that can be correlated with various types of processing. Components are generally labeled by reference to their polarity (P for positive-going and N for negative-going activity), and their latency, or when they occur relative to either the onset of the stimulus event or to other ERP components.

In a classic ERP language experiment, Kutas & Hillyard (Kutas & Hillyard, 1980) contrasted ERPs elicited by visually presented sentences that ended congruously, as in (12), with ERPs elicited by sentences that ended incongruously, as in (13).

(12) I take my coffee with cream and sugar.

(13) I take my coffee with cream and socks.

They found a negativity in the brainwaves that was much larger for incongruous sentence completions than the congruous ones. Because it peaks about 400 milliseconds after the onset of a visually presented word, this negativity is called N400.

Over thirty years of research have revealed reliable relationships between the nature of various stimulus and task manipulations designed to alter participants' cognitive state, and corresponding modulations of ERP components. For example, the P1 component is the first positive deflection in the ERP elicited by visually presented words. This component, evident 70-100 milliseconds after the word is shown (or post-word onset), reflects early sensory and vision-related attentional processing. Mangun, Hillyard,

& Luck (Mangun, 1993) have proposed that the P1 component reflects a gating mechanism responsible for modulating the width of the attentional spotlight. P1, N1, and P2 components elicited during reading probably reflect the visual feature extraction necessary to relate the visual stimulus to information in memory (Kutas & King, 1996).

Most ERP studies of language processing have focused on longer latency components. In particular, much language research has focused on the N400 component. Because the N400 was initially reported as a brain response to incongruous sentence completions (Kutas & Hillyard, 1980), many people mistakenly believe it is only elicited by semantic anomalies. However, research indicates it is a far more generally elicited ERP component associated with the integration of a word into the established context. In fact, N400 is elicited by *all* words, spoken, signed, or read, and its size, or amplitude is an index of the difficulty of lexical integration. The best predictor of N400 amplitude is a word's cloze probability in a particular sentence: N400 is small for high cloze expected completions like *sugar* in (12), large for low cloze completions like *socks* in (13), and intermediate amplitude for sentence final words of intermediate cloze probability.

Besides providing an on-going record of brain activity during language processing, ERP data can complement reaction time data such as that collected in the self-paced reading and eye movement registration paradigms discussed above. These two kinds of data are often complementary as reaction time data can provide an estimate of how long a given processing event took, while ERP data can suggest whether distinct processes were used in its generation. An experimental manipulation that produces a reaction time effect might produce two or more ERP effects, each of which is affected by different sorts of manipulations. To the extent that ERP effects can be identified with specific cognitive processes (i.e., the N400 and lexical integration), they provide some evidence of *how* processing differs in the different conditions (King, 1995).

With this in mind, Coulson & Kutas (2001) recorded participants' brainwaves as they read sentences that ended either as jokes or with unexpected straight endings (as in Coulson & Kutas, 1998). In an ERP study of the brain response to jokes like those discussed in the previous section, Coulson & Kutas found that ERPs to the joke endings differed in several respects from those to the straight endings, depending on contextual constraint as well as participants' ability to get the jokes. In poor joke comprehenders,

jokes elicited a negativity in the ERPs between 300 and 700 milliseconds after the onset of the disjunct. In good joke comprehenders, high but not low constraint endings elicited a larger N400 (300-500 ms post-onset) than the straights. Also, in this group, both sorts of jokes (high and low constraint) elicited a positivity in the ERP (500-900 ms post-onset) as well as a slow, sustained negativity over left frontal sites. Multiple ERP effects of frame-shifting suggests the processing difficulty associated with joke comprehension involves multiple neural generators operating with slightly different time- courses.

3.4 Summary

Taken together, these three studies of frame-shifting in jokes are far more informative than any one study alone. The self-paced reading time studies suggested that frame-shifting needed for joke comprehension exerts a processing cost that was especially evident in high constraint sentence contexts (Coulson, 1998). The eye movement study of the same stimulus set confirmed that the processing cost of frame-shifting was evident under more natural reading conditions, and replicated the finding that differences between reading times for joke and straight endings were much larger for high constraint sentences. Moreover, the eye movement study suggested that the processing cost was not at the level of word recognition (indexed by the length of a reader's initial fixation of a word), but was related to higher-level processing indexed by the total amount of time spent looking at the word (that is the sum of the time that elapsed during the initial fixation as well as all subsequent fixations). Coulson & Urbach (in prep) also found that people were more likely to make regressive eye movements when they read the joke than the straight endings, suggesting they wanted to re-examine earlier parts of the sentence for clues to which alternative frames should be retrieved.

