3 The contribution of linguistics

BETH LEVIN

3.1 Introduction

The lexicon has come to occupy an increasingly central place in a variety of current linguistic theories, and it is equally important to work in natural language processing. The lexicon – the repository of information about words – has often proved to be a bottleneck in the design of large-scale natural language systems, given the tremendous number of words in the English language, coupled with the constant coinage of new words and shifts in the meanings of existing words. For this reason, there has been growing interest recently in building large-scale lexical knowledge bases automatically, or even semi-automatically, taking various on-line resources such as machine readable dictionaries (MRDs) and text corpora as a starting point, for instance, see the papers in Boguraev and Briscoe (1989) and Zernik (1989a). This chapter looks at the task of creating a lexicon from a different perspective, reviewing some of the advances in the understanding of the organization of the lexicon that have emerged from recent work in linguistics and sketching how the results of this work may be used in the design and creation of large-scale lexical knowledge bases that can serve a variety of needs, including those of natural language front ends, machine translation, speech recognition and synthesis, and lexicographers' and translators' workstations.

Although in principle on-line resources such as MRDs and text corpora would seem to provide a wealth of valuable linguistic information that could serve as a foundation for developing a lexical knowledge base, in practice it is often difficult to take full advantage of the information these existing resources contain. Dictionaries, for example, might seem particularly well-suited as a basis for automatic lexicon construction, since the information they provide is structured within the entry, and it would seem possible to extract certain information, for example, part of speech, fairly trivially. However, this is only a fraction of the information available in a dictionary. Dictionaries are designed for human users by humans. Human users are native speakers of language who know at least

Much of the material on verbs of sound in this chapter grew out of a series of discussions with Sue Atkins that accompanied the writing of Atkins and Levin (1988). I would like to thank Sue Atkins, Chuck Fillmore, Bob Ilson, and Annie Zaenen for their comments on an earlier version of this chapter.

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

implicitly how the lexicon of their language is structured, and lexicographers exploit the lexical knowledge of potential users in writing dictionary entries. Consequently, dictionary entries only need to say enough about a word to allow native speakers of a language to tap into their general knowledge. Thus entries often leave much implicit or unsaid, something that would be unacceptable in a lexical knowledge base for a natural language system. The missing information must be filled in from somewhere, and linguistic studies into lexical organization can contribute to this task. Even learner's dictionaries, which are intended for learners of a language, take advantage of general properties of language, although typically they do provide fuller information than dictionaries intended for native speakers of that language about syntactic properties, as well as a range of example sentences illustrating word use. For more discussion of this issue see Atkins, Kegl, and Levin (1986, 1988), Boguraev and Briscoe (1989), and McCawley (1986).

These considerations aside, the value of using dictionaries as a starting point for building a lexical knowledge base is diminished by the limitations of dictionary-making itself. Dictionaries are written by lexicographers, who are themselves humans working within rigorous time and space constraints. Consequently, not all words receive the attention they deserve (see Atkins [this volume] for discussion). Even the best dictionaries have flaws; for instance, they are often incomplete and inconsistent (Atkins, Kegl, and Levin, 1988; Boguraev and Briscoe, 1989; Neff and Boguraev, 1989; among others). For instance, words that pattern in the same way are often not given parallel treatment in dictionaries, due either to time and space limitations or to the failure of the lexicographer to recognize the pattern.[1]

The goal of this chapter is to sketch the contribution that linguistics can make to the task of building lexical knowledge bases. Specifically, results of linguistic research into lexical organization have implications for the design of a lexical knowledge base: they both suggest the overall structure of the knowledge base and delineate the type of information that must be available in this knowledge base. This framework in turn should facilitate the extraction of as much information as possible from on-line resources. Specifically, efforts to build lexical knowledge bases automatically or semi-automatically could use template entries for verbs of particular semantic types motivated by linguistic research to guide attempts to extract information about specific verbs from existing on-line resources such as dictionaries and corpora.

Section 3.1 outlines the nature of a native speaker's lexical knowledge from

[1]On-line dictionaries are unlikely to serve as a lexical knowledge base, even if, as suggested by some researchers, several dictionaries were merged on the assumption that the result will be more complete than any single dictionary. As pointed out in Atkins and Levin (1988), the process of merging dictionary entries faces many obstacles. Furthermore, there is no guarantee that the result of merging the entries for a given word would be an entry that is substantially better than the entries of individual dictionaries; such an entry is unlikely to approximate a linguistically motivated lexical knowledge base entry for that word.

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

the perspective of linguistics. Section 3.2 provides a case study of the lexical knowledge of a native speaker of English by focusing on what a native speaker of English knows about verbs of sound. Section 3.3 exemplifies the design of some aspects of linguistically motivated entries for verbs of sound in a lexical knowledge base, drawing to a large extent on earlier work presented by Atkins and Levin (1988). Section 3.4 discusses how the properties of lexical organization can be used to take maximal advantage of the information available in existing lexical resources.

3.2 The nature of lexical knowledge

A prerequisite for the design and creation of lexical knowledge base is an understanding of what constitutes the lexical competence of a speaker of English, since a lexical knowledge base must make this knowledge explicit. Therefore, this section reviews what research in linguistics has uncovered concerning the nature of lexical knowledge, focusing almost exclusively on lexical knowledge associated with verbs, since they are typically the most complex lexical items to deal with. Similar issues arise for lexical items from other parts of speech; for example, Pustejovsky (1990) discusses some of the problems posed by nouns.

First, consider a problem that computational linguists engaged in building natural language systems might face, a problem that arises because of the complex nature of lexical knowledge.[2] The problem involves the two interchanges below.

(1) Speaker 1: Sally ate a peach.

Speaker 1: Did Sally eat?

Speaker 2: Yes.

(2) Speaker 1: David dressed the baby.

Speaker 1: Did David dress?

Speaker 2: I don't know.

If asked *Did Sally eat?* after having been told that Sally ate a peach as in interchange 1, speakers of English would not hesitate to answer *Yes*. But the same speakers would not answer *Yes* if asked *Did David dress?* after being told that David dressed a baby (see interchange 2). Here the appropriate answer would be *I don't know*. Two sequences consisting of a statement followed by a question that appear to be parallel syntactically (transitive use of a verb in the initial statement, intransitive use of the same verb in the question) elicit quite different

[2]For more discussion of this problem and related ones, as well as an attempt to design a natural language system that can deal with such problems, see Katz and Levin (1988). The examples in (1) and (2) are taken from this chapter.

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

responses from speakers of English. The simple syntax of each pair is unlikely to present a challenge to the parsers used in most existing natural language systems. The difficulty is that the intransitive uses of the two verbs receive very different interpretations. The intransitive use of *eat* found in the question *Did Sally eat?* implies the existence of an understood but unexpressed object, some type of food or meal. The question *Did David dress?* on the other hand does not mean 'Did David dress something one typically dresses?'; it means 'Did David dress himself?'. Speakers of English draw on their knowledge of the properties of these verbs to determine the appropriate interpretation of their intransitive use. Similarly, the lexicons of natural language systems must be built to allow these systems to recognize that the relationship between transitive and intransitive *dress* is not the same as that between transitive and intransitive *eat,* so that they can handle interchanges of the type sketched here appropriately.

