A GRAMMAR AND A LEXICON FOR A TEXT-PRODUCTION SYSTEM

Christian M.I.M. Matthiessen
USC/Information Sciences Institute

ABSTRACT

In a text-production system high and special demands are placed on the grammar and the lexicon. This paper will view these components in such a system (overview in section 1). First, the subcomponents dealing with semantic information and with syntactic information will be presented separately (section 2). The problems of relating these two types of information are then identified (section 3). Finally, strategies designed to meet the problems are proposed and discussed (section 4). One of the issues that will be illustrated is what happens when a systemic linguistic approach is combined with a KL-ONE like knowledge representation - a novel and hitherto unexplored combination.[1]

1. THE PLACE OF A GRAMMAR AND A LEXICON IN PENMAN

This paper will view a grammar and a lexicon as integral parts of a text production system (PENMAN). This perspective leads to certain requirements on the form of the grammar and that of the subparts of the lexicon and on the strategies for integrating these components with each other and with other parts of the system. In the course of the presentation of the components, the subcomponents and the integrating strategies, these requirements will be addressed. Here I will give a brief overview of the system.

PENMAN is a successor to KDS ([12], [14] and [13]) and is being created to produce multi-sentential natural English text. It has as some of its components a knowledge domain, encoded in a KL-ONE like representation, a reader model, a text-planner, a lexicon, and a sentence generator (called NIGEL). The grammar used in NIGEL is a Systemic Grammar of English of the type developed by Michael Halliday -- see below for references.

For present purposes the grammar, the lexicon and their environment can be represented as shown in Figure 1.

The lines enclose sets; the boxes are the linguistic components. The dotted lines represent parts that have been developed independently of the present project, but which are being implemented, refined and revised, and the continuous lines represent components whose design is being developed within the project.

The box labeled *syntax* stands for syntactic information, both of the general kind that is needed to generate structures (the grammar; the left part of the box) and of the more specific kind that is needed for the syntactic definition of lexical items (the syntactic subentry of lexical items; to the right in the box -- the term lexicogrammar can also be used to denote both ends of the box).

[1]This research was supported by the Air Force Office of Scientific Research contract No. F49620-79-c-0181. The views and conclusions contained in this document are those of the author and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research of the U.S. Government. The research represents a joint effort and so do the ideas stemming from it which are the substance of this paper. I would like to thank in particular William Mann, who has helped me think, given me many helpful ideas and suggestions and commented extensively on drafts of the paper; without him it would not be. I am also grateful to Yasutomo Fukumochi for helpful comments on a draft and to Michael Halliday, who has made clear to me many systemic principles and insights. Naturally, I am solely responsible for errors in the presentation and content.

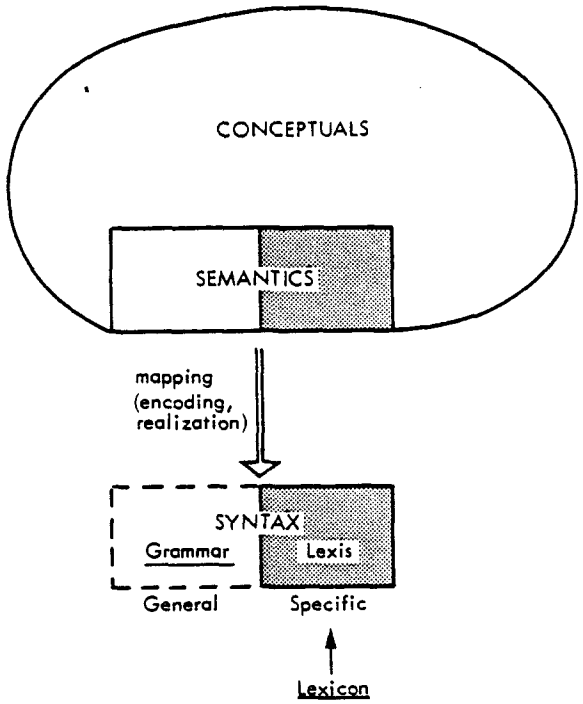

Figure 1-1: System overview.

