Hyperlexis: a hypertext system for representing lexical information at discourse level

Francisco J. García Marco
(Universidad de Zaragoza)
Francisco J. Ruiz de Mendoza Ibáñez (Universidad de La Rioja)
José Luis Otal Campo (Universitat Jaume I)

Abstract

Hyperlexis is a multirelational database based on hypertext. In this paper we offer a summary of some of its technical possibilities but the main discussion focuses upon the theoretical basis of the lexicological model which lies at its foundation. The model is intended to be discourse-sensitive as well as functionally and cognitively oriented. We discuss the form of nominal and verbal entries, and explore and illustrate the relational system chosen.

0. Introduction

Hyperlexis is a multirelational database. In the same way as other lexical models based upon hypertext (see Bui, 1989), Hyperlexis combines the advantages of a conventional linear-access dictionary with those of a thesaurus. To this we add the creation of computer-based links between data and the possibility for the user of interacting with the database by creating his own nodes, links and link routes, or by modifying those already existing.

At the present stage of development, Hyperlexis is able to provide definitions for some two hundred terms of common usage. The definitions are based on a lexicological model whose main features shall be discussed below. It also gives the user quick access facilities to all the component terms of definitions and to a vast network of word-relations. It provides an alphabetic index, "find path" utilities -which allow the user to trace back his own steps in the relational network- and the possibility of studying "empirical sets" of relations suggested by the occurrences of a certain term in the dictionary. Navigation options can be broadened or narrowed down at will to make them work within the bonds of specific sets[1].

In this project the development of the program and of the lexicological model run parallel. This provides the benefit of mutual feedback and constant readaptation. This paper will be mainly concerned with the justification of the lexicological model, which will be done from the point of view of both conceptual and relational theories of

knowledge organization. Once the database has taken its final form, future expansions are intended to go in the direction of wider lexical coverage and multilingualism[2].

1. *Discourse relevance*

The Hyperlexis database is being designed with the overall idea of making explicit, in a way which is compatible with present theories of cognition and pragmatics, whatever links a language user has to make in order to connect pieces of discourse coherently (see Petöfi, 1985). There are two basic ways in which we can talk of discourse coherence: one is based on lexical and grammatical resources developed by each individual language; the other is based on that part of our common knowledge of the world which has not found its way directly into the linguistic system, seems to remain in the realm of our cultural idiosyncracy but is still conceptually tied to language. To provide some illustration consider the following examples:

(1)

(a) It was hot. The ice melted.

(b) ?It was hot. The table melted.

In example (1a) we would seem to have a fairly straightforward case of lexically-motivated coherence, based on the semantic relationship between hot/melt/ice. But, strictly speaking, the lexical relationship is rather more complex:

hot-heat-melt-solid substance-ice

The connection can be spelled out as follows: Something becomes hot if it is heated. Some solid substances can melt when heated under certain conditions. Ice is such a substance. Therefore heated ice can melt.

Of course, our everyday experience allows us to take a short-cut in the inferential process, perhaps by means of a proposition like:

(2) Ice melts when it is heated

If we think of melting we may also associate it with cheese, butter, wax and some metals (typically gold and iron). It is not the lexicon itself (that is, the definition of *melt* as 'become liquid through heating') but our knowledge of the world that provides us with this sort of information and allows coherence to take place.

Consider now example (1b). Even if we think of a table made of metal (e.g. iron), (1b) evinces some sort of semantic implausibility. Once again it is our knowledge of the world that reveals the reason for such implausibility: hot weather is not enough to make iron melt[3].

Since we had the goal of making Hyperlexis sensitive to discourse-related problems like the ones pointed out above, we needed a way to incorporate a lot of encyclopedic knowledge in the database. However, we also worked under the assumption -held for a long time in Artificial Intelligence and in Cognitive Psychology- that knowledge is stored in our minds in an economical, well-organized way[4]. A sheer listing of facts -

after the fashion of an encyclopedia- would have run counter to these assumptions. A better solution was to devise a relational system where all facts relevant to a word could be accessed. In order to do so, we first differentiated between lexical and grammatical words, the former being of no avail for our purposes since their role is one of acting upon the information given by the latter, as is the case with articles, deictics, prepositions and adverbial particles. Among lexical (or content) words at the present stage we have focused our attention on verbs and nouns, leaving adjectives and adverbs aside for a future stage, since these have the role of providing additional information about processes and the entities set in relation by processes.

2. *Processes in Hyperlexis*

A rough idea of the way processes are dealt with in Hyperlexis can be gained from the following example. Take the case of the verb "kill". The entry for kill contains the information that it is an action verb where a certain entity brings to an end the life of another (living) entity:

(3) Mary killed John

(4) The poison killed John (=Someone killed John with poison/John killed himself by taking poison, whether accidentally or not)

(5) The rock (fell and) killed John

(6) The cat killed the mouse

Since "kill" is an action verb, the feature [+action] will be part of its semantic specification. Associated with this feature are a number of roles: in this case they are Agent (A), Patient (P) and Instrument (I). For the verb "kill", the term taking the role Agent can be either an animate or inanimate entity as can be seen from examples (3)-(6) above, with the restriction that it will always be animate if an Instrument attains separate lexical expression:

(7) *The poison killed John with a sword

It must be noted that sometimes what is grammatically coded as an Agent can be as a matter of fact the Instrument, as is the case with "the poison" in (4), but not necessarily so (compare example (5) where "the rock" is not Instrument). This poses no problem whatsoever for the organization of Hyperlexis since it is a conceptual database and takes no account of the syntactic exploitation of the linguistic system. The semantic specification of any action verb will contain the three associated roles A, P, I, whether these are all actually expressed in the sentence or not. In relation to this it must be noted that the role Instrument can be omitted in English, but not so the other two action roles. This difference has been taken account of in some linguistic theories, like Functional Grammar (Dik, 1989), where a division is made between predicate, arguments and satellites, the role of Instrument being a satellite role rather than an argument role[5]. This division is immaterial from the point of view of semantic organization.

