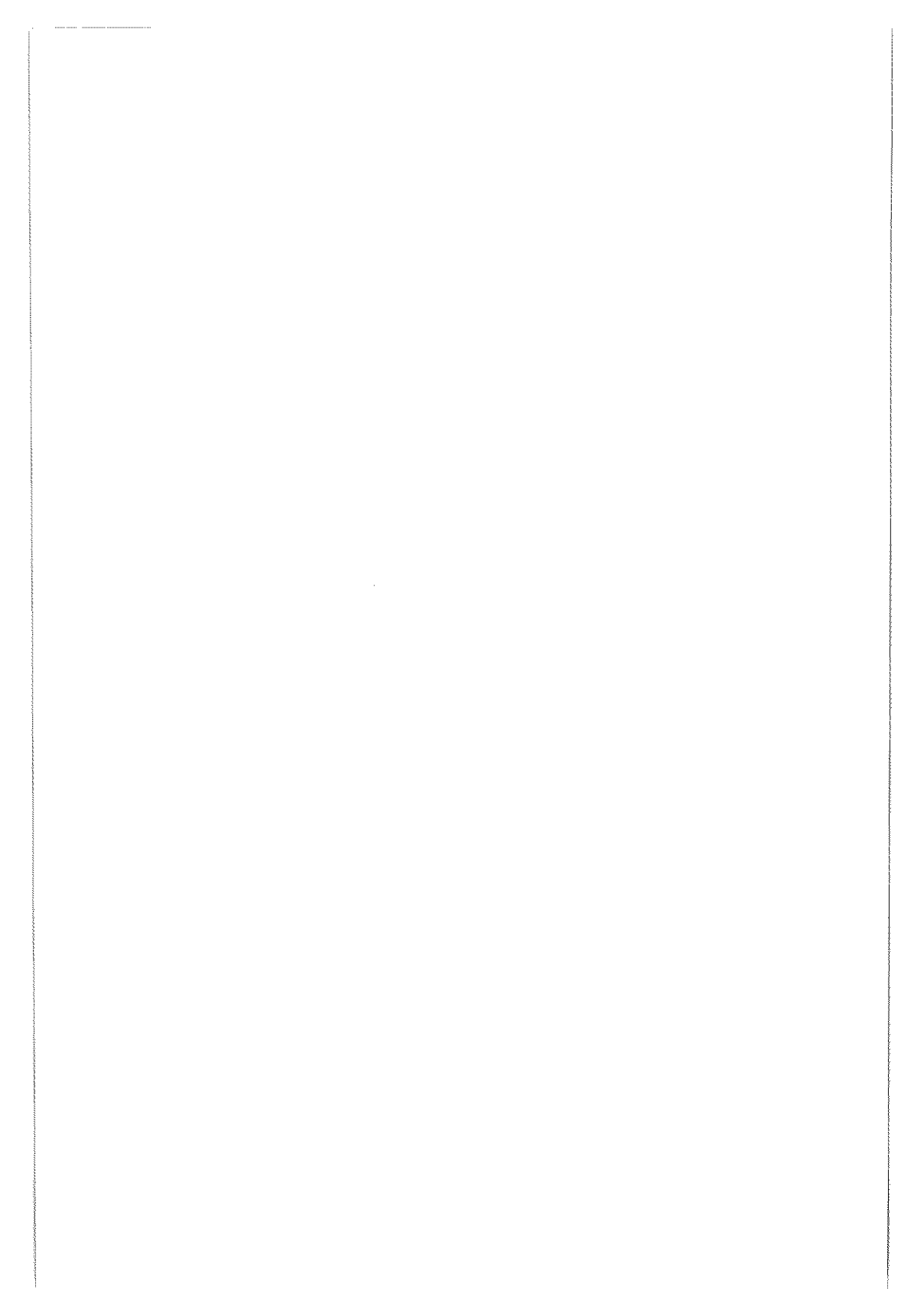*Dª. Pamela Faber & D. Ricardo Mairal Usón*

DEFINITIONAL ANALYSIS IN THE FUNCTIONAL-LEXEMATIC LEXICOGRAPHIC MODEL

UNIVERSIDAD DE GRANADA
U. N. E. D.

