Event categorization: A cross-linguistic perspective

Asifa Majid (asifa.majid@mpi.nl)
Max Planck Institute for Psycholinguistics, Postbus 310 Nijmegen, 6500 AH The Netherlands

Miriam van Staden (m.vanstaden@uva.nl)
Department of Theoretical Linguistics, Spuistraat 210, 1012 VT Amsterdam, The Netherlands

James S. Boster (james.boster@uconn.edu)
Department of Anthropology, 354 Mansfield Road Storrs, CT 06269-2176 USA

Melissa Bowerman (melissa.bowerman@mpi.nl)
Max Planck Institute for Psycholinguistics, Postbus 310 Nijmegen, 6500 AH The Netherlands

Abstract

Many studies in cognitive science address how people categorize objects, but there has been comparatively little research on event categorization. This study investigated the categorization of events involving material destruction, such as "cutting" and "breaking". Speakers of 28 typologically, genetically, and areally diverse languages described events shown in a set of video-clips. There was considerable cross-linguistic agreement in the dimensions along which the events were distinguished, but there was variation in the number of categories and the placement of their boundaries.

Introduction

Categorization research in cognitive science has focused overwhelmingly on the mental representation of **objects**. Behavioral studies with adults, neuropsychological studies with patient populations, cross-cultural comparisons, and acquisition evidence provide converging evidence about how objects are represented. For example, objects are stored according to semantic domains, with natural kinds represented distinctly from artifacts. Within these categories there are subdivisions: animals are stored separately from fruits, while musical instruments are stored separately from furniture (Shallice, 1988). Objects are organized not only by semantic domain but also hierarchically, with categories at the superordinate, basic, and subordinate levels (Rosch, 1978). Basic level categories are cognitively privileged, in the sense that they are labeled with shorter words, they constitute the preferred level of naming, they can be verified faster than superordinate and subordinate categories in judgment tasks, and they are acquired earlier by children (Brown, 1958; Rosch et al., 1976). There also appears to be considerable cross-cultural consensus in the organization of object representations (Berlin, 1992; Malt, 1995).

In contrast to all the work on objects, relatively little has been done on the mental representation of **events**. One line of research, with roots in social psychology, has investigated how people segment events (Newtson & Engquist, 1976; Newtson, Engquist, & Bois, 1977; Zacks et al., 2001). Another important line of work on event representation, originating in cognitive psychology and artificial intelligence, has examined the organization of event knowledge in scripts, frames, and schemas (Minsky, 1975; Schank & Abelson, 1977).

Neither of these approaches to event representation has examined how everyday activity types are categorized. Studies of event segmentation do not ask which event segments are regarded as being "of the same kind". Script and frame research concentrates on scenarios like "going to the movies", "going to a restaurant", "sports", or "housework" (Morris & Murphy, 1990; Rifkin, 1985). These scenarios are often culture-specific, and so do not lend themselves to cross-cultural research. They are also complex, consisting of sequences of finer-grained events such as "walking into the restaurant", "sitting down", "ordering", "eating", and "paying the bill". Little is known about how uniformly people categorize such finer-grained units, but it has been widely assumed – certainly by developmentalists – that there is a universal core set of everyday event types and that children learn basic verbs such as *have, hit, move, put*, and *give* by linking them directly to these concepts (Gleitman, 1990; Pinker, 1989).

In the present study, we focus on the linguistic categorization of a set of everyday events of "cutting and breaking" – more formally known as events involving a "separation in the material integrity" of objects (Hale & Keyser, 1987).[1] This domain was chosen because such events are universal and do not rely on specialized knowledge; they are accessible to everyone. The manufacture and use of tools for purposes of cutting and breaking has been dated back to at least 2.5 million years ago in the East African Rift area. Modern humans (*homo sapiens sapiens*) appear to be distinctive for making and using particular tools for "cutting", such as pressure-flaked knives (Toth & Schick, 1993). "Cutting" and "breaking" can, then, be taken as human activities that are central to human language and cognition.

We examine the categorization of "cutting and breaking" events by looking at how speakers of

---

[1] The terms "cutting" and "breaking", with quotes, designate actions of the type that speakers of English typically label with verbs like *cut* and *break*; other languages may or may not have words with closely similar meanings. Throughout this paper, words in quotation marks point to actions of a certain general type, and words in italics designate linguistic forms.

genetically, typologically, and areally diverse languages describe a set of actions shown in video-clips. Do speakers of all languages make the same distinctions when they are talking about such events?