The other box (*semantics*) represents that part of semantics that has to do with our conceptualization of experience (distinct from the semantics of interaction -- speech acts etc. -- and the semantics of presentation -- theme structure, the distinction between given and new information etc.). It is shown as one part of what is called *conceptuals* -- our general conceptual organization of the world around us and our own inner world; it is the linguistic part of conceptuals. For the lexicon this means that lexical semantics is that part of conceptuals which has become lexicalized and thus enters into the structure of the vocabulary. There is also a correlation between conceptual organization and the organization of part of the grammar.

The double arrow between the two boxes represents the mapping (realization or encoding) of semantics into syntax. For example, the concept SELL is mapped onto the verb *sold*.[2]

The grammar is the general part of the syntactic box, the part concerned with syntactic structures. *The lexicon* cuts across three levels: it has a semantic part, a syntactic part (lexis) and an orthographic part (or spelling; not present in the figure).[3] The lexicon

[2]I am using the general convention of capitalizing terms denoting semantic entries. Capitals will also be used for roles associated with concepts (like AGENT, RECIPIENT and OBJECT) and for grammatical functions (like ACTOR, BENEFICIARY and GOAL). These notions will be introduced below.

[3]This means that an entry for a lexical item consists of three subentries, viz. a semantic entry, a syntactic entry and an orthographic entry. The lexicon box is shown as containing parts of both syntax and semantics in the figure (the shaded area) to emphasize the nature of the lexical entry.

consists entirely of independent lexical entries, each representing one lexical item (typically a word).

This figure, then, represents the part of the PENMAN text production system that includes the grammar, the lexicon and their immediate environment.

PENMAN is at the design stage; consequently the discussion that follows is tentative and exploratory rather than definitive. -- The component that has advanced the farthest is the grammar. It has been implemented in NIGEL, the sentence generator mentioned above. It has been tested and is currently being revised and extended. None of the other components (those demarcated by continuous lines) have been implemented; they have been tested only by way of hand examples. This paper will concentrate on the design features of the grammar rather than on the results of the implementation and testing of it.

2. THE COMPONENTS

2.1. Knowledge representation and semantics

The knowledge representation

One of the fundamental properties of the KL-ONE like knowledge representation (KR) is its intensional -- extensional distinction, the distinction between a general conceptual taxonomy and a second part of the representation where we find individuals which can exist, states of affairs which may be true etc. This is roughly a distinction between what is conceptualizable and actual conceptualizations (whether they are real or hypothetical). In the overview figure in section 1, the two parts are together called conceptuals.

For instance, to use an example I will be using throughout this paper, there is an intensional concept SELL, about which no existence or location in time is claimed. An intensional concept is related to extensional concepts by the relation Individuates: intensional SELL is related by individual instances of extensional SELLs by the Individuates relation. If I know that Joan sold Arthur ice-cream in the park, I have a SELL fixed in time which is part of an assertion about Joan and it Individuates intensional SELL.[4] A concept has internal structure: it is a configuration of roles. The concept SELL has an internal structure which is the three roles associated with it, viz. AGENT (the seller), RECIPIENT (the buyer) and OBJECT. These roles are slots which are filled by other concepts and the domains over which these can vary are defined as value restrictions. The AGENT of SELL is a PERSON or a FRANCHISE and so on.

In other words, a concept is defined by its relation to other concepts (much as in European structuralism). These relations are roles associated with the concept, roles whose fillers are other concepts. This gives rise to a large conceptual net.

There is another relation which helps define the place of a concept in the conceptual net, viz. SuperCategory, which gives the conceptual net a taxonomic (or hierarchic) structure in addition to the structure defined by the role relations. The concept SELL is defined by its place in the taxonomy by having TRANSACTION as a SuperCategory. If we want to, we can define a concept that will have SELL as a SuperCategory (i.e. bear the SuperCategory relation to SELL), for example SELLOB 'sell on the black market'. As a result, part of the taxonomy of events is TRANSACTION --- SELL --- SELLOB.

If TRANSACTION has a set of roles associated with it, this set may be inherited by SELL and by SELLOB -- this is a general feature of the SuperCategory relation. In the examples involving SELL that follow, I will concentrate on this concept and not try to generalize to its supercategories.