Introduction

Clearly a very important reason for studying lexical structure is the insight it gives us into cognitive processing. According to Leibniz (apud Wierzbicka 1992, 18), language is the best mirror of the mind. Human language with its design features of duality of patterning, displacement, and open-endedness makes us unique among living creatures (Trask 1995). Not surprisingly, there is a world of meaning encapsulated even in a single word, and a good definition necessarily contains many different types of information about the lexeme in question.

Just as certain features are shared by all the languages of the world, others make them different from each other. For example, each language is a selective representation of reality, and the fact that we, as speakers of a certain language, have chosen to name some things and not others is meaningful. Weinrich (1961, 142) writes:

> "The semantic mapping of the universe by a language is, in general, arbitrary, and the semantic `map' of each language is different from those of all other languages."

In other words, the best way we have of understanding our system of

conceptual structure is to examine our linguistic representations of these concepts and their interrelationships. The lexicon in this sense can be said to contain information about how we perceive, categorize, and impose order on our impressions of the world.

1. The Functional-Lexematic Lexicographic Model

The Functional-Lexematic Lexicographic Model elaborated by Prof. Leocadio Martín Mingorance (cf. Martín Mingorance, 1984; 1985a,b; 1987a,b,c; 1990) gives us a window into the design of semantic space. This model integrates Coseriu's **Lexematics** (Coseriu 1977) and S.C. Dik's **Functional Grammar** (Dik 1978a,b; 1989), and has the following objectives:
  (i) The structuring of the semantic architecture of the lexicon of a language.
  (ii) The representation / organization of our knowledge based on the linguistic encoding found in our dictionary entries.

This approach offers a solution to the problems of Lexical Field Theory by giving the concept of *lexical field* a more systematic and dynamic meaning. When describing possible structures for the lexicon, certain authors (e.g. Langacker 1987) have postulated domains of TIME and SPACE. However, such domains are impossible for the simple reason that they permeate all the fields, and to a greater or lesser degree form the coordinates from which all lexical structure is derived.

Ruhl (1989, 74) states that a semantic field is a superordinate with its hyponyms, but offers no formal criteria of how to determine field membership. We have postulated that a lexical field is a hierarchy of lexemes all of which share the same nuclear meaning. The *definiens* or act-nucleus marks the semantic territory covered by a specific field or dimension and becomes the factor which determines lexical membership.

Consequently, the lexeme is the central unit of linguistic description. We use the terms, *lexical unit* and *lexeme* interchangeably in the sense of "a word taken in one well-specified sense and supplied with all the information specifying its behaviour when it is used in this sense" (Mel'cuk 1988). Each lexical unit is conceived as a process in itself, containing a variety of different types of information at various levels. The lexicon thus becomes a dynamic storehouse of words which codifies and explains how we create and use language. This inventory of information is the mental dictionary of the NLU (natural language user).

220

2. Dictionary structure

The dictionary of contrastive lexical fields which we are presently in the process of elaborating, not only codifies how lexemes are arranged on the axis of selection, but also organizes them onomasiologically into a hierarchy of fields, dimensions and subdimensions. Traditional onomasiological dictionaries order lexical items in terms of conceptual meaning, while semasiological dictionaries (of which alphabetical dictionaries are the most prominent sub-type) classify lexemes according to form. Onomasiological organization has many advantages, not the least of which is its similarity to the organization of our mental lexicon.

Even though the majority of dictionaries are structured alphabetically, this type of order is not psychologically relevant, as it has nothing to do with how words are stored within our mind. Psycholinguistic experiments have shown that semantic relations are a function of memory, and that words with related meanings are stored near each other in the mental lexicon (Aitchison 1987). A relational approach to the lexicon focuses on semantic fields and accepts the supposition that there are common properties that bind the items in a field together, as well as properties that differentiate items from each other. Such a model represents human semantic memory as a network in which each node is a concept and concepts are linked together by a variety of semantic and lexical relations (Iris et al. 1988, 263).

A more useful type of dictionary would thus endeavour to reflect conceptual order by arranging words in terms of the broad areas of meaning they belong to. The question is how to find the type of onomasiological organization most in consonance with that within our mind. This obviously cannot be done by concocting an inventory of meaning areas *ad hoc* in order to fit words into a lexical straitjacket. Since direct observation of our mental processes is impossible, one way to have an idea of the way cognition works is to use the patterns in lexical structure as a window into the mind.