Although dictionaries are a rich source of information about words, the information needed in dealing with problems of the type described here is often not explicitly signaled, if it is included at all. Most dictionaries indicate whether verbs have a transitive use, an intransitive use, or both, but relationships between transitive and intransitive uses of verbs such as *eat* and *dress* are not as a rule explicitly indicated. However, such relationships are often encoded using a variety of cues in the dictionary entry that involve the grammatical codes, the wording of the definitions, and properties of the example sentences (Atkins, Kegl, and Levin, 1986, 1988). Thus although the relevant information can sometimes be found in a dictionary, it is not trivially accessible, but will require queries formulated in terms of the specific cues in dictionary entries, a problem complicated by the fact that the same cues are not used consistently across the entries of verbs that pattern in the same way.

As the *eat/dress* example illustrates, some verbs may express their arguments in more than one way, sometimes with slightly different semantic interpretations. Any natural language system that aims at substantial coverage of English must be able to handle correctly not only these but the entire range of possible relationships between alternate expressions of the arguments of verbs. The understanding of the lexical organization of English verbs of the type that emerges from linguistic investigations can contribute to the realization of this goal.

Although the lexicon has been considered the domain of the idiosyncratic, there is much evidence that the relationship between the meaning of verbs and their syntactic behavior is governed by quite general principles, with evidence coming from studies in both lexical semantics and syntax (Bresnan and Kanerva, 1989; Carter, 1976, 1988; Fillmore, 1968; Foley and Van Valin, 1984; Gruber, 1976; Hale and Keyser, 1986, 1987; Jackendoff, 1983, 1990; Levin, 1985; Marantz, 1984; Pustejovsky, 1990; Rappaport, Levin, and Laughren, 1988; Talmy, 1985; and many other works in various theoretical frameworks). The *eat/dress* example shows that certain verbs have both transitive and intransitive uses, and that the relationship between the uses is not uniform across all verbs.

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

However, such a relationship is not merely an idiosyncratic property of a verb; rather it is to a large extent predictable from the verb's meaning. Interchanges parallel to the one described for *eat* are possible with a wide range of verbs, including *type, sew, sweep,* and *read.* These verbs are all activity verbs; most of them describe typical occupations. Another set of verbs including *bathe, change, shave, shower,* and *wash* – all verbs of grooming or bodily care – behave like *dress.*

Linguists have extensively studied a wide range of linguistic phenomena involving the expression of the arguments of verbs, such as the alternations in transitivity exhibited by the verbs *eat* and *dress.* These studies reveal that English verbs are organized into classes on the basis of shared components of meaning. The members of these classes have in common a range of properties, specifically properties concerning the possible expression and interpretation of their arguments, as well as the extended meanings that they can manifest (Levin, 1985).

The long-term goal of much current linguistic research is explaining what a native speaker of a language knows about the lexical properties of verbs, focusing on those aspects of lexical knowledge related to argument structures, the semantic and syntactic properties of verbs tied to their status as argument-taking lexical items. A central concern of linguistic research on the lexicon is the study of the meanings of verbs and the elaboration of a theory of the representation of lexical entries in which the meaning of a verb is properly associated with the syntactic expressions of its arguments. Ideal lexical entries of verbs should embody the full range of linguistic knowledge possessed by an English speaker in relation to those verbs. At the same time, however, any given entry should supply the minimum amount of information necessary to account for the native speaker's linguistic knowledge of it. This dual requirement naturally leads to the investigation of those aspects of the linguistic behavior of lexical items that are determined by general principles of grammar.

Currently, an important part of this research is the rigorous study of diathesis alternations, alternations in the expression of the arguments of verbs. As the discussion of the verbs *eat* and *dress* illustrates, since diathesis alternations reflect the interaction between a representation of the meaning of a verb and the principles that determine the syntactic realization of its arguments, they can be used to probe into both the lexical representation of meaning and the relationship between syntax and semantics. As the distinctive behavior of verbs with respect to diathesis alternations arises from their lexical properties, specifically their meaning, the exploration of the ways in which diathesis alternations distinguish among verbs should reveal semantically coherent verb classes. Once identified, these classes can be examined to isolate the components of meaning common to verbs participating in particular alternations. These components of meaning would be expected to figure prominently in the lexical representation of the meaning of these verbs. Attempts to formulate the principles according to which these elements of meaning determine the syntactic behavior of verbs then be-

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

come possible. For some work along these lines, see Hale and Keyser (1986, 1987), Laughren (1988), Levin and Rappaport (1988).

For these reasons, the study of diathesis alternations can make a significant contribution to the elucidation of the lexical representation of meaning. These studies have established a range of diathesis alternations relevant to the lexical organization of English and have identified a number of essential semantically coherent classes of verbs, as well as the central properties characterizing verbs of each type. (See Levin [1989] for a description of the lexical properties of English verbs taking noun phrase and prepositional phrase complements.) Nevertheless, much basic research remains to be done in this area.

3.3 A case study of lexical knowledge: verbs of sound

Linguistic investigations into lexical organization of the type described here can be used to form the basis for the systematic treatment of lexical items, which in turn will ensure the completeness and consistency of verb entries in a lexical knowledge base. As an example of what such a treatment entails, this section will provide a case study of the lexical knowledge that a native speaker of English possesses with respect to a certain class of verbs, the verbs of sound. The focus will be on the syntactic and semantic properties of the verb *whistle* and related verbs; morphological and phonological properties, which are also part of a native speaker's lexical knowledge, will be ignored. This section will not attempt to provide an exhaustive analysis of these facets of these verbs, even in the senses considered here. Rather the aim is to provide a picture of the form that the lexical knowledge associated with the members of a given verb class takes. Section 3.4 will begin to explore how this knowledge translates into a lexical knowledge base entry. The discussion of verbs of sound in this and subsequent sections draws and elaborates on the discussion of these verbs in Atkins and Levin (1988).

As discussed in the previous section, current linguistic research on the lexical knowledge of native speakers of a language shows that the human lexicon is highly structured. Verbs fall into classes on the basis of shared meaning components, and the members of these classes have in common a range of properties concerning the expression and interpretation of their arguments. Verbs of sound are no exception. In this section, the structure of the verb lexicon will be explored by focusing primarily on the verb *whistle,* chosen as a representative member of the class of verbs of sound. English possesses a large class of verbs that denote in their basic (simplest) sense the emission or production of a sound in some way. This class includes the verbs *whistle, whine, groan, grunt, snort, hum, hoot, howl, moan.* In their basic sense, the members of this class share the meaning 'emit a sound'; however, the meanings of the members of this class differ from each other in several well-defined ways. The most obvious differences involve the physical properties of the sound produced ('shrill' for

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

whistle, 'low' for *grunt*, and 'high' for *bleep*) and the means of producing this sound ('electronically' for *bleep*, 'blowing' for *whistle*, 'vibrating the vocal cords' for *grunt*).

