

J. Pustejovsky S. Bergler (Eds.)

Lexical Semantics and Knowledge Representation

First SIGLEX Workshop
Berkeley, CA, USA, June 17, 1991
Proceedings

Springer-Verlag

Berlin Heidelberg New York
London Paris Tokyo
Hong Kong Barcelona
Budapest

Chapter 18

The Autonomy of Shallow Lexical Knowledge

Kathleen Dahlgren

Abstract

The question of what is “purely linguistic” is considered in relation to the problem of modularity. A model is proposed in which parsing has access to world knowledge, and both contribute to the construction of a discourse model. The lexical semantic theory of naive semantics, which identifies word meanings with naive theories, and its use in computational text interpretation, demonstrate that a shallow, constrained layer of knowledge which is linguistic can be identified.

18.1 Introduction

My work has involved both the development of a theory of lexical semantic representation and the implementation of a computational text understanding system which uses this lexicon to interpret English text (Dahlgren, 1976, Dahlgren, 1988, Dahlgren, McDowell and Stabler, 1989). The question posed by the workshop is a theoretical one: is there any justification for distinguishing between lexical semantics and world knowledge? My affirmative answer is based upon both theory and practice. Section I addresses the theoretical issues and makes a claim, and Section II supports the claim with results from computational linguistics. The experience of developing a wide-coverage natural language understanding system suggests a methodology for distinguishing between lexical knowledge and knowledge in general.

18.2 Theoretical Considerations

The first question to consider is where and how the line is drawn between linguistic knowledge and knowledge in general. Arguments from both psychology and theoretical linguistics are relevant here. Cognitive psychologist Fodor (1987) argues that linguistic knowledge is contained in a module which automatically (deterministically) processes the speech input and produces a logic form without any access to other cognitive components. In this modularity hypothesis, other cognitive components have no access to the intermediate representations produced in the linguistic module. Its formal properties are thought to be the result of innate features of human intelligence. It operates in a mandatory way without access to "central processes." It does not "grade off insensibly into inference and appreciation of context" (1987). One of the chief arguments for the autonomy of the linguistic model is the speed with which it must operate. In contrast, inference is assumed to be open-ended and time-consuming.

While many agree that syntax has special formal properties which indicate the existence of a distinct module which operates under the constraints reflected in these formal properties, its autonomy is controversial. In particular, Marslen-Wilson and Tyler (1987) have conducted a number of experiments to show that the production of a discourse model which may require pragmatic knowledge is just as fast as the production of logical form, and that there is influence from the discourse level on syntactic choice. The experiment showing the latter point involved a contrast between adjective and gerund readings of expressions like "visiting relatives". They had subjects read texts with either an adjectival

or gerund bias to establish a context, as in:

Adjectival Bias: If you want a cheap holiday, visiting relatives . . .

Gerund Bias: If you have a spare bedroom, visiting relatives . . .

They gave subjects a probe of "is" or "are" in these contexts. Response times for "is" were significantly slower in the gerund bias context, and response times for "are" were significantly slower in the adjectival bias case. Thus there are context effects at the earliest point that can be measured. Since central processes are already demonstrably active during processing of the third word in these ambiguous clauses, Marslen-Wilson and Tyler argue that the empirical facts are just as consistent with a model which "allows continuous or even predictive interactions between levels" as they are with a hypothesis in which an encapsulated module produces all possible readings for a higher-level module which sorts them out using inference. They suggest a model in which an independent syntactic processor and an inferencing mechanism contribute in parallel to the construction of a discourse model. Where syntax fails to determine an element unambiguously, the inferencing mechanism fills the gap or makes the choice.

Let us assume a Marslen-Wilson type model in which pragmatics affects sentence interpretation and disambiguation. The pragmatics they consider ranges from selectional restrictions to knowledge that visitors stay in spare bedrooms. Whether or not this knowledge is linguistic knowledge is a question for theoretical linguistics. Is knowledge linguistic only if it is an element of the machinery required for syntactic processing?

Another possibility is that the world knowledge required for interpretation is restricted in principled, predictable ways. Such constraints would indicate a level or module responsible for providing just the information required for linguistic interpretation. This could be called lexical semantic knowledge, as distinguishable from knowledge in general. A separate issue is whether such knowledge has a form which is different from the cognitive models which are the output of linguistic interpretation or from memory. It is possible that lexical semantics could provide the world knowledge necessary for interpretation via a specialized formal language, or via representations which are employed in other types of inference. My claim is that there is an identifiable, constrained layer of linguistic semantic knowledge, but that its form does not differ from the form of general conceptual knowledge.