The verbs c*ut* and *break* have been widely discussed in the linguistics literature. One influential approach has suggested that "cutting"-type verbs and "breaking"-type verbs can be universally distinguished on the basis of their semantic and syntactic behavior (Guerssel et al., 1985). This suggests that speakers of different languages should recognize similar distinctions.

Other work, however, suggests that there may be significant differences in the way languages categorize "cutting" and "breaking" events; for example, English speakers use *break* for actions on a wide range of objects (e.g., a plate, a stick, a rope), while speakers of K'iche' Maya must choose from among a set of "breaking" verbs on the basis of properties of the object; e.g., *-paxi:j* 'break a rock, glass, or clay thing' (e.g., a plate); *-q'upi:j* 'break (other kinds of) hard thing' (e.g., a stick); *-tóqopi'j* 'break a long flexible thing' (e.g., a rope) (Pye, 1996; Pye, Loeb, & Pao, 1995). Differences in the categorization of "cutting and breaking" events might also be expected due to variation in cultural tools and techniques; for example, Americans and Europeans chop vegetables by holding them still and bringing a knife down on them from above, whereas Punjabi speakers in rural Pakistan and India often move the vegetables against a stationary curved knife.

In studying the categorization of "cutting and breaking" events, it is not obvious a priori what the domain of investigation should be taken to encompass. Whereas speakers of English do not use *cut* and *break* for actions like peeling a banana or pulling paper cups apart, and they do not use *open* for events like breaking the stem off an apple, perhaps such categorizations occur in other languages. Children learning English in fact make such overextensions (Bowerman, in press; Schaefer, 1979), which suggests that the boundaries of the "cutting and breaking" domain may not be cognitively obvious, and therefore not universally shared. One important goal for the present study, then, is not only to examine the categorization of "cutting and breaking" events by speakers of different languages, but also to discover the extent to which "cutting and breaking" events hang together in the first place as a relatively coherent semantic domain, as distinct from events involving other kinds of separations.

Method

Participants

Event descriptions were collected from speakers of 28 typologically, genetically and areally diverse languages. For each language there were between one and seven consultants. Twenty researchers collaborated in this effort, all of them experts on the language they worked on – a critical point for the validity of the coding of the data (see Results section). Data collection was carried out in the language being studied, not a contact language. Details of the languages, language affiliations, and researchers responsible for the collection and coding of the data are given in Table 1.

Materials

The data were collected using a set of 61 video-clips that depicted a wide range of events (Bohnemeyer, Bowerman, & Brown 2001). The majority of these clips showed an event in which an actor brought about a change of state in an object – specifically, some kind of destruction of the object's material integrity. Some clips depicted state-change events that involved separation but not material destruction, such as opening a pot or pulling paper cups apart. Still others depicted "peeling" events, which share properties with events of both material destruction and simple separation. Stimuli were constructed by varying the agent, the instrument used, the object acted upon, the manner of the destruction, and the prototypicality of the event (see Figure 1).

Figure 1: Example stills from video clips

Procedure

Consultants saw one video-clip at a time on a laptop. The clips were presented in a fixed order. The consultants' task was to describe what the agent did. After free description they were asked what other descriptions could be applied felicitously to each clip. They were also asked whether other descriptions would be infelicitous.

Results

Coding

We defined the target event we were interested in as the change in an object from a state of integrity to a state of separation or material destruction. For each of the languages, the researcher who collected the data identified those constituent(s) of a speaker's description which

Table 1: Language details and associated researchers

Language	Language affiliation	Country	Researcher
Biak	Austronesian	Indonesia	W. van de Heuvel
Chontal	Isolate	Mexico	L. O'Connor
Dutch	Indo-European	Netherlands	M. van Staden
English	Indo-European	UK, USA	M. Bowerman, A. Majid, C. Wortmann
Ewe	Niger-Congo	Ghana	F. Ameka
German	Indo-European	Germany	M. van Staden
Hindi	Indo-European	India	B. Narasimhan
Jalonke	Niger-Congo	Guinea	F. Lüpke
Japanese	Isolate	Japan	S. Kita
Kilivila	Austronesian	Papua New Guinea	G. Senft
Lao	Tai	Laos	N. Enfield
Likpe	Niger-Congo	Ghana	F. Ameka
Mandarin	Sino-Tibetan	China	J. Chen
Miraña	Witotoan	Colombia	F. Seifart
Otomi	Otomanguean	Mexico	E. Palancar
Punjabi	Indo-European	Pakistan	A. Majid
Spanish	Indo-European	Spain, Mexico	M. Bowerman, E. Palancar
Sranan	Creole	Surinam	J. Essegbey
Swedish	Indo-European	Sweden	M. Gullberg
Tamil	Dravidian	India	B. Narasimhan
Kuuk Thaayorre	Pama-Nyungan	Australia	A. Gaby
Tidore	West Papuan Phylum	Indonesia	M. van Staden
Tiriyó	Cariban	Brazil	S. Meira
Touo	Papuan Isolate	Solomon Islands	M. Dunn, A. Terrill
Turkish	Altaic	Turkey	A. Özyürek
Tzeltal	Mayan	Mexico	P. Brown
Yélî Dyne	Papuan Isolate	Rossel Island	S. Levinson
Yukatek	Mayan	Mexico	J. Bohnemeyer

