

Lexical Rules is Italicized

**Stephen Helmreich
David Farwell**

Computing Research Laboratory
PO Box 30001/3CRL
New Mexico State University
Las Cruces, New Mexico

Fax: (505) 646-6218

Phone: (505) 646-2141

E-mail: david@crl.nmsu.edu/shelmrei@crl.nmsu.edu

1.0 Introduction

In this paper we present three types of arguments for a pragmatics-based approach to treating the phenomenon of lexical polysemy (lexical sense extension) and rebut three types of arguments that have been used against pragmatics-based approaches.

In particular, we argue that what have been claimed to be rules of lexical sense extension operating on lexical items are instead to be accounted for by a process of reference transfer by which the referents of lexical items (and phrases) can be used to refer to related things, as governed by pragmatic rules of conversation and contextual knowledge. By contextual knowledge we include that information provided by the surrounding language, by the salient facts about the non-linguistic context and by the general knowledge about the world that can be presumed to be shared by the participants in the conversation.¹

We except morphological rules, which apply to lexical items, but we argue that these rules are more productive and the semantics less precise than that suggested by the lexical rule approach, which often treat a single morphological process as a number of rules, subdivided according to the semantic category of the base or derived item.

The structure of the paper is as follows. Section 2 provides some background material on lexical rules and their hypothesized grammatical status and contrasts this with our approach to such phenomena. Section 3 presents arguments against the lexical nature of sense extensions rules and for their status as reference transfer rules. These arguments show that the domain of polysemous usage is not lexical items, but rather their referents. Section 4 examines the arguments made for their lexical status, which we find wanting. We argue that it is indeed possible to separate semantic from morphological phenomena, that blocking or pre-emption can be explained pragmatically, and that, barring syntactic and cultural differences, the range of polysemous usage is not language specific. Section 5 provides a conclusion.

2.0 Background

As the field of NLP has matured, the nature of the problem of sense disambiguation for lexical items has become clearer (Copestake & Briscoe, 1991; Levin, 1993; Nirenburg 1996; Pustejovsky & Bergler, 1992;

1. In general, we prefer to classify context not on the basis of its type, but rather on its use. So there is a general context, consisting of everything, and a subset of that context which is relevant to the interpretation of an utterance.

Saint-Dizier & Viegas, 1995; to name just a few). First, the wide variety of different uses to which a single lexical item can be put has been better described. Second, the semantic relatedness of these uses is more thoroughly understood.

For example, the lexical item *newspaper* can be used to refer to a physical published object, a particular edition, a newspaper company, an editorial staff, an article in the newspaper, or a reporter for the newspaper. While this represents a wide variety of different ontological entities (people, physical objects, conceptual objects, corporations), it is clear that all of these entities are related to each other by the process of creating, producing, selling, buying and reading newspapers. Finally, the same set of uses can be found not just for *newspaper* but for other words referring to similar congeries of relationships: *journal*, *newsletter*, *The New York Times*, *tabloid* and so on.

One computational approach to dealing with this multiplicity of senses is through the use of lexical rules. Within this class would fall the Lexical Implication Rules (LIRs) of Ostler and Atkins (1991), the lexical rules of Copestake and Briscoe (1991), the Generative Lexicon of Pustejovsky (1995), and the ellipsis recovery procedures of Viegas and Nirenburg (1995). In all these approaches, derived senses are not listed in the lexicon, but are generated on the basis of rules that operate on a class of lexical items defined by a conjunction of semantic and syntactic/lexical information (e.g. LIRs) or on the basis of patterns in the semantic structure of the lexical item itself (Pustejovsky, Copestake and Briscoe).¹

In this paper, we take issue not with the computational claims of this approach (i.e., that it allows easier acquisition of lexical items or more effective disambiguation in context), but rather with its linguistic claims about the status of lexical rules within a formal description of a language.

It is our position that the wide variety of uses that any one lexical item may have is not a result of formal mechanisms within the grammar of the language, but rather follow from general knowledge about the things to which the lexical item is used to refer. Knowing how a lexical item may be used, then, is not so much a matter of knowing the language, but of knowing the subject matter under discussion and the cultural conventions of the discussants. In this respect, our approach aligns with that of Hobbs et al. (1987) and of Nunberg (1978) in regarding the process as one which is pragmatic and inferential in nature rather than lexical and semantic.

We cannot propose here a complete computationally tractable solution. It is likely that use of lexical rules will continue to be valuable in the construction of lexica and within NLP systems. However, it is not difficult to show how these senses might be derived, nor how they might then be selected in context. The difficulty is, of course, limiting this process to a computationally tractable level. However, as indicated above, our critique is of the linguistic claims, not the computational ones.

This position stems from our basic views about the nature of language. We distinguish the *rules of language* which include the syntactic, phonological, and morphological regularities within which language users operate, the interface *conventions of meaning* which relate the elements of the language system (lexical items, morphological patterns and syntactic structures) with something outside the system (e.g., use *pineapple* to refer to pineapples), and the language-independent rules of inference, principles of communication, and world-knowledge which make up the background against which linguistic utterances are interpreted.²

1. We find less antithetical to our viewpoint the emphasis on the computational value of lexical rules (Viegas et al., 1996) and on ellipsis recovery as triggered by syntactic/lexical cues and accomplished through contextual means (Viegas & Nirenburg, 1995).