Following the criteria above, part of the lexical entry for "kill" in Hyperlexis would have this form:

(8) kill [+action] (x_A, y_P, z_I): x causes y to die by means of z

/y_P [+entity], [+animate] / z_I [+entity, typ. weapon or killing instrument] or [+action][6]

Now, if we consider sentence (5) again we shall see that there is no apparent Instrument. What happens in cases like this, from a conceptual point of view, is that the roles of Instrument and Agent are fused into one which is expressed as the subject of a sentence following a syntactic principle of English according to which agents attain lexical and syntactic expression with preference over other roles.

The semantic framework specified in (8) is intended to relate arguments at the highest level of genericity. Also, the arguments supplied by the definition are the most generic abstract entities that can fit in the framework. Thus, "x" must be read as "any entity", "y" as "any animate entity" and "z" as "any entity or action". Note that such an interpretation will make the following strange utterances conceptually feasible:

(9)

(a) John killed the goose with an elephant

(b) The elephant killed the blade of grass with a goose

(c) The Martian killed the bacteria with a hacksaw

However, the strangeness of these sentences has to do not so much with their semantic impossibility as with their implausibility and since they are discourse relevant they should be allowed by the information provided by Hyperlexis. Lexical entries denoting entities (not processes) of any sort will need to have the indicator [+/-animate]. Other restrictions should derive from the definitions themselves: for example, an elephant or a goose are not typical killing instruments, and they will not be found in the network for "weapons"; a gun or a sword are typical killing instruments and they match perfectly the requirement for z_I; a hacksaw is not intended for killing but can be used for such purpose. Hyperlexis makes provision for this fact in the combination between the definition for "hacksaw" as a cutting instrument and the definition of "weapon" as any instrument which kills by means of cutting, piercing or otherwise destroying.

3. *Semantic roles*

As we have seen, a substantial part of Hyperlexis is dependent on an analysis of semantic roles. Consequently, we need to establish a taxonomy of such roles. The idea behind our proposal will be based upon Ruiz de Mendoza (1992). Here it is suggested that the description of semantic roles is derived from a typology of predications at a purely conceptual level, the syntactic expression of roles not being a determining taxonomic criterion.

We have provided some evidence for this second idea in our analysis of example (5) above. Then, the first idea is related to Halliday's (1985) contention that classifying the different process types will yield the different categories of the participants: a material process has an Actor and an Object, a mental process has a Senser and a Phenomenon, and so on. In Dik's FG we find a similar point of view, but Dik does not speak of process types but of predications consisting in the instantiation of a predicate frame (a process) by means of the insertion of terms (or arguments). A predication denotes a State of Affairs (or SoA) which can be defined according to a number of parameters which are not exclusively derived from the verbal expression of process in the language. For example, an SoA can be [+dynamic] when we find an action predicate (like "paint"), but still we have to examine whether the process is seen as attaining completion or not. If it does, the SoA will also be defined as [+telic], as in the sentence

(10) John was painting a portrait [+telic]

in contrast to

(11) John was painting [-telic]

There are other relevant features. An SoA is [+control] if the first argument of the predication can determine whether the SoA will obtain; an SoA is [+dynamic] if there is change in an SoA.

Using these features, Dik (1989) classifies SoAs as follows:

•Situation [-dynamic]

Position [+control]

State [-control]

•Event [+dynamic]

Action [+control]

Accomplishment [+telic]

Activity [-telic]

Process [-control]

Change [+telic]

Dynamism [-telic]

For Dik, semantic functions specify the roles which entities play within an SoA. Therefore, his classification of argument roles depends on the previous typology of SoAs (Dik, 1989: 101 ff.):

Agent: the entity controlling an Action

Positioner: the entity controlling a Position

Force: the non-controlling entity instigating a Process

Processed: the entity that undergoes a Process

Zero: the entity primarily involved in a State

Goal: the entity affected or effected by the operation of some controller (Agent/Positioner) or Force

Recipient: the entity into whose possession something is transferred

Location: the place where something is located

Direction: the entity towards which something moves

Source: the entity from which something moves

Reference: the second or third term of a relation with reference to which the relation is said to hold.

It will be noted that this description of semantic functions or roles only takes into account the four intermediate levels of specification of an SoA: Position, State, Action, Process. The feature [+/-telic] becomes irrelevant, since it is a feature that depends on the nature of the arguments of the predication rather than on the predicate itself. As a result, Dik's background assumptions for the classification of semantic functions converge naturally with Halliday's and can be reduced to an analysis of process types.