encoded the event. For example, the event of "a boy cutting a carrot", at the top left of Figure 1, can be expressed in English as *The boy cut the carrot*. Here the caused state-change event is expressed solely by the transitive verb *cut*.

Languages differ in whether information about the state change is typically located in a single verb or is spread out across a number of constituents, such as additional verbs or particles. For example, speakers of Mandarin use verb compounds to describe many of the events; e.g., *qie1-duan4* 'cut-break.long.thin.object' for the scene of someone karate-chopping a carrot shown in the lower left corner of Figure 1. For purposes of the present study, we concentrated on how the stimuli were categorized by the **verbs** of a language. Every verb in the data that described the target event was input to the analysis.

Analysis

Speakers' event descriptions can be treated as analogous to the data obtained in sorting tasks designed to study categorization. In a typical sorting task, a subject might receive a set of cards, each depicting a different stimulus, and be asked to sort them into piles of objects that are similar. Speakers in the present study received no metalinguistic instructions; they were simply asked to describe what they saw in the video-clips. But each

different verb they applied to the target events was taken to define a category ("pile"). Across languages (and of course also within individuals or across individuals within the same language), stimuli that are often described with the same verb ("are sorted into the same pile") can be taken to be more similar to each other than stimuli that typically fall under different verbs (Bowerman, 1996). Multivariate statistics can then be used to explore the similarity structure of the data set as a whole.

To extract the most important dimensions organizing the similarity space of our stimuli, we used correspondence analysis (Greenacre, 1984). Correspondence analysis provides a dual factoring of a rectangular matrix in which the column scores and row scores are projected into the same low dimensional space. To perform the correspondence analysis, we first transformed the linguistic data for each language into a similarity matrix. This was done by determining, for all scenes taken pairwise, whether each member of the pair was ever described by the same verb. If so, the pair was assigned a similarity score of one; if not, zero.[2]

---

[2] This technique was adopted rather than a more graded approach to similarity based on the number of speakers within each language who used the same description, so as not to bias the results toward the categorizations favored by languages for which we happened to have more speakers.

The similarity matrices from all the languages were then stacked one on top of another to build a matrix with 61 columns (the stimuli) and 28*61 (language*stimuli) rows. This matrix was submitted to correspondence analysis to find the dimensions that are cross-linguistically the most important in structuring the similarity space of the stimulus set. The analysis extracts first the dimension that accounts for the most variance, then the dimension that accounts for the next most variance, and so on. Each stimulus scene is positioned in this multidimensional space in such a way that the distance between any two scenes reflects the degree to which, across languages, people described them with the same verbs. Scenes often described with the same verb are positioned close together, while scenes that are rarely or never described with the same verb are positioned far apart.

The major dimensions

The first and most important dimensions extracted in our analysis distinguished between events of material destruction and other events involving separation. There was widespread consensus across languages that events of "taking apart" (e.g., separating paper cups), "opening" (e.g., opening a box) and "peeling" (fruit) should be described with different verbs than events of "cutting and breaking". "Cutting and breaking" events are distinguished as a group from other kinds of separation, and so form a coherent semantic domain.