The Semantic Subentry

In the overview figure (1-1), the semantics is shown as part of the conceptuals. The consequence of this is that the set of semantic entries in the lexicon is a subset of the set of concepts. The subset is proper if we assume that there are concepts which have not been lexicalized (the assumption indicated in the figure). The assumption is perfectly reasonable; I have already invented the concept SELLOB for which there is no word in standard English; it is not surprising if we have formed concepts for which we have to create expressions rather than pick them ready-made from our lexicon. Furthermore, if we construct a conceptual component intended to support say a bilingual speaker, there will be a number of concepts which are lexicalized in only one of the two languages.

A semantic entry, then, is a concept in the conceptuals. For *sold*, we find sold with its associated roles, AGENT, RECIPIENT and OBJECT. The right part of figure 4-1 below (marked "se:"; after a figure from [1] gives a more detailed semantic entry for *sold*: a pointer identifies the relevant part in the KR, the concept that constitutes the semantic entry (here the concept SELL).

The concept that constitutes the semantic entry of a lexical item has a fairly rich structure. Roles are associated with the concept and the modality (necessary or optional), the cardinality of and restrictions on (value of) the fillers are given.

Through the value restriction the linguistic notion of selection restriction is captured. *The stone sold a carnation to the little girl* is odd because the AGENT role of SELL is value restricted to PERSON or FRANCHISE and the concept associated with *stone* falls into neither type.

The strategy of letting semantic entries be part of the knowledge representation would not have been possible in a notation designed to capture specific propositions only. However, since KL-ONE provides the distinction between intension and extension, the strategy is unproblematic in the present framework.

So what is the relationship between intensional-extensional and semantic entries? The working assumption is that for a large part of the vocabulary, it is the concepts of the intensional part of the KR that may be lexicalized and thus serve as semantic entries. We have words for intensional objects, actions and states, but not for individual extensional objects etc. with the exception of proper names. They have extensional concepts as their semantic entries. For instance, *Alex* denotes a particular individuated person and *The War of the Roses* a particular individuated war.

Both the SuperCategory relation and the Individuates relation provide ways of walking around in the KR to find expressions for concepts. If

[4] It should be emphasized that calling the concept SELL says nothing whatsoever about the English expression for it; the reasons for giving it this name are purely mnemonic. The only way the concept can be associated with the word *sold* is through being part of a lexical entry.

we are in the extensional part of the KR, looking at a particular individual, we can follow the Individuates link up to an intensional concept. There may be a word for it, in which case the concept is part of a lexical entry. If there is no word for the concept, we will have to consider the various options the grammar gives us for forming an appropriate expression.

The general assumption is that all the intensional vocabulary can be used for extensional concepts in the way just described: expressability is inherited with the Individuates relation.

Expression candidates for concepts can also be located along the SuperCategory link by going from one concept to another one higher up in the taxonomy. Consider the following example: *Joan sold Arthur ice-cream. The transaction took place in the park.* The SuperCategory link enables us to go from SELL to TRANSACTION, where we find the expression *transaction*.

Lexical Semantic Relations

The structure of the vocabulary is parasitic on the conceptual structure. In other words, lexicalized concepts are related not only to one another, but also to concepts for which there is no word-encoding in English (i.e. non-lexicalized concepts).

Crudely, the semantic structure of the lexicon can be described as being part of the hierarchy of intensional concepts -- the intensional concepts that happen to be lexicalized in English. -- The structure of English vocabulary is thus not the only principle that is reflected in the knowledge representation, but it is reflected. Very general concepts like OBJECT, THING and ACTION are at the top. In this hierarchy, roles are inherited. This corresponds to the semantic redundancy rules of a lexicon.

Considering the possibility of walking around in the KR and the integration of lexicalized and non-lexicalized concepts, the KR suggests itself as the natural place to state certain text-forming principles, some of which have been described under the terms lexical cohesion ([8]) and Thematic Progression ([6]).

I will now turn to the syntactic component in figure 1-1, starting with a brief introduction to the framework (Systemic Linguistics) that does the same for that component as the notion of semantic net did for the component just discussed.