In the traditional thesaurus, macro-areas of human experience are established *a priori* by the lexicographer who then groups words accordingly. Since it is based on a top-down (or concept-driven) type of processing, we have no assurance that the system of descriptors elaborated by the lexicographer are really inherent cognitive universals. We believe that this uncertainty can be resolved by using a bottom-up (or data-driven) type of analysis. In this way, the areas of meaning or lexical fields established have a more solid basis and are not the product of intuition.

The method of lexical organization that we have followed in our model is based on definitional structure. Words with the same central meaning

component, belong to the same lexical field or dimension. The structure of the lexicon is thus arrived at by working upwards from words, not downwards from concepts. We believe that this type of lexical organization, if done correctly, should have psychological validity. Givón (1984:13) affirms:

> "The functional realm of lexical semantics pertains primarily to the storage of generic, culturally-shared knowledge largely embodied in the lexicon. This knowledge pertains to relatively stable phenomena, concepts or points of reference which constitute an intricate network that is part and parcel of our cognitive map of the phenomenological universe."

The structure of the lexicon should thus ideally be based on conceptual invariants.

3. Definitional analysis

To obtain the meaning components through which the definition of each lexical unit is elaborated, we have consulted the most widely-used dictionaries. This procedure can be justified for the following reasons:

(i) Standard dictionaries contain the body of knowledge gathered by lexicographic tradition.

(ii) Such definitions have the status of referential authority for users of the language in question.

(iii) Generally speaking, dictionary definitions provide a basis for extracting the stocks of more generic terms, which are intuitively felt by most speakers to be close the status of archilexemes. (cf. Martin Mingorance 1984, 229).

In this way, structured definitions are obtained, which in turn form the meaning hierarchies that are the basic architecture of our lexical fields. Although the majority of standard dictionaries are not consistent in their methods of defining words, they nevertheless are a treasure house of information for the factorization of the semantic components of lexemes. Many of the inconsistencies found in dictionary definitions are at least partly due to the fact that lexicographers have gone about their work with little recourse to any theoretical framework. Atkins et al (1988, 100) write:

"A theory of lexical organization is needed in order to provide the context for building the entry. And it is the lack of theory in the semantic-syntactic interdependency area that makes these dictionary entries less than adequate."

However, dictionaries are the first place one must look in order to find

the information needed for more coherent definitions, which will reflect meaning relationships between words. Levin (1991), Atkins (1991), and Felbaum (1990) also ascribe dictionaries a fundamental role for the construction of a lexical database. Levin (1991, 220) writes:

> "Monolingual collegiate-size 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."

The search for the most pertinent meaning components of each lexeme is a complex process since obviously a word cannot be defined in terms that are more complex than the word itself.

The meaning components we have used are natural language phrases found in definitions. Each definition in our lexicon is conceived as a syntactic frame with slots, having certain selection restrictions and default values. Each verb in the lexicon can be said to activate its own scenario, which determines its semantic participants/ actants to a fairly specific degree. In order to decide what kind of frame a verb has, as well as the number and type of participants inherent in its subcategorization, lexical entries in various monolingual dictionaries are analyzed. These definitions are segmented and put under the following headings: *nuclear meaning, subject, direct object, adverbial modification, usage label*, and *pragmatic information*. The different meaning components are compared and contrasted in order to find those which are the most appropriate, and semantically, the least complex.

4. Segmentation of dictionary information: *Embezzle*

As an example of the way in which this is done, we shall take the verb, *embezzle*, and show how its definition was arrived at in terms of its meaning, prototypical complementation pattern, and selection restrictions for the elements that can appear in association with it. This is also a good example of how a verb and the linguistic context associated with it can determine an entire expression.

Embezzle is an example of a verb that automatically brings its own frame with it, and determines its arguments to a fairly specific degree. It can even be said to presuppose a certain situation within a given sociocultural context. As a hyponym of *steal* in the dimension, *to cause to stop having*, it belongs to the field of verbs of POSSESSION, and like all negative dimensions in the lexicon, it is highly lexicalized.