Leaving aside the events of "taking apart", "opening", and "peeling", we next focused specifically on the similarity structure of the remaining 46 events. These stimuli were analyzed with the same procedure outlined in the previous section.
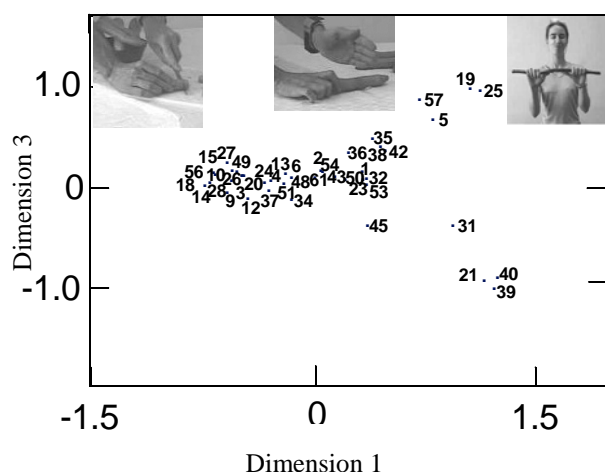

Figure 1: Plot of scenes, based on all languages, along Dimensions 1 and 3. Dimension 1 distinguishes events with precise control over the locus of separation (cutting a carrot with a knife) from scenes with intermediate control (karate-chopping a carrot) and imprecise control (breaking a stick with the hands).

The first and most important dimension of this analysis distinguishes among events on the basis of how precisely the agent controls the locus of the separation in the object. The events are distributed continuously on this dimension. (See Figure 1 for the placement of the scenes along Dimension 1. Each scene is represented by a number.) Events involving relatively precise control (e.g., cutting a carrot with a knife, scene 10) is positioned to the left, events with imprecise control (e.g., breaking a stick with the hands, scene 19) to the right, and events with intermediate degrees of control (e.g., karate-chopping a carrot, scene 32) in between. Events intermediate on this dimension are treated variably across languages, with some languages grouping them with the "precise control" events positioned to the left, others with the "imprecise control" events positioned to the right, and still others assigning them to categories of their own.

Dimension 2 distinguishes just two scenes from the rest – those showing an agent tearing a piece of cloth (a two-dimensional flexible object) partially (scene 36) or completely (scene 1) with the hands. These events were labeled *tear* in English, as distinct from *cut* and *break*. Nineteen out of the 28 languages have a verb that was used to categorise these and only these scenes. The remaining 9 languages did not distinguish these scenes, but grouped them in various ways with other scenes.

Within the group of scenes pulled out on Dimension 1 as lacking precise control over the locus of separation, Dimension 3 makes a further distinction between "snapping" and "smashing" events (see Figure 2a). The "snapping" cluster comprises events in which a one-dimensional rigid object is separated into two pieces by applying pressure to both ends (scenes 25, 19, 57, 5), while the "smashing" cluster is made up of events in which a rigid object is fragmented into many pieces by applying a blow, e.g., with a hammer (40, 39, 21, 31). The Dimension 3 distinction between "snapping" and "smashing", like the Dimension 2 distinction between "tearing" and separations of other kinds, is respected by speakers of many languages – cf. the distinction in Likpe between events described with *f3s3* (the snapping scenes) and those described with *ba* (the smashing scenes) (see Figure 2b). But this distinction is not made in all languages; colloquial Tamil, for example, collapses these two categories (along with a few additional scenes) into a single event type, denoted by the verb *oDai* (see Figure 2c).

Discussion

Speakers of a variety of typologically, genetically and areally diverse languages agree to a surprising extent in their linguistic categorization of events of material destruction of objects ("cutting and breaking" events). First, they agree on treating such events as a relatively coherent semantic domain. A priori, it is not obvious that languages will distinguish "cutting and breaking" events as a group from events involving other kinds of separations of objects or object parts, such as "taking apart", "opening", and "peeling"; after all, learners of English make a number of errors suggesting that the
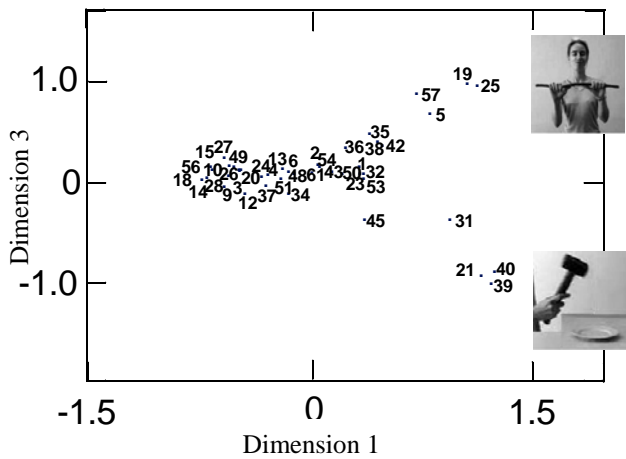
Figure 2a: Plot of scenes, based on all languages, along Dimensions 1 and 3, showing the distinction between "snapping" and "smashing" events.
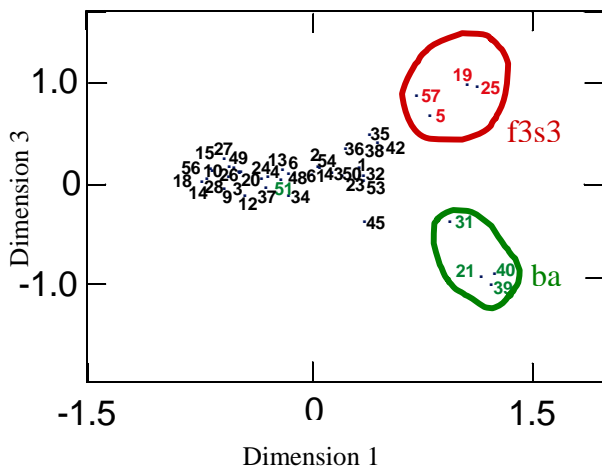