Halliday (1985) distinguishes three types of processes as encoded in language: (i) material (or processes of 'doing'), (ii) mental (or processes of 'sensing') and (iii) relational (or processes of 'being'). Material processes are events very much in the sense of Dik's dynamic predicates, but no further distinction seems to be made by Halliday between controlled or non-controlled dynamic predicates. For him, a material predicate will typically have an Actor and a Goal in a transitive relationship which is captured by the grammar. However, sometimes as the process becomes more abstract the distinction between Actor and Goal becomes harder to draw and we may only have a 'happening' rather than a 'doing' as in

(12) The tourist collapsed

where the Actor is as involuntary as if it were a Goal. The linguistic system, Halliday notes, is often capable of perspectivizing the predicate in such a way that apt use is made of intransitivization for different shades of meaning. Thus, according to his analysis, "the two schools" is Actor in (13) but Goal in (14):

(13) The two schools combined

(14) The two schools were combined

But what this analysis fails to reveal is that there is a clear difference between the Actor in predications like (13) and in others where it takes on a more active value as in

(15) The two schools combined their resources

It can hardly be argued that the difference rests on degrees of activity because, while this may be truly a semantic aspect of certain predications, it does not seem to explain the passive role alloted to the first argument in (13). The question is rather one of control of the first argument over the coming about of the State of Affairs designated

by the predication: in (13) there is no such control in contrast to what happens in (15). Then, in (14) what we have is that the controlling entity has been suppressed by grammatical means ('someone' combined the two schools). Therefore, it may be safely suggested that an adequate distinction between 'doing' and 'happening' is not to be made on the basis of types of Actor but rather on different types of process and that Dik is right in pointing out that a dynamic event is an action if it has the feature [+control] and a process if it has the feature [-control]. The roles Agent and Goal are to be maintained for Action predicates but not so for processes, where the role Processed seems to be particularly apposite.

Halliday differentiates mental from material processes on the basis that it is difficult to see how certain clauses of feeling, thinking and perceiving can be said to have actors and goals. He also points out as distinctive to mental processes that they are two-way. Consider:

(16) John likes the present

(17) The present pleases John

According to Halliday, someone might be tempted to simplify and say that "John" is an Actor in (16) and that "the present" is the Goal. But (16) is semantically related to (17) where we can see that it is somewhat artificial to say that "the present" is Actor. However, if we consider "John" as the Senser and "the present" as the "Phenomenon" (whatever is sensed) the analysis is able to relate both expressions of a same process and turns out to be more natural. The analysis would seem to apply to other pairs like

(18)

I fear it	it frightens me
I wonder at it	it amazes me
I don't understand it	it puzzles me
I enjoy it	it delights me
I admire it	it impresses me

One more piece of evidence furnished by Halliday to defend the separation between material and mental processes comes from the possibility of saying (19) but not (20):

(19) What John did to the mouse was kill it

(20) *What John did to the present was like it

There are problems with this analysis, however. First, it is true that it is typical of mental predicates to stand in such relationships as the ones in (18). But it must be noted that the pattern has important exceptions: *think*, *see*, *worry*, *know*, *guess*, *ponder*, and *shock* are some examples. Then consider cases like the following:

(21)

(a) I admire it

(b) I admire John

(22)

(a) It impresses me

(b) John impresses me

We may say that "it" and "John" are sensed phenomena in (21), but are they so in (22)?. By all appearances both terms acquire a more active (or less passive) value in (22) than in (21), especially in (22b) since we know John to be necessarily animate. It is as if there was something purposeful about these terms that provokes the mental reaction on the Senser. Moreover, it would be perfectly feasible to say

(23) What John did was impress me

but not to say

(24) *What I did to John was admire him

Conversely, there are seemingly material processes where the do-substitution test fails and whether the test fails or not does not depend on the nature of the predicate but of the whole predication:

(25) John painted a portrait

(26) *What John did to the portrait was paint it

(27) John painted the house

(28) What John did to the house was paint it

In (27) "the house" is not the result of the painting activity but in (25) "the portrait" is so and that is why (26) turns out to be impossible. In a sense, a predication like (25) could be described as lying somewhere between an action and a process, and it is to some extent arguable whether we can think of "a portrait" as a Goal or merely as part of the process itself. Notice that there are more extreme cases in which the second argument of the predication is usually made redundant, as in *He sang (a song)*.

The way out of these problems is to be found in positing the feature [+control]. Compare:

(29) I like Mary

(30) Mary pleases me

In (29) Mary has no control over my liking her, but in (30) it is suggested that she may indeed have some control. In (30) "me" is more of a Goal than of a Senser, and "Mary" has a certain agentive quality which is certainly missing in (29).

What all this discussion comes to is to the positive assessment of Dik's distinction between Position, State, Action and Process. Mental processes denote situations whose first term can be either a controlling or a non-controlling entity. They are necessarily [-dynamic], in contrast to events which are always [+dynamic]. Wherever there is a [+control] feature involved there will be a Goal term or at least a Goal-like term (as is the case of "paint a portrait" or "sing a song"). In this analysis, "John" in (22b) is a

Positioner, while the value of "me" is that of Goal. A sentence like (29) can be said to describe a state (it is [-dyn, -control]) whose first term has the Zero role (the entity primarily involved in a state) and whose second term is an object of experience rather than a Goal in a strict sense (since there is a [-control] feature involved). Thus, apart from the role Instrument that we discussed before, we still have to add at least two more roles to Dik's inventory: Object (for the second term of some stative predications) and Result (for cases where the second term of a controlled event is the result of the activity rather than the goal).