Whatever the specific status of the conventions of meaning, we expect them to be vague, minimal, and idiosyncratic. It is these very qualities that make it unlikely, in our opinion, that these interface conventions actually function as part of the rules of language. We do not expect these interface conventions to be robust, universal, or specific enough to sustain the complex structure of semantic features required by most lexical rules. This is not to say that there is no reference to semantic information in any of the rules of the language. For example, the plural morpheme indicates that whatever is referred to is both countable and that there is more than one. For the most part, though, these are basic distinctions that reflect gross semantic properties and are only partially correlated with independently verifiable semantics, rather than features that distinguish small classes of lexical items, directly related to external semantic categories.

This position is not, however, that which lies behind the concept of lexical rule. Ostler and Atkins state, for example: “We contend that they [Lexical Implication Rules] are bona fide rules in the formal grammar of languages involved; it is wrong to see them as mere reflections, quasi-systematic effects of applying the formal structure of a language in a real-world context.” (Ostler & Atkins 1991:76).

Copestake and Briscoe state, “We argue that such processes [of morphological derivation and metonymic and metaphoric sense extension] must be treated as, in part, ‘linguistic,’.... Our account contrasts with that of, say, Hobbs et al. (1987) who posit an underspecified and impoverished lexical semantic representation which is enriched through an open-ended process of abductive or deductive reasoning with world (or domain) knowledge” (Copestake & Briscoe, 1991:88).

Pustejovsky is somewhat more careful in his claims about the grammatical nature of the generative lexicon, and at the same time, somewhat more ambivalent. On the one hand, he states that “grammatical distinctions are a useful metric in evaluating competing semantic theories” (1995:6) which may imply that semantics, and specifically lexical semantics, is grammatical in nature. On the other, he states that “the semantics of natural language should be the image of nonlinguistic conceptual organizing principles, whatever their structure” (1995:6) implying that semantics, and specifically lexical semantics, is extragrammatical in nature. Or again, in stating that “the representation of lexical semantics, therefore, should be seen as just one of many levels in a richer characterization of contextual structure,” (1996:7) he appears to be stating that lexical semantics is extragrammatical. At the same time, he appears to be stating that lexical semantics is grammatical in suggesting that “lexical items... have a richer typed semantic representation than conventionally assumed; semantic operations of lexically-determined type changing (e.g., type coercions) operate under well-defined constraints. Different subclasses of polymorphic operations are well-defined, each with independent properties and conditions on their applications.” (1995:57). Although not exactly lexical rules, his *lexical conceptual paradigms* (lcp's) (1996:72) function similarly.¹

2. It may be plausible as well to distinguish yet another variety of convention: conventions of use. In this category would fall generalizations such as *speakers of Australian English often use the term ‘Good day’ as an introductory greeting* or *speakers of German preferentially use a dative-subject construction with experiencer verbs to indicate sensations of heat or cold*. Those who regard all these rules and conventions as equivalent in type will not be persuaded by all our arguments here. But for them, there is only one system of knowledge, which contains both linguistic and world knowledge. They do not distinguish between types of knowledge, and so the claim that certain knowledge is of a particular subtype (lexical or ontological) has no significance.

1. They do, for example, allow for multiple references in the same predicative situation, something lexical rules don't. However, note that Pustejovsky's account of the differing referential properties of *book* and *newspaper* (for example, the fact that *newspaper* can be used to refer to the publishing organization while *book* in general cannot) hinges on small differences in the qualia of the lexical entry. Nunberg (1978) proposes more intuitively that a newspaper identifies its publisher in a way that a book does not.

Common to all these approaches, however, is the claim that lexical rules are necessary because of a crucial interaction between grammatical and semantic behavior that can only be accounted for by positing such combined syntactic/semantic rules.

In contrast, we suggest that there is no need for such rules in a description of the grammar of English, or, indeed, any natural language. Specifically, we propose that while syntactic and morphological features and rules will indeed be part of the grammar and part of one's knowledge of the language, the polysemic use of lexical items is entirely separate from the morphological or syntactic rules that may generate the forms and structures which are to be interpreted.

We outline briefly here how such an approach would work, following Nunberg (1978). Given that a lexical item, *X*, is known to be used frequently by speakers of language *Z* to refer to (some concept, activity, or object) *Y*, and that *Y* is known to be connected in some way to (concept, activity, or object) *W*, then lexical item *X* can be used to refer *W* in (relevant) context *C*, given certain constraints that derive from principles of effective communication. For example, among such constraints would be that any proposition about a referent, *R*, must be relevant to *C* and that the proposition to be predicated of *R* must be rationally predicatable of *R*. If these constraints do not hold for referent *Y*, but do hold for referent *W*, for example, then *X* may be used to refer to *W*, and not otherwise.

To use a concrete example, the lexical item, *Bach* is known to be generally used by speakers of English to refer to the composer, Johann Sebastian Bach. Bach is known to be a composer and composers are known to compose music. Therefore, the lexical item *Bach* can be taken to refer to Bach's music, given the relevant context: *I was listening to some Bach last night*. This approach has been computationally implemented on a small scale in the Metalel program (Iverson and Helmreich, 1992), a reworking of the meta5 program (Fass, 1988).

In the subsequent sections we argue for this pragmatics and knowledge-based approach and against a lexical approach to the problem of polysemy and sense extension.