Figure 2b: Likpe is a good example of a language which distinguishes "snapping" from "smashing" events.
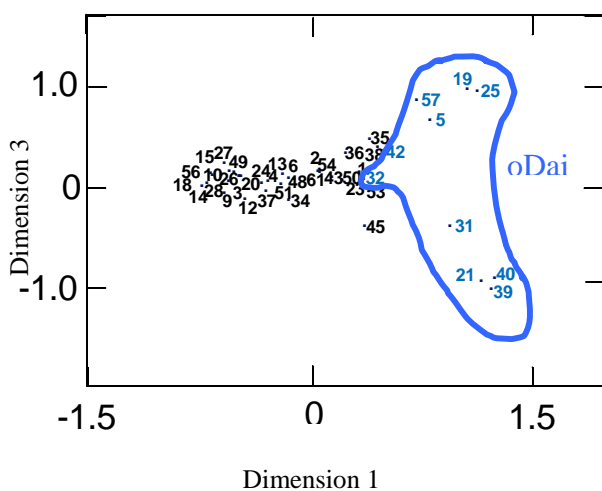

Figure 2c: Tamil collapses the "snap-smash" distinction.

boundaries of these event types are not obvious. For this reason our set of events to be described included not only scenes of "cutting and breaking", but also of various other kinds of separations. But these other separations were rarely described with the same verbs that were applied to the core set of "cutting and breaking" events. The "cutting and breaking" events were treated as far more similar to each other than they were to the other kinds of separations, in the sense that they were much more often described by the same verbs.

Second, speakers of different languages also showed considerable agreement in the kinds of distinctions they drew **within** the domain of "cutting and breaking" events. Although their societies ranged from industrial urban-dwelling to rainforest-dwelling swidden agriculturist, and they varied in their tools and techniques for cutting and breaking things in their daily lives, they converged on a shared similarity space for events of "cutting and breaking". The most important dimension for the set of 28 languages taken as a group distinguishes events featuring precise control over the locus of separation from those with imprecise control (roughly, "cutting" events vs. "breaking" events). Further, "tearing" events are very often distinguished from among other events with an intermediate degree of control (Dimension 2), while "snapping" and "smashing" events are often distinguished among the events involving imprecise control (Dimension 3).

Despite this cross-linguistic agreement there were also many differences – language-learners clearly have something to learn. Speakers of different languages varied in the number of categories of "cutting and breaking" they recognized and in where they placed the category boundaries. For example, speakers of most of the languages respected the distinction between "tearing" and other actions of material destruction, but some did not; speakers of many languages rigorously distinguished between actions of "snapping" and "smashing", but some did not (see Figures 2a-c); and languages differed in where they placed the boundary between "precisely" and "imprecisely" controlled acts of separation. These differences respected the overall structure of the semantic space; for example, no speakers described events at the far left of Dimension 1 with the same verb(s) as events at the far right, while describing the events falling between them with different verbs.

One topic we have not yet mentioned is how a language's semantic categories of "cutting and breaking" are related to one another. For instance, English clearly organizes its "cutting and breaking" terms hierarchically, with the high-frequency verbs *break* and *cut* each encompassing a number of more specific subtypes, such as *snapping* and *smashing* for *break*, and *slicing* and *chopping* for *cut*. This kind of organization is less apparent in many of the other languages in our sample. For example, Dutch has no verbs for "cutting and breaking" with as wide an application as English *cut* and *break*. "Cutting" events are obligatorily subdivided according to whether they involve a single-bladed tool like a knife or a double-bladed tool like scissors (*snijden* vs. *knippen*), and there is also no cover term for a wide range of "breaking" events; e.g., *breken* – cognate with English *break* – is used only for "snapping" events. It is unclear, then, whether the hierarchical organization found

across languages in words for **objects** will also be characteristic of words for **events**.