(1) Extract from the subdimension *to cause to stop having* (verbs of POSSESSION)

TO CAUSE TO STOP HAVING

take (away) from to cause sb/sth to stop having sth.

> **steal** to take sth away from sb without their permission and not intending to return it (unlawfully).
> **rob** to steal sth (usu. money/property) from sb/institution.
>
> **thieve** to steal (old-fashioned).
>
> **embezzle** to steal money placed in your care for your own purposes.
>
> **pocket** to take sth, usu. money without having the right to do so.
>
> **filch** to steal things that are small/of little value in a very secretive way (informal).
>
> **pilfer** to steal things that are small /of little value esp. continuously over a period of time.
>
> **purloin** to steal sth usu. small (formal).
>
> **pinch** to steal sth usu. small directly off sb (informal).
>
> **nick** to steal (informal).
>
> **lift** to steal (informal).
>
> **swipe** to steal sth by removing it quickly.
>
> **snatch** to steal sth (usu. a purse) on the street.
>
> **rustle** to steal cattle/horses (AmE).
>
> **plunder** to steal property from sb, using force and causing damage.
>
> **loot** to steal things from buildings (shops, churches, houses), usu. during a violent event (riot/battle) or after a natural catastrophe (hurricane/typhoon).

The principal semantic parameter of differentiation is that of the object or what is stolen. Verbs in this dimension show us what is considered to be valuable in the language community (i.e. money, livestock, and property).

The meaning of *embezzle* is constituted as much by the features which differentiate from its co-hyponyms as by its subcategorization pattern. The meaning of a lexical unit is a product of its lexico-semantic structure and its combinatorial properties, in other words, the intersection of the paradigmatic and syntagmatic axes.

In order to construct this situation/frame we begin with the analysis of

lexical entries in various monolingual dictionaries to arrive at the definition of the verb and the number and type of participants inherent in its subcategorization. The dictionaries used for this analysis were *Collins Cobuild* (CC), *The American Heritage Dictionary* (AHD), *Longman Lexicon of Contemporary English* (LLCE), *Oxford Advanced Learners' Dictionary* (OALD), *Longman Dictionary of Contemporary English* (LDCE), *Webster's New World Dictionary for Young Readers* (WDYR), and *The Random House Dictionary of the English Language* (RHD). The definitions we considered were the following:

(2)

EMBEZZLE

CC:	*If someone embezzles money, they take or use it illegally for their own purposes, when it belongs to a company or organization that they work for.*
AHD:	*to take (money, for example) for one's own use in violation of a trust.*
LLCE:	*to use (money placed by others in one's care) for one's own purposes, usu. unlawfully.*
OALD:	*to use (money placed in one's care) in a wrong way to benefit oneself.*
LDCE:	*to take and use for oneself in the wrong way (money that is placed in one's care).*
WDYR:	*to steal money that has been placed in one's care.*
RHD:	to appropriate fraudulently to one's own use, as money or property entrusted to one's care.

These definitions are segmented and put under the following headings: *nuclear meaning, subject, direct object, modification, usage label/ pragmatic information*. In this way, they are better able to be analyzed and contrasted in order for us to derive the most relevant information from them.

(3)

Nuclear Meaning$_1$ ->   to take/use/take and use
Adverbial modification-> illegally/unlawfully/in the wrong way/in violation of trust [CC, OALD, LDCE, LLCE, AHD].
Nuclear Meaning$_2$ ->   to appropriate
Adverbial modification-> fraudulently [RHD]
Nuclear Meaning^3 ->   to steal [WDYR]

Subject (A) ->	Somebody (+Human) works for company/organization [CC] position of trust/direct responsibility [*placed in /* *entrusted to one's care*: LLCE, OALD, WDYR, LCDE, RHD] for money (usu. large sum).
Dir. Object (B) ->	money belonging to company/organization [CC] placed/entrusted (by others) in/to one's care [*placed* *in / entrusted to one's care*: LLCE, LDCE, OALD, WDYR, RHD]
from (C) ->	location company/organization etc. where money is [CC]
for (D) ->	beneficiary for own purposes/benefit/oneself [CC, LLCE, OALD, RHD, LDCE]
Usage label/Pragmatic information ->	[Cultural context: modern society in which money has an important role]

Regarding its argument structure, *embezzle* is a verb with four possible slots, which have the following semantic roles:

A [agent] *embezzles* B [affected] *from* C [source] *for* D [beneficiary].