ERP results from the study by Coulson & Kutas (2001) also suggest the processing cost associated with frame-shifting is related to higher-level processing. In the case of the high constraint jokes, the difficulty includes the lexical integration process indexed by the N400, as well as the processes indexed by the late-developing ERP effects. In the case of the low constraint jokes, the difficulty was confined to the processes indexed by the late-developing ERP effects. The added difference in lexical integration indexed by the N400 may explain why the joke effects on both reading times and gaze durations were more pronounced for high constraint sentences than for low. Because the late developing ERP

effects were only evident for good joke comprehenders who successfully frame-shifted, they are more likely to be direct indices of the semantic and pragmatic reanalysis processes involved in joke comprehension. The temporally extended nature of these effects – lasting at least 400 ms – is also consistent with the idea that they index the construction of the message-level representation.

These experiments suggest that the relationship between a word and its surrounding context is multifold. This relationship involves the ways that individual words add to the cognitive models active in working memory and how individual words can prompt the construction of new models. This view correctly predicts that scenarios which occasion frame-shifting present a challenge to the processor which differs from that presented by lexical violations consistent with the currently active frame. In contrast to the impoverished notion of context in psycholinguistics as something that is important only insofar as it facilitates processes that are clearly linguistic, the difficulty of frame-shifting in jokes demonstrates the need for a model of message-level processing prompted by language. Moreover, it suggests that message-level representations are amenable to fairly substantive changes with minimal linguistic input.

4. Frame-Shifting and Cognitive Science

Researchers in artificial intelligence have long acknowledge the importance of this sort of reanalysis for joke comprehension. Minsky notes the prominence of frame-shifting in many kinds of jokes:

"The element that seems to me most common to all the different kinds of humor is that of unexpected frame-substitution, in which a scene is first described from one viewpoint and then suddenly -- typically by a single word -- one is made to view all the scene-elements in another, quite different, way." (Minsky, 1980)

Indeed, the computational challenge of connecting an initial interpretation to the reinterpretation seems to require essential properties of frames, including the representation of causal and relational information, attribute/value (or slot/filler) organization, and the existence of default values.

Ironically, traditional implementations of frames are completely inadequate for modeling the semantic reorganization involved in frame-shifts. Wilensky (1986), for example has argued that scripts are rigid data structures that cannot accommodate events that are out of the ordinary. While knowledge of typical scenarios represented in scripts and frames is *necessary* for comprehension of many jokes, it is far from *sufficient*. For example, in (14), the word *water* is surprising, not because it is unusual to put water in a swimming pool, but because it would be unusual not to.

(14) Everyone had so much fun diving from the tree into the swimming pool we decided to put in a little water.

Presumably, the swimming pool frame constructed to understand (14) has a **Contains(x)** slot that has been filled by a default value *water*. Interpretation of (14) indeed relies on knowledge of the typical backyard swimming pool. The first clause (“Everyone had so much fun diving from the tree into the swimming pool,”) evokes a model of people having fun diving from a tree into a backyard swimming pool. Moreover, cloze data collected by Coulson & Kutas (1998) suggests the second clause (“we decided to put in a little...”) is initially interpreted as referring to the owner’s decision to install a piece of equipment commonly found near backyard swimming pools that might function in an analogous way to the tree (e.g. a diving board). However, the disjunct *water* prompts the reader to revise a default assumption of the Backyard swimming frame, namely that there was water in the pool. Revising this simple assumption has substantial implications for the consequences of diving from the tree into the pool, and for the mindset of those who enjoy such activities. However, it is unlikely that these implications are represented in generic frames for Backyard Swimming Pools, and less likely that these implications can be logically derived from them.

On the one hand, the interpretation of (14) suggests that in the absence of information to the contrary, readers assume that typical assumptions obtain. However, when given information to the contrary, they are nonetheless able to adapt their models in order to incorporate new information. While a traditional script- or frame- based system can generate a new slot in response to an unexpected event, it is unable to compute the relationship between unexpected and normal events, because its inferencing capacity is based on knowledge represented in the script itself.