A final topic that we also leave to future work is the intriguing question of how the categorization of events imposed by language is related to categorization as studied with nonlinguistic techniques such as similarity ratings. For the object domain of "containers", speakers of different languages classified nonlinguistically more similarly than they classified linguistically (Malt et al., 1999). Whether the same will be true for event categories remains to be seen.

Acknowledgments

References

Berlin, B. (1992). *Ethnobiological classification: Principles of categorization of plants and animals in traditional societies*. Princeton, NJ: Princeton University Press.

Bohnemeyer, J., Bowerman, M., & Brown, P. (2001). Cut and break clips, version 3. In S. C. Levinson & N. Enfield (Eds.), *Field Manual 2001*. Language & Cognition Group, Max Planck Institute for Psycholinguistics.

Brown, R. (1958). How shall a thing be called? *Psychological Review*, *65*, 14-21.

Bowerman, M. (1996). Learning how to structure space for language: A crosslinguistic perspective. In P. Bloom, M. Peterson, L. Nadel, & M. Garrett (Eds.) *Language and Space*. Cambridge MA: MIT Press.

Bowerman, M. (in press). Why can't you 'open' a nut or 'break' a cooked noodle? Learning covert object categories in action word meanings. In L. Gershkoff-Stowe & D. Rakison (Eds.), *Building object categories in developmental time*. Mahwah, NJ: Lawrence Erlbaum.

Gleitman, L. (1990). The structural sources of verb meanings. *Language acquisition 1,* 3-55.

Greenacre, M. J. (1984). *Theory and applications of correspondence analysis*. Academic Press: London.

Guerssel, M., Hale, K., Laughren, M., Levin, B., White Eagle, J. (1985). A cross-linguistic study of transitivity alternations. In W. H. Eilfort, P. D. Kroeber, & K. L. Peterson (Eds.), *Papers from the Parasession on Causatives and Agentivity at the Twenty-First Regional Meeting*. Chicago, IL: Chicago Linguistics Society.

Hale, K., & Keyser, S. J. (1987). *A view from the middle*. Lexicon Project Working Papers 10. Cambridge, MA: MIT, Center for Cognitive Science.

Malt, B. C. (1995). Category coherence in cross-cultural perspective. *Cognitive psychology*, *29*, 85-148.

Malt, B. C., Sloman, S. A., Gennari, S., Shi, M. Y., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language*, *40*, 230-262.

Minsky, M. (1975). A framework for representing knowledge. In P. Winston (Ed.), *The psychology of computer vision.* New York: McGraw-Hill.

Morris, M. W., & Murphy, G. L. (1990). Converging operations on a basic level in event taxonomies. *Memory & Cognition*, *18*, 407-418.

Newtson, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology*, *12*, 436-450.

Newtson, D. & Engquist, G., & Bois, J. (1977). The objective basis of behavior units. *Journal of Personality and Social Psychology*, *35*, 847-862.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.

Pye, C. (1996). K'iche' Maya verbs of breaking and cutting. *Kansas Working Papers in Linguistics, 21* (part II).

Pye, C., Loeb, D. F., & Pao, Y.-Y. (1995). The acquisition of breaking and cutting. In E. Clark (ed.), *Proceedings of the Twenty-seventh Annual Child Language Research Forum*. Stanford: CSLI Publications.

Rifkin, A. (1985). Evidence for basic level in event taxonomies. *Memory & Cognition*, *13*, 538-556.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B.B. Lloyd (Eds.) *Semantic factors in cognition*. Hillsdale, NJ: Lawrence Erlbaum.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *7*, 573-605.

Schaefer, R. (1979). Child and adult verb categories. *Kansas Working Papers in Linguistics, 4,* 61-76.

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.

Schank, R. C. & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures.* Hillsdale, NJ: Lawrence Erlbaum.

Toth, N. & Schick, K. (1993). Early stone industries and inferences regarding language and cognition. In K. R. Gibson & T. Ingold (Eds.) *Tools, language and cognition in human evolution*. Cambridge: Cambridge University Press.

Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., Buckner, R. L. & Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience*, *4*, 651-655.