Evidence for constraints upon the world knowledge used during sentence processing comes from both psycholinguistic research and my own work in computational linguistics.

The protracted debate over the existence of semantic primitives resulted in their ultimate rejection and provided evidence that lexical knowledge does not differ from other knowledge in form of representation. Addressing first the question of form, I will sketch the debate and its outcome.

The classical theory of word meaning decomposed words into semantic primitives which had the force of truth conditions. The word *water* meant "clear, tasteless, liquid". A sentence such as "That is water" was true only of a substance for which the predicates "clear", "tasteless" and "liquid" were true. The implication was that word meanings were required to have the force of scientific theories. True sentences couldn't be uttered unless the speakers had knowledge of the true properties of objects. This led Putnam (1975), and others to separate the theory of word meaning from the theory of reference. The relationship between sentences and states of the world sentences was given by a theory of reference. The theory of word meaning became a theory of the knowledge required to be competent in a language, and this knowledge was of prototypes of objects.

Convergent with this development in the philosophy of language, Rosch (1976) and other cognitive psychologists questioned the assumption that conceptual knowledge took the form of semantic primitives. They found that categories have gradient properties, rather than the all-or-none membership predicted by the classical theory. As a result of these findings, the assumption that word meanings are decomposable into a small fixed set of primitives has been rejected (Smith and Medin, 1981). Where does that leave lexical semantics? Fillmore (1985) has proposed that word meanings are frames of default knowledge. Recent studies in cognitive psychology show that some concepts involve theories of the world (Murphy and Medin, 1985). Building on these findings, Dahlgren (1988) suggests that word meanings are naive theories of objects and events.

The theory of lexical semantics proposed in Dahlgren (1988), naive semantics, takes lexical representations as concepts. A word names a concept, and also plays a role in a discourse model which can be subjected to formal reasoning for purposes of determining the truth conditions of a discourse. However, the concept the word names constitutes a naive theory of the objects of reference, so that reasoning with word meanings must be non-monotonic. Furthermore, the naive theory has much more information in it than would be included in a representation formed from a stock of semantic primitives. Thus the representation of water includes information such as:

Water is typically a clear liquid, you find it in rivers, you find it in streams,
you drink it, you wash with it.

Furthermore, the knowledge places objects in a classification scheme (or ontology) which is intended to correspond to English speakers' conceptions of distinctions such as real versus abstract, and animate versus inanimate. The scheme is based upon psycholinguistic evidence, the classes required to represent verb selection restrictions, and the philosophical arguments concerning the distinction between sentient and all other objects. Study of protocols from experiments in the prototype theory reveal patterns of properties in naive theories. For example, artifacts have function, parts and operation features, animals have habitat and behavior features, while roles have function and status features. These patterns, called kind types, form constraints upon the feature types which are evident at nodes in the ontology.

Knowledge of verb meanings consists of the implications of events. Cognitive psychological studies show that verbs are not conceived in terms of classes such as motion, exchange, but rather in terms of the other events surrounding the event the verb denotes (Graesser and Clark, 1985, Trabasso and Sperry, 1985). The typical causes, consequences, goals, instruments, locations of events are the main components of conceptual knowledge for verbs.

When word meanings are identified with conceptual knowledge, a proliferation of mental representational types in the semantic lexicon is predicted. Color concepts have been shown to relate directly to the organization of color perception. Thus this theory predicts that words naming colors have meanings which include mappings to color perceptors. Words naming foods have meanings which include taste representations, along with some verbal elements. Some words are fully represented in terms of other words. (At this stage of computational linguistics, of course, we are in a position to represent only the verbal elements of word meanings.)

Thus the main assumptions of naive semantics are that words name concepts which are naive theories of objects and events. The content of these theories is not limited to a set of primitive features. Elements of meaning representations belong to a variety of sensory types. There is no difference in form between word meanings and cognitive representations.