Finally we have relational processes. Halliday includes here three types of English clause:

(i) Intensive: 'x is a'

(ii) Circumstantial: 'x is at a'

(iii) Possessive: 'x has a'

each of which may be expressed in two ways:

(i) Attributive: 'a is an attribute of x'

(ii) Identifying: 'a is the identity of x'

The general roles pertaining to a relational process are Token and Value, but these can be further specified into Carrier and Attribute for the attributive mode, and Identified and Identifier for the identifying mode. For example:

(31) Mary is tall (Token/Value; Carrier/Attribute)

(32) Mary is the tall one (Token/Value; Identifier/Identified)

(33) Mary is the leader (Value/Token; Identified/Identifier)

(34) Tomorrow is the tenth (Value/Token; Identified/Identifier)

(35) Mary owns the bag (Value/Token; Identified/Identifier)

(36) The bag is Mary's (Value/Token; Identified/Identifier)

In a general sense, the sentences above denote states where the first term of the predication fulfils the role Zero, in much the same way as it occurs with some mental processes. The second term, when it is present, is often a Referent but with circumstantial predicates it can be a Location:

(37) Mary is at the desk

Dik's Zero and Referent roles are but alternative labels for Halliday's Token and Value, with perhaps the exception (apart from the one noted in example (37)) that attributes are not roles but part of the predicate itself, since they are not terms referring to some world entity. For this reason, we shall stick to Dik's analysis and terminology with respect to relational processes.

4. *The treatment of entity-denoting terms*

In language use world entities are set in relation through acts of predication: we say something about one or more entities and the way we see or want somebody to see them related. If this is so, it is not unreasonable to think that the number and nature of these relations has to do with the different process types. Following Dik, we have broadly distinguished four such categories:

[+dyn, +control] = Actions

[+dyn, -control] = Processes

[-dyn, +control] = Positions

[-dyn, -control] = States

This classification allows us to enquire into the nature of entities in some organized way. Not all entities can take part, for example, in an action or in a position (e.g. non-controlling entities). All entities can potentially be in a certain state or suffer some sort of process (change). Some entities can be instrumental in carrying out an action and some can be the cause (or the instigators) of a process.

Linguists, psychologists and lexicologists have tried to discover semantic relations empirically either by examining corpora of data to derive their hypotheses from them or by having subjects to produce associated pairs to calculate strength of bonds (see Clark, 1970), relation types (Chaffin & Hermann, 1988) or prototypicality judgements (Rosch, 1976). Resulting classifications are consequently overlapping and too heavily tied to the limited nature of the samples under scrutiny. Also some of the classifications seem to be too rich and detailed but at the same time perhaps not very revealing about the way we handle relationships in our minds. We must not ignore that any processing system, in order to be optimally efficient, should be able to control the largest amount of information with the least expenditure of resources. The human processor need not be any different (see Sperber & Wilson, 1986). For that reason, we have attempted to establish a set of basic, primitive relations for entity-denoting terms which provide for every such item in the data-base what we may call its prototypical core. This is the information that presumably comes to mind in the first instance as we invoke a term without any particular effort. It is also the information which is considered to square in best with what we believe about the entity referred to by the term. To give but one simple example, think of the animal we call *cat*. A cat usually has a tail, pointed ears, sharp claws a certain oval shape of eyes, and whiskers. These and maybe some other features are highly idiosyncratic with cats. We also associate certain actions with them, like licking milk from a small dish, chasing mice, purring or moving about in a swift, silent way. Of course we may find cat-like creatures which do not match the model in all its aspects but as long as they do not depart from it in too significant a way, we are sure to call them cats. These model features are to be listed in the entry for cat. Other features which are compatible with cats but not typical of them will be derivable from the relational paths provided by Hyperlexis. For example, a cat has a liver, a heart and

kidneys. This information is general for a large number of other creatures and can be found under the entry for *mammal.*

In talking about genericity we have come to another major feature of Hyperlexis. The data-base is hierarchical in the sense that all entries are assigned a level of genericity-specificity, like animal-mammal-cat-Siamese, artefact-tool-saw-hacksaw, etc. The scale of genericity is a top-down/bottom-up one, whereas prototypical relations are to be envisaged as in a horizontal axis. If we consider prototypical relations to be part of a single schema, any item we connect to another in such a manner can be considered to be internal to the schema. Other relations, like hyponymy, will be external.