3.0 Arguments for a pragmatics-based approach to the phenomena

In this section we present arguments in favor of treating polysemy as determined in context on the basis of world knowledge and principles of communication, and against treating it as lexical a lexical phenomenon. First, we present two types of putative sense extensions that cannot be accounted for lexically, but which can be accounted for under our approach. We argue that an approach which does not rely on lexical rules is therefore more general and thus preferable.¹ We then present evidence that the formulation of rules is best stated not in terms of a function from lexical items to lexical items, but in terms of a function from referent to referent, independent of the lexical items that are used to refer to those objects or concepts.

3.1 Non-lexical regular sense extensions

A lexical rule can only operate if its input and output can be explicitly stated (or at least if the relation between them is specifiable). However, there are innovative uses of language that can be accounted for only by inferencing in context,² since novel uses are necessarily non-conventional and a priori non-identi-

1. There is also a weak (Occam's razor) argument that if a pragmatic hypothesis can account for the data, then it must be accepted, as that account will stand independently of any language-specific account. That is, the rules of inference and the facts upon which they operate exist independently of any language-specific semantics. At a simpler level, the argument would conclude that lexical rules require the duplication in the lexicon of semantic information that is independently available in the ontology, and are therefore unnecessary.

fiable. Thus they cannot be the results of the operation of a lexical rule. As these novel uses parallel ordinary lexical rule sense extensions, we claim that our pragmatic reference transfer approach covers a broader variety of cases, and thus provide a more unified solution to the problem of lexical interpretation. We present briefly two such sense extensions: novel metaphors and novel metonymy.

3.1.1 Novel metaphors

This includes both novel elaborations of well-known metaphors, as suggested in of Lakoff & Johnson (1980) and Martin (1990) and completely new metaphors. We hesitate to provide any examples of novel metaphors, as they will quickly become semi-conventional, but it should be clear that such metaphors can be interpreted only through the use of non-linguistic (non-conventional) means. If these same means can account for extensions of conventional metaphors, it would seem reasonable to account for both classes in the same way. In this category (of novel elaborations of well-known metaphors) are such utterances as (1).

(1) The ideas were bouncing around in her head.

This utterance exploits a conventional metaphor for talking about beliefs, i.e., the mind (head) as (physical) container and ideas as the (physical) objects contained, and extends that in a novel way to describe intense mental activity (possibly either creative or chaotic).¹ To arrive at an interpretation of *were bouncing around*, one must both know about objects bouncing around inside containers (i.e., the usual sense of *bounce around*) and about the mental process of integrating new ideas into a consistent world view and infer (possibly) that what is being described is a situation in which the ideas are not getting integrated.

While it is true that this is a productive metaphor, so that most verbs of physical activity can be used to refer to mental activities, the actual communicated intent of the reference is not specifiable. For example, what it would mean for ideas to *congeal*, and how that would differ from ideas *freezing* is not determinable from information about the lexical items *congeal* and *freeze* but only from their use in a context that allows us to relate their literal use with what we know about mental activity. That is, we need to reason simultaneously within two non-linguistic domains to arrive at a meaning for *bounce around*. It is unlikely that any two people would come up with the same description (paraphrase) of what the expression means without further context. The point here is that since metaphors must be interpreted and interpretable within a specific context and without any assistance from some lexicalized metaphorical sense extension, the interpretation of *head* as mind can be treated in the same way without recourse to lexical rules.

Copestake and Briscoe (1991) mention that they regard the metaphoric extension involving use of a word denoting an animal to refer to humans as involving a core component that should be expressed as a lexical rule (1991:89). We note that there is no need for any novel syntactic apparatus to generate such sentences as (2).

(2) He is a pig.

We would argue strongly that such animal-metaphors are creative in that which aspect (realistic, stereotypical, or phonological) of the animal is applied to the human referent is determined by context. Thus (1)

2. We are not claiming that all reference transfers are computed during the act of interpretation. The criterion suggested by Grice (1975) for conversational implicature holds here, namely that it is not that it is computed, but that it is computable. Similarly, the arguments here against lexical rules apply whether or not the lexical rules are precomputed and stored or apply at run time.

1. Ostler and Atkins exclude such cases from their lexical rules, agreeing with our analysis.

could mean that he is messy, that he overeats, or that he is selfish and inconsiderate (each a stereotypical attribute of pigs). Which ones are appropriate depends on the context.

3.1.2 Novel metonymies

These occur frequently in speech and are readily understood given the particular context in which they occur. However, the terms involved in a specific metonymy are often not semantically related, but simply pragmatically related and, even more often, very contingently so. Moreover, the relation between original and transferred referent is quite complex, insusceptible to easy formulation as a lexical rule.

For instance, in (3), the speaker intends the referent of *my grandmother* to be his grandmother's house.

(3) My grandmother was broken into. She was on vacation at the time.

To arrive at this interpretation, one must reason from one's knowledge of break-ins, that the usual motive is theft, that the usual scene is some living quarters (a house, an apartment, a room) or some other secured object (a vehicle, a shed, a barn), and so on and their knowledge of grandmothers, or possibly the speaker's grandmother specifically, that she is elderly, that she lives in a house, and so on, in order to establish a coherent connection between the speaker's grandmother and a break-in.

Again, the resultant interpretation can vary greatly depending on what the addressee knows about break-ins and about the speaker's grandmother or even grandmothers in general. A lexical rule extending, say, things that possess to things possessed, is difficult to implement, very specific, and insufficient in any case since it still leaves the problem of inferring what the possession is.

Other examples of novel metonymies include (4)--(7)

- (4) All the speakers are Thursdays except Soames.
- (5) Let's see. Sue Ann is in this pocket and I'm in the other.
- (6) That gentleman is a salad, right?
- (7) Green means aisle.