From the various definitions given for the verb, the following information about all of these slots can be obtained.

4.1. Nuclear meaning

Within the category of nuclear meaning, there is a wide range of possible choices. On the one hand, we have a choice between the verbs, *take* and *use*, as well as a combination of the two (*take and use*), all of which are modified by *illegally/unlawfully*. On the other hand, we have *appropriate* modified by *fraudulently*, and *steal*.

If we begin by first eliminating those terms which are less suitable, *appropriate* can be discarded, since it violates our criterion of linguistic simplicity. Words must only be defined in terms of other words whose total meaning is included in theirs.

For example, if *embezzle* were to be defined in terms of *appropriate*, this

would mean that its definition would include all the meaning components (stylistic register included) of *appropriate*, something which is manifestly untrue. Although *appropriate* can mean *to take sth illegally*, it cannot constitute the base meaning of *embezzle*, because firstly it is possible to appropriate other things besides money, (such as land, for example), and secondly, when what one appropriates is <u>money</u>, there are two potential meanings. The first possible meaning is *to take the money illegally* (which does fit the meaning for *embezzle*), but the second meaning, *to put the money aside for a special use*, has nothing to do with illegal activity.

Consequently, this second possible definition disqualifies *appropriate* as part of the definition of *embezzle*, since *embezzle* is <u>always</u> classified as a dishonest action.

At the same time the choice of *fraudulently* is not a good one, because although it means *deceitfully/illegally*, the purpose of this type of deceit referred to in the definition of *fraudulently* is to obtain <u>money or goods</u>. As we have said, one can only *embezzle* money.

Furthermore, *appropriate* is classified as a formal word in *Collins Cobuild*, whereas *embezzle* has no such usage label.

A second possibility is *take, use*, or a combination of the two. However, *take* seems to be the essential action within *embezzle*, not only because intuition and common sense tell us so, but also because its four-slot subcategorization pattern is typical of those verbs within the general area of POSSESSION/ TRANSACTION. The option of combining the two meanings together into one lump *take and use* is basically unsatisfactory, since although money is embezzled for the purpose of using it in a certain way, this use is not included in the nuclear part of the definition.

In this way, we deduce that the essential action involved is *take*. This general verb would require the specification of *illegally* in order to differentiate it from other verbs having *take* as part of their base definition. *Illegally*, though similar in meaning to *unlawfully*, would be the best choice, because in the dictionaries consulted, *unlawfully* is defined in terms of *illegally*, and it is considered to have a more formal register (OALD).

However, if we select *to take illegally* as the nuclear meaning of *embezzle*, thus satisfying our simplicity requirement, we find this definition coincides with the one already given to the verb *steal* (*to take sth not belonging to you, usu. secretly, illegally*).

If we analyze the meaning components of *steal* as given in the same dictionaries, we find them to be the following:
(4)

    Nuclear Meaning->   To take

Subject (**A**) ->	Somebody
Dir. Object (**B**) ->	Something belonging to someone else
from source (**C**)->	Somewhere/Somebody location/possessor
Modification ->	usu. secretly
	usu. illegally/without the right to do so.
Usage label/	Pragmatic information -> None

As can be seen, the frame for *embezzle* is essentially the same as that of *steal*. The basic difference between them is in their respective levels of specificity under all headings, something which we will now look at in greater detail.

4.2. Subject

The subject of *embezzle* [the same as *steal*] must always be human. The only possible specification for the Subject slot in *steal* would be the necessarily negative evaluation of anyone who would engage in such an activity. Prototypical fillers for this verb would be agents such as a thief, robber, or burglar, whereas we have to bridge a considerable credibility gap before being able to accept slot fillers with high plus values on the axiological scale such as a nun, a two-year-old child, or a volunteer worker in a leper colony.

