

- Bresnan J (ed.) 1982 *The Mental Representation of Grammatical Relations*. The MIT Press, Cambridge, MA
- Bresnan J 2000 *Lexical Functional Syntax*. Blackwell, London
- Bresnan J, Kanerva J 1989 Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry*, 20(1): 1–50
- Carnie A 2001 *Syntax*. Blackwell, London
- Dalrymple M, Kaplan R M, Maxwell J T III, Zaenen A (eds.) 1995 *Formal Issues in Lexical-Functional Grammar*. Center for the Study of Language and Information, Stanford University
- Falk Y 2001 *Lexical-Functional Grammar: An Introduction to Parallel Constraint-based Syntax*. Center for the Study of Language and Information, Stanford
- Kaplan, R.M. and Zaenen A 1989 Long-distance dependencies, constituent structure, and functional uncertainty. In: Baltin M, Kroch A, eds. *Alternative Conceptions of Phrase Structure*. Chicago University Press, Chicago, pp. 17–42. [Reprinted 1995, 137–165]
- Nordlinger, R 1998 *Constructive Case: Dependent Marking Nonconfigurationality in Australia*. CSLI Publications, Stanford, CA

P. K. Austin

Lexical Processes (Word Knowledge): Psychological and Neural Aspects

He doth not neglect the profit of Lexicons (wherein all sayings and speeches are numbred)
E. Topsell, *The historie of foure-footed beastes*, 1607

If one looks in any psychology or psycholinguistics textbook for a definition of the mental *lexicon*, the superficial description given does not differ to one of the first uses of the word quoted above from the Oxford English Dictionary. Most modern researchers would agree that the lexicon refers to our knowledge of the words in our vocabulary. There is, however, a very broad range of opinion if one asks the next obvious questions: what is this knowledge and how is it represented in the brain?

As with many areas of higher brain function, studies of the mental lexicon generally follow one of two levels of explanation: descriptive and computational. There are currently no theories of word knowledge at the level of individual neurons. Descriptive explanations try to outline the nature of the lexicon in detail, attempting to establish both what information is represented by the lexicon and how this information is accessed. Computational explanations provide more explicit descriptions of word knowledge and lexical processes through either formal mathematical treatments or in the form of computer-based models that

take a specific input and transform it to the appropriate output.

Studies on the lexicon can be grouped into four categories, each employing a number of different research methodologies:

(1) *Normal, intact subjects*: most research here relies principally on reaction time, and to a lesser extent accuracy, as the critical dependent measures. Participants are typically required to give some form of response to spoken or written words (word recognition) or to give a spoken, single name (word production). A large number of different tasks are employed in each of these modalities. These include lexical decision (deciding whether a spoken or written stimulus is a real word or not), word monitoring (listening to spoken input and indicating when a target word has been presented), picture naming, and word naming (single word reading aloud). All these techniques aim to capture variations in reaction times in terms of differences between words of different types and/or with different characteristics. Alternatively the experimenter aims to manipulate a baseline reaction time through different presentation conditions (for example, the experimenter may investigate the effect of semantic priming—presenting a target word soon after a semantically related item).

(2) *Patients with brain injury*: the critical assumption underpinning this area of research is that different aspects of language, or in this case word knowledge, are supported by different areas of the brain (concentrated almost entirely in the left, dominant hemisphere). Therefore by studying patients with damage to one of these areas, it should be possible to reveal the underlying characteristics of how the human brain supports language, including word knowledge. Unlike studies with intact, normal participants, most research in this area relies most heavily on accuracy rather than reaction times. Studies are conducted on groups of patients (comparing to normal subjects and/or different types of patient group) or individual cases. A number of recent studies have also used a technique known as case-series methodology in which patients are selected on the basis of having the same type of deficit but at different levels of severity.

(3) *Functional neuroimaging*: for those techniques that concentrate more on spatial than temporal resolution, the underlying assumption for functional neuroimaging is the same as for neuropsychological studies with aphasic patients. In this way, functional magnetic resonance imaging (fMRI) or positron emission tomography (PET) compare words that vary according to a specific characteristic either using a parametric, subtraction methodology (looking for areas of the brain that are more active for one set than another) or a correlational technique (looking for areas in the brain where changes in activation are correlated with a changing characteristic of the presented words). Other imaging techniques are able to investigate how the time-course of activation com-

pare for words of different type. Such methods include event-related potentials (ERP) and magnetoencephalography (MEG).