Internal relations are tied to the four major process types identified above. We distinguish provisionally[7]:

Action:

makes (A shoemaker makes shoes)
is made by (Shoes are made by a shoemaker)

is used for (A saw is used for cutting wood)
is performed by means of (Cutting wood is performed by means of any cutting instrument, typically a saw)

causes (A hurricane causes destruction)
is caused by (Destruction is caused by any destructive agent like a hurricane)

Process:

originates in (Ice originates in water)
is converted into (Water is converted into ice, vapour, etc)

Position:

has (=possesses) (A richman has riches)
belongs to (Riches belong to a richman)

State:

is made of (A table is typically made of wood or metal)
is the material of (Wood is typically the material of furniture)

contains (The sea contains water)
is in (Water is in the sea, the oceans, the rivers, etc.)

is a part of (A leg is part of the body)
consists of (The body consists of head, limbs, and trunk)

External relations are of four major types:

Identifying:

is (menace/threat)

Classifying:

is a type of (penguin/bird)

is classified into (bird/penguin, ostrich, robin, etc.)

Contrasting:

is the opposite of (blessing/curse)
is in contrast to (arm/leg)
is similar to (analogy relationship: foot/hoof)

These relations generalize over a number of other finer subrelations. For example, if you take the relation IS A PART OF, this may apply to a large number of items:

(38)

leg-body

cowboy-rodeo

soldier-army

brother-fraternity

tree-forest

pint-gallon

slice-pie

These entities are not related to each other in exactly the same way even though it might be appropriate in some instances to establish analogies like the following:

(39)

leg is to body as cowboy is to rodeo as tree to forest, etc.

but compare:

(40)

leg is to body as handle is to cup

where both "leg" and "handle" are functional parts of a whole, serving a definite purpose. This functional relationship cannot be found in the other pairs: brother-fraternity, tree-forest, pint-gallon, slice-pie, etc. But the idiosyncracy of these relations derives from the semantic nature of the entity-denoting terms involved and it should not be specified in the network but rather in the definitions of the terms. As part of the definition for *leg*, therefore, together with the physical description we should make mention of its purpose within the body.

In cognitive linguistics it has been noted that meanings do not exist independently by themselves but in the context of other cognitive structures -whether lexicalized or not- which constitute domains. A linguistic form derives its meaning by highlighting (or profiling) a certain region within a domain (see Langacker, 1991). The concept of leg is understood within the domain of the human body in the same way as the concept of "Monday" is to be understood against the domain of the seven-day week, whose

domain in turn is the day-night cycle, whose ultimate domain is the general concept of time. This characterization has one advantage over the traditional assumptions of structural linguistics, where the concept "leg" would be considered meaningful only by contrast with other items like "arms", "head" and "back". One might argue (see Taylor, 1989: 83) that we can learn and think of each of these concepts without referring to the others, but still we are able to relate all of them as functional parts of a larger whole. Hyperlexis provides cognitive descriptions of terms and at the same time incorporates them into a large network of relations where they acquire full meaning. Consider the following entry:

(41)

leg /leg/:

entity type: physical object

description: long part of the body of some animals that extends in humans from the hip to the ankle and in other animals from hip/shoulder to hoofs (typ. horse, cattle)/ paws (typ. dogs, cats)/ claws (bears).

prototype:

in humans: a pair
in quadrupeds: four

Internal relations:

is used for- moving the body in some direction (typ. walking)
is performed by means of
is made of- flesh, bones, veins, arteries
contains- cells
is a part of- body
consists of- thigh, knee, calf, foot/paw (typ. cats, dogs)/hoof (typ. horses) /claw (typ. bears)

External relations:

is a type of- limb
is in contrast to- arm, wing (in birds)

As can be seen from this entry, not all relations are activated for every possible entity-denoting term. Only a subpart of the whole basic set is instantiated. It is this subpart, together with the definition, that constitutes the semantic schema for the term. The concept of semantic schema is much richer than Langacker's concept of domain and more adequate for a semantic description: it is not only a unitary domain of reference that has cognitive relevance but also the vast network of possible relations which we are able to establish for each concept.

It must also be noted that the definition provided for *leg* includes the part-whole and purpose relationships which are listed with the Internal Relations. But in this case what we have done is to give the domain for understanding the concept "leg", the resulting

linguistic expression being rather haphazardly that of the relationship in question. There are other cases where a part-whole relationship is to be specified in the relation network but not in the definition: for example, we can have the following definition of *tree* :

(42) tree /tri:/

entity type: physical object

description: type of plant with a trunk made of wood

prototype: tall, with branches and leaves some distance above the ground

A tree can be part of a forest but a forest is not its reference domain, so the part-whole relationship does not belong to the definition. However, in this case it is the material (IS-MADE-OF) relationship that is to be specified, for nature does not produce trees made of any other material than wood. In other words, being made of wood is a necessary condition for an entity to be called a tree.

Another thing to be noted is that, whenever it is possible, we avoid redundancy in the network of relations by providing the most generic term as a link-up to another more specific ones. Above we can see that the purpose relationship leads to the concept "moving". Then, the lexical entry for "move" will need to cover the different types of action we can call "movement", some of them by using the legs. Here is part of the entry:

(43)

$move_1$ [+action] (x_A, x_P, z_I): x causes y to change position by means of z/ z = entity or action carried out in a certain way w.