Without any context for interpreting them, they are gibberish. If the relevant meanings can be arrived at by the application of a lexical rule, that should not be the case. On the other hand, in context they normally pass unnoticed, inferencing from knowledge of the world instantly providing the interpretation.

3.2 Lexical rules are not lexical

In this subsection, we present two types of argumentation. First, we argue that certain proposed or plausible lexical rules are actually best explained in terms of a larger syntactic process, applying to lexical items only as exemplars of syntactic categories. Second, we argue that the true domain of lexical rules is not, in fact, the lexical item, but rather the referent of the lexical item.

3.2.1 Lexical rules as syntactic processes

One of the interesting facts about the use of any natural language is that it can be self-referring. Any natural language serves as its own meta-language. Thus, anyone can use any lexical item not only to refer to some extra-linguistic state of affairs, but to refer to the lexical item, qua symbol. This symbol, of course, has many representations: phonetic, phonemic, syntactic, or orthographic, as well as more exotic representations such as Morse code, ASCII code, spectrographic, and so on, but it also has many variant forms within each of those representation (e.g., handwriting, neon lights, type), as in examples (8)--(16).

- (8) *Cake* is a four-letter word.
- (9) *Brownie* is a seven-letter word.
- (10) *Cake* is a noun.
- (11) *Cake* starts with a voiceless stop.
- (12) *Stress* ends with 7 dots in Morse code.
- (13) *Able was I ere I saw Elba* is a palindrome.
- (14) *I* is a first person singular pronoun in English.
- (15) You just said “cow”.
- (16) Schwarzenegger’s “Hasta la vista, baby” is without question a dramatic masterpiece.

Note that this use is not without its syntactic implications: whereas (8) and (10) could be generated by standard rules of English, (9) and (14) could not.

However, the rule schema in (17) will produce the correct syntactic structures:

- (17) NP → X, where X is any category in the grammar.

This will produce the correct results for phrases and lexical items alike. Since this is a rule of syntax, it is possible that there is a semantics associated with it, following normal GPSG/HPSG practice.¹

Another example of lexical rules as syntactic processes is the proposed LIR that allows days of the week to be used as adverbs as in (18).

- (18) They will arrive Tuesday.

It is our claim that this involves a more general syntactic rule which permits a noun phrase to be used adverbially with a temporal interpretation, as in (19).

- (19) They will arrive February 24th, at 11:06 pm, on flight 103.

The point is that this sense extension applies not simply to lexical items, but to the referents of noun phrases. Thus, we find no reason to propose that the interpretation of (18) is to be accounted for by lexical rule, that of (19) by means of (presumably) a rule of compositional semantics, and that of (20) by pragmatic coercion.

- (20) They arrive St. Crispin’s.

3.2.2 Lexical items as reference transfer rules

In this section, we show that similar sense extensions apply to referents that meet the semantic criteria for the extensions, even if the lexical item used to refer to it would not ordinarily meet the semantic criteria, or if the referent is referred to by a compositional phrase rather than a single lexical item.

For example, Ostler and Atkins propose an LIR which provides an additional sense for lexical items whose semantics indicate that they are purpose-built containers. The semantics of this additional sense is the material contained in the container, as in (21).

1. Note that it is not our claim that if a process is syntactic it cannot be lexical. Following Dowty (1979) and others, we would treat most regular morphological processes as syntactic that is, as part of a set of regular syntactic/semantic rules. The claim rather is that the phenomenon applies to more than just lexical items, but to phrases and sentences as well.

(21) He ate three boxes of cereal before I could stop him.

The purpose-built constraint seems somewhat narrow to us, but even relaxing this constraint will not account for examples such as (22) and (23).

(22) She drank three slippers of champagne.

(23) The fairy offered him a thimble of dew.

It is unlikely that some feature such as **+container** would be in the actual or inherited properties of the lexical item *slipper* in any lexicon. Nor, given ordinary ways in which slippers are categorized, would it be likely that some teleological qualia of *containing* would be in the generative lexical entry for *slipper*.

In fact, even simple inferencing from slippers to containers may not be easy for current common-sense KB systems.¹ Drinking champagne from a slipper probably falls in the same category as wearing a lampshade, stereotypical but unusual behavior at a party or at least the parties such knowledge based systems appear to attend.

However, given a reasonable representation of the physical artifact, it is likely that some form of analogical reasoning could map the shape of a slipper onto that of a container, and similarly with thimble.² Certainly this would be a more plausible solution than to include **+container** in some way as part of the lexical semantics of *slipper*. If this is the case, it becomes difficult to see why a lexical rule is needed at all. The pragmatic inferencing can occur just as easily (in fact more easily) when the object in question is definitely a container.

Thus, the constraints on this LIR appear to us to be similar to those they state for the LIR which applies to produce, e.g., *thimbleful* from *thimble*. They state the semantic constraint as applying to containers (especially all purpose-built ones) as well as some other items that can be viewed as containers (e.g., *spade*, *spoon*, *fork*, *hand*, *fist*, *belly*) (Ostler and Atkins, 1991:78). The very inclusion of the phrase *viewed as* indicates to us that it is a context-dependent situation which allows the interpretation of the *-ful* suffix. It would be our claim that this morphological rule is completely productive. If examples of suffixation to certain lexical items have not been found, it is not because of restrictions of application of the rule, but rather because of the (real world) improbability of contexts in which the referent of that lexical item can be viewed as a container.