(4) *Computational modeling*: this is a formal technique designed to instantiate a proposed computational description of the processes that underlying a specific set of behaviors. Most computational models do this by specifying an arrangement of simple processing units (known as a *network architecture*). The architecture describes how simulated activation can potentially be passed between these units. Information is either represented in localist form (the activation of a single unit corresponds to a specific piece of information—e.g., a word) or in a distributed manner (a number of pieces of information are represented by different patterns of activation across the same set of units—therefore, the activation of a particular unit is not unique to a specific datum). These computational models also differ along another dimension. In some models, the strength of the connection between units (that governs how much activation is passed between them) is specified explicitly by the modeler. Other networks use mathematical algorithms such that the model can extract its own set of weights to performance optimally. In doing so, these latter types of network can also be used to compare patterns of development as well as mature performance between the model and human subjects. Both types of model can be used to simulate disordered behavior (e.g., through brain injury) by various means including the removal of units and/or their connections, or by adding noise to the activation of units.

These four methods have given rise to a vast literature on the nature of the lexicon and its interaction with the rest of the language system. Some of the major themes arising from this work are summarized below.

1. *The Traditional Conception of a Lexicon*

Many theories of lexical representation are not that different from Topsell's definition of 1607: the lexicon is a list of all the words in a person's vocabulary. By some form of look up, all or part (in the case of models with multiple lexicons, see below) of our knowledge about a word is retrieved. An early functional and mathematical description of this process was given by Morton with his influential logogen theory (1969), which was originally applied to explain word recognition but was extended later to address spoken and written recognition and production. The key elements to this approach are as follows. Each word is represented by a separate unit, or logogen, which gathers evidence for the presence of that word in the input (in the form of sensory input for recognition or semantic memory for production). Each logogen does this relatively independently of the others. A word is recognized or produced when there is sufficient posi-

tive evidence for that word in the nature of the input (e.g., the letters in the written input match those for the target word). At this point, the logogen makes other information about that word available—such as the word's meaning or its pronunciation. With reference to signal detection theory, Morton also included the notion of a 'threshold' which governs how much positive evidence is required before the logogen makes a response. In this way, word frequency effects in recognition and production are described in terms of different thresholds for different logogens (high frequency words have lower thresholds and thus require less positive evidence to respond).

2. *How Many Lexicons are there?*

If one considers the literate adult, the sum total of the acquired knowledge can be classified into three forms: the meaning of words; the sound of each word; and, the written form or orthography. With reference to the traditional view of the lexicon (see Sect. 1), some researchers have argued that these three forms of information are represented in separate lexicons: lexical semantics, a speech lexicon, and an orthographic lexicon. Although most models have interconnections between these different lexicons (at some level they need to combine to reflect all aspects of word knowledge), the notion is that they are discrete entities both at a functional and neural level. Functionally, each lexicon can then have its own dedicated form of representation and associated processes. At the neural level (both for neuropsychological and neuroimaging studies) this means there are distinct regions in the brain that can be damaged/activated relatively independently of each other. The cognitive neuropsychology literature, in particular, has also encouraged a further division in the mental lexicon—for input and output processes. The notion here is that there might be four separate lexicons to cover the four combinations of recognition and production for spoken and written words (for an overview, see Ellis and Young 1988). The same division is not proposed for meaning, which in a sense would be nonsensical anyway (for effective communication, the meaning driving the speaker/writer's output must be the same as that decoded by the listener/reader). The need to have separate input and output representations is still contested, however. Some neuropsychological data would seem to favor this position—for example, there are patients with word-finding difficulties (anomia) that do not have trouble understanding words (word deafness), and other patients with the opposite pattern. However, more studies are needed to rule out the possibility that the patterns reflect modality-specific nonlexical deficits or the fact that expression is always more vulnerable to impairment than receptive abilities.