$move_2$ [+process] (y_A): y changes position by means of z/ z = entity or action carried out in a certain way w.

y = head --- nod, shake

= eyelids --- wink, blink

z = legs --- (in humans) walk
run
jog
jump
hop
leap
dance

--- (in other animals)
walk
run
trot
gallop

Each of the relations is a lexical entry where we obtain a definition and other related terms. Here are some partial definitions which hark back to the entry for *move*:

(44)

walk: $move_2$ (y_A, z_I, w/ y_A [+human, +legged animal], z_I = legs,
w = along, putting forward each foot in turn
run: $move_2$ (y_A, z_I, w/ y_A [+human, + legged animal], z_I = legs,
w = with quick steps, faster than walking
jog: run (y_A, z_I, w/ y_A [+human, +legged animal], z_I = legs,
w = slowly, steadily, with an up and down movement
jump: $move_2$ (y_A, z_I, w/ y_A [+human, + legged animal], z_I = legs,
w = up into the air, quickly and suddenly, using own strength
hop: jump (y_A, z_I, w/ y_A [+ human], z_I = legs,
w = using one foot
leap: jump (y_A, z_I, w/ y_A [+human, + legged animal], , z_I = legs,
w = a long distance
dance: $move_2$ (y_A, z_I, w/ y_A [+human], z_I = legs,
w = following a rhythm
trot: $move_2$ (y_A, z_I, w/ y_A typ. [horse], z_I = legs,
w = with quick, short steps
gallop: $move_2$ (y_A, z_I, w/ y_A typ. [horse], z_I = legs,
w = faster than trotting, with all four legs off the ground in each stride

Each definition provides selection restrictions with respect to the entity type which can fulfil the relevant roles. The selection restrictions are entries themselves, which allows the user to enquire more about the nature of each entity involved.

4. *Final remarks*

Lack of space has prevented us from dealing with other terms which appear in all dictionaries, like non-content words, and adjectives and adverbs. However, with respect to the former it should be noted that a relational system is essentially focused on content words since in forming networks of relations these are the ones that reflect our knowledge of the world. Non-content words, on the other hand, are used for the expression of certain grammatical operations like definiteness, deixis, person, number, and others and should not be of immediate concern for our data-base.

Adjectives and adverbs make up networks on their own and their description hinges entirely on the nature of the entity-denoting terms and processes to which they can apply. For that reason, a most profitable study of adjectives and adverbs should be reserved for a later stage.

References

Bui, K. P. 1989. "Hyperlexicon, a hypermedia-based lexicon for vocabulary acquisition", in Goos, G. & Hartmanis, J. (eds.), *Lecture Notes in Computer Science*, Springer-Verlag.

Chaffin, R. & Herrmann, D. J.1988. "The nature of semantic relations: a comparison of two approaches", in Evens, M. W. (ed.), *Relational Models of the*

Lexicon. Representing Knowledge in Semantic Networks , Cambridge: Cambridge University Press.

Clark, H. H. 1970. "Word associations and linguistic theory", in Lyons, J. (ed.), *New Horizons in Linguistics*, Harmondsworth, Middlesex: Penguin.

Dik, S. C. 1989. *The Theory of Functional Grammar. Part I: The Structure of the Clause*. Dordrecht: Foris.

Evens, M. W. (ed.). 1988. *Relational Models of the Lexicon. Representing Knowledge in Semantic Networks* , C. U. P.

Halliday, M. A. K. 1985. *An Introduction to Functional Grammar*. London: Edward Arnold.

Langacker, R. 1991. "Cognitive grammar", in Droste, F. G. & Joseph, J. E. (eds.), *Linguistic Theory and Grammatical Description*, Amsterdam/Philadelphia: John Benjamins.

Leech, G. 1981. *Semantics. The Study of Meaning*. Harmondsworth, Middlesex: Penguin.

Petöfi, J. 1985. "Lexicon", in van Dijk, T. (ed.), *Handbook of Discourse Analysis*, Vol. 2, *Dimensions of Discourse*, London, etc.: Academic Press.

Rosch, E. 1976. "Structural bases of typicality effects", *Journal of Experimental Psychology: Human Perception and Performance* 2: 491-502.

Ruiz de Mendoza, F. 1992. "Predicate types and semantic functions: the role of roles", *Proceedings of the 16th Congress of AEDEAN*, University of Valladolid, forthcoming.

Ruiz de Mendoza, F. y Otal, J. 1993. "La operatividad discursiva del paradigma esquemático-procedimental en el hecho discursivo" *Revista Complutense de Estudios Ingleses*; forthcoming.

Ruiz de Mendoza, F. 1993. "Lexical categories, term structure and metaphor", *Proceedings of the XVII AEDEAN Congress*, University of Córdoba, forthcoming.

Ruiz de Mendoza, F. 1994. "La formación de predicaciones", in Martin, J. (ed.) *Functional Grammar*, University of Zaragoza, forthcoming.

Sperber, D. & Wilson, D. 1986. *Relevance. Communication and Cognition*. Oxford: Basil Blackwell.

Taylor, J. R. 1989. *Linguistic Categorization. Prototypes in Linguistic Theory*, Oxford: Clarendon.