The nature of the differences in meaning among the verbs of sound is not unusual, but reflects one facet of English lexical organization. There are a variety of classes of verbs in English whose members share a basic sense but differ from each other in terms of a means or manner component (Carter, 1976; Fellbaum, 1990). For instance, English has a large class of verbs of manner of motion, which includes *jump, hop, run, skip, walk, waddle*, and a class of verbs of means of removal, which includes *erase, mop, shovel, sweep, vacuum, wipe*. Furthermore, some of the classes of means or manner verbs are opposed to a class of verbs that share the same basic meaning, but combine it with what might be called a 'result' component. For example, English verbs of removal fall into two classes, the class of means of removal verbs just mentioned, and a smaller class of verbs denoting only the result (and not the means) of removal, which includes *clear, clean, empty;* the two subclasses of verbs of removal behave very differently. In fact, the presence or absence of a means or manner component often plays a part in determining the properties of a word.[3] The importance of the notions of means and manner in the organization of the English verb lexicon is also reflected in the verb component of WordNet (Miller et al., 1988, 1990; Miller and Fellbaum, in press), which as described by Fellbaum (1990) is to a large extent structured around means or manner relations between verbs.

An important part of an English speaker's knowledge of the verbs of sound is knowledge of their argument-taking properties. In their basic sense – as verbs denoting the emission or production of a particular sound – these verbs denote events with a single participant, the emitter of the sound. This participant is expressed as the subject of these verbs, so that in this sense these verbs are intransitive. (Some of these verbs can take objects, as in "*he honked the horn*", but such uses represent different senses of the verbs; see below.)[4]

(3) a. "*I hissed and snarled and ground my teeth at them.*"
b. "*the line of wart-hogs moved snuffling and grunting across the trail*"
c. "*the wet candles hissed and went out one by one*"
d. "*cars honked and hummed in the road*"

Most of the verbs of sound impose very tight restrictions on the choice of possible subjects, since the subject must be something that is capable of inherently producing a sound with the appropriate characteristics. An examination of verbs of sound reveals that these verbs may denote the production of that sound

[3]For further discussion with respect to verbs of motion see Levin and Rappaport Hovav (in press a); for discussion with respect to verbs of removal see Levin and Rappaport Hovav (in press b).

[4]The italicized citations in double quotes are taken from the Cobuild Corpus, which is part of the Birmingham Collection of English Text, held at the University of Birmingham, England, and jointly owned by the University of Birmingham and by Collins Publishers Ltd.

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

by a human (*"I put back my head and howled"*), another animate entity (*"I could hear a wolf howl"*), or an inanimate entity (*"the wind howled in the trees"*). However, no verb of sound can have an abstract noun as a subject (**the despair growled, *the clarity cooed*). Furthermore, verbs of sound differ in the selectional restrictions on their subjects. That is, it appears to be a property of a particular verb of sound which of the possible subjects it allows, with members of this class differing from each other in this respect. The verb *whistle* permits all three types of subject, *bleep* is rarely if ever found with animate subjects, and *grunt* is rarely if ever found without them.[5] As we shall see, the choice of subject plays a part in determining whether the verb shows certain extended senses.

Verbs of sound in their basic sense do not show a variety of ways of expressing their arguments, probably because they take a single argument. However, they manifest a range of extended senses – additional senses that are systematically related to the basic sound emission sense. As discussed further below, individual verbs vary as to which of the possible extended senses they manifest, so that not all possible senses are shown by every verb in the class. Taking the verb *whistle* as an example where possible, we see the following possibilities:

(4) a. 'emit a sound in a particular way'
e.g., *the girl whistled / grunted / hissed*

b. 'emit the sound as a reaction'
e.g., *the boy whistled / grunted / shrieked at the dog*

c. 'utter by emitting the sound'
e.g., *he whistled / hummed / yodelled a tune*

d. 'signal by emitting the sound'
e.g., *they whistled / grunted / hissed a warning*

e. 'express (an emotion) by emitting the sound'
e.g., *she whistled / grunted her disgust*

f. 'communicate verbally by emitting the sound'
e.g., *he grunted / hissed that the meeting was over*

g. 'move while causing the sound to be emitted'
e.g., *the rope whistled / hissed through the air*

h. 'cause to emit the sound'
e.g., *the postman rang / buzzed the doorbell*

i. '(of a place) be full of the sound'
e.g., *the air whistled / hummed with bullets*

[5]Selectional restrictions are sometimes violated under particular conditions. For instance, in the context of a science fiction movie it might be possible to say *The man bleeped.* In general, acceptable violations of selectional restrictions arise (i) in contexts where certain beliefs about the word are suspended or modified or (ii) through the operation of specific lexical processes such as metonymy or coercion, whose operation crucially takes advantage of the underlying selectional constraints (Pustejovsky, 1990). Therefore, violations of selectional restrictions should not be taken as evidence for the point of view that stating the restrictions in the first place is futile.

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

This set of senses was identified because of properties that differentiate them from other senses (selectional restrictions on arguments, expression of arguments, other syntactic properties). Some of these properties will be discussed further in the remainder of this section. Except that the basic sense of a verb of sound is listed first, no particular importance should be attached to the order of presentation of these senses.

What is important for this chapter is that there is a set of senses that can be manifested by verbs belonging to a particular semantic class, since their existence can be exploited in the design of a lexical knowledge base as discussed in Section 3.5. It is likely that further study will show that some of the senses distinguished here might prove to overlap with or be subsumed under others, but as a first approximation it seems preferable to distinguish too many senses than too few. Senses (4b)–(4f), in particular, could use further refinement, possibly using corpus-based lexicography techniques of the type described by Atkins (this volume) and Fillmore and Atkins (in press).

A full account of the extended senses of a verb should identify the factors that allow that verb to show a specific extended sense, whether or not it is actually attested. Some of the senses identified here may be interdependent; that is, their existence may depend on the existence of other senses. Although some factors that determine the availability of particular extended senses are identified here, a full study of this issue is outside the scope of this chapter.

Several of the extended senses of verbs of sound are available only to verbs denoting sounds that can be made by animates, as in (4b)–(4f). However, even when the sound is emitted by an inanimate entity, verbs of sound in their basic sense, as well as in some of their extended senses, are not necessarily verbs of communication. That is, in some senses they simply describe the emission of a sound by some entity without any intent to communicate something through the emission of this sound. Those verbs that may take animate subjects may denote the intentional emission of a sound, and the sound may be emitted with communicative intent. Sometimes, as in extended sense (4b) above, the sound is simply intended as a reaction directed at a certain target. The target may be expressed in an *at* phrase, as in "*A marmot whistled at us and withdrew*" or "*they didn't whistle at her*". Although the target is the intended recipient of the sound, there is no indication that the actual emitter of the sound succeeds in conveying anything to the target; it is in this respect that verbs of sound are not necessarily verbs of communication in all their senses.

However, among the extended senses of verbs of sound that allow human subjects, there is often a sense that might be characterized as in (4f): 'communicate verbally by emitting a sound'. On this sense, which is not typically observed with the verb *whistle* since physically a whistle precludes speech, the verbs denote what Zwicky (1971: 223) calls "intended acts of communication by speech'. When used as verbs of manner of speaking, these verbs take a variety of sentential complements.