2.2. Lexicogrammar

Systemic Linguistics stems from a British tradition and has been developed by its founder, Michael Halliday (e.g. [7], [9], [10]) and other systemic linguists (see e.g. [5], [4] for a presentation of Fawcett's interesting work on developing a systemic model within a cognitive model) for over twenty years covering many areas of linguistic concern, including studies of text, lexicogrammar, language development, and computational applications. Systemic Grammar was used in SHRDLU [15] and more recently in another important contribution, Davey's PROTEUS [3].

The systemic tradition recognizes a fundamental principle in the organization of language: the distinction between *choice* and the *structures* that express (realize) choices. Choice is taken as primary and is given special recognition in the formalization of the systemic model of language. Consequently, a description is a specification of the choices a speaker can make together with statements about how he realizes a selection he has made. This realization of a set of choices is typically linear, e.g. a string of words. Each choice point is formalized as a *system* (hence the name Systemic). The options open to the speaker are two or more *features* that constitute alternatives which can be chosen. The preconditions for the choice are *entry conditions* to the system. Entry conditions are logical expressions whose elementary terms are features.

All but one of the systems have non-empty entry conditions. This causes an interdependency among the systems with the result that the grammar of English forms one network of systems, which cluster when a feature in one system is (part of) the entry condition to another system. This dependency gives the network depth: it starts (at its "root") with very general choices. Other systems of choice depend on them (i.e. have a feature from one of these systems -- or a combination of features from more than one system -- as entry conditions) so that the systems of choice become less general (more *delicate* to use the systemic term) as we move along in the network.

The network of systems is where the control of the grammar resides, its non-deterministic part. Systemic grammar thus contrasts with many other formalisms in that choice is given explicit representation and is captured in a single rule type (systems), not distributed over the grammar as e.g. optional rules of different types. This property of systemic grammar makes it a very useful component in a text-production system, especially in the interface with semantics and in ensuring accessibility of alternatives.

The rest of the grammar is deterministic -- the consequences of features chosen in the network of systems. These consequences are formalized as feature *realization statements* whose task is to build the appropriate structure.

For example, in independent indicative sentences, English offers a choice between declarative and interrogative sentences. If interrogative is chosen, this leads to a dependent system with a choice between wh-interrogative and yes/no-interrogative. When the latter is chosen, it is realized by having the FINITE verb before the SUBJECT.

Since it is the general design of the grammar that is the focus of attention, I will not go through the algorithm for generating a sentence as it has been implemented in NIGEL. The general observation is that the results are very encouraging, although it is incomplete. The algorithm generates a wide range of English structures correctly. There have not been any serious problems in implementing a grammar written in the systemic notation.

Before turning to the lexico- part of lexicogrammar, I will give an example of the toplevel structure of a sentence generated by the grammar. (I have left out the details of the internal structure of the constituents.)

[1]	LOCATION	ACTOR	PROCESS	BENEFICIARY	GOAL
[2]		SUBJECT	FINITE		
[3]	THEME				
	In the park	Joan	sold	Arthur	ice-cream

The structure consists of three layers of function symbols, all of which are needed to get the result desired. -- The structure is not only functional (with function symbols labeling the constituents instead of category names like Noun Phrase and Verb Phrase) but it is multifunctional.

Each layer of function symbols shows a particular perspective on the clause structure. Layer [1] gives the aspect of the sentence as a representation of our experience. The second layer structures the sentence as interaction between the speaker and the hearer; the fact that SUBJECT precedes FINITE signals that the speaker is giving the hearer information. Layer [3] represents a structuring of the clause as a message; the THEME is its starting point. The functions are called experiential, interpersonal and textual respectively in the systemic framework: the function symbols are said to belong to three different metafunctions. In the rest of the paper I will concentrate on the experiential metafunction, partly because it will turn out to be highly relevant to the lexicon.

The syntactic subentry.

In the systemic tradition, the syntactic part of the lexicon is seen as a continuation of grammar (hence the term lexicogrammar for both of them): lexical choices are simply more detailed (delicate) than grammatical choices (cf. [9]). The vocabulary of English can be seen as one huge taxonomy, with Roget's Thesaurus as a very rough model.