APPENDIX

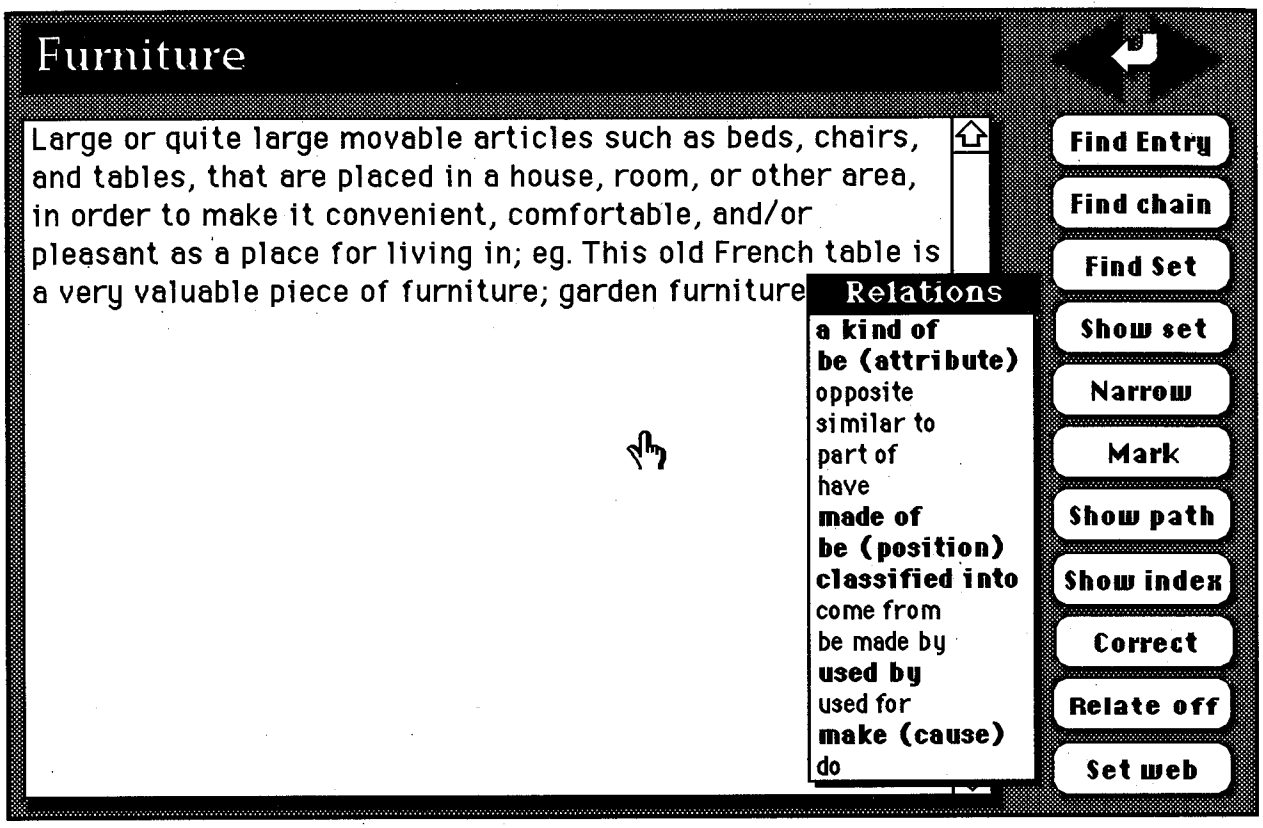

Fig. 1. Hyperlexis is an interactive dictionary with a set of built-in utilities —mainly a parser able to find the entry from all its different forms—, and different navigational tools —some fixed, like the main index, and others to be adapted by the user, like paths with the words you have already visited, sets of related words, search sets and so on, which the user can clear, save, modify and sort.

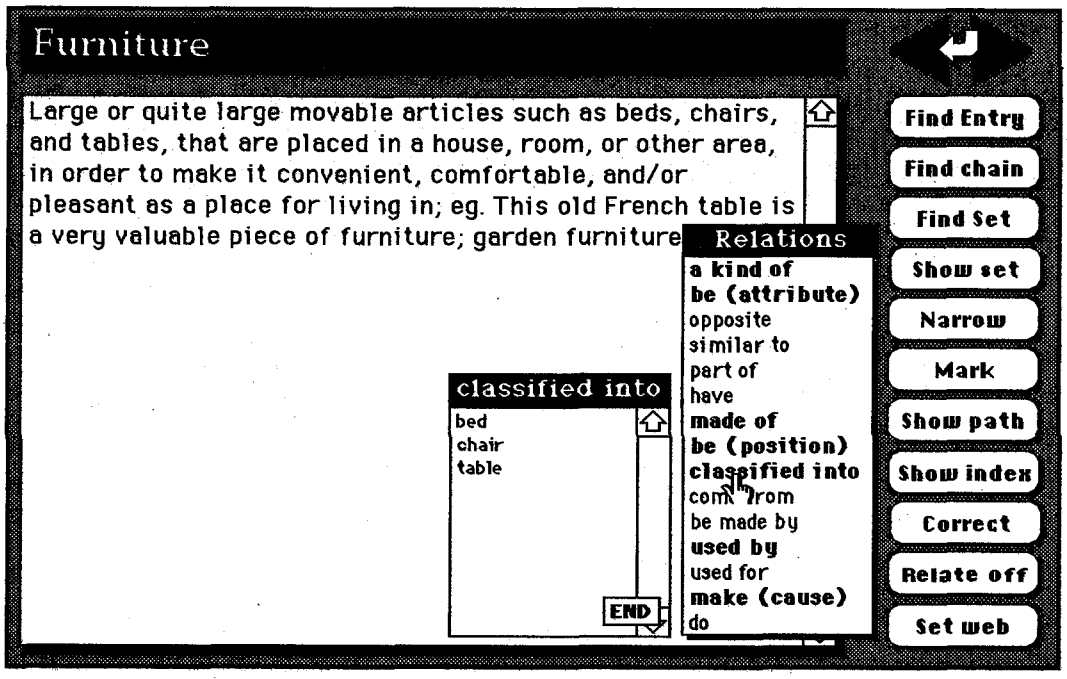

Fig. 2. The main feature of Hyperlexis is, nevertheless, its relational capabilities. The user is able to look for words related to one entry, choosing from a very specific set of relations, based on an elaborated theory of semantic relations (see text). The system of relations is predetermined throughout the system, and Hyperlexis tells the user which relations are available.