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

(5) a. Susan bellowed that the party would be tonight.
b. Susan bellowed how to avoid the crowd.
c. Susan bellowed for me to come.
d. Susan bellowed to Rachel to come.
e. Susan bellowed to come.
f. Susan bellowed, 'Come'.
g. 'Come', bellowed Susan.

As (5d) illustrates, verbs of sound may also take a *to* phrase indicating the recipient of the communication on this sense. The use of a *to* phrase with this sense contrasts with the use of the *at* phrase found in sense (4b), where there is no necessary implication that the target is a cooperating participant in the event. See Zwicky (1971) and Mufwene (1978) for a detailed exposition of the properties of verbs of sound when used as verbs of manner of speaking.[6]

Many verbs of sound have extended senses as verbs of motion, as in "*A rocket whistled by, missing the hill . . .*", where *whistle* means 'move while causing a whistling sound to be emitted'. This type of meaning extension is relatively productive across verbs of sound, although it is restricted to sounds that are made when a physical entity – whether it be human, animate, or inanimate – moves (and is not manifested with sounds such as *grunt,* which can only be made by the vocal tract). The verbs showing this extended sense are those verbs that can take inanimate subjects (although animate subjects are possible in this extended sense precisely when they make these sounds by moving rather than through their mouth or nose).

Some of those verbs of sound that denote sounds that can be emitted by inanimate objects, including those in (6), manifest a causative meaning, 'cause to emit a particular sound'. This possibility is open to those sounds that can be brought about by an outside cause and are not only emitted due to the inherent properties of the entity producing them.

(6) bang, beat, beep, buzz, clatter, crack, crunch, jangle, jingle, knock, rattle, ring, roar, rustle, toll, toot, twang, . . .

Many of the extended senses manifested by verbs of sound are not limited to the members of this class. Certain sounds are associated with particular emotions and are emitted to express the associated emotion. For instance, a whistle can indicate surprise or admiration, whereas a hiss might indicate anger. These associations give rise to extended sense (4e), characterized as 'express (an emotion) by emitting the sound'. This extended sense is really an instantiation of the more general meaning 'express (an emotion) in some way'; one way of expressing an emotion is through the emission of a sound that expresses that emotion.

[6]Although the complement-taking properties of verbs of sound used as verbs of manner of speaking parallel those of simple verbs of speaking such as *say,* there are some differences between the two in syntactic behavior. Specifically, verbs of manner of speaking, unlike the verb *say,* are what are known as 'non-bridge' verbs (Erteschik, 1973; Stowell, 1981, among others).

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

This extended sense is found not only with verbs of sound but also with verbs of gestures and signs such as *nod* (*nod one's approval*) or *frown* (*frown one's dismay*), as the examples below illustrate. Gestures and signs, like certain sounds, can also be used to convey certain emotions.

(7) a. "*the men in the audience whistled their appreciation of her figure*"
b. Marlowe roared approval in his characteristic way, . . .
(D. Sayers, *The Documents in the Case,* Perennial, 1987, p. 94)

(8) a. She sniffed her disapprobation of the police in general . . .
(M. Grimes, *The Man with a Load of Mischief,* 1981, p. 132)
b. Instead, he snorted derision.
(L. Grant-Adamson, *The Face of Death,* Scribner's, 1986, p. 108)

The process that is involved in creating many of these extended senses is also responsible for creating extended senses of English verbs associated with other semantic classes. This process, called 'lexical subordination' by Levin and Rapoport (1988), involves 'subordinating' the meaning of the verb associated with the basic sense under an additional component of meaning to give the extended sense; typically the meaning associated with the basic sense is a means of bringing the additional meaning about (Jackendoff, 1990; Levin and Rapoport, 1988; as well as Talmy, 1985). For instance, in the extended sense described as 'express (an emotion) by emitting the sound', the basic sense 'emit a sound' is subordinated under the meaning 'express'. The process of lexical subordination is also manifested outside of the verb of sound class. For example, the cooking verb *bake,* which is basically a change of state verb, has an extended sense as a creation verb, with the meaning 'create by means of change of state *bake*' (Atkins, Kegl, and Levin, 1988; Ilson and Mel'čuk, 1989).

Another sense of *whistle,* '(of a place) be full of the sound', is exemplified in *The air whistled with bullets,* which appears to be a near paraphrase of the basic sense of the verb *whistle* found in *The bullets whistled in/through the air.* In this extended sense, the sound is attributed to the location in which it is emitted, with the actual entities emitting the sound expressed in a *with* phrase. Such pairs of sentences are instances of the intransitive variant of the locative alternation characteristic of such verbs as *swarm* (*The garden is swarming with bees/Bees are swarming in the garden*).[7] This sense of *whistle* represents an extended sense of a verb of sound that is manifested very productively across the whole class.

[7]Not every use of a verb of sound in its basic sense can be paraphrased using this extended sense; a variety of factors determine when such pairs are possible. For an extensive discussion of the properties associated with this sense see Salkoff (1983). This sense is not present in any of the dictionaries I have examined, nor are there citations for it among the uses of *whistle* in the Cobuild Corpus. The alternative expressions of the arguments of *swarm* manifested in *The garden is swarming with bees/Bees are swarming in the garden* are the intransitive version of the locative alternation found with verbs like *spray* or *load* (*spray the wall with paint/spray paint on the wall*). This alternation is discussed by Jeffries and Willis (1984), Rappaport and Levin (1988), Schwartz-Norman (1976), among many others.

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

In their basic sense, the verbs of sound are part of a larger class of verbs of stimulus emission. This larger class includes three additional major subclasses, which can be distinguished by the nature of the stimulus: smell, light, or substance. Representative verbs from these subclasses are listed below:

(9) LIGHT: flicker, gleam, glitter, glow, shimmer, shine, sparkle, twinkle
(10) SMELL: reek, smell, stink
(11) SUBSTANCE: bubble, gush, ooze, puff, spew, spout, squirt

Many of the properties of verbs of sound are shared with the members of the larger class of verbs of stimulus emission. For instance, all are single argument verbs, taking only a subject in their basic sense. Members of all the subclasses are open to the causative extended sense described above (e.g., *The driver flashed the headlights*), although because of the constraints on this meaning only a limited number of verbs in any subclass exhibit this sense. A more general form of the extended sense '(of a place) be full of the sound' – '(of a place) to be full of the stimulus' – is also manifested by verbs from other subclasses of verbs of stimulus emission (e.g., *The sky twinkled with stars*). However, the verbs of sound have a wider range of extended senses than other verbs of stimulus emission: since only sounds can be used for communication, only the verbs of sound have the extended senses associated with communication.

This section has exemplified the close correlation between the basic semantic class of a verb and the way it behaves in the language (expression of arguments, selectional restrictions, extended senses, etc.). As the next section demonstrates, again with verbs of sound, research into lexical organization clarifies not only the basic properties of the verbs, as described in this section, but also the necessary organization and components of lexical entries for verbs of particular types.

