Lexicon and Grammar

Dr. Anna Herwig

Lexicon and Grammar

Dr. Anna Herwig

Trinity College Dublin

Abstract

Over the past decades it has become generally acknowledged that lexicon and grammar are inseparably linked, constituting "a continuum of symbolic structures" (Langacker 1990:2). Yet, a comprehensive integration of the two realms of knowledge appears to be a difficult task. The present article offers a unified psycholinguistic perspective, which is centred on the mental lexicon, considering grammatical knowledge as part of the information structure of lexical items. It aims to model the complexity of lexical knowledge such that its perceived psychological reality, including various levels of linguistic description, becomes discernible.

1. Lexical Information Structure[1]

In the light of recent research undertaken in different linguistics fields (cf., e.g., Bybee 1988, Ellis 1997, Langacker 1990, Singleton 1999), the mental lexicon can be described as that domain of language where the various dimensions of linguistic information meet. A comprehensive description of lexical knowledge must therefore take account not only of formal and semantic but also of grammatical knowledge. It follows that the elements of the lexicon need to be modelled as highly complex entities, including information on representational substance (conceptual, perceptual, and articulatory patterns) and combinatorial potential. Combinatorial knowledge relates to an item's collocation and colligation, i.e., its valency structure (cf., e.g., Langacker 1987, Lutjeharms 1994, Singleton 1999). It has various facets and is relevant for phrasal construction. Combinatorial knowledge is associated with specific types of semantic and formal relations, which reflect the distributional properties of lexical elements (cf., e.g., Bybee 1988, Ellis 1997).

The different knowledge components are mutually dependent and interact in our use of language. I will aim to integrate representational substance and combinatorial potential of lexical items in an all-embracing psycholinguistic component structure model, which coordinates the different levels of description. The model provides a framework for discussing grammatical processing with reference to lexical knowledge. It also draws a unified picture of lexical items, which provides the grounds for illustrating the perceived psychological reality of lexical networks.

---

[1] For a more differentiated discussion of lexical information structure and applications of the proposed model cf. Herwig 1994.

2. Lexical Items as Mental Models

In order to illustrate how the various dimensions and aspects of linguistic knowledge merge in the information structure of lexical items, I suggest that words should be viewed as mental models, similar to conceptual frames employed by cognitive semantics to describe conceptual structure (cf., e.g., Barsalou 1992, Lakoff/Johnson 1999, Minsky 1975, Mervis/Rosch 1981). Following Cohen (1990:316), mental models are defined as extended conceptual frames, "consist[ing] of constructed groups of concepts which constitute the generic knowledge about events, scenarios, actions, or objects … [and of] relations together with slots, or variables, which can be filled with optional values". Lexical information structure can be compared to that of complex cognitive events like scenarios etc., and the organizational structure of mental models appears to be tailor-made for explaining the complexity of lexical knowledge. I will therefore describe a lexical (semantic[2]) item in terms of a multidimensional knowledge unit in the form of a conceptual frame consisting of a number of interrelated and dynamically interacting domains of knowledge. It illustrates grammatical knowledge by elucidating lexical valency and dependency structures and renders possible a comprehensive psycholinguistic perspective on lexical organization and processing. This is considered useful for explaining the principles of phrasal construction and for elucidating the foundations of production errors.

Figure 1 displays the dimensions of lexical knowledge as a set of seven attributes, five of which, *semantic quality*, *semantic and syntactic valency*, and *phonological and (ortho)graphic layout*, relate to intellectual aspects, the remaining two (*pronunciation structure* and *written output structure*) to motor patterns. These latter are connected to their respective domains of formal knowledge and are primarily relevant for language production, while the other dimensions are essential for all situations of language use. Motor knowledge will not be discussed any further here.

---

[2] The following model is primarily designed to illustrate the information structure of *semantic* items. It may therefore not be entirely suitable to represent *grammatical* items, in particular their semantic structure.
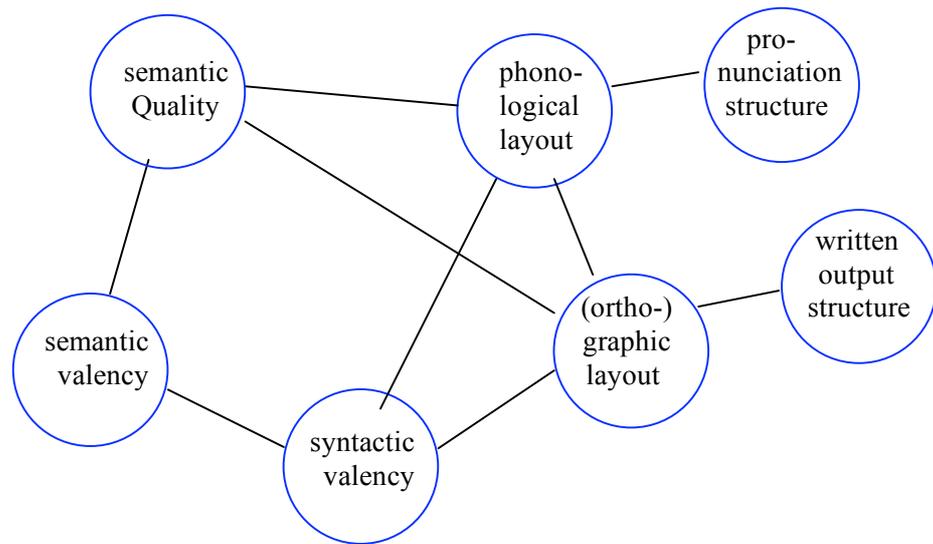
Figure 1 Lexical information structure: frame of attributes