A taxonomic organization of the relevant part of the vocabulary of English is intended for PENMAN, but this organization is part of the conceptual organization mentioned above. There is at present no separate lexical taxonomy.

The syntactic subentry potentially consists of two parts. There is always the class specification -- the lexical features. This is a statement of the grammatical potential of the lexical item, i.e. of how it can be used grammatically. For *sold* the class specification is the following:

```
verb
class 10
class 02
benefactive
```

where "benefactive" says that *sold* can occur in a sentence with a BENEFICIARY, "class 10" that it encodes a material process (contrasting with mental, verbal and relational processes) and "class 02" that it is a transitive verb.

In addition, there is a provision for a configurational part, which is a fragment of a structure the grammar can generate, more specifically the experiential part of the grammar.[5] The structure corresponds to the top layer (# [1]) in the example above. In reference to this example, I can make more explicit what I mean by fragment. The general point is that (to take just one class as an example) the presence and character of functions like ACTOR, BENEFICIARY and GOAL -- direct participants in the event denoted by the verb -- depend on the type of verb, whereas the more circumstantial functions like LOCATION remain unaffected and applicable to all types of verb. Consequently, the information about the possibility of having a LOCATION constituent is not the type of information that has to be stated for specific lexical items. The information given for them concerns only a fragment of the experiential functional structure.

The full syntactic entry for *sold* is:

```
PROCESS • verb
          class 10
          class 02
          benefactive
ACTOR   •

GOAL    •

BENEFICIARY •
```

This says that *sold* can occur in a fragment of a structure where it is PROCESS and there can be an ACTOR, a GOAL and a BENEFICIARY. The usefulness of the structure fragment will be demonstrated in section 4.

3. THE PROBLEM

I will now turn to the fundamental problem of making a working system out of the parts that have been discussed.

The problem has two parts to it, viz.

1. the design of the system as a system with integrated parts and

2. the implementation of the system.

I will only be concerned with the first aspect here.

The components of the system have been presented. What remains -- and that is the problem -- is to design the missing links; to find the strategies that will do the job of connecting the components.

Finding these strategies is a design problem in the following sense. The strategies do not come as accessories with the frameworks we have used (the systemic framework and the KL-ONE inspired knowledge representation). Moreover, these two frameworks stem from two quite disparate traditions with different sets of goals, symbols and terms.

I will state the problem for the grammar first and then for the lexicon. As it has been presented, *the grammar* runs wild and free. It is organized around choice, to be sure, but there is nothing to relate the choices to the rest of the system, in particular to what we can take to be semantics. In other words, although the grammar may have a part that faces semantics -- the system network, which, in Halliday's words, is semantically relevant grammar -- it does not make direct contact with semantics. And, if we know what we want the system to encode in a sentence, how can we indicate what goes where, that is what a constituent (e.g. the ACTOR) should encode?

The lexicon incorporates the problem of finding an appropriate strategy to link the components to each other, since it cuts across component boundaries. The semantic and syntactic subparts of a lexical entry have been outlined, but nothing has been said about how they should be matched up with one another. The reason why this match is not perfectly straightforward has to do with the fact that both entries may be structures (configurations) rather than single elements. In addition, there are lexical relations that have not been accounted for yet, especially synonymy and polysemy.

---

[5] This configurational part does not stem from the systemic tradition, but is an exploration in the present design.

4. LOOKING FOR THE SOLUTIONS

4.1. The Grammar

Choice experts and their domains.

The control of the grammar resides in the network of systems. Choice experts can be developed to handle the choices in these systems.

The idea is that there is an expert for each system in the network and that this expert knows what it takes to make a meaningful choice, what the factors influencing its choice are. It has at its disposal a table which tells it how to find the relevant pieces of information, which are somewhere in the knowledge domain, the text plan or the reader model.