Two separate strands of research have argued for another lexicon—the semantic lexicon or lemmas, that mediate between meaning and other language processes and representations (neuropsychology—e.g., Butterworth et al. 1984; speech production—e.g., Levelt 1992). In terms of neuropsychology, an impaired separate lexicon that normally acts as a gateway between meaning and language processes explains the pattern found in a number of patients who have impaired production and comprehension of speech and writing but much better understanding of nonverbal material (again, this position is contended, for a recent overview, see Lambon Ralph and Howard 2000). In speech production, Levelt (1992) has argued for a two-stage process within lexical access. First the meaning of the intended utterance is used to select a single abstract lexical form, a lemma (that also codes for syntactical markings such as gender and number), and then this in turn activates a lexeme, which makes the sound of the word available both in terms of the constituent phonemes, their arrangement and associated stress pattern. There is, however, a general concern with this direction of theorizing in that a proliferation of multiple lexicons means the potential for redundancy of representations is much higher, and tends to make the theories unfalsifiable.

3. An Alternative Conception of the Nature of Lexical Knowledge?

A quite different approach has been adopted by some parallel distributed processing (PDP) computational models in which there is knowledge of the words in our vocabulary but not a lexicon—at least, in the traditional sense of that word. A landmark publication in this respect was the word reading and recognition model of Seidenberg and McClelland (1989). In this framework word knowledge is represented as distributed patterns of activation across three ‘primary’ systems corresponding to orthography, phonology and semantic memory. Through a gradual learning process, in which the weights of the connections within and between these representational layers are adjusted, models of this type extract word ‘knowledge’ from the vocabulary they are trained on. Note this is quite different from the traditional metaphor of looking up an entry in dictionary or Morton’s logogen theory. The Seidenberg and McClelland model does not have a list of all the known words (i.e., a lexicon) so the terms ‘access’ and ‘retrieval’ are redundant. Instead the model takes one form of input (e.g., the written word) and then computes the appropriate representation reflecting the meaning and/or the pronunciation. In this way the model can be said to have word knowledge but not a lexicon. This is further reinforced by the fact that the model processes novel words (nonwords that, by definition, cannot be in a lexicon) through the very same set of units corresponding to

the written or spoken form. Another potential allure of this approach to word knowledge is that the formulation may help when future research (see Sect. 5) attempts to bridge the gap between brain and behavior. The neural basis for the primary systems of semantics, phonology, and orthography is much more obvious. It is possible, for example, that orthography reflects a stable set of representations extracted through visual perception (Behrmann et al. 1998) and semantic memory, or conceptual knowledge, reflects high-level abstraction based on all forms of verbal and nonverbal experience, perhaps supported primarily by regions within the temporal lobes (Tranel et al. 1997).

This alternative approach has also been used to address some of the data that has lead researchers, using a traditional view of the lexicon, to posit the need for multiple lexicons (see Sect. 2). For example, patients with both poor comprehension and production of words (whether spoken or written) are described from a traditional viewpoint in terms of damage to a semantic lexicon that mediates between conceptual knowledge and the language system. The same data are adequately captured by PDP computational principles in terms of the nature of the mapping between representations for words vs. pictures/objects (Lambon Ralph and Howard 2000). The fact that there is an arbitrary mapping between phonology/orthography and conceptual knowledge, unlike the mapping for pictures or objects, means that verbal comprehension and production are highly vulnerable to damage to the system that represents the meaning of both words and pictures.

4. How Does Word Knowledge Give Rise to Specific Language Activities?

Within the repertoire of language activities there are some that are assumed to depend directly on word knowledge or lexical information. For example, understanding words or producing words is assumed to proceed with reference to lexical information. So, as described above, within a logogen framework speech production occurs by the target meaning acting as an input to a set of logogens, which selects the correct word and then makes the spoken form available. There is also a direct use of word knowledge in PDP computational models as well—the knowledge extracted through learning the mapping between semantics and phonology is used to produce the correct spoken form in speech production. Other language activities are assumed to depend upon a combination of lexical knowledge and other nonlexical processes. So, for example, Levelt (1992) proposes that speech production is based upon a ‘formulator’ that specifies the syntactical frame for a sentence into which lexical representations (lemmas) and their associated knowledge are fitted. Two other well-known examples are reading aloud and the past tense transformation of