Semantic quality, semantic and syntactic valency, and phonological and orthographic layout are regarded as the central attributes of lexical knowledge, which set the outline conditions for language use. These attributes, which corresepond to specific dimensions of knowledge, are particularised for a range of values, some of which are further differentiable.
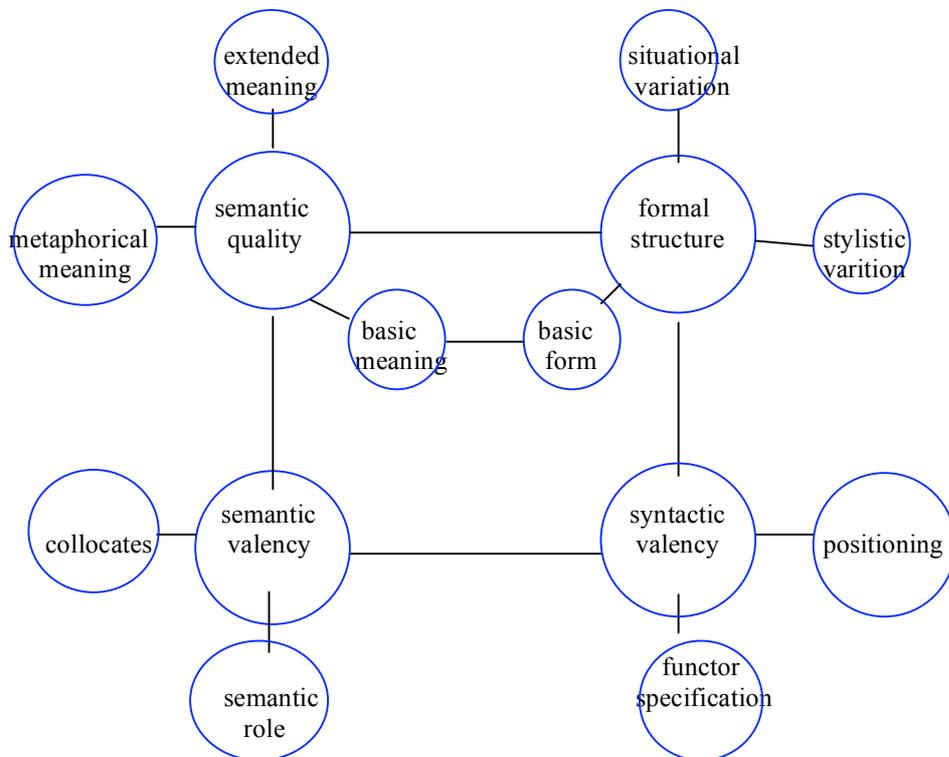
Figure 2 Lexical information structure: detailed Frame Model

Figure 2 illustrates a selection of values associated with the different attributes. Each of them is seen as being individually pronounced for a specific lexical item. Phonological and orthographic knowledge have, for the sake of simplicity, been subsumed under the heading of *formal structure*. The correlation of these attributes is highly variable, and even in alphabetical systems the degree of correspondence between sounds and letters can vary considerably.

We are now left with a frame consisting of four attribute-value sets, four dimensions of lexical knowledge, which interact in putting lexical items to use. They are part of an individual item's information structure, but their impact extends beyond the word level. Before discussing their dynamic relation, I will outline the function of the attribute-value sets. They are seen as representing the following knowledge structures:

Semantic quality relates to aspects of meaning, comprising a *basic meaning*, context-dependent *modified* or *extended* shades of this meaning, and *metaphorical* meanings.

Semantic valency refers to meaning-related combinatorial aspects, which include the range of associative connections with other semantic items (*collocates*), and to an item's semantic role within such configurations. In other words, the attribute specifies the way lexical items are connected on the basis of their conceptual content. It covers, on the one hand, concrete information about possible alliances, as emerging from an item's conceptual frame. **Collocates** relate to different types of associations with other semantic items, which can be distinguished as being either *natural, conventional, personal,* or *situational*. Semantic valency is further seen as including semantico-grammatical information about an item's semantic role in a complex setting. It refers to the relational status or function of an entity in a given scene. Semantic role thus relates to the notion of syntactic category as the grammatical specification of the semantic status of lexical items.