Embezzle, however, has a subject which has very definite specifications. Such an entity would necessarily be human, (prototypically someone responsible for somebody else's money, such as a bank employee, treasurer or company executive). Although such a person generally does not fit into the classification of *thief*, the fact that in this case, he acts in the same way as one, places him in the same category. The negative evaluation of the *embezzler* is enhanced not so much by the illegality of the action (which is part of the nuclear meaning of the verb anyway), but because of his deviation from the norm of responsible behaviour which is an implicit part of his position of authority and trust.

4.3. Direct Object

The object of *steal* would have very few specifications. It could be any person or any thing (both concrete and abstract), since it is possible to steal someone else's boyfriend, pet boa constrictor or idea about how to get rich quick. The only selection restriction to fill this slot pertains to the ownership of the thing in question; it must be something that belongs to someone else, or is perceived as belonging to someone else.

228

In reference to *embezzle*, the object slot must be filled by something much more specific. The direct object of this verb is always a sum of money the Subject is responsible for (because of his job) and which he takes when he should not.

4.4. Source (*from* + O)

Both *steal* and *embezzle* have as part of their frame the source from which they take the object in question. In *steal*, the source can be a place or person. For example:

(5) The thief stole the diamonds from the castle
(6) The thief stole the diamonds from the duchess.

In contrast, in *embezzle* the source would typically be a company, bank or organization which would function both as the location [where the money is] and group of people [whose possession it is in]. The source could not be a single person.

(7) The vice-president embezzled $25,000 from the company.
(8) * The vice-president embezzled $25,000 from the company treasurer.

4.5. Beneficiary (*for* + O)

Another possible slot for semantic participants is that of beneficiary. This is often left unspecified, because its *default value* would be something which personally benefits the subject. When the beneficiary is given, this would be either to present it as a deviation from the norm (if it is not something that personally benefits the subject), or to reaffirm the default value by giving it greater specificity.

(9) The vice-president embezzled $25,000 from the company for
     (a)   Mother Theresa's Hospital in Calcutta [deviation from norm]
     (b)   his annual vacation spree in the Bahamas [specification of default value]

Our definition of *embezzle* would be the following:
(10)
  embezzle: <u>to steal money placed in one's care for one's own purpose</u>
  S =   (Agent) [Human] [Adult], working in large organization/ company, position of responsibility, lacking/in need of money.
  O =   (Affected) Money S is responsible for (usually a large sum).
  from + O =   (Source) Organization/company that S works for.
  for + O =   (Beneficiary) own purpose, personal benefit.

This is an improvement over the definitions which appear in the dictionaries consulted. As we have seen, the RHD uses a defining vocabulary which is semantically more complex than the word being defined. The AHD definition (*to take (money, for example) for one's own use in violation of a trust*) is the least accurate. *Money* is more than just an example of what one can embezzle; it is really the only possibility. Furthermore, added information needs to be included about the semantic participants, as this lack of specification causes situations that are definitely not cases of embezzlement to be classified as such. If a 9-year-old boy decides he wants to buy a bag of Fritos, and takes money out of his sister's piggy bank against her wishes after she has trusted him with the key, this would be *taking money for one's own use in violation of a trust*. It would thus fulfil all the conditions set in the AHD for *embezzlement*, yet it is obviously not embezzlement, because the features of the other sentence elements do not fit into the verb frame.

Throughout this analysis, it cannot be forgotten that all language is embedded within a larger context. A verb such as *embezzle* obviously presupposes a certain type of society in which there are organizations where people place their money in the care of others with the ultimate purpose of keeping it safe or of acquiring more money. In a different type of society (i.e. Indian tribes in the Amazon rain forest) with an economy based on bartering, such a verb would obviously not exist, since such an activity would be inconceivable.

5. Conclusion

This paper has been a demonstration of definitional analysis, one of the basic principles of the Functional-Lexematic lexicographic model. Through this intricate process, it is possible to formulate definitions for lexemes, the hierarchical structure of which form the basic architecture of lexical fields. Taking the verb *embezzle* as a case in point, we have outlined how dictionary definitions can be used to discover meaning components relevant to the lexical hierarchy in question, and ultimately to knowledge representation.

BIBLIOGRAPHY

AITCHISON, J. 1987. *Words in the mind: an introduction to the mental lexicon*. Oxford: Basil Blackwell.