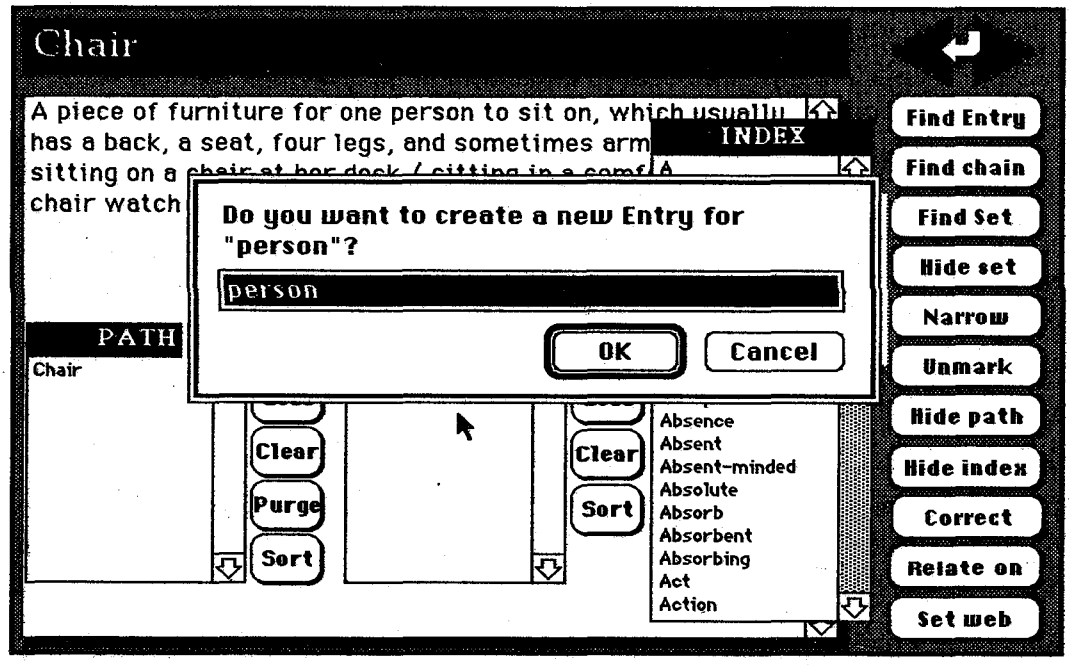

Fig. 3. The vocabulary grows dinamically. When the user asks for a word that is not an entry in the Dictionary, Hyperlexis tells him that the word is not available and asks him whether he wants to create a new entry. Thereafter, the system creates a new card and opens the definition field to allow editing.

NOTAS:

[1].- Hyperlexis is being developed on the basis of version 2.0. of Hypercard for Macintosh. For further information on Hyperlexis utilities and other technical details the reader is referred to the Hyperlexis User's Guide, which can be obtained from the authors at the following address: jgarcia@cc.unizar.es (Javier García), or Javier García, Departamento de Biblioteconomía, Facultad de Filosofía y Letras, Universidad de Zaragoza, c/Pedro Cerbuna 12, 50009, Zaragoza, Spain.

[2].- For information on the actual layout of Hyperlexis, see the explanations and illustrations in the Appendix.

[3].- It must be noted that here we are dealing with non-metaphorical uses of terms. As is well known, some metaphors have found their way into the language, a factor which is to be accounted for by the organization of a lexicon. Others, however, are innovative. We intend Hyperlexis to be sensitive to novel metaphors as well, in the sense that the dictionary needs to provide all the necessary information for metaphorical meaning to be worked out. See Ruiz de Mendoza (1993)

[4].- The reader may be referred to the well-known notions of 'schemata', 'frames', 'scripts', 'scenarios' and others, elaborated in the mid 1970s. For a review, see Ruiz de Mendoza & Otal (1993).

[5].- In Ruiz de Mendoza (1994) some arguments for and against the distinction between arguments and satellites are discussed. The idea defended by the author is that there is no neat division between both but rather a continuum of "argumenthood" and "satellitehood" of all terms which instantiate what in FG is termed the "predicate frame".

[6].- In a user-friendly version of Hyperlexis formalizations like this can be spelled out easily as follows:

kill: an action in which a certain entity causes another (living) entity to die by means of (i) another (physical/non-physical) entity, typically a weapon or other killing instrument, (ii) performing an action.

[7].- Note that each relation works in two directions. By way of illustration, if a tree is part of a forest, it is because a forest consists of trees.