

An Argument about the Composition  
of Conceptual Structure

Ray Jackendoff  
Brandeis University

In order for people to be able to talk about what they perceive, there must be a level of mental representation at which information conveyed by language is compatible with information from sensory systems such as vision, nonverbal audition, touch, and so forth. I will call this level conceptual structure. Though the existence of conceptual structure has been more or less taken for granted (especially by the AI community), the need to consider it seriously has been brought to the attention of linguists rather recently, by such works as Fodor (1975) and Miller and Johnson-Laird (1976). This paper will present a combination of linguistic and visual evidence which bears on the nature of conceptual structure.

1. General properties of a theory of conceptual structure

A linguist's questions about conceptual structure can be separated into two major issues. The first, which the linguist shares with many branches of psychology, concerns the form of conceptual structure itself; the second, particular to linguistics, concerns the mapping between conceptual structure and syntactic structure. An answer to the first question, within the theoretical paradigm I will assume, will consist of a set of well-formedness rules for conceptual structure. The second will be answered by a set of correspondence rules which relate some subset of conceptual structures (the verbally expressible concepts) into some subset of syntactic structures (the meaningful sentences).

It seems reasonable to assume for a first approximation that the well-formedness rules for conceptual structure are universal and innate, i.e. that everyone is born with the capacity to develop the same concepts. However, the actual concepts that one does develop will depend to some extent on experience--including possibly linguistic experience, so there is room for a certain amount of "Whorfian" variation if necessary.

On the other hand, this position is not consistent with what I gather is the strongest version of Piagetian developmental theory, which could be construed in the present framework as a claim that certain conceptual well-

formedness rules must be learned. Rather, the development of the child's conceptual ability must be attributed to increasing richness and interconnection of concepts, or to growth either in the well-formedness rules or in computational capacity, over which the child and the environment have little or no control. (The kind of growth I have in mind here is akin to the growth of bones and muscles: the environment must furnish nourishment, but it hardly can be said to control the interesting aspects of structure. See Chomsky (1975) for discussion.)

In addition to the assumption of universality and innateness of the conceptual well-formedness rules, I will make three other assumptions about the theory of conceptual structure and its relation to language. First, a theory of conceptual structure must be observationally adequate: as the level linking language and other perceptual systems, it must at least be able to express all the conceptual distinctions made by natural language. In practice, this calls for an attempt to account for a lexically and grammatically significant fragment of the language without artificial assumptions about the semantics (such as restriction to a microworld).

Second, a theory of conceptual structure must provide some principled way for the meanings of the parts of a sentence to be combined into the meaning of the whole sentence. This requirement of compositionality may be taken more or less strongly, depending on whether or not one requires each syntactic constituent (or even each word) of a sentence to correspond to a well-formed concept.

Related to this second assumption is a third, that the correspondence rules relating syntax and conceptual structure be relatively simple. As motivation for this constraint, we observe that the language learner must relate syntactic form to understood meanings--in fact he must probably learn many aspects of syntactic form in part by figuring out from context what meaning is intended by other speakers. Since syntactic form varies from language to language, the correspondence rules must be at least partly learned. In order to be able to explain how the child manages to acquire language, we should strive for a theory in which at least the language-particular part of the correspondence rules is fairly straightforward.

A second argument for the simplicity of correspondence rules is more heuristic. Language is, after all, an information transmission system, conceptual structure being the information which language conveys. It would be perverse not to take as a working assumption that language is a relatively efficient and accurate encoding of the information it conveys--despite generations of philosophers who have assured us that language is impossibly unsystematic and vague. To give up this assumption is to refuse to look for principles in natural language semantics. Accepting it entails that all deviations from efficient encoding be rigorously justified; what appears to be a quirky relationship between syntax and conceptual structure may turn out to be merely a bad theory of conceptual structure. (See Goldsmith and Woisetschlaeger (1976), Jackendoff (1978) for arguments to this effect.)

2. The argument: Figure formation and pragmatic anaphora

The preliminaries complete, we turn to the main argument, which begins with a discussion of one aspect of visual perception before turning to linguistic matters. We then will draw consequences for conceptual structure, where visual and linguistic information interact.

One of the most important and well-studied phenomena of visual perception is the emergence of a figure against a background. Intuitively, the figure is what attention is directed to; coherence or "thingness" inheres in the figure. It is often reported that the figure seems to stand out from the ground or to be imbued with greater vividness than the ground. The study of the figure-ground opposition has been one of the major preoccupations of the school of gestalt psychology (see e.g. K hler (1947), Koffka (1935)).