3.4 The design of a lexical knowledge base entry for verbs of sound

A lexical knowledge base should make explicit what a native speaker knows about a word. To ensure that a full and systematic presentation of relevant facts is provided, this presentation must be informed by an understanding of lexical organization. Leaving aside phonological and morphological information, a verb entry must give information about each of its sense categories. For each of these, it must give information and examples, when appropriate, about:[8]

(12) a. Semantic class, aktionsart, and arguments
b. Selectional restrictions on the arguments
c. Syntactic properties and expression of arguments
d. Related extended senses

[8]Many people have discussed the issue of what types of information should be included in an ideal dictionary entry. For another recent list of such information, which overlaps substantially with the list provided here, see Hudson (1988).

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

e. Morphologically related nouns, adjectives, and verbs
f. Related idiomatic uses
g. Collocates
h. Pragmatic force

It is worth noting that only some of these facts about a verb can be learned from its dictionary entry. Existing dictionaries might provide a starting point for the construction of a lexical knowledge base entry for a verb, but they do not provide all the information that such a knowledge base should ideally make available.

This section describes lexical knowledge base entries that handle the first four types of facts. The entries developed in this section and the next section again draw on the study of verbs of sound described in Atkins and Levin (1988), which tries, by studying certain aspects of this class, to take the first steps toward defining some of the structure of a template for a lexical knowledge base entry for such verbs.

An essential part of the lexical knowledge base entry for any lexical item is a description of its meaning. Prototypical dictionary definitions consist of two parts, a genus word and differentiae; the genus word is that part of its meaning that a word shares with other hyponyms of the particular genus term selected, and the differentiae are properties that distinguish that word from these co-hyponyms, and this particular sense of the word from other senses of the same word. The discussion of the meaning of verbs of sound in their basic sound emission sense shows that definitions of their meaning conform to this scheme. With verbs of sound, the genus is 'emit a sound' and the differentiae describe the type of sound produced (low, high, loud, soft, shrill, etc.), the means of producing the sound (by blowing, vibration, impact, etc.), what produces the sound (a person, animal, device), and, if a person produces the sound, whether or not an instrument is used to produce it. The first two of these, sound-type and means, were discussed in the previous section. They could be indicated in a lexical knowledge base entry as follows:[9]

	WHISTLE	GRUNT	BLEEP
genus	emit sound	emit sound	emit sound
diff-1: sound-type	shrill	low	high
diff-2: means	by blowing	by vibration	electronically

[9]The lexical knowledge base entries given here are intended only to be suggestive, and actual entries might look rather different. The use of items such as 'shrill' or 'by blowing' as fillers of the differentiae slots may seem a cause for concern to some readers, since they are not themselves 'primitive' notions. However, the concern in this chapter is with the representation of linguistically relevant information, and, from this point of view, it is the existence of specific components of meanings (the nature of the differentiae slots in an entry), rather than the specific instantiation of the components (the fillers of the slots), that is important. For instance, as mentioned, earlier, a verb whose meaning involves a notion of means or manner demonstrates certain linguistic properties independent of the specific means or manner associated with that verb.

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

As discussed in the previous section, the verbs of sound differ in the types of subjects they allow; therefore, the knowledge base will need to identify the potential subjects of these verbs. For simplicity, the possibilities are indicated using features to encode the possibilities.

	WHISTLE	GRUNT	BLEEP
genus	emit sound	emit sound	emit sound
diff-1: sound-type	shrill	low	high
diff-2: means	by blowing	by vibration	electronically
select-restrns subj	+concrete (+/−animate) (+/−human)	+animate (+/−human)	+concrete −animate

There is one additional differentia that was not discussed in the previous section. When a person produces a sound, an instrument may or may not be involved. A person can whistle either by blowing air through the lips or by forcing the air through a device (*a whistle*), but a person does not use a device to grunt. Electronic devices bleep, but people typically do not, so the use of an instrument is not relevant to the verb *bleep*. However, because the use of an instrument is pertinent to many verbs of sound, this possibility, which is really a refinement of the means differentia, must be incorporated into a lexical knowledge base entry.

	WHISTLE	GRUNT	BLEEP
genus	emit sound	emit sound	emit sound
diff-1: sound-type	shrill	low	high
diff-2: means	by blowing	by vibration	electronically
diff-3: instrument	+/−device	—	—
select-restrns subj	+concrete (+/−animate) (+/−human)	+animate (+/−human)	+concrete −animate

Finally, it is important for a verb's entry to contain information regarding its argument-taking properties, and the syntactic expression of these arguments. As discussed in the previous section, in their basic sense, verbs of sound are intransitive. There are a variety of ways in which the syntactic argument-taking properties can be expressed. The entries below represent these properties via a subcategorization frame (Chomsky, 1965).

	WHISTLE	GRUNT	BLEEP
genus	emit sound	emit sound	emit sound
diff-1: sound-type	shrill	low	high
diff-2: means	by blowing	by vibration	electronically
diff-3: instrument	+/−device	—	—

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

	WHISTLE	GRUNT	BLEEP
select-restrns subj	+concrete (+/−animate) (+/−human)	+animate (+/−human)	+concrete −animate
subcategorization	+[—]	+[—]	+[—]

Alternatively, the argument-taking properties could be represented in terms of the various argument structure representations that have replaced subcategorization frames in recent work in syntax. One possibility is to use the notions of external, internal, direct, and indirect arguments (Williams, 1981; Rappaport, Levin, and Laughren, 1988, among others); on this approach, verbs of sound in the sound emission sense would be characterized as taking only an external argument (the analogue of a subject). Alternatively, the notion of argument classification recently introduced into Lexical Functional Grammar (LFG) could be used (Bresnan and Kanerva, 1989).

In order to introduce the structure of the lexical knowledge base, the focus has been on a single major sense of the verb *whistle,* setting out similarities and differences between this verb and other verbs of sound. Since the goal of this chapter is not to present an exhaustive analysis of these verbs, many additional facts that an ideal lexical knowledge base entry must include for such verbs, even in the sense represented in the template, have been omitted.

3.5 The contribution of on-line resources to a lexical knowledge base

Although the design of an effective and successful lexical knowledge base should be informed by work in linguistics, the knowledge base itself does not need to be completely built up from scratch. The construction of such a knowledge base should take advantage of existing on-line resources, both MRDs and text corpora, extracting whatever information possible from them. Having outlined the contribution that linguistics can make to the design of lexical knowledge base entries of verbs of a particular type, this section will consider how linguistics can also enhance the contribution that on-line resources can make to the creation of lexical knowledge base entries. Specifically, a comprehensive model of the lexicon will enable the designers of a knowledge base to predict the type of facts about a verb that might be expected to be found in dictionaries and corpora once a verb's class has been established. Using such a model, the knowledge base designers can define accurately the type of information that might be expected to be retrieved.

As Section 3.3 makes clear, knowing a verb's semantic class membership is crucial to understanding the properties of a given word and to determining its relation to other words. To state this differently, certain aspects of a verb's meaning serve as a pointer to its place in the organizational scheme of English verbs. Once this place is identified, various properties of the verb can be determined. For example, turning again to verbs of sound, it is clear that there are

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

certain properties that would be expected to be recorded in the lexical knowledge base entry of any verb of this type. These properties can be used to construct a template lexical knowledge base entry for all verbs of sound. An instance of this template entry can then be instantiated for each verb using on-line lexical resources as a source of information, with the template entry-serving in some sense as a guide to the required information.