3.2.3 Usage extension applies to referents referred to by compositional means

The interpretation of the following sentences relies on an item-ordered-for-orderer lexical rule.

(24) The hamburger wants his check.

(25) The burnt hamburger wants his check.

(26) The blonde hamburger wants her check.

Assuming an LIR in this case, we know that the word *hamburger*, which ordinarily refers to a food item, can also, in certain situations, be used to refer to the person who ordered it, who is eating it, who wants it, or who is in some way connected to it³.

1. We actually tried and failed to find a connection between *slippers* and *containers* in the CYC data base.

2. Many other means of making the connection are possible. In fact, one of the virtues of a pragmatics-based approach is that it does not require that any specific piece of information be present in a particular form. As long as the inferences can be made (in whatever way), the account succeeds.

The paradox is that while both the second and third examples are easily interpretable, only one of them is explicable if the *orderer* sense is a rule-generated extension of *hamburger*. This is, if both *hamburger* and *hamburger orderer* are word senses of *hamburger* then the third example can be interpreted in a straightforward manner by choosing the *hamburger orderer* sense of *hamburger*. Since *hamburger orderers* are people and people may have blonde hair, the *blonde hamburger* is understood as the blonde person who ordered a hamburger.

However, in the second example no such simple choice is possible. If the *hamburger* sense is chosen then *the burnt hamburger* will compositionally refer to a uniquely specifiable burned hamburger. But the interpretation we want is the person who ordered the burnt hamburger. If, on the other hand, we choose *hamburger orderer* as the interpretation of *hamburger* as we did above, then *burnt hamburger* must be non-referential (unless the *hamburger orderer* his/herself is burnt).

The reference transfer function approach, on the other hand, which is not lexical and can operate on parts of a phrase or on whole phrases, offers an appropriate solution. Thus *the burnt hamburger* is taken to refer to a specific burnt hamburger in the domain of discourse, and since it is known to have been ordered, it can be related to the person who ordered it.

Note that this is not a problem just for lexical rules designed to capture a relatively ad hoc relationship such as *ordered-for-orderer*. It can occur with the *container-for-contained* relationship as well. Thus, examples (27) and (28) are fine where *glass* is interpreted as *the contents of the glass*.

(27) Drink a glass of wine before retiring.

(28) Drink this glass of water, rinse your mouth and spit it into this bowl.

(29) Drink the blue glass of wine.

but for (29) we have the same problem as above. Following a lexical rule approach, the interpretation of *glass* in *the blue glass* must be as a container since blueness is a possible attribute of glasses and not wines (contents) and yet the interpretation we need is the contents of the blue glass of wine not the blue glass itself.

Yet another example to show that lexical rules are really reference transfer rules is that of scalar adjectives, whose senses vary with the type of object or activity modified. To deal with this problem, Pustejovsky would say that the meaning of *good* intersects with some aspect of the lexical meaning of the head noun it is modifying in order to identify the scale along which *good* is measured (e.g., good weather, good music, good knife, good ball-game). Though compositional in nature and not treated as a lexical by Copestake and Briscoe (1995), it clearly involves searching the semantic entry for a lexical item to determine the meaning of an adjacent one. However, this process cannot be limited just to the head noun, since the interpretation of the scalar depending on the whole phrase which it modifies, and not simply the head noun. For example, the meaning of *good* in (30)--(32) differs markedly from their respective counterparts in (33)--(35).

(30) good [Heavy Metal] music

(31) good weather [for skiing]

(32) a good [hunting] knife

3. The difficulty here in specifying the exact relationship between the original referent and the referent of the output of the lexical rule would seem to indicate that the exact nature of that relationship is context dependent. Thus, in Vie gas' work, although the notion that there is a missing predicate may be derived on the basis of lexical and syntactic cues, the exact nature of that predicate is dependent on context---though there may be typical predicates, e.g., in *I forgot my keys*. the typical understood predicate is *bring* though it might be *leave* or *pick up* or *give to the housesitter*.

- (33) good [Country-Western] music.
- (34) good weather [for fishing]
- (35) a good [butcher] knife or a good [bread] knife

It is *music*, *weather* and *knife* that have qualia, not *Heavy Metal* or *Country-Western*, at least not qualia that will enable us to determine what *good Heavy Metal {music/clothing/decor}* is. While it is possible to formulate more complicated compositional rules, it seems to us implausible that it is some aspect of the combination of the meanings of the lexical items *bread* and *knife*, or *skiing* and *weather* that allows people to know what a *good bread knife* or *good weather for skiing* are. Rather it is knowing about how temperature, sun, and wind affect skiing and how the bread should look when it is cut.

In summary then, we have argued first that there are processes that look much like lexical sense extensions rules but which cannot be, as the extension must be determined on the basis of pragmatics and world-knowledge. Second, we have argued that what look like lexical rules are really reference rules and thus to be handled pragmatically in context, rather than by lexical means.

4.0 Rebutting arguments for the lexical nature of sense extension data

In this section, we examine the arguments that have been put forward in favor of treating sense extension phenomena by means of lexical rules. These are most directly presented in Ostler and Atkins (1991). In Pustejovsky (1995), most of the argumentation is directed towards the opposite position to the one taken here, namely that the lexicon consists of an exhaustive listing of all senses. There appear to be three main arguments for lexical rules: (1) the existence of apparently idiosyncratic and therefore language specific phonological and morphological constraints on lexical extensions; (2) the phenomenon of *blocking* or *pre-emption* (both morphological and semantic); (3) the language specific nature of lexical extensions.