230

ATKINS, B.T.S. 1991. "Building a lexicon. The contribution of lexicography". *International Journal of Lexicography* 4-3: 167-204.

ATKINS, B.T.S. et al 1988. "Anatomy of a verb entry: From linguistic theory to lexicographic practice". *International Journal of Lexicography* 1:-84-126.

COSERIU, E. 1977. *Principios de semántica estructural.* Madrid: Gredos.

DIK, S.C. 1978a. *Functional Grammar.* Dordrecht: Foris Publications.

-- 1989. *The theory of Functional Grammar. Part I: the structure of the clause.* Dordrecht: Foris Publications.

EVENS, M. (ed.) 1988. *Relational models of the lexicon: representing knowledge in semantic networks.* Cambridge: Cambridge University Press.

FABER, P. 1994. "The semantic architecture of the lexicon". *Lexicographica. Series Maior* 57: 37-50.

FELLBAUM, C. 1990. 'English verbs as a semantic net'. *International Journal of Lexicography* 3-4: 279-301.

GIVÓN, T. 1984. *Syntax. A Functional Typological Introduction. Volume I.* Amsterdam/Philadelphia: John Benjamins.

IRIS, M. et al. 1988. "Problems of the part-whole relation". In Evens, M. (ed.) 1988: 261-287.

LANGACKER, R.W. 1987. *Foundations of Cognitive Grammar.* Stanford: Stanford University Press.

TOMASZCZYK, J. AND LEWANDOWSKA-TOMASZCZYK, B. (eds.) 1990. *Meaning and Lexicography.* Amsterdam: John Benjamins.

LEVIN, B. 1991. 'Building a lexicon: The contribution of linguistics'. *International Journal of Lexicography* 4-3: 205-226.

MARTÍN MINGORANCE, L. 1984. "Lexical fields and stepwise lexical decomposition in a contrastive English-Spanish verb valency dictionary". In: Hartmann, R. (ed.) (1984): 226-236.

-- 1985a. "La semántica sintagmática del adjetivo. Parámetros para la organización de un lexicón inglés/español de valencias adjetivales". In: Monroy-Casas, R. (ed.): 329-340.

-- 1985b. "Bases metodológicas para un estudio contrastivo del léxico derivado". *Revista Española de Lingüística Aplicada* 1: 37-54.

-- 1987a. "Classematics in a Functional-Lexematic grammar of English". In: Olivares, C. (ed.) 1987: 377-382.

-- 1987b. "Semes, semantic classemes, and dimensions: The lexicological and lexicographic perspectives. *XIVth International Congress of Linguists*: 10-15.

-- 1987c. "Pragmatic features in the lexicon of a Functional Grammar". *International Pragmatics Conference*: 17-22.

-- 1990. "Functional Grammar and Lexematics". In: Tomaszczyk, J.; and Lewandowska-Tomaszczyk, B. (eds.) 1990:-227-253.

-- 1988. "Semantic description or lexical units in an explanatory combinatorial dictionary: basic principles and heuristic criteria". *International Journal of Lexicography* 1, 3: 165-188.

MONROY-CASAS, R. (ed.) 1985. *Actas del II Congreso Nacional de la Asociación Española de Lingüística Aplicada.* Madrid: Sociedad General Española de Librería.

OLIVARES, Carmen (ed.) 1987. *Actas del X Congreso de la Asociación Española de Estudios Anglo-Norteamericanos.* Zaragoza: Universidad de Zaragoza.

RUHL, C. 1989. *On Monosemy: A Study in Linguistic Semantics.* Albany: State University of New York Press.

TRASK, R.L. 1995. *Language: the basics.* London: Routledge.

WEINRICH, U. 1961. "On the semantic structure of language" in Greenberg, J. (ed.), *Universals of language.* Cambridge, Mass: M.I.T. Press.

WIERZBICKA, A. 1992. "In search of tradition: The semantic ideas of Leibniz". *Lexicographica* 8: 10-25.

DICTIONARIES:

American Heritage Dictionary

Collins Cobuild

Longman Dictionary of Contemporary English

Longman Lexicon of Contemporary English

Oxford Advanced Learners' Dictionary

Random House Unabridged Dictionary

Roget's Thesaurus

Websters Dictionary for Young Readers