Given that one of the main difficulties of previous frame-based systems is in defining criteria for frame-selection (see Allen, 1987), the psychological reality of frame-shifting would seem to pose a serious problem for a frame-based account. In many cases, there is no previously stored frame which can be recruited to relate events to one another. The challenge of frame-shifting is to create a new superframe and to adapt previously created structure accordingly. Given the rigidity of the frame as a data structure (e.g. Allen, 1987; Wilensky, 1986), one might question whether frame-based models can accommodate the demands of frame-shifting.

The need for a sufficiently flexible implementation of frames has driven some researchers to explore the adequacy of sub-symbolic processing in neural networks (see McClelland, 1986 for review). The propensity of these networks to display (a) graceful degradation, viz. arriving at a best guess given imperfect information, (b) spontaneous generalization, that is, accommodating inputs that do not conform to previously instantiated schemas, and (c) the ability to arrive at a compromise solution to mutual constraint satisfaction problems is compatible with the flexibility people show in their interpretation of jokes.

The early promise of this approach can be seen in a model proposed by Rumelhart and colleagues (Rumelhart, 1986) that classifies rooms in a house based on their contents (e.g. whether they have beds, chairs, refrigerators, and so on). Units in the network represent semantic micro-features, and the weights between units encode correlations between those micro-features. The network is set up to promote excitatory weights between micro-features that co-occur, and inhibitory weights between features that do not. If the network has experienced a high correlation between the mutual activations of stove, refrigerator, and counter, when the *stove* unit is activated, the network (using a gradient descent algorithm) activates correlated micro-features (e.g., refrigerator, counter) until it settles into a kitchen frame.

While the model by Rumelhart and colleagues reveals the flexibility of probabilistic approaches, it is incapable of representing information needed to get the jokes discussed above. This is because it contains no mechanisms for generating the high-level inferences that relate frames to one another. A more sophisticated model by St. John is able to use co-occurrence frequencies in its input to infer default information, and to

modify its predictions about upcoming events in a way that is sensitive to context (St. John, 1992). However, St. John's model is limited in much the same way as symbolic implementations: information that deviates too much from stored frames cannot be accommodated. Because it is unable to compute the relationships between different higher-level representations, St. John's (1992) model is incapable of combining information from different scripts in any sensible way.

Lange & Dyer propose a structured connectionist model called ROBIN (role-binding network) that explicitly attempts to capture inferential revisions in frame-shifting (Lange, 1989). ROBIN uses connections between nodes to encode semantic knowledge represented in a frame type data structure. Each frame has one or more slots, and slots have constraints on the type of fillers to which they can be bound. The relationships between frames are represented by excitatory and inhibitory connections between nodes and the pathways between corresponding slots. Once initial role assignments have been made, ROBIN propagates evidential activation values in order to compute inferences from the information the programmers have given it.

Inference is understood as resulting from the spread of activation across the connections between related frames and competing slot-fillers. For example, connections between frames for Transfer-Inside and Inside-of allow the system to 'infer' **Inside-of (Pizza, Oven)** from **Transfer-Inside(Seana, Pizza, Oven)**. In this model, frame selection is entirely a matter of spreading activation. Because each slot has a number of binding nodes, all of the meanings of an ambiguous word can serve as candidate bindings. Candidate bindings can be activated simultaneously, and the binding node with the greatest evidential activation eventually wins out. Because multiple frames are activated in parallel, contextual information can further activate an already highly activated node (or set of nodes), thus confirming an initial interpretation. Alternatively, contextual information can activate a previously less-active interpretation, thus implementing frame-shifting.

While the use of spreading activation to implement frame-shifting looks promising (as does a related model presented by Shastri, 1993), Lange & Dyer's model is limited by its inability to represent scenarios that require recursion, as well as its inability to learn new concepts. Perhaps more seriously, ROBIN is unable to handle metaphoric or metonymic uses of language, and has limited ability to represent the shading of meaning

found in natural language. Further, while ROBIN can cope with lexical ambiguity, it is completely unable to deal with the semantic indeterminacy that constitutes the main problem of language comprehension.