We examine each of these arguments in turn below. In each case, we present in general (and often in specific), how we would account for the data and argue against its use to support lexical rules.

4.1 Lexically specific constraints

The first argument is that in order to account accurately for the data of sense extension, it is necessary to propose rules which are storable only in terms of a combination of phonological, morphological, and syntactic constraints on lexical items, on the one hand, and specific semantic constraints on the meaning of those items, on the other hand. As indicated above, we are not arguing against morphological lexical rules, as long as they are accompanied by what we have termed minimal semantics. Thus, in cases where lexical morphological rules are proposed with complex semantic constraints, our approach is to argue for a more general and unconstrained application of the morphological derivation, while arguing that the semantics are more general, with the more specific readings being determined by contextual information.

For example, Copestake and Briscoe (1991) discusses primarily the *grinding* rule, which creates a mass noun denoting substance out of a count noun. With their unification-based approach, they claim to account for the more specific *meat* reading with edible animals (a specialized subrule) as well as the less specific *substance* readings. Ostler and Atkins propose four different LIRs, all of which change count nouns into mass nouns (Tree--Wood; Animal--Fur; Animal--Meat: Food Item--Mass).

Our claim in general is that it is possible to distinguish the phonological and morphological constraints that exist within the language from the semantic rules that govern their interpretation. For example, instead of suggesting that there are several lexical rules that relate a count noun sense to a mass noun sense, we assume there is a general morphological rule that turns count nouns into mass nouns, with a semantic rule

that indicates no more than that whatever the mass noun refers to, it is some sort of substance related to the interpretation of the count noun. As a result, the rule proposed by Copestake and Briscoe to cover grinding, along with its accompanying specialized meat and fur/skin rules (*That coat is rabbit/Those boots are alligator*), and the more general pulverizing rule (*There was rabbit all over the road*), as well as the additional LIRs proposed by Ostler and Atkins can be replaced by a single morphological rule¹ with a minimal semantic correlate. The specialization occurs because of specific world knowledge about the referent of the noun itself (people eat eggs, many animals have edible flesh, the skin/fur of some animals can be used to make clothing or accessories, trees have woody trunks), what is the subject of discussion (clothing, cuisine, roadkill, lumbering) and what specifically is being predicated of the referent (it is being eaten, worn, used for construction, splattered over the road).

Note that this approach accounts for creative uses as well as conventional. Though badgers and moles are not marked (+edible) in this culture, one can enthymemically attribute that feature to them in Chinese culture, accounting for the acceptability and understandability of *mole* as foodstuff as in (36).

(36)... mole is eaten in many parts of Africa. (LOB corpus via Briscoe and Copestake, 1991:97)

Ostler and Atkins point to the exceptions to the food--substance rule such as *pea*, *bean*, and *lentil* (that is, it is not usual to say, *Have some more pea*). They claim this is a semantic class, namely, pulses, which defines a class of lexical exceptions. We suggest that it is rather that the speaker's view of the referents of these items that accounts for this exception. They are small, and a normal portion usually includes more than one. Thus, even when mashed up, a small portion of peas would contain more than one pea. We note a similar exception to the animal--meat and hunting plural rules, where things that are small and normally eaten more than one to a portion are not normally referred to by a mass term. Thus, (37) is fine, while (38) and (39) are questionable. (31) They are out after lobster.

(37) They are out after lobster./Have some more lobster.

(38) They are out after clam./Have some more clam

(39) They are out after oyster./Have some more oyster.

In addition, other materials (not pulses) resist being viewed as substances in eating situations, such as *minnows*, *sardines*, and *oats*, which are also small and generally eaten more than one to a portion.

Another claimed example of intricate relationships between phonological, morphological, and semantic data is that of adjectives of nationality being used to refer to people of that nationality. That is, certain adjectives of nationality, such as *French*, *British*, *Welsh*, can be used as plural mass nouns denoting the people of that country: *the French*, *the British*, *the Welsh*. The claim is, however, that this extension is possible only for national adjectives ending in a sibilant. Thus, it is possible to talk in general about *the Dutch*, but not about *the German*.

Here again, it seems that this LIR is part of a larger morphological process in English whereby adjectives can function as mass plural nouns denoting that set as in *The poor you will always have with you*, *The rich are usually powerful*, and so on. It appears that some adjectives which are themselves already derived from nouns fail to permit this re-nominalization. In particular, adjectives derived by using the suffix *-ian* (*-an*) do not appear to undergo this rule or are actually nouns used adjectivally, such as words ending with *-ist*,

1. Pustejovsky proposes a process of type coercion rather than a morphological rule. The only difference in our approaches is that for him the subsequent specialization of substance into meat, skin, fur, wood, food substance, is constrained by information in the qualia of the lexical item. For us, the constraining information is located in the belief systems of the speaker and hearer.

or, possibly, *-an*). Thus with religions you get *the Amish* and *the Orthodox*, but not *the Lutheran* or *the Methodist*.¹

4.2 Blocking

With respect to *blocking*, termed *pre-emption* by Ostler and Atkins, two kinds are distinguished: semantic and lexical. Lexical blocking occurs when a sense extension (with morphological or syntactic variation -- this is important) does not occur because there is already another lexical item of that form with an entirely different meaning. For example, the nouns referring to dances can be regularly used as verbs referring to dancing that dance resulting in “to waltz” from “the waltz” and so on. But this rule does not apply to “the reel” since a verb “to reel” already exists with a different meaning². We find no difficulty with this reading in the appropriate context. In addition, it appears that the verb *pair* was derived from the noun *pair*, despite the existence of the homophonic verb, *pare*.