Closely associated with semantic valency is **syntactic valency**. It relates to aspects of formal representation, namely to the arrangement of lexical forms as reflective of the relations that hold between their conceptual contents. It thus refers to *grammatical* or *morpho-syntactic adaptations*, in particular to *functor specification* and to the determination of an item's position in phrase and sentence (*positioning*). The interdependence and interaction of semantic and syntactic valency will be discussed in more detail later.

The remaining attribute-value set (*formal structure*) relates to lexical-formal knowledge. It includes a *basic form*, and information on possible *situational* or *stylistic variations*, as, for example, allophonic variation in spoken language.

I will now take a closer look at the relations within the lexical frame, and at the interdependence of the different attributes and their values.

To begin with, the given quadrangle can be sectioned in two ways:

1. The two dimensions on the left relate to aspects of conceptual-semantic structure, those on the right to their formal linguistic representation. This division singles out the complementary poles of a symbolic structure.
2. The upper two attributes relate to an items representational substance, the lower two to its contextual behaviour.

The most fundamental relation within the lexical frame holds between *basic meaning* and *basic form*. Basic form relates to a non-inflected form; basic meaning is used to refer to its most readily associated conceptual content, which may vary across individuals but often relates to central category members (e.g., the lexical form 'horse' is usually associated with a prototypical horse as opposed to its wider meaning which would include Shetland ponies and other more peripheral members). This primary word knowledge could be considered the *core entry* of the mental lexicon and might be compared to the basic entry in a dictionary.

Semantic quality and semantic valency structure are seen as being mutually dependent. The precise meaning of an item emerges from its contextual embedding, i.e., its collocational environment (e.g., the meaning of horse in the context of riding is different from that in the context of handicraft). Conversely, a specific meaning can put restraints on the collocations an item can enter into. An item's status or function relative to its collocates is situationally specified. It finds its formal expression in morpho-syntactic modifications, which are associated with syntactic valency structure (e.g., whether the mouse is afraid of the elephant or *vice versa* determines their sentence position and inflectional adaptations (if applicable)).

Semantic and syntactic valency provide the outline framework for an item's integration in larger contexts, i.e., for combining lexical items into meaningful stretches of language. As such, their combined information can be regarded as the main body of grammatical knowledge. Morpho-syntactic information is also related to an item's formal structure. The latter determines the shape that morphological modifications will take (e.g., type of plural suffix). Conversely, morpho-syntactic specifications can give rise to formal adaptations (e.g., shift of stress pattern).

Following this outline of the suggested dependencies within the lexical frame, I will proceed with a closer look at the information structure of semantic and syntactic valency and give a basic outline of their interrelation.

Semantic valency is defined as the information about an item's combinatorial potential at the conceptual level, which is formally expressed in its syntactic valency (Lutjeharms 1994:150). It relates to two aspects of collocational information; to possible co-occurring elements and to the item's possible positions in relation to these elements. An item's relational status, i.e., its position or function relative to other items, is referred to as *semantic role*, and it is seen as being conventionally associated with morpho-syntactic specifications to express this role. This means that, for nouns, for example, their thematic role is linked with certain grammatical modifications (as associated with syntactic category), which involve determination of sentence position, inflection, or use of prepositions, as appropriate. Verbs would be identified as to their specific predication in relation to their arguments and adapted accordingly (inflection according to person, number, tense, aspect, etc.). Beyond that, they have an influence on the grammatical behaviour of other words. The valency, or argument structure of relational items, in particular verbs and other predicative elements, includes information about required modifications of their collocates in a given context. They can determine grammatical structures (as, for example, infinitive or gerund constructions) and morpho-syntactic modifications of their arguments (such as word order or case structure in many languages, or the use of prepositions). The syntactic valency of predicative items can thus be described as having an *internal* dimension, relating to their own grammatical behaviour, and an *external* one, relating to modifications of other elements. Also of importance is the fact that predicative items are seen as linking up grammatical items, i.e., that these are regarded as part of their information structure.

In sum, the valency structure of lexical items is seen as holding the main body of grammatical information. This supports the view that the separation of lexicon and grammar is an artificial distinction, and that language production springs from lexical knowledge, rather than from a pre-lexical grammatical scaffolding to be filled out with words, as has been suggested by the Chomskyan tradition. The grammatical scaffolding of a phrase or sentence arises from the information structure of the lexical items used and is specified on the basis of situational conditions.