The partial lexical knowledge base entry for the verb *whistle* presented in the previous section identified certain properties that this verb shares with other verbs of sound and that, therefore, must be taken into account in the design of a knowledge base entry for such verbs. For example, all verbs of sound have a basic sense that can be characterized as emitting or producing a particular sound, and many of them will also have some or all of the extended senses listed in Section 3.3. Information about each of these senses must be included in the knowledge base.

The properties common to all verbs of sound can be incorporated directly into a template lexical entry generated for such verbs. The entry for the basic sense ('emit a sound') must indicate that the genus is of a particular nature ('emit sound') and that the possible differentiae are of predictable types (sound-type, means, instrument), their exact nature depending on the properties of each individual verb. The template entry must also specify that such verbs require a non-abstract subject (with individual verbs imposing additional restrictions). The following is a possible template for the basic sound emission sense of verbs of sound:

genus	emit sound
diff-1: sound-type	
diff-2: means	
diff-3: instrument	
select-restrns subj	−abstract
subcategorization	+[—]

The lexical knowledge base entry for a specific verb of sound in this basic sense can be completed by the addition to this template entry of the differentiae and, if required, further selectional restrictions on the subject. On-line resources would be used to furnish this data in order to complete the description, as well as to provide corroboration for the existence of specific options for specific words.

Similar template entries will have to be constructed for the extended senses of the verbs of sound. The existence of these senses depends in part on certain properties of the basic sense. To the extent that the factors determining the availability of each extended sense are known, the extended senses can be linked to the template entry of the basic sense in such a way that if information obtained about the basic sense precludes a certain extended sense, there would be no need to look for evidence that it exists. For example, a verb of sound can only show an extended sense as a verb of verbal communication if it can denote a sound made

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

by a human; therefore, if the verb under consideration does not take a human subject, the potential existence of this extended sense does not need to be posited for that verb.

The use of a template entry makes it possible to define the type of information being sought from on-line resources, whether MRDs or text corpora, before an attempt is made to extract data. The template defines the specific queries that are being made for any particular word. Given the structured basis of dictionary entries, this type of semi-automatic lexical knowledge base construction is more likely to be successful than attempts to create lexical knowledge bases in some other ways.

3.5.1 The contribution of dictionaries

A model of lexical organization can suggest the properties that a given verb might be expected to show, and hence it can give rise to expectations about what properties might be illustrated in the dictionary entry of that verb. These expectations will help make sense of the information in dictionary entries, allowing information in entries to be exploited beyond what might have been possible in the absence of such a model. Such a model can help overcome certain shortcomings that arise because words in existing dictionaries are treated in isolation. Existing dictionary entries do not fit words into the larger organizational scheme of the English lexicon, nor do they impose any internal structure on the sense categories given in their entries. Learners' dictionaries often present properties that reflect a verb's semantic class through the use of a combination of grammar codes, stylized definitions, and example sentences, but information about the class itself is implicit in an entry (Atkins, Kegl, and Levin, 1986). A model of lexical organization can contribute such information, which is a central component of what a native speaker knows about words.

There are various types of dictionaries, and every dictionary has its own contribution to make to the construction of a lexical knowledge base due to its individual nature and design. Among other things, a dictionary entry includes some or all of the following elements, depending on its individual nature and design: a definition, prototypically structured into genus and differentiae sections; examples of usage, with or without glosses or translations; and metalinguistic information relating to subcategorization and selectional restrictions. Differences among dictionaries should be exploited to obtain as much information as possible from them.

Monolingual collegiate-sized dictionaries of a language can contribute information about a word's possible senses, and, for each sense, the relevant semantic class and the genus and differentiae that constitute its definition. Entries in the dictionaries designed for the foreign learner of English typically contain very specific grammar information, usually in the form of codes, as well as many examples of use of the headword, carefully constructed in order to enrich the

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

information provided by the entry. Some MRDs also contain explicit indicators of semantic domains and of selectional restrictions.

Certain information required for filling in the template entry for a verb of sound is available in existing dictionaries, where, although this information is not overtly identified, it is often presented in a reasonably structured way, a prerequisite for automatic extraction and transfer into the lexical knowledge base entry. Dictionary definitions frequently use formulaic expressions that allow information of certain types to be identified relatively easily (Fox et al., 1986; Markowitz et al., 1986, among others). For instance, various researchers have shown that genus terms can be identified relatively successfully (Amsler, 1980; Byrd et al., 1987; Chodorow et al., 1985; Klavans, 1988; Markowitz et al., 1986, among others), allowing the initial identification of a verb's semantic type. Once a verb of sound has been identified in this way, it should then be possible to use automatic methods to pinpoint within a parsed MRD entry the means differentia, since it is usually expressed in a *by* phrase, as in sense 1 of the *Collins English Dictionary* (CED) (Hanks, 1986) entry for the verb *whistle,* which includes the phrase 'by passing breath through a narrow constriction'. Similarly, the contents of examples (e.g. in CED *whistle* 1 "*he whistled a melody*") or the metalanguage (e.g. in CED *whistle* 3 'of a kettle, train etc.' or CED *whistle* 5 'of animals, esp. birds') provide more detailed information about the selectional restrictions associated with the various senses of the word.

Certain extended senses can also be identified and extracted fairly easily. For instance, the sense of *whistle,* characterized as 'express (an emotion) by emitting the sound', can be found by searching for definitions including the string 'express'. In fact, a search of *The Longman Dictionary of Contemporary English* (Procter et al., 1978) for verbs having both transitive and intransitive uses whose definition includes the word 'express' yields a number of verbs of sound, including *babble, cluck, giggle, growl, howl, moan, purr, roar, scream, snort.* The extended sense of a verb of sound as a verb of motion ('move while causing the sound to be emitted') can also be identified, as it is occasionally signaled in the dictionary entries of such verbs. However, an examination of entries for verbs of sound reveals that this sense has posed a problem for lexicographers, who are not always sure whether to define this sense as a sound or a movement, leading to inconsistencies in and across dictionaries in the treatment of this sense (sense 1c of the *Webster's Ninth* [Mish, 1986] entry for the verb *whistle* is 'to make a shrill clear sound, esp. by rapid movement ⟨the wind whistled⟩'; sense 4 in CED is 'to move with a whistling sound caused by rapid passage through the air'). Although these inconsistencies might complicate the process of identifying this sense, once the source of the inconsistency is understood, they need not prevent its automatic identification in a MRD, assuming that the appropriate clues are searched for.

Bilingual dictionaries are also a rich source of material for monolingual lexical knowledge base entries. A monolingual dictionary intended for native speakers, such as the CED or *Webster's Ninth,* is obliged to contain real definitions; there-

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

fore, it provides explicit information regarding genus terms and differentiae. The parallel part of a bilingual dictionary entry contains target language equivalents of the headword. Thus a bilingual dictionary, although it does not provide genus terms or differentiae, does distinguish between two non-identical target language equivalents. This differentiation is often achieved by indicating selectional restrictions on the subject and/or object of verbs, and these indications in turn provide systematic information for the knowledge base. As an example consider section 3 of the *Collins-Robert English–French Dictionary* (Atkins et al., 1987) entry for *whistle:*

(13) 3 *vi* [*person*] siffler, (*tunefully, light-heartedly*) siffloter; (*blow a–*) donner un coup de sifflet; [*bird, bullet, wind, kettle, train*] siffler . . .

In this part of the entry, the material in square brackets sets out typical subjects of the verbs that follow, showing that all are possible subjects of the English headword *whistle*. This information may also be found in monolingual dictionaries but is often implicitly held in the form of examples. The explicit treatment that it often receives in dictionaries for the foreign learner renders such works a valuable source of material for the semi-automatic acquisition of information for a lexical knowledge base.

Bilingual dictionaries can also be used to provide information about whether certain verbs of sound have extended senses as verbs of motion. French differs from English in not allowing verbs of sound, and, in fact, verbs from a variety of other classes, to become verbs of motion (Carter, 1988; Talmy, 1985; Levin and Rapoport, 1988). For example, in French, *to roar down the driveway* would be expressed as 'to go down the driveway roaring'. Consequently, an English–French dictionary is likely to provide information about this sense. As discussed by Boguraev et al. (1989), this information does, in fact, tend to be provided in a fairly structured way, allowing the verbs that show this sense to be readily identified.

3.5.2 The contribution of text corpora

Text corpora of both written and spoken language also provide a rich resource that has much to offer to the construction of a lexical knowledge base. Tools for taking advantage of these resources are only now beginning to be developed (Church and Hanks, 1989; Hindle, 1989; Zernik, 1989b,c, among others), but it is clear that corpora will eventually offer much more detail about the properties of words than can be found in their dictionary entries, although dictionaries will still be essential to any attempt to construct a lexical knowledge base automatically.

Text corpora are an excellent source of information about the selectional restrictions on the arguments of verbs. The information provided by corpora would be enhanced if the corpus citations were tagged with a sense from the lexical

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

knowledge base, but it is possible that the selectional information derived from the corpus might itself provide a basis for the identification of the senses of the occurrences of a given word in a corpus. Also interesting are the examples of the possible uses of a word that a corpus can provide, since these might exemplify some of the syntactic properties of that word that are necessary to pigeonhole it appropriately. Even if a word's semantic type has already been identified, these facts could be used at least to confirm a classification previously reached on the basis of the genus term in a dictionary definition, and possibly also to fill in additional properties that were not noted in the dictionary. For instance, the citations in the Cobuild Corpus show that the verb *whistle* is found with both *at* and *to* phrases.

(14) "*A man whistled AT her*"
"*we hate being whistled AT in the street*"
(15) "*He whistled TO his partner*"
"*I whistled TO her*"

The presence of *to* phrase complements to the verb, for example, could be used either to confirm or to determine that *whistle* does indeed have an extended sense as a verb of verbal communication, whereas the presence of *at* phrase complements signals the 'emit a sound as a reaction' sense. Similarly, a corpus could be used to determine whether a verb of sound has an extended sense as a verb of motion by searching for instances of intransitive uses of the verb followed by directional phrases, such as the following.

(16) "*A shot whistled past Bond's head.*"
"*A rocket whistled by, missing the hill . . .*"

3.6 Conclusion

This chapter explores the use of results of linguistic research into the lexicon in the construction of a lexical knowledge base by showing how the information regarding a word is structured. The chapter proposes that a linguistically-motivated template entry should be designed for each type of word, and that these entries might then be instantiated for individual words. The use of such template entries does not preclude the use of on-line dictionaries or text corpora to aid the process of lexical acquisition. Rather the information available in such on-line resources is essential for building lexical knowledge base entries for specific verbs from the appropriate template. However, the template entries will guide the search for information within these resources, allowing them to be exploited as fully as possible.

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

References

Amsler, R. A. (1980). *The Structure of the Merriam-Webster Pocket Dictionary,* Ph.D. thesis, University of Texas, Austin, TX.

Atkins, B. T., Duval, A., Milne, R. M., et al. (1987). *Collins-Robert English-French Dictionary.* London and Glasgow: Collins Publishers. Second edition.

Atkins, B. T., Kegl, J., and Levin, B. (1986). 'Implicit and Explicit Information in Dictionaries', *Advances in Lexicology,* Proceedings of the Second Conference of the Centre for the New OED, University of Waterloo, 45–63.

Atkins, B. T., Kegl, J., and Levin, B. (1988). 'Anatomy of a Verb Entry: from Linguistic Theory to Lexicographic Practice', *International Journal of Lexicography* 1: 84–126.

Atkins, B. T. and Levin, B. (1988). 'Admitting Impediments', *Information in Text,* Fourth Annual Conference of the Centre for the New OED, University of Waterloo, 97–113. Also to appear in Zernik, U. (ed.). *Using On-line Resources to Build a Lexicon.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Boguraev, B. and Briscoe, T. (eds.). (1989). *Computational Lexicography for Natural Language Processing.* Harlow: Longman.

Boguraev, B., Byrd, R. J., Klavans, J. L., and Neff, M. (1989). 'From Structural Analysis of Lexical Resources to Semantics in a Lexical Knowledge Base', position paper prepared for the Workshop on Lexicon Acquisition, IJCAI, Detroit.

Bresnan, J. and Kanerva, J. (1989). 'Locative Inversion in Chichewa: A Case Study of Factorization in Grammar', *Linguistic Inquiry* 20: 1–50.

Byrd, R. et al. (1987). 'Tools and Methods for Computational Lexicology', *Computational Linguistics* 13: 219–240.

Carter, R. J. (1976). 'Some Constraints on Possible Words', *Semantikos* 1: 27–66.

Carter, R. (1988). *On Linking: Papers by Richard Carter* (edited by Levin, B. and Tenny, C.), Lexicon Project Working Papers 25, Center for Cognitive Science, MIT, Cambridge, MA.

Chodorow, M. S., Byrd, R. J., and Heidorn, G. E. (1985). 'Extracting Semantic Hierarchies from a Large On-line Dictionary', *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics:* 299–304.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Church, K. C. and Hanks, P. (1989). 'Word Association Norms, Mutual Information and Lexicography', *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics:* 76–83.

Erteschik, N. (1973). *On the Nature of Island Constraints,* Ph.D. thesis, MIT, Cambridge, MA.

Fellbaum, C. (1990). 'English Verbs as a Semantic Net', *International Journal of Lexicography* 3.

Fillmore, C. J. (1968). 'The Case for Case'. In Bach, E. and Harms, R. T. (eds.). *Universals in Linguistic Theory.* New York: Holt, Rinehart and Winston, 1–88.

Fillmore, C. J. and Atkins, B. T. (in press). 'RISK: The Challenge of Corpus Lexicography'. In Atkins, B. T. and Zampolli, A. (eds.). *Computational Approaches to the Lexicon: Automating the Lexicon II.* Oxford: Oxford University Press.

Foley, W. A. and Van Valin, R. D., Jr. (1984). *Functional Syntax and Universal Grammar.* Cambridge: Cambridge University Press.

Fox, E. A. et al. (1986). 'Building the CODER Lexicon: *The Collins English Dictionary* and its Adverb Definitions', Technical Report 86–23, Department of Computer Science, Virginia Tech, Blacksburg, VA.