So far, naive semantics seems most consistent with a model in which there is no distinction between lexical semantics and world knowledge. This is the view of Hobbs, et al, (1986), as well as all of the computational linguistic theories which use frames and scripts to encode domain knowledge. In the Hobbs method all of the commonsense knowledge associated with a word is encoded. For example, extremely detailed levels of naive physics are represented with the expression "wear down". However, experience with naive seman-

tics in the development of a computational text understanding system indicates that an extremely shallow layer of knowledge is sufficient to provide the information for successful disambiguation and anaphor resolution. A theory which identifies word meanings as just the knowledge needed for linguistic interpretation flows from this experience.

Furthermore, a theory which constrains lexical semantic knowledge to a very shallow layer would explain the real-time speed with which the discourse model is constructed. Fodor's fear of a universe of fridgeons is groundless. Interpretation does not involve an endless chain of inferences, but instead employs a short sequence of predictable inferences. This must be the case because cognitive psychologists have repeatedly demonstrated that inferences are drawn during discourse interpretation, and we know that many of these inferences are drawn in real time while hearing the utterance, rather than later. McKoon and Ratcliff (1990) have conducted a series of experiments to tease out the differences between the effects of discourse context and test questions in recall experiments, and to trace the time course of interpretive inferences. The experiments separate the variables of time and degree of familiarity of semantic information. They have found that well-known information contributes to interpretive inference within 250 ms of reading a sentence, while less-well-known information contributes only after 650 ms. One experiment involved the following context sentence:

The old man loved his granddaughter and she liked to help him with his animals; she volunteered to do the milking whenever she visited the farm.

When subjects were asked to recognize the word "cow" as having occurred in the sentence, the effect of the typically highly familiar association between cows and milking was evident within 250 ms. However, when the association between the context and test word was not highly familiar, the effect was not observed. Given the following sentence,

The director and the cameraman were ready to shoot close-ups when suddenly the actress fell from the 14th story.

When subjects were asked to recognize the word "dead," the effect of context was not evident after 250 ms (though it was when subjects were given 650 ms). This experiment shows that highly typical information with strong association to words is employed during the construction of the interpretation of a sentence (milking and cows). Information which requires more inferencing is not employed during the interpretation, but can be called upon later (falling and death). McKoon and Ratcliff conclude that "inferences mainly establish local coherence among immediately available pieces of information and there is

only minimal encoding of other kinds of inferences." These findings support a theory which isolates the highly associated, highly typical knowledge as linguistic semantic knowledge.

Part of the explanation for the shallowness of interpretive inference lies in the cooperative principles of communication. If a speaker doesn't believe a hearer shares a piece of knowledge required for the interpretation, then the speaker will include that information in the utterance. Another part of the explanation for shallowness lies in the intuition that knowledge of one's language includes knowledge of the naive theories of other speakers of the language. We know the culture-wide theory of certain objects, even when we are experts on those objects and have a completely different personal theory. An example would be the word "computer", which is a keyboard, monitor and printer to the non-technician, and is a central processing unit plus peripherals to the technician. The point is that technicians either know the naive theory, or they fail to communicate with non-technicians.

Thus my claim is that a shallow layer of commonsense knowledge is sufficient to disambiguate and build a discourse model in real time. Furthermore, this shallow layer has a constrained range of feature types, if not of feature values.

18.3 Experience with Naive Semantics in Computational Text Understanding

In computational linguistic research, I have been involved in the development of a system which reads and "understands" text sufficiently to answer questions about what the text says and to extract data from the text. The system contains interpretive algorithms which disambiguate the text, assign anaphors to antecedents, and connect events with coherence inferences. Each of these algorithms draws upon lexical semantic information. In the course of building the algorithms and testing them against corpora in three domains (geography, finance and terrorism), two results have emerged:

1. All of the algorithms use the same knowledge base of lexical information.
2. The algorithms succeed with only a very shallow layer of knowledge.