In other words, the part of the grammar that is related to semantics is the part where the notion of choice is: the choice experts know about the semantic consequences of the various choices in the grammar and do the job of relating syntax to semantics.[6]

The recognition of different functional components of the grammar relates to the multi-functional character of a structure in systemic grammar I mentioned in relation to the example *In the park Joan sold Arthur ice-cream* in section 2.2. The organization of the sentence into PROCESS, ACTOR, BENEFICIARY, GOAL, and LOCATIVE is an organization the grammar imposes on our experience, and it is the aspect of the organization of the sentence that relates to the conceptual organization of the knowledge domain: it is in terms of this organization (and not e.g. SUBJECT, OBJECT, THEME and NEW INFORMATION) that the mapping between syntax and semantics can be stated. -- The functional diversity Halliday has provided for systemic grammar is useful in a text-production system; the other functions find uses which space does note permit a discussion of here.

Pointers from constituents.

In order for the choice experts to be able to work, they must know where to look. Assume that we are working on *in the park* in our example sentence *In the park Joan sold Arthur ice-cream* and that an expert has to decide whether *park* should be definite or not. The information about the status in the mind of the reader of the concept corresponding to *park* in this sentence is located at this concept: the trick is to associate the concept with the constituent being built. In the example structure given earlier, *in the park* is both LOCATION and THEME, only the former of which is relevant to the present problem. The solution is to set a pointer to the relevant extensional concept when the function symbol LOCATION is inserted, so that LOCATION will carry the pointer and thus make the information attached to the concept accessible.

4.2. The lexicon and the lexical entry

I have already introduced the semantic subentry and the syntactic subentry. They are stated in a KL-ONE like representation and a systemic notation respectively. The question now is how to relate the two.

In the knowledge representation the internal structure of a concept is a configuration of roles and these roles lead to new concepts to which the concept is related. A syntactic structure is seen as a configuration of function symbols; syntactic categories serve these functions -- in the generation of a structure the functions lead to an entry of a part of the network. For example, the function ACTOR leads to a part of the network whose entry feature is Nominal Group just as the role AGENT (of SELL) leads to the concept that is the filler of it. The parallel between the two representations in this area are the following:

KNOWLEDGE REPRESENTATION	SYNTACTIC REPRESENTATION
role	function
filler	exponent

(Where *exponent* denotes the entry feature into a part of the network (e.g. Nominal Group) that the function leads to.)

This parallel clears the path for a strategy for relating the semantic entry and the syntactic entry. The strategy is in keeping with current ideas in linguistics.[7] Consider the following crude entry for *sold*, given here as an illustration:

Subentries:

semantic		syntactic			orthographic
		Functions		Lexical features	
SELL-concept	=	PROCESS	=	verb class 10 class 02 benefactive	"sold"
AGENT	=	ACTOR			
OBJECT	=	GOAL			
RECIPIENT	=	BENEFICIARY			

where the previously discussed semantic and syntactic subentries are repeated and paired off against each other.

This full lexical entry makes clear the usefulness of the second part of the syntactic entry -- the fragment of the experiential functional structure in which *sold* can be the PROCESS.

Another piece of the total picture also falls into place now. The notion of a pointer from an experiential function like BENEFICIARY in the grammatical structure to a point in the conceptual net was introduced above. We can now see how this pointer may be set for individual lexical items: it is introduced as a simple relation between a grammatical function symbol and a conceptual role in the lexical entry of e.g. SELL. Since there is an Individuates link between this intensional concept and any extensional SELL the extensional concept that is part of the particular proposition that is being encoded grammatically, the pointer is inherited and will point to a role in the extensional part of the knowledge domain.

At this point, I will refer again to the figure below, whose right half I have already referred to as a full example of a semantic subentry ("se:"). "sp:" is the spelling or orthographic subentry; "ge:" is the syntactic subentry.

We have two configurations in the lexical entry: in the semantic subentry the concept plus a number of roles and in the syntactic subentry a number of grammatical functions. The match is represented in the figure above by the arrows.

---

[6] A possible definition of the full semantics of the grammar is, as a result of this approach, "semantics = what the grammatical choice experts look at". In the present discussion, I have focused on the knowledge domain only, partly because this is the area most relevant to lexical semantics.

[7] The mechanism for mapping has much in common with one developed for Lexical Functional Grammar (see e.g. [2]), although the levels are not the same. The entry also resembles a lexical entry in the Pan-Lexicalism framework developed by Hudson in [11].