For a simple and hopefully illustrative example, consider the contents of this page, in particular the geometric configuration in Figure 1.

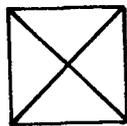


Figure 1

The whole of this configuration can form a visual figure seen against the background of the page. Parts of it can also emerge spontaneously as figures; probably the most prominent are a square and an X, each of which can seen against the rest of the page (including the rest of the configuration) as background. Among less natural figures, which emerge only with more deliberate effort from Figure 1, are such configurations as these (in order of decreasing salience, from left to right):



Figure 2

A number of important observations can be drawn from this simple example.

1. The number of possible figures that can be perceived in a given configuration is very large, perhaps unlimited; however,
2. Only a small number of these are particularly salient.
3. Relative salience of perceived figures is a function of both features of the physical signal and properties of the visual system.
4. Features of the visual context can affect relative salience of figures. For example, the configurations in Figure 2 become much more likely to emerge from Figure 1 upon presentation of Figure 2; certain other possible figures (such as other arrow-shaped configurations) undoubtedly become somewhat more salient than before, while certain other possible figures not presented here remain as unlikely as before.
5. Features of the visual signal interacting with the viewer's intention can make certain figures more salient than they otherwise might be. For example, more figures will emerge from Figure 1 if the viewer is instructed to find all possible figures. Similarly, this aspect of figure formation is crucial in children's puzzles which ask the reader to find three rabbits and two bears hidden in the forest, or in the Hirschfeld cartoons in the New York Times in which the reader in on the joke is to find a stipulated number of instances of the configuration NINA. These figures would not emerge at all were it not for the reader's intention to find them.
6. Features of the visual signal interacting with the viewer's knowledge may make certain figures more salient. Someone who has worked with automobile engines will perceive more distinct figures upon looking under the hood than a mechanical novice; a botanists may see a number of distinct plants where a layman sees only a tangled confusion.

It is important to distinguish the conscious decomposition of a figure into parts from the process of figure formation itself. The former is available to awareness, and is in fact governed by principles of figure formation: as was seen in the example above, the perceived parts are themselves figures. On the other hand, the mental processes which bring about the emergence of a figure are themselves not open to awareness, and the aspects of the configuration which are relevant to the character of the perceived figure may have little to do with the intuitive (i.e. conscious) decomposition of the figure. To make this clearer, consider the not atypical example of facial recognition. Though one can recognize and distinguish thousands of faces, one cannot in most cases consciously decompose faces and say specifically what makes them recognizable and different from each other. In present terms, each recognized face forms a remembered figure, but many of the distinctive features of faces do not themselves form figures and hence are not available to conscious awareness (see Carey (1976), and also Helmholtz's (1885) brief but insightful remarks (p. 369)).

The observations we have made about the figure-

ground phenomenon have a bearing on the nature of conceptual structure. They suggest that there is a privileged set of conceptual structures which encode figures as unitary entities, and that this set is somehow related to conscious awareness. As far as consciousness is concerned, there is no representation of how figures are composed; though to other, unconscious, processes the composition of figures is accessible. Let us call this set of structures figural expressions.

We now relate figural expression in conceptual structure to language. Suppose someone points and simultaneously utters (1).

(1) I bought that yesterday.

What must the hearer do to fully understand the speaker? He must of course understand the words and the syntactic structure, and be able to use the correspondence rules involved in interpreting the sentence; but he also must interpret the word that. In this particular utterance, that is a case of what Hankamer and Sag (1976) call "pragmatic anaphora." In order to understand the intended referent of a pragmatically controlled pronoun like that in (1), the hearer must pick something out of his visual field, perhaps aided by the speaker's pointing gesture.

To make clearer the process of interpreting pragmatic anaphora, consider an example where no figure emerges which can correspond to a pragmatically controlled pronoun. Suppose speaker A utter (1) and points to a blurry photograph: "I bought that yesterday--isn't it gorgeous?" Speaker B, unable to make out anything in the picture, doesn't fully understand the utterance and responds "What are you talking about?" Suppose A then says, "That boat!" B peers at the picture and sure enough the figure of a boat emerges. He has a minor aha-experience: "Oh, that! How could I miss it?" He now has received the message and discourse can continue.