English verbs. Researchers studying the process of reading aloud have classified words into three basic types: (a) most words, known as regular or consistent words, are pronounced in line with their spellings (examples of these are YARD and MINT); (b) a smaller but still substantial number of irregular or inconsistent words deviate from the typical spelling-to-sound translations (examples of these are YACHT and PINT); and (c) we are also able to read new or novel words without extra training (e.g., EMAIL or WUG). There is a very similar classification for describing the past tense of an English verb. There are many regular verbs, all of which are turned into the past tense by the addition of -ED (e.g., *talk* → *talked* or *need* → *needed*), and a smaller number of irregular verbs which mark the past tense in another way (e.g., *speak* → *spoke* or *think* → *thought*). We are also able to produce the past tense for new or novel verbs (e.g., *email* → *emailed*, *wug* → *wugged*), again without any specific training on these novel forms. The traditional theories of verb inflection (the dual-mechanism account, Pinker 1999) and reading aloud (the dual-route theory, Coltheart et al. 1993) have both used this tripartite division of word type to argue that each language activity is underpinned by two distinctly different processes. The two theories are sufficiently similar that they can be summarized together. Irregular words/verbs (e.g., PINT → 'pint' for reading aloud and 'brought' → 'brought' for verb tense) are said to be translated by reference to a lexicon—or indeed, in the case of reading, two lexicons: one orthographic and the other phonological—that stores the whole specific word form for every irregular item in the language. The ability to read or inflect novel forms is explained in these theories by positing a task-specific, generative process encapsulated by a set of rules: three rules for the three allomorphs required for the inflection of regular verbs (/t/, /d/, /@d/) and an extensive series of rules for the translation of letters/graphemes into phonemes. Neuropsychological investigations of impairments to both reading aloud and past tense production have highlighted two contrasting patterns that have been taken to support these dual-mechanism theories (some patients have specific difficulties with irregular forms and others with novel forms: see Ullman et al. 1997 for verb morphology; an overview of acquired dyslexia can be found in Ellis and Young 1988).

Again, PDP computational models have taken a different approach (verb morphology, Joanisse and Seidenberg 1999; reading aloud, Plaut et al. 1996). In these frameworks word knowledge is not reflected in a lexicon and thus the strict division between the lexicon and the generative rules does not apply. Instead the interaction between and within the primary systems (semantics, phonology, and orthography) not only give rise to word knowledge, the models are also capable of transforming one form of representation into another (e.g., reading is orthography → phon-

ology) for all three word types (irregular, regular, and novel). Critically, however the division of labor amongst the primary representations is not even for the different word types, and dissociations in performance can occur if one of the primary systems is damaged (for an overview of this perspective on acquired dyslexias, see Patterson and Lambon Ralph 1999).

5. Future Directions

Research on the psychological aspects of word knowledge will, no doubt, continue in the future. One potential area for new research must be on the neural aspects of word knowledge. There are already signs that neuropsychological and computational investigations of people with acquired language disorders (aphasia) are being used to address both psychological and more neural aspects of word knowledge (examples can be found in Ullman et al. 1997, Patterson and Lambon Ralph 1999). While there has been some use of positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) to look at word knowledge (for an overview, see Price 1998), there is much more for these techniques to tell us about the neural aspects of word knowledge in the future.

See also: Lexical Access, Cognitive Psychology of; Lexical Semantics; Lexicon; Word Meaning: Psychological Aspects; Word Recognition, Cognitive Psychology of

Bibliography

- Behrmann M, Plaut D C, Nelson J 1998 A literature review and new data supporting an interactive account of letter-by-letter reading. *Cognitive Neuropsychology* **15**: 7–52
- Butterworth B, Howard D, McLoughlin P 1984 The semantic deficit in aphasia: The relationship between semantic errors in auditory comprehension and picture naming. *Neuropsychologia* **22**(4): 409–26
- Coltheart M, Curtis B, Atkins P, Haller M 1993 Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review* **100**(4): 589–608
- Ellis A W, Young A W 1988 *Human Cognitive Neuropsychology*. Lawrence Erlbaum, London
- Joanisse M F, Seidenberg M S 1999 Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences of the United States of America* **96**: 7592–7
- Lambon Ralph M A, Howard D 2000 Gogi aphasia or semantic dementia? Simulating and assessing poor verbal comprehension in a case of progressive fluent aphasia. *Cognitive Neuropsychology* **17**: 437–66
- Levelt W J M 1992 Accessing words in speech production: Stages, processes and representations. *Cognition* **42**: 1–22
- Morton J 1969 The interaction of information in word recognition. *Psychological Review* **76**: 340–54
- Patterson K, Lambon Ralph M A 1999 Selective disorders of reading? *Current Opinion in Neurobiology* **9**: 235–9