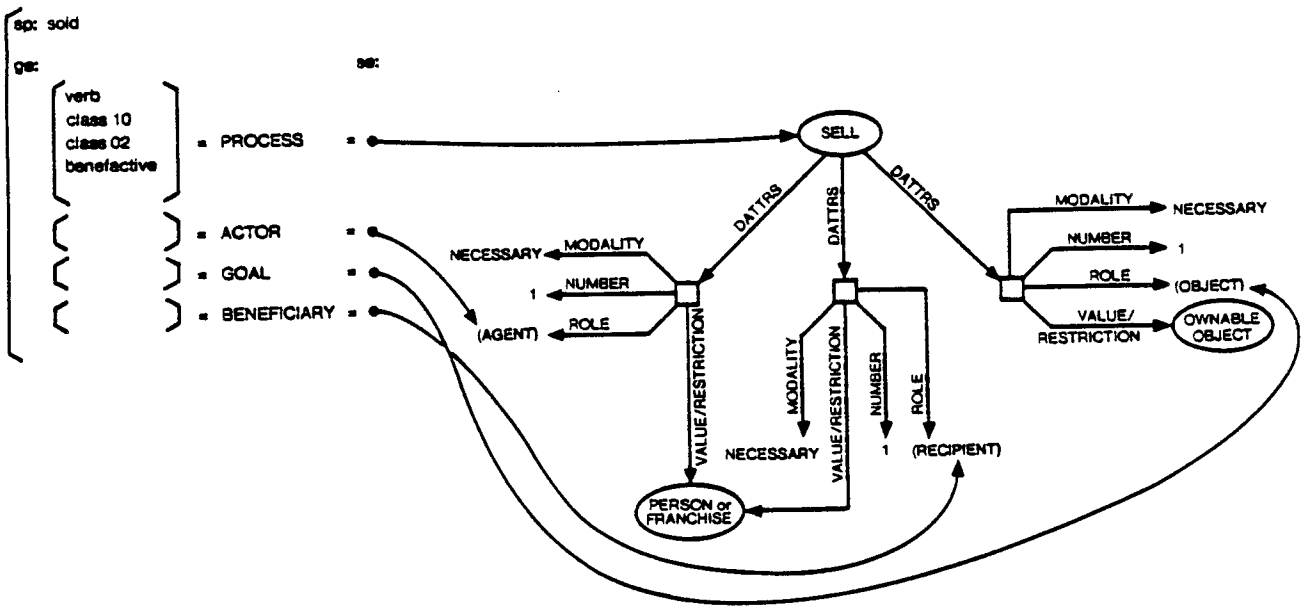

Figure 4-1: Lexical entry for *sold*

All three roles of SELL have the modality "necessary". This does not dictate the grammatical possibilities. The grammar in Nigel offers a choice between e.g. *They sold many books to their customers* and *The book sold well*. In the second example, the grammar only picks out a subset of the roles of SELL for expression. In other words, the grammar makes the adoption of different perspectives possible.[8] I can now return to the observation that the functional diversity Halliday has provided for systemic grammar is useful for our purposes. The fact that grammatical structure is multi-layered means that those aspects of grammatical structure that are relevant to the mapping between the two lexical entries are identified, made explicit (as ACTOR BENEFICIARY etc.) and kept separate from principles of grammatical structuring that are not directly relevant to this mapping (e.g. SUBJECT, NEW and THEME).

In conclusion, a strategy for accounting for *synonymy* and *polysemy* can be mentioned.

The way to capture synonymy is to allow a concept to be the semantic subentry for two distinct orthographic entries. If the items are syntactically identical as well, they will also share a syntactic subentry. Polysemy works the other way: there may be more than one concept for the same syntactic subentry.

5. CONCLUSION

I have discussed a grammar and a lexicon for PENMAN in two steps. First I looked at them as independent components -- the semantic entry, the grammar and the syntactic entry -- and then, after identifying the problems of integrating them into a system, I turned to strategies for relating the grammar to the conceptual representation and the syntactic entry to the semantic one within the lexicon.