5. Integration

We examined the use of frames in the context of reading jokes that require semantic and pragmatic reanalysis known as *frame-shifting*. Frame-shifting seems to occur when it is necessary to represent the relationship between two or more objects, actions, or events. If the disjunct, or frame-shifting trigger, cannot be sensibly incorporated into existing structure, the words that served to evoke that structure are reanalyzed to provide a coherent bridge between the initial and the revised representations. The relationship between the disjunct and the connector can be suggested by grammatical clues, conceptual relationships, or a combination of the two.

Of course, the experiments reviewed above do not substantiate all of the theoretical claims we have made. They merely establish that people spend longer reading jokes than straight versions of the same sentences, that jokes prompt more regressive eye movements, that jokes elicit slightly different brainwaves, and that those differences are far more pronounced in people who get jokes than in those who don't. However, these findings *are* consistent with the psychological reality of the frame-shifting process, and in that sense they argue against the pragmatics as disambiguator model so often implicitly assumed by psycholinguists. Clearly, the role of background knowledge goes way beyond that of filling in gaps as the integration of current experience with background knowledge is the essence of comprehension.

The appeal of the frame is that the same type of data structure can be used to represent information at the level of words, sentences, texts, and even grammatical constructions (Fillmore, 1968; Fillmore, 1976; Fillmore, 1982; Fillmore, 1988). Interestingly, none of the frame-based models reviewed in section 4 could adequately account for the reanalysis needed to understand the simple jokes used in our experiments. Regardless of whether these were classical symbolic models or whether they were connectionist models that detect statistical patterns, the problem was similar. In these models, a given frame enables inference about elements within that frame, and, in the case of ROBIN (Lange & Dyer, 1989), even enables inferences about closely related

frames. What these models lack, however, is the capacity to creatively combine frames, to draw inferences that require an understanding of the relationship between frames, and to construct novel frames in response to contextual demands.

In accordance with the frame semantics program, we have argued that linguistic utterances cue the retrieval of abstract grammatical frames which speakers unify with frames evoked by lexical and contextual information. Meaning construction thus consists of constructing a series of simple cognitive models while keeping track of common elements and relations in successive models. But, rather than just retrieving and instantiating frames, speakers are continuously and creatively building and blending cognitive models to yield new concepts and construals. Indeed, the demands of joke comprehension suggest the models we build and revise so quickly derive from perceptual symbols. Perceptual symbols are schematic representations of perceptual experience stored around a common frame that promotes schematized simulations (Barsalou, 1999). We suggest that with frames built from perceptual symbols, one could maintain the representational advantages of hierarchically organized slot-filler structures, as well as explaining how speakers might construct a simulation of a parachute that was used, but never opened, in order to infer the origin of its stain.

6. References

- Allen, J. (1987). *Natural Language Processing*. Philadelphia, PA: Benjamin/Cummings Publication.
- Barsalou, L.W. (1992). Frames, concepts, and conceptual fields. In A.L.E.F. Kittay (Ed.), *Frames, Fields, and Contrasts: New Essays in Lexical and Semantic Organization*. Hillsdale, NJ: Erlbaum.
- Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577-609.
- Clark, H.H. (1983). Making sense of nonce sense. In G.B.F.D.A.a.R.J. Jarvella (Ed.), *The Process of Language Understanding* (pp. 297-332). Chichester: John Wiley & Sons.
- Coulson, S., & Kutas, M. (1998). *Frame-shifting and Sentential Integration* (Cognitive Science Technical Report No. 98.02). UCSD.