Semantic blocking is a different matter. In such cases, it is claimed, a lexical rule fails to produce an extended sense if that new sense is already taken by another lexical item. Thus, the grinding rule of Copestake and Briscoe which associates lexical items referring to edible animals with a mass sense referring to their flesh is blocked in English for such words as *cow*, *sheep*, *deer*, and *pig* since lexical items referring to their meat already exist: *beef*, *mutton*, *venison*, *pork*. This phenomenon clearly exists. It is our claim, however, that the preferred analysis would be to view the process as fully productive, and the blocking is due to the pragmatics of referring.

From Grice's (1975) principle of *Be perspicuous*, a speaker should use a word which is conventionally used to refer to what they wish to refer to unless they wish to make some specific effect, as in (40), where the use of *cow* to refer to *beef* is likely to draw attention to the animal origin of the meat.

(40) I'm a vegetarian, I won't eat cow.

Similarly, the use of *pig* in the context below draws attention to the cooking of the whole animal as in (41).

(41) At the luau we had roast pig.

4.3 Language Specificity

The final argument against using pragmatics to account for sense extensions and in favor of the use of lexical rules is that they are language specific. If they were language independent, then we would expect all languages to have the same set of, say, metonymic references available. What is curious is that this is, by and large, the case.

We are aware of only a handful of specific claims which have been made that certain sense extensions are language specific: Na (1986), Wakao & Helmreich (1993), and the two examples cited in Ostler and Atkins.³ Without a good background in the languages involved, it is difficult to discuss the specific data, but we present below several plausible counterarguments to the claims exemplified with this data.

1. An exception here is *the Jewish* -- though this may be an example of blocking in that *the Jews* is a more typical way to refer to the people.

2. Actually, we find this claim questionable, and see no reason why “reel” cannot mean “to dance the reel.”

3. There are also apparently examples in Nunberg & Zaenen (1992), but we did not have access to this paper at the time of writing.

In general, differences in possible sense extensions between languages may well be due to differences in the syntactic and morphological structures permissible in the language. For example, the unavailability in English of counterparts to the Ancient Greek examples cited in Ostler and Atkins seems to depend on the nonexistence of a rule allowing modified pronouns in English. Certainly the temporal adjectives listed (e.g., *nocturnal*) can produce adverbial interpretations as in (42) which can be paraphrased as (43).

(42) A nocturnal visitor appeared at the door.

(43) A visitor appeared at the door during the night.

Second, apparent differences in sense extensions may be due to a usage preference for a different manner of referring to the object. Thus, the data in Na (19xx) appears to show a preference for certain kinds of construction that differ from those typically used in English. Thus while (44) is not ungrammatical, nor uninterpretable, it is less preferred than the more explicit (45).

(44) Harry-nin khita.

Harry-TOP big

(45) Harry-nin momcip-i khita.

Harry-TOP body size-NOM big

Yet a third explanation may be that due to culture specific differences, some interpretations are not possible in one language, but are in another. For example, the data in Wakao & Helmreich (1993) finds that certain common sense extensions in English are not so frequent in Chinese and Japanese, and that one, *reading Mao*, is completely unacceptable in Chinese. This data strikes us as somewhat suspect as other instances of reading authors are acceptable in Chinese, and the awkwardness of the *Mao* example may have something to do with the special status accorded to Mao Tse-Tung.

Finally, dialect difference may be due not to differences in lexical rules, but to differences in other pragmatics aspects. For example, Ostler and Atkins second example is of the dialect of SBE which does not permit the days of the week to be used alone adverbially. For speakers of this dialect, example (46) is ungrammatical.

(46) We arrive Tuesday.

The claim is that if this usage was pragmatically controlled, it would not be possible for (46) to be fine in most dialects of English, but not in all, given that culture and world knowledge is presumably nearly equivalent in all. In our explanation, we note again that this sense extension is part of a much larger syntactic process by which any temporal noun phrase can be used alone adverbially, as in (47)–(49).

(47) We arrived last year.

(48) We go there every Tuesday.

(49) We'll travel Christmas Eve, and arrive Christmas Day.

We are not familiar with the SBE dialect, but in Standard American it is also not usual to use months alone as temporal adverbials as in (50).

(50) We arrive February.

Years specified as numbers are also unusual as in (51).

(51) We came 1995.

Noun phrases, however, denoting the same time period are acceptable, as well as conjoined months and/or years as in (52)--(55).

- (52) We came last year.
- (53) We were there February and March.
- (54) We arrived March of '96.
- (55) We stayed in Stockholm 1994 through 1996.

Rather than positing a lexical rule which combines a reference to the special semantic classes of days, months, and years, it seems more plausible to us to propose a general syntactic rule, whereby any NP may be generated as a verbal modifier with a temporal significance. Thus, it is not that there is a category shift that turns certain nouns into adverbs with a temporal adverbial sense, but rather they still remain nouns, but are placed in a syntactic context which requires a temporal adverbial interpretation. The minor exceptions of bare days of the week, single months, and numerical years could be the result of something similar to the pragmatic feature **+heavy** that plays a role in **Heavy NP-Shift** (see Postal, 1975). This is the phenomenon that makes (56) perfectly acceptable while the nearly similar (57) is awkward and unusual.