- Pinker S 1999 *Words and Rules: The Ingredients of Language*. Weidenfeld & Nicolson, London
- Plaut D C, McClelland J L, Seidenberg M S, Patterson K 1996 Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review* **103**(1): 56–115
- Price C J 1998 The functional anatomy of word comprehension and production. *Trends in Cognitive Sciences* **2**: 281–8
- Seidenberg M, McClelland J L 1989 A distributed developmental model of word recognition and naming. *Psychological Review* **96**: 523–68
- Tranel D, Damasio H, Damasio A R 1997 A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia* **35**(10): 1319–27
- Ullman M T, Corkin S, Coppola M, Hickok G, Growdon J H, Koroshetz W J, Pinker S 1997 A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience* **9**(2): 266–76

M. A. Lambon Ralph

Lexical Semantics

The ‘core’ areas of linguistics are generally agreed to be: phonology, the study of the sounds and sound systems of language; syntax, the study of grammar; and semantics, the study of linguistic meaning. Words figure in all three areas, but it is in their semantic aspect that they offer the most distinctive field of inquiry. This article presents a survey of the topics which form the principal foci of interest within the domain of lexical semantics.

1. Introduction

To the layman, words are the prime bearers of meaning in language, and formulating a message is essentially a matter of choosing the right words. This view is of course an oversimplification, and seriously understates the roles of grammar and of context, but there is more than a germ of truth in it. The branch of linguistics whose domain is the systematic study of word-meanings is lexical semantics. But before proceeding any further, it is necessary, first, to mark off an area of the slippery notion of ‘meaning’ which is appropriate for our purposes, and second, since other people, notably lexicographers, also concern themselves with the meanings of words, something needs to be said about what is distinctive about the activities of lexical semanticists.

If John says to Mary *Have you been to the bank and seen the manager?* and Mary replies *Yes*, there is an obvious and intuitively valid sense in which Mary’s reply ‘means’ ‘I have been to the bank and seen the

manager.’ Yet it is also fairly obvious that ‘I have been to the bank and seen the manager’ would not be a satisfactory definition of the meaning of the word *Yes*. It is therefore necessary to make a distinction between the ‘inherent meaning’ of a word and its ‘meaning-in-context.’ (The inherent meaning of *Yes* might be glossed: ‘The proposition you present to me for confirmation is true.’) A similar point (less dramatic) can be made about *The child ran to his mother*. Here it is clear that *the child* refers to a boy; but this, too, is meaning-in-context—the inherent meaning of *child* is something like ‘immature human being.’ As a first approximation, it can be said that lexical semantics focuses on inherent meanings of words: meanings-in-context are the concern of pragmatists.

Lexicographers also are interested mainly in the inherent meanings of words. So what is the difference between a lexicographer’s interest in word meaning and that of a lexical semanticist? Basically, this is a matter of theoretical vs. practical concerns, parallel in many ways to the relation between physicist and engineer, or botanist and farmer/gardener. A lexicographer aims to impart the maximum of information useful to the dictionary-user with maximum economy; a lexical semanticist wants to understand the principles underlying lexical semantic phenomena. In an ideal world there would be considerable cross-fertilization between the two fields: in reality, the cross-domain traffic has been somewhat disappointing.

2. Approaches to Word Meaning

There are numerous ways of approaching the study of word meaning, depending on certain basic assumptions one makes about it.

2.1 The Componential/Atomistic Approach

One influential approach, much criticized but constantly reborn, is to think of the meaning of a word as being a more or less complex structure, built up out of combinations of simpler, or more primitive, units of meaning, a simple example would be the analysis of *woman* as [ADULT] [HUMAN] [FEMALE]. Such ‘componential’ or ‘atomistic’ approaches come in a variety of types, depending on what is claimed regarding the nature of the analysis, the identity and nature of the units, how they are combined, and what aspects of word meaning are to be explained by them.

Some componential analyses (probably the majority) aim at a reduction in the number of combinatorial elements, that is to say, they aim to explain the large number of word meanings in terms of combinations of units drawn from a smaller set, parallel to the way that the phonological forms of all