- Coulson, S. (2000). *Semantic Leaps: Frame-shifting and Conceptual Blending in Meaning Construction*. Cambridge, UK: Cambridge University Press.
- Coulson, S.U., T. (In Preparation). Joke Comprehension: An Eye Tracking Study. .
- Fauconnier, G. (1994). *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge: Cambridge University Press.
- Fauconnier, G., & Turner, M. (1998). Conceptual integration networks. *Cognitive Science*, 22, 133-187.
- Fauconnier, G., & Turner, M. (2002). *The Way We Think*. New York: Basic Books.
- Fillmore, C.J. (1968). The case for case. In E.B.R.T. Harm (Ed.),, *Universals of Linguistic Theory* (pp. 1-90). New York: Holt, Rinehart, and Winston.
- Fillmore, C.J. (1976). The need for frame semantics within linguistics. *Statistical Methods in Linguistics*, 12(5), 5-29.
- Fillmore, C.J. (1982). Frame Semantics. In L.S.o. Korea (Ed.),, *Linguistics in the Morning Calm* (pp. 111-137). Seoul: Hanshin.
- Fillmore, C.J. (1988). The mechanisms of construction grammar. In A.J. Shelley Axmaker, & Helen Singmaster (Ed.),, *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society* (pp. 35-55). Berkeley: Berkeley Linguistics Society.
- Frazier, L. (1995). Issues in representation in psycholinguistics. In J.M.a.P. Eimas (Ed.),, *Speech, Language, and Communication: Handbook of Perception and Cognition* (pp. 1-27). San Diego: Academic Press.
- Gazdar, G. (1979). *Pragmatics: Implicature, Presupposition and Logical Form*. New York: Academic Press.
- Johnson-Laird, P.N. (1993). The mental representation of the meaning of words. In A.I. Goldman (Ed.),, *Readings in Philosophy and Cognitive Science* (pp. 561-584). Cambridge, MA and London: MIT Press.
- Katz, J.J., & Fodor, J.A. (1963). The structure of a semantic theory. *Language*, 39, 170-210.
- King, J.W., & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7(3), 376-395.

- Kutas, M., & Hillyard, S. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203-205.
- Langacker, R.W. (1987). *Foundations of Cognitive Grammar: Theoretical Prerequisites*. Stanford, CA: Stanford University Press.
- Lange, T.E., & Dyer, M.G. (1989). High-level inferencing in a connectionist network. *Connection Science*, 1, 181-217.
- Mangun, G.R., Hillyard, S.A., & Luck, S.J. (1993). Electrocortical substrates of visual selective attention. In D.E.M.S. Kornblum (Ed.), *Attention and Performance 14: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience* (pp. 219-243). Cambridge, MA: MIT Press.
- McClelland, J.L., & Rumelhart, D.E. (Ed.). (1986). *Explorations in Parallel Distributed Processing: A Handbook of Models, Programs, and Exercises*. Cambridge, MA: MIT Press.
- Minsky, M. (1980). *Jokes and the logic of the cognitive unconscious* (AI Memo No. 603). MIT.
- Minsky, M. (1975). Frame system theory. In P.N.J.-L.a.P.C. Wason (Ed.), *Thinking: Readings in Cognitive Science* (pp. 355-376). Cambridge: Cambridge University Press.
- Norman, D.A. (1974). Some observations on mental models. In D.G.A. Stevens (Ed.), *Mental Models* (pp. 7-14). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rumelhart, D.E., Smolensky, P., McClelland, J.L., & Hinton, G. (1986). Schemata and sequential thought processes in PDP models. In D.E.R.J.L. McClelland (Ed.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Vol. 2, pp. 7-57). Cambridge, MA: MIT Press.
- Sanford, A., & Garrod, S. (1981). *Understanding Written Language: Explorations Beyond the Sentence*. Chichester: John Wiley & Sons.
- Schank, R.C., & Abelson, R.P. (1977). *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Shastri, L., & Ajjanagadde, V. (1993). From simple associations to systematic reasoning: A connectionist representation of rules, variables, and dynamic bindings using temporal asynchrony. *Behavioral & Brain Sciences*, 16(3), 417-494.
- Simpson, G.B. (1994). Context and the processing of ambiguous words. In E. Morton Ann Gernsbacher (Ed.), *Handbook of psycholinguistics*. (pp. xxii, 1174): Academic Press, Inc.
- Spivey-Knowlton, M., & Tannenhaus, M. (1994). Referential context and syntactic ambiguity resolution. In a.L.F. Charles Clifton, and Keith Rayner (Ed.), *Perspectives on Sentence Processing*. Hillsdale, NJ: Lawrence Erlbaum.
- St. John, M.F. (1992). The story gestalt: A model of knowledge intensive processes in text comprehension. *Cognitive Science*, 16, 271-306.
- Wilensky, R. (1986). Points: A theory of the structure of stories in memory. In K.S.J. B.J. Grosz, & B.L. Weber (Ed.), *Readings in Natural Language Processing* (pp. 459-473). Los Altos, CA: Morgan Kaufman Publishing.