In other words, the highly typical, strongly associated knowledge is the knowledge that is used to build an interpretation of just what the text says, as McKoon and Ratcliff found, and this is confirmed in the application of a shallow layer of naive semantics to

disambiguation and discourse reasoning tasks. This layer is sufficient for the following interpretive components:

- word sense disambiguation
- relative clause attachment disambiguation
- prepositional phrase attachment disambiguation
- nominal compound disambiguation
- anaphor resolution
- coherence relation assignment

Three of the components (word sense disambiguation, PP attachment and anaphor resolution) have been tested against large corpora of text, and have been found to prefer the correct interpretation in over 95% of the cases. This statistical result is important because it is always possible to find examples which require more knowledge than is included in the naive semantics. The result shows that shallow layer of knowledge is sufficient in all but a few real cases. Again, the explanation for this sufficiency is that if the speaker believes that the hearer lacks an element of a naive theory, and that element is necessary for interpretation, the speaker is obligated to express it. Extrapolating to the naive semantic representations for English, the conceptual information subjects produce when asked to rapidly volunteer characteristics of objects and implications of events tends to be shared across a subculture. If information is not widely shared, speakers tend to state it explicitly. The use of naive semantic information containing only the shared knowledge has resulted in broad success statistically in the disambiguation and discourse algorithms; this would be the expected outcome if we assume that the writers of the test corpora followed the cooperative principle.

The text understanding system *Interpretext* has been under development for over five years. The early system parsed English text, producing one parse per sentence. This parse was then subjected to disambiguation algorithms which reformed the parse to correctly attach prepositional phrases and disambiguate word senses. (At present we are building a new wide coverage parser which will use naive semantic information to disambiguate structure during the parse, reflecting adherence to a model in which the parser has access to and uses lexical semantics). The formal semantic component of the system translates the disambiguated parse into a Discourse Representation Structure (DRS) (Kamp, 1981b). Each new sentence adds new predicates to the DRS. A discourse

reasoning module finds the antecedents of anaphors (Dahlgren, Lord and McDowell, 1990) and assigns coherence relations between discourse events (Dahlgren, 1989). The resulting representation is a shallow cognitive model of the text content. It represents only the inferences which must be drawn in order to ensure that one syntactic structure is selected, that word senses are disambiguated, that the individuals or events which are the same are given the same reference markers, and that each discourse event is connected to some other in the discourse. The cognitive model is translated to first order logic, and thence to Prolog. Text retrieval is accomplished with a standard Prolog problem-solver.

To illustrate the functioning of Interpretex, consider the following short text and the cognitive model produced by Interpretex (Figure 18.1).

Guatemala was charged by the US with terrorist attacks. They cited treatment of suspected guerrillas.

```

e1,us,g,a1,a2,e2,a3,r1
-----
charge6(e1,us,g)
with(e1,a1)
attacks(a1)
terrorist(a1)
cite(e2,us,a2)
treatment(a2)
of(a2,a3)
guerrillas(a3)
r1 before now
e1 included in r1
e2 included in r1
constituency(e2,e1)

```

Figure 18.1: Sample DRS produced by Interpretex

The parser produces a labelled bracketing for the first sentence which has the prepositional phrase "with terrorist attacks" attached to the noun phrase dominating "Guatemala". The disambiguation step finds that the prepositional phrase "with terrorist attacks" modifies the verb "charge", rather than the object noun phrase. In addition, word senses are chosen: the legal sense of "charge" rather than the monetary or physical senses, the social sense of "treatment" rather than the medical sense, and the social sense of "attack" over the physical or medical sense. The formal semantic module translates the disambiguated

parse into a DRS. The DRS has a set of reference markers, which stand for each of the entities and events or other abstract types which have been introduced into the discourse (e1, a1, etc., above), and a set of conditions, which stand for the relations and properties of these entities asserted by the discourse. The DRS provides a framework for interpretation of discourse semantics, such as pronoun resolution. After parsing and semantic translation of the second sentence, the anaphor resolution module identifies "they" with the US rather than Guatemala or "attacks". The coherence relation module assigns the coherence relation of "constituency" between the events in the two sentences, so that "citing" is seen as part of "charging". Temporal equations place the charging and citing within the same time interval, r1. The resulting representation is a cognitive model. It is a collection of predicates derived from the text itself expressing the properties of the entities introduced in the text, relations between them, and added inferred coherence relations between the segments of the text. All of the components of this analysis are presently prototyped and running in Prolog. A number of implemented formal semantic treatments such as the handling of plurals, modal contexts, questions, and negation are not shown in the example.

The naive semantics which is needed for the algorithms is limited to certain feature types. In general, ontological knowledge is used everywhere, especially the sentient/non-sentient distinction. This is because many verbs have selectional restrictions involving sentients, and verb selectional restrictions are frequently in the disambiguation algorithms as well as in anaphor resolution. As for the generic knowledge for verbs, the "cause", "goal", "consequence", and "instrument" features are used by all of the algorithms. For nouns, the features "function", "rolein", "partof", "haspart", "sex", "tool" and some others are used by the algorithms, but others, like "exemplar" and "internal trait" are not.