Every reader has probably had an experience like this; its relevance in the present setting is as follows: in order for a pragmatically controlled pronoun to be understood, its intended referent must emerge as a figure in the mind of the hearer, that is, it must have a representation as a figural expression in conceptual structure. Thus we have established an important connection between the figure-ground phenomenon and pragmatic anaphora.

So far we have dealt only with figures that correspond to things (or their shapes). By and large this has been the kind of figure that has been investigated. But as Hankamer and Sag (1976) point out, there are many sorts of pragmatic anaphora.

- (2) Here and there:  
Your coat is here (pointing) and your hat is there (pointing).
- (3) Do it:  
(Hankamer attempts to stuff a 9" ball through a 6" hoop)  
Sag: 'It's not clear you'll be able to do it.
- (4) It happen:  
That (pointing) better not happen again.

- (5) Nominal identity-of-sense anaphora:  
a. (Sag produces an apple)  
Hankamer: Did you bring one for me?  
b. Those (pointing, e.g. to a (single) Cadillac) are expensive.

(6) Manner adverb:  
You shuffle cards { thus } (demonstrating)  
                                  so }  
                                  this way }

(7) Measure phrase:  
The fish that got away was { this }  
  that }  
  yay }  
(demonstrating) long.

There was more than that much (pointing) in the jar when I left.

The same conditions hold on the comprehensibility of these sorts of pragmatic anaphora as on that in (1). For example, if the hearer is unable to see or figure out what goes on the speaker is pointing at in (4), he will not fully understand the utterance (in the sense of having received all the information he is intended to receive).

Given that the existence of an appropriate figural expression in conceptual structure, supplied by the visual system, is necessary for the comprehension of pragmatic that-anaphora in (1), we must conclude similarly that a figural expression is necessary for all the sorts of pragmatic anaphora in (2)-(7) as well. But from the selectional restrictions involved in these constructions, we see that the figures involved cannot be things or shapes. Rather, each corresponds to a different sort of figure, distinct from things. Roughly, here and there correspond to places; do it to actions; it happen to events; nominal identity-of-sense anaphora to categories or kinds; manner adverbials to manners; and measure phrases to amounts. Each of these types of figures represents a different organization of the visual field than do figures corresponding to physical objects.

The existence of this variety of types of pragmatic anaphora suggests three points. First, the mind has the capacity to form figures of a number of distinct types on the basis of visual perception. Second, conceptual structure can represent such entities as places, actions, events, etc. as figural expressions, and this is why we can talk about them. Third, by the criterion of simplicity of correspondence rules, these entities are conceptually simple, since they correspond to something syntactically simple. More explicitly, that, a maximally simple NP, represents a minimally specified thing in (1), and the visual field is the source of the remaining information about the intended message. Similarly, the other expressions of pragmatic anaphora are maximally simple PPs, VPs, etc., and therefore should likewise correspond to minimally specified entities of the proper type; again, the remainder of the intended message is conveyed through the visual system.

### 3. The alternative to psychological and philosophical reductionism

One might object that all these different types of entities should be reduced by the theory of conceptual structure to concurrences of physical objects over time (a four-dimensional space-time map, for example), and that such entities as places and events should play only a derivative role in linguistic semantics. But such a view generally assumes that the psychological notion thing can be fairly simply correlated with physical objects via patterns of retinal stimulation; and this assumption is patently false. Most of the literature of perception is concerned with how we manage the remarkable feat of construing the world as full of more or less stable things, given constantly shifting patterns of retinal stimulation, and with how the things we see are or are not correlated with actual physical facts. It's not easy.

What seems to me a more productive approach is to abandon the goal of reduction and to claim that the types of entities referred to by the anaphoric expressions in (1)-(7) are all present as primitives of conceptual structure. Formally, this means that the well-formedness rules for conceptual structure must allow for figural expressions which correspond to each type. Furthermore, the well-formedness rules must provide an algebra of relationships among the types: a thing can be in a place, an event may have a certain number of things and places as constituents, some events consist of an action performed by a thing (the agent), and so forth. Under this approach, linguistic semantics is not concerned with reducing out events, places, and so forth, but with clarifying their psychological nature and with showing how they are expressed syntactically and lexically.