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

Gruber, J. S. (1976). *Lexical Structures in Syntax and Semantics*. Amsterdam: North-Holland.

Hale, K. L. and Keyser, S. J. (1986). 'Some Transitivity Alternations in English', Lexicon Project Working Papers 7, Center for Cognitive Science, MIT, Cambridge, MA.

Hale, K. L. and Keyser, S. J. (1987). 'A View from the Middle', Lexicon Project Working Papers 10, Center for Cognitive Science, MIT, Cambridge MA.

Hanks, P. (ed.). (1986). *Collins English Dictionary*. London and Glasgow: Collins Publishers. (CED)

Hindle, D. (1989). 'Acquiring Disambiguation Rules from Text', *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics:* 118–125.

Hudson, R. (1988). 'The Linguistic Foundations for Lexical Research and Dictionary Design', *International Journal of Lexicography* 1: 287–312.

Ilson, R. and Mel'čuk, I. (1989). 'English *Bake* Revisited', *International Journal of Lexicography* 2: 325–346.

Jackendoff, R. S. (1983). *Semantics and Cognition*. Cambridge, MA: MIT Press.

Jackendoff, R. S. (1990). *Semantic Structures*. Cambridge, MA: MIT Press.

Jeffries, L. and Willis, P. (1984). 'A Return to the Spray Paint Issue', *Journal of Pragmatics* 8: 715–729.

Katz, B. and Levin, B. (1988). 'Exploiting Lexical Regularities in Designing Natural Language Systems', *Proceedings of the 12th International Conference on Computational Linguistics:* 316–323.

Klavans, J. (1988). 'Building a Computational Lexicon Using Machine-Readable Dictionaries', paper presented at the EURALEX Third International Congress, Budapest, Hungary.

Laughren, M. (1988). 'Towards a Lexical Representation of Warlpiri Verbs'. In Wilkins, W. (ed.). (1988), *Thematic Relations*. New York: Academic Press 215–242.

Levin, B. (1985). 'Introduction'. In Levin, B. (ed.). *Lexical Semantics in Review*, Lexicon Project Working Papers 1, Center for Cognitive Science, MIT.

Levin, B. (1989). 'Verbal Diathesis', Lexicon Project Working Papers 32, Center for Cognitive Science, MIT, Cambridge, MA.

Levin, B. (in press) 'Approaches to Lexical Semantic Representation'. In Walker, D., Zampolli, A., and Calzolari, N. (eds.). *Automating the Lexicon*. Oxford: Oxford University Press.

Levin, B. (unpublished ms.) *Towards a Lexical Organization of English Verbs*, Northwestern University, Evanston, IL.

Levin, B. and Rapoport, T. (1988). 'Lexical Subordination', *Proceedings of the 24th Annual Meeting of the Chicago Linguistic Society*, 275–289.

Levin, B. and Rappaport, M. (1989). 'Approaches to Unaccusative Mismatches', *Proceedings of the 19th Annual Meeting of the North-Eastern Linguistics Society:* 314–328.

Levin, B. and Rappaport Hovav, M. (in press: a). 'The Lexical Semantics of Verbs of Motion: The Perspective from Unaccusativity'. In Roca, I. (ed.). *Thematic Structure: Its Role in Grammar*. Dordrecht: Foris.

Levin, B. and Rappaport Hovav, M. (in press: b). 'Wiping the Slate Clean: A Lexical Semantic Exploration', *Cognition*.

Marantz, A. P. (1984). *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press.

Markowitz, J., Ahlswede, T., and Evens, M. (1986). 'Semantically Significant Patterns in Dictionary Definitions', *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics:* 112–119.

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press

McCawley, J. (1986). 'What Linguists Might Contribute to Dictionary-making if They Could Get Their Act Together'. In Bjorkman, P. C., and Raskin, V. (eds.). *The Real-World Linguist*. Norwood, NJ: Ablex, 1–18.

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., and Miller, K. (1990). 'Introduction to WordNet: An On-Line Lexical Database', *International Journal of Lexicography* 3.

Miller, G. A., Fellbaum, C., Kegl, J., and Miller, K. (1988). 'WORDNET: An Electronic Lexical Reference System Based on Theories of Lexical Memory', *Revue Quebecoise de Linguistique* 17: 181–211.

Miller, G. A. and Fellbaum, C. (in press). 'Semantic Networks of English', *Cognition*.

Mish, F. C. (ed.). (1986). *Webster's Ninth New Collegiate Dictionary*. Springfield, MA: Merriam-Webster.

Mufwene, S. (1978). 'English Manner-of-speaking Verbs Revisited'. *Papers from the Parasession on the Lexicon*. Chicago: Chicago Linguistic Society, 278–289.

Neff, M. and Boguraev, B. (1989). 'Dictionaries, Dictionary Grammars and Dictionary Entry Parsing', *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*.

Procter, P., et al. (eds.). (1978). *Longman Dictionary of Contemporary English*. London: Longman Group. First edition. (LDOCE)

Pustejovsky, J. (1990). 'Towards a Generative Lexicon', *Computational Linguistics* 17.

Rappaport, M. and Levin, B. (1988). 'What to Do with Theta-Roles'. In Wilkins, W. (ed.). (1988), *Thematic Relations*. New York: Academic Press 7–36.

Rappaport, M., Levin, B., and Laughren, M. (1988). 'Niveaux de representation lexical', *Lexique* 7: 13–32. To appear in English as 'Levels of Lexical Representation'. In Pustejovsky, J. (ed.). *Semantics and the Lexicon*. Dordrecht: Kluwer.

Salkoff, M. (1983). 'Bees Are Swarming in the Garden', *Language* 59: 288–346.

Schwartz-Norman, L. (1976). 'The Grammar of "Content" and "Container"', *Journal of Linguistics* 12: 279–287.

Stowell, T. (1981). *Origins of Phrase Structure*, Ph.D. thesis, MIT, Cambridge, MA.

Talmy, L. (1985). 'Lexicalization Patterns: Semantic Structure in Lexical Forms'. In Shopen, T. (ed.). *Language Typology and Syntactic Description 3*, Grammatical Categories and the Lexicon. Cambridge: Cambridge University Press, 57–149.

Wilkins, W. (ed.). (1988). *Thematic Relations*. New York: Academic Press.

Williams, E. (1981). 'Augment Structure and Morphology', *The Linguistic Review* 1: 81–114.

Zernik, U. (ed.). (1989a). *Proceedings of the First International Lexical Acquisition Workshop*, Detroit, Michigan.

Zernik, U. (1989b). 'Paradigms in Lexical Acquisition'. In Zernik, U. (ed.). *Proceedings of the First International Lexical Acquisition Workshop*, Detroit, Michigan.

Zernik, U. (1989c). 'Lexical Acquisition: Learning from Corpus by Capitalizing on Lexical Categories', *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, Detroit, Michigan.

Zwicky, A. M. (1971). 'In a Manner of Speaking', *Linguistic Inquiry* 2: 223–232.

https://doi.org/10.1017/CBO9780511659478.005 Published online by Cambridge University Press