Interpretext contains algorithms for structural and word sense disambiguation which use naive semantics. In this section two algorithms are cited to illustrate the power of shallow naive semantics in a computational text understanding task. As explained above, all language understanding occurs in the context of some knowledge. Within a subculture there is a body of common knowledge that is shared among participants. There is a relatively shallow layer of that common knowledge ("lexical semantic knowledge") which the hearer/reader employs in discourse interpretation; this shallow knowledge is accessed as "default values" in the absence of relevant contextual information to the contrary. Two processes central to discourse interpretation are anaphor resolution and the structural interpretation of prepositional phrases. The following examples illustrate the use of lexical semantic knowledge as a default in prepositional phrase disambiguation and anaphor

resolution.

Sentences with prepositional phrases are well-known for their multiple possibilities for syntactic interpretation. Consider:

- (1) Radio Marti reports that guerrillas are shooting villagers with Chinese rifles.

The complement clause is syntactically ambiguous. Plausible interpretations for the prepositional phrase "with Chinese rifles" include:

- (1a) Guerrillas are using rifles to shoot villagers.
(1b) Villagers who have Chinese rifles are being shot by guerrillas.

If (1) is the first line of a news story, the most likely interpretation is (1a). People know that shooting is typically done with guns, and that guerrillas are probably more likely to have guns than villagers are. However, suppose the same clause occurs in another news story but in a different immediate linguistic context:

- (2) Radio Marti reports that Chinese rifles have been given to villagers cooperating with the government. In retaliation, guerrillas are shooting villagers with Chinese rifles.

Here the text tells the reader that villagers have rifles. The immediate salience of this fact overrides the general knowledge expectation about who is more likely to have guns, making it more likely that the reader will choose interpretation (1b). The default interpretation favors VP attachment for the prepositional phrase, but the context in (2) favors NP attachment. If a speaker/writer suspects that the hearer/reader might have difficulty interpreting a message, the speaker/writer usually provides clarifying information according to a principle of cooperation in discourse. Consequently, where a correct interpretation goes against the expected default interpretation, there are usually contextual cues. In (2), the first sentence "sets the stage" so that the VP attachment default is overridden.

These assumptions about lexical semantic knowledge and sentence interpretation are built into the Interpretex system. The idea that shooting is typically done with guns is part of the naive semantic knowledge encoded in the lexical entry for the verb "shoot". A rifle is identified as a gun, and a feature in the entry for "rifle" indicates that a typical operation performed with a rifle is shooting. The knowledge that guerrillas typically

use guns is part of the naive semantic knowledge about guerrillas; the lexical entry for "villager" does not mention guns. The representation of this shallow level of knowledge is sufficient for the Interpretex system to choose interpretation (1a), VP attachment, for (1). This knowledge would also favor an incorrect VP attachment interpretation for (2), unless the system recognizes that, as a result of having been given rifles, the villagers now have them, and a discourse entity of "villagers having Chinese rifles" is established and available for access in the next sentence.

In the Interpretex system, the shallow knowledge in the lexical entry for the verb "give" includes the fact that, as a consequence of the event of giving, the Recipient has the Object—i.e., the villagers have rifles. Thus, the shallow knowledge about the consequences of "giving" in one sentence can be used to override the knowledge about rifles and shooting in the next sentence (reasoning of this sort has not yet been implemented in the Interpretex system, but it is entirely feasible). It follows from the principle of cooperation that inferences established from the interpretation of the previous linguistic context will be favored by the system if they conflict with default inferences.

The principle of cooperation also accounts for contextual information overriding default knowledge in anaphora resolution, as in the following examples:

(3) The doctor looked up and recognized the nurse. She smiled.

Plausible interpretations of (3) include:

(3a) The nurse smiled—i.e., "she" = the nurse.

(3b) The doctor smiled—i.e., "she" = the doctor.

Although doctors can be men or women, and nurses can be men or women, the current typical default for these roles is to expect doctors to be men and nurses to be women, favoring interpretation (3a). However, these expectations can be altered by previous discourse, as in (4).

(4) Nurse Roger Smith was nervous as he entered Dr. Mary Brown's office.

The doctor looked up and recognized the nurse. She smiled.