If any reduction is to take place, it will be in the theory of perception, which now must explain the relation of retinal (and auditory, etc.) stimuli to event- and place-perception as well as thing-perception. If this view is correct, one would expect these other aspects of perception to have many of the same gestalt properties as thing-perception: dependence on proximity, closure, "good form," and so forth. In fact, the few pieces of work I know of on perception of entities other than things (Michotte (1954) on causation, Jenkins, Wald, and Pittenger (1976) on events, remarks of Kbhler (1947, pp. 89-90) on temporal grouping) do reveal just what we are led to expect. This suggests that there is no fundamental new difficulty for perception in admitting entities other than things into conceptual structure--just more of the old problem of how we perceive anything at all.

The argument of section 2 also explicitly addresses the important philosophical issue of the ontological commitment of natural language semantics: what entities should semantic theory allow language to talk about? The predominant philosophical tradition, modeling itself after mathematics, tries to minimize primitives and axioms. (It is perhaps not insignificant that Frege and Russell, the founders of modern logic, were most deeply concerned with foundations of mathematics.) Thus logicians mostly confine themselves to semantic systems which contain only

things and sets as primitive ontological types. A few, such as Davidson (1967, 1969), have tried to argue that events and actions are necessary as well. But Davidson himself has some qualms about such entities, because criteria of individuation are not easy to define.

The view of language taken here, however, takes the ontological commitment of natural language semantics not to be a question of elegance, but an empirical question: what possible ontological types are psychologically real? The argument presented here, based on the interaction of language and visual experience, provides simple preliminary evidence for a relatively rich ontology. As further justification, one would hope to show that this ontology is necessary on both language-internal and perception-internal grounds independently. (Jackendoff (1978) gives linguistic arguments to justify the notion place.) One would also hope to come to terms with the philosophical traditions concerning reference, which will be drastically affected by the expansion. I am not prepared to reconstruct the entire edifice at this point; some arguments will appear in Jackendoff (in preparation).

### 4. Conclusion

The main argument of this paper combined perceptual and linguistic evidence to show that figural expressions in conceptual structure must include entities of a great number of ontological types. I take this to be a prototype for a novel sort of linguistic argumentation--one that treats descriptive semantics as fundamentally a psychological rather than logical discipline, and which seeks to account for the nature of thought and of human experience through grammatical structure. It is not clear that this is linguistics in the usual sense any more. Rather it is an attempt to use linguistic theory as a tool of cognitive psychology. This seems to me to be a promising way to go.

### References

- Carey, Susan (1976) "A Case Study: Face Recognition," in Explorations in the Biology of Language, Report of the MIT Work Group in the Biology of Language, mimeo, pp. 229-264.
- Chomsky, Noam (1975) Reflections on Language, New York, Pantheon.
- Davidson, Donald (1967) "The Logical Form of Action Sentences," in N. Rescher, ed., The Logic of Decision and Action, Pittsburgh, Univ. of Pittsburgh Press.
- (1969) "The Individuation of Events," in N. Rescher et al., eds., Essays in Honor of Carl G. Hempel, Dordrecht, Reidel.
- Fodor, Jerry (1975) The Language of Thought, New York, Crowell.
- Goldsmith, John, and Erich Woisetschlaeger (1976) "The Logic of the Progressive Aspect," Bloomington, Indiana University Linguistics Club.

Hankamer, Jorge, and Ivan Sag (1976) "Deep and Surface Anaphora," Linguistic Inquiry 7.3, 391-428.

Helmholtz, Hermann (1885) On the Sensations of Tone, New York, Dover reprint, 1954.

Jackendoff, Ray (1978) "Grammar as Evidence for Conceptual Structure," in M. Halle, J. Bresnan, and G. Miller, eds., Linguistic Theory and Psychological Reality, Cambridge, MIT Press.

----- (in preparation) Semantics and Cognition.

Jenkins, James J., Jerry Wald, and John B. Pittenger (1976) "Apprehending Pictorial Events: An Instance of Psychological Cohesion," in Minnesota Studies in the Philosophy of Science, Vol. 9.

Koffka, Kurt (1935) Principles of Gestalt Psychology, New York, Harcourt, Brace.

Köhler, Wolfgang (1947) Gestalt Psychology, 2d ed., New York, Liveright; Mentor Books reprint.

Michotte, A. (1954) La Perception de la Causalité, 2d ed., Louvain, Publications Universitaires Louvain.

Miller, George, and Philip Johnson-Laird (1976) Language and Perception, Cambridge, Harvard University Press.

Wittgenstein, Ludwig (1953) Philosophical Investigations, Oxford, Blackwell.

---

This research was supported in part by a Fellowship for Independent Study and Research from the National Endowment for the Humanities.