In the first step I introduced the KL-ONE like knowledge representation and the systemic notation and indicated how their design features can be put to good use in PENMAN. For instance, the distinction between intension and extension in the knowledge representation makes it possible to let lexical semantics be part of the conceptuals. It was also suggested that the relations SuperCategory and Individuates can be used to find expressions for a particular concept.

The second step attempted to connect the grammar to semantics through the notion of the choice expert, making use of a design principle of systemic grammars where the notion of choice is taken as basic. I pointed out the correlation between the structure of a concept and the notion of structure in the systemic framework and showed how the two can be matched in a lexical entry and in the generation of a sentence, a strategy that could be adopted because of the multi-functional nature of structure in systemic grammars. This second step has been at the same time an attempt to start exploring the potential of a combination of a KL-ONE like representation and a Systemic Grammar.

Although many aspects have had to be left out of the discussion, there are a number of issues that are of linguistic interest and significance. The most basic one is perhaps the task itself: designing a model where a grammar and a lexicon can actually be made to function as more than just structure generators. One issue related to this that has been brought up was that different parts external to the grammar find resonance in different parts of the grammar and that there is a partial correlation between the conceptual structure of the knowledge representation and the grammar and lexicon.

As was emphasized in the introduction, PENMAN is at the design stage: there is a working sentence generator, but the other aspects of what has been discussed have not been implemented and there is no commitment yet to a frozen design. Naturally, a large number of problems still await their solution, even at the level of design and, clearly, many of them will have to wait. For example, selectivity among terms, beyond referential adequacy, is not addressed.

---

[8] The strategy of letting the functional syntactic entry pick up different parts of a concept and adopt different perspectives finds many uses, e.g. in the treatment of pairs like *buy* vs. *sell* and *give* vs. *receive* and in the account for nominalizations.

In general, while noting correlations between linguistic organization and conceptual organization, we do not want the relation to be deterministic: part of being a good verbalizer is being able to adopt different viewpoints -- verbalize the same knowledge in different ways. This is clearly an area for future research. Hopefully, ideas such as grammars organized around choice and choice experts will prove useful tools in working out extensions.

REFERENCES

1. Brachman, Ronald, *A Structural Paradigm for Representing Knowledge*, Bolt, Beranek, and Newman, Inc., Technical Report, 1978.

2. Bresnan, J., "Polyadicity: Part I of a Theory of Lexical Rules and Representation," in Hoekstra, van der Hulst & Moortgat (eds.), *Lexical Grammar*, Dordrecht, 1980.

3. Davey, Anthony, *Discourse Production*, Edinburgh University Press, Edinburgh, 1979.

4. Fawcett, Robin P., *Exeter Linguistic Studies*. Volume 3: *CognitiveLinguistics and Social Interaction*, Julius Groos Verlag Heidelberg and Exeter University, 1980.

5. Fawcett, R. P., Systemic Functional Grammar in a Cognitive Model of Language. University College, London. Mimeo, 1973

6. Danes, F., ed., *Papers on Functional Sentence Perspective*, Academia, Publishing House of the Czechoslovak Academy of Sciences, 1974.

7. Halliday, M. A. K., "'Categories of the theory of grammar'," *Word 17*, 1961.

8. Halliday M. A. K. and R. Hasan, *Cohesion in English*, Longman, London, 1976. English Language Series, Title No. 9

9. Halliday, M. A. K. , *System and Function in Language*, Oxford University Press, London, 1976.

10. Hudson, R. A., *North Holland Linguistic Series*. Volume 4: *English complex sentences*, North Holland, London and Amsterdam, 1971.

11. Hudson, R. A., DDG Working Papers. University College, London. Mimeo, 1980

12. Mann, William C., and James A. Moore, *Computer as Author--Results and Prospects*, USC/Information Sciences Institute, Research report 79-82, 1980.

13. Mann, William C. and James A. Moore, Computer Generation of Multiparagraph English Text, 1979. AJCL, forthcoming.

14. Moore, James A., and W. C. Mann, "A snapshot of KDS, a knowledge delivery system," in *Proceedings of the Conference, 17th Annual Meeting of the Association for Computational Linguistics*, pp. 51-52, August 1979.

15. Winograd, Terry, *Understanding Natural Language*, Academic Press, Edinburgh, 1972.