In (4) the default interpretation (3a) is overridden by the information in the previous sentence.

Shallow information encoded in the Interpretex lexical entries makes possible the correct default interpretation for (3): the entry for "doctor" includes the information that doctors are typically (but not inherently) male, and the entry for "nurse" specifies that nurses are typically (but not inherently) female. For discourse (4), the (3a) interpretation needs to be overridden by identifying the definite noun phrase anaphors "doctor" and "nurse" with their respective antecedents, and by accessing shallow lexical knowledge about names indicating that a "Roger" is typically male and a "Mary" is typically female, so that "she" can be only Dr. Mary Brown. A shallow level of lexical semantic knowledge provides enough information to correctly interpret (1) and (3), but in (2) and (4) this information is overridden by inferences from the shallow information in the immediately preceding context.

Lexical-level shallow knowledge is sufficient for correct interpretation of most instances of sentence structure and anaphor resolution—the immediate representation of text meaning. It is less likely to be adequate for remote bridging inferences. In example (5), shallow naive semantics can bridge between "go" to transportation as an instrument of going, and from transportation to "car" because the inherent function of a car is transportation. However, in (6), shallow knowledge would not be sufficient to bridge from "pregnant" to "surprise" to "swallow gum".

(5) Ed decided to go to the movies. He couldn't find his car keys.

(6) Susan told Ralph that she was pregnant. He swallowed his gum.

18.4 Conclusion

A model which permits interaction between a syntactic module, a formal semantic module and world knowledge is theoretically attractive, and justified in psycholinguistic studies. World knowledge can be separated into a shallow linguistic layer and knowledge in general. The linguistic layer contains just the information required for discourse interpretation. In the naive semantic approach to word meaning, words name concepts, and concepts are naive theories of objects and events. Experience building a computational text understanding system demonstrates that a constrained, predictable portion of these naive theories is sufficient to disambiguate words and structure, and to build a discourse model with anaphors resolved and coherence relations assigned.

sparse. As a result ITP has chosen to handle collectives with syntactic and generic features. The syntactic features handle properties like singular agreement with no determiner, and the semantic features control quantification, and verb agreement.

Acknowledgements: This ontology was developed with the assistance of Claire Chyi, Vic Liptak, Carol Lord, Joyce McDowell, Susan Mordechai, Hajime Wada and Karen Wallace.

Bibliography

- Asher, N. 1987. A Typology for Attitude Verbs and Their Anaphoric Properties. *Linguistics and Philosophy*. 10:125-198.
- Berlin, B. Speculations on the Growth of Ethnobotanical nomenclature. *Language in Society* 1:51-86.
- Cantor N. and W. Mischel. 1979. Prototypes in Person Perception. *Advances in Experimental Social Psychology*. 12:3-52.
- Dahlgren, K. and J. McDowell. 1986. Kind Types in Knowledge Representation. *Proceedings COLING86*.
- Dahlgren, K. 1988. *Naive Semantics for Natural Language Understanding*. Boston: Kluwer Academic Press.
- Gelman, R. and E. Spelke. Thoughts about Animate and Inanimate Objects. in J.H. Flavell and L. Ross, eds., *Social Cognitive Development*. Cambridge, England: Cambridge University Press.
- Hayes, P. J. 1985. The Second Naive Physics Manifesto. In J. R. Hobbs and R. C. Moore, eds., *Formal Theories of the Commonsense World*, Norwood, NJ: Ablex.
- Kamp, H. 1981. A Theory of Truth and Semantic Representation. In J. Groenendijk, Th. Janssen, and M. Stokhof, eds., *Formal Methods in the Study of Language*, Amsterdam: Mathematisch Centrum, 277-322.
- Keil, F. 1989. *Concepts, Kinds and Categories*. Cambridge, Massachusetts: Harvard University Press.
- MacDonald, D. 1982. *Understanding Noun Compounds*. Carnegie-Mellon
Doctoral Dissertation.
- Quinton, A. 1975. Social Objects. *Proc. Aristotelian Society*. 75:1-27.
- Rosch, E., C. B. Mervis, W. D. Gray, D. M. Johnson, and
P. Boyes-Braem. 1976. Basic Objects in Natural Categories. *Cognitive Psychology*. 8:382-439.
- Strawson, P. C. 1953. *Individuals*. London: Methuen.