- (56) They attributed to arson the fire which destroyed 70,000 acres in the Gila wilderness.
- (57) ?They attributed to arson the fire.

This judgment difference is not due to any formal feature involved in the generation of either sentence, but rather to a pragmatic judgment about appropriateness of use. Example (57) is not ungrammatical, it is simply not used, since the shift to focus position of the direct object is not justified by its brief and simple linguistic description. It seems plausible that something similar might account for the difference between (50) and (51) on the one hand, and (53) and (55) on the other.

In summary, we would claim that there are adequate pragmatic, cultural, and syntactic explanations for alleged language specificity of sense extension rules.

5.0 Conclusion

In this paper we have taken issue with the concept of lexical rule as a rule of grammar combining syntactic, morphological, and phonological references on the one hand, with reference to specific semantic classes on the other. We argue instead for a separation of the two, with the grammar specific rules firmly within the grammar, and information about semantic classes firmly in the encyclopedia of general world knowledge. Lexical semantics is primarily a matter of pointing the lexical items toward the appropriate locations in the ontology/encyclopedia, without requiring any except the most basic formal semantic properties (mass/count, number, gender, etc.) as directly referenced by rules of the grammar, properties which have semantic correlates, but which function independently in the grammar. While we have not addressed all claims made on behalf of lexical rules, we have provided an alternative explanation in several cases.

REFERENCES

- Copetake, Ann, and Ted Briscoe. 1991. Lexical operations in a Unification-based Framework. In James Pustejovsky and Sabine Bergler, eds., *Lexical Semantics and Knowledge Representation, Proceedings of a Workshop Sponsored by the Special Interest Group on the Lexicon of the ACL*: pp. 88-101. 17 Jun 1991. University of California, Berkeley, CA.
- Dowty, David. 1979. *Word Meaning in Montague Grammar*. Reidel, Dordrecht, Holland.

- Fass, Dan. 1988. Collative Semantics: A Semantics for Natural Language Processing. *Memoranda in Computer and Cognitive Science*, MCCS-88-118, Computing Research Laboratory, New Mexico State University, Las Cruces, NM.
- Grice, H. Paul. 1975. Logic and Conversation. In Peter Cole and Jerry Morgan, eds., *Syntax and Semantics 3: Speech Acts*, pp. 51-58.
- Iverson, Eric, and Stephen Helmreich. 1992. Metallel: an Integrated Approach to Non-literal Phrase Interpretation. *Computational Intelligence*, 8(3): pp. 477-493.
- Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live By*. University of Chicago Press, Chicago, IL.
- Levin, Beth. 1993. *English Verb Classes and Alternations*. University of Chicago Press, Chicago, IL.
- Martin, James. 1990. *A Computational Model of Metaphor Interpretation*. Academic Press, Cambridge, MA.
- Na, Younghee. 1986. The Conventionalization of Semantic Distinctions. In *Proceedings of the 22st Regional Meeting of the Chicago Linguistics Society*, pp. 166-178. University of Chicago, Chicago, IL.
- Nirenburg, Sergei and Victor Raskin. 1995. Lexical Semantics of Adjectives: a Microtheory of Adjectival Meaning. *Memoranda in Computer and Cognitive Science*, MCCS-95-288, Computing Research Laboratory, New Mexico State University, Las Cruces, NM.
- Nunberg, Geoffrey. 1978. *Pragmatics of Reference*. Indiana University Linguistics Club, Bloomington, IN.
- Nunberg, Geoffrey, and A. Zaenen. 1992. Systematic Polysemy in Lexicology and Lexicography. In H. Tommola, K. Varantola, T. Tolonen and J. Schopp (eds.), *Proceeding of Euralex II*, Tampere, Finland.
- Ostler, Nicholas, and B.T.S. Atkins. 1991. Predictable Meaning Shifts: Some Linguistic Properties of Lexical Implication Rules. In James Pustejovsky and Sabine Bergler, eds., *Lexical Semantics and Knowledge Representation, Proceedings of a Workshop Sponsored by the Special Interest Group on the Lexicon of the ACL: 76-87*. 17 Jun 1991. University of California, Berkeley, CA.
- Postal, Paul. 1974. *On Raising*. The MIT Press: Cambridge, MA.
- Pustejovsky, James. 1995. *The Generative Lexicon*. The MIT Press, Cambridge, MA.
- Pustejovsky, James and Sabine Bergler, eds. 1992. *Lexical Semantics and Knowledge Representation*, Springer Verlag, Berlin, Germany.
- Saint-Dizier, Patrick, and Evelyn Viegas. 1995. *Computational Lexical Semantics*. Cambridge University Press, Cambridge, UK.
- Viegas, Evelyne, and Sergei Nirenburg. 1995. The Semantic Recovery of Event Ellipsis: its Computational Treatment. In *Working Notes of the IJCAI-96 Workshop on Context in Natural Language Processing*, pp. 155-163. August, 1996, Montreal, Canada.
- Viegas, Evelyne, Boyan Onyshkevych, Victor Raskin and Sergei Nirenburg. 1996. From *Submit* to *Submitted* via *Submission*: On Lexical Rules in Large-scale Lexicon Acquisition. Paper to be presented at the 34th Annual Meeting of the Association for Computational Linguistics, Santa Cruz, CA. June, 1996.
- Wakao, Takahiro and Stephen Helmreich. 1993. Translation of Metonymy in an Interlingual MT System. In *Proceedings of the Pacific Association for Computational Linguistics (PACLING)*. Vancouver, British Columbia.