3. Phrasal Construction

Utterance formulation can be defined as the selection and meaningful arrangement of lexical items. Different schools of thought distinguish different procedural mechanisms, sometimes involving rule-based grammatical operations that appear to be complex beyond cognitive reality (cf. Herwig 1994:66ff). A useful delineation is provided by Aitchison's (1989) model of sentence planning, which posits two major processing stages, labelled as *outline planning* and *detailed planning*. They are seen as relating to the selection of certain key words

(semantic items) and the syntactic structures they determine, and to the activation and arrangement of supplementary elements. Aitchison's perspective can be set in relation to Garrett's (e.g., 1988) model, which distinguishes two classes of processing systems, lexical selection and phrasal construction. Phrasal construction includes the syntactic specifications assigned to Aitchison's outline planning and the operations involved in her account of detailed planning.

I will now discuss the mechanisms of phrasal construction in the light of lexical information structure. They are seen as involving the interaction and interdependence of different knowledge components and operational mechanisms. It needs to be emphasised that phrasal construction is here not seen as a distinct processing phase, but as a description of the kind of operations traditionally associated with grammatical processing. These mechanisms and their underlying organisational principles can be elucidated with reference to the above frame model.

The semantic items selected to represent a given conceptual event are here considered to contain, by virtue of their component properties, all the information required to express this concept in a complex phrase or sentence. The specific configuration of the conceptual scenery, i.e., the nature and relational status of its component parts, determines the semantic role of their referring expressions and is seen as initiating morpho-syntactic specifications. These include their sequential activation and morphological modification and the triggering of function words. This operation is likely to involve two processing steps, because morphological adaptations and the use of grammatical words depend on the nature and arrangement of the selected content words (cf. *Peter's book* versus *the book of Peter*, where the serial order of *Peter* and *book* determines the use of different functors). Perhaps more important to emphasise, however, is the fact that phrasal construction is here seen as involving the parallel and interactive processing of the different dimensions of knowledge of all the selected semantic items. The resulting phrase or sentence structure is then posited to be a product of the joint information of the different items involved.

Much of the process of phrasal construction can be related to the retrieval of flexible formulaic sequences. Such prefabricated linguistic patterns determine a specific grammatical construction and provide slots for certain groups of lexical items (cf., e.g., Perkins/Wray 2000). The most fundamental formulaic sequences are seen in this perspective as being governed by the valency structure of predicative items. These determine the grammatical construction they appear in, as well as the morpho-syntactic specification of obligatory and possible arguments. Carrying the main body of information about sentence structure, predicative items are regarded on this view as the key items of phrasal construction. Their

function is to specify the relation that holds between certain events, not only by virtue of their meaning, but also by providing a structural framework for verbalising this relation. This framework offers a range of possibilities, which are situationally realised depending on the nature of the arguments and on the perspective taken on a given scene. An example may serve to illustrate these explanations:

The inner state of fear can be expressed by the adjective *afraid* or by the verb *fear*. Both items require two arguments, an experiencer and a cause. *Fear* determines that the experiencer has to be in the subject position and the cause in the object position. The latter can take the shape of a thing or an action. *Afraid* also requires a form of *be* to link up inner state and experiencer, and the preposition *of* to link up the cause. It then offers two possibilities: if the cause is a thing, a noun can be slotted in; in the case of an action, a gerund is required. Let us assume Peter cannot swim. The situation provides experiencer (Peter), relation (fear), and cause (water or swimming). If the item *afraid* is selected to express Peter's fear, it automatically triggers the linking functors and determines the sequential arrangement of the different elements. The realisation of *be* as *is*, additionally depends on the fact that the experiencer is a single person, and on the speaker/writer's present viewpoint. We thus arrive at one of the following utterances:

- *Peter is afraid of water*
- *Peter is afraid of swimming*

A selection of the verb *fear* to verbalise Peter's feeling would have resulted in one of the formulations:

- *Peter fears water*
- *Peter fears swimming*

Here, the syntactic valency of the predicate does not specify any linking elements to express the interrelation of the different semantic items.

The proposed mechanism of phrasal construction as outlined here suggests that the semantic items selected to express a given conceptual configuration immediately and interactively trigger the grammatical items consistent with them. This hypothesis bears important implications for the organization of grammatical items (cf., e.g., Bates/MacWhinney 1989, Green 1993). Elements with a mere grammatical function, such as many bound morphemes or semantically empty words like postpositions, could be bound up in the lexical network primarily on the basis of formal connections to semantic items. Their selection may – at least in many instances – depend entirely on the semantic items they modify, and they may not

have in such instance an independent semantic quality. An investigation of the processing of grammatical items could shed further light on this question.

4. Multidimensional Connectivity of Lexical Items

Lexical connectivity relates to semantic and formal associations of different types. At the conceptual level, an item is linked to semantically related items, at the formal level to formally similar or grammatically associated items. This complexity is traditionally described with reference to different types of lexical networks (cf., e.g., Aitchison 1994). In order to illustrate the perceived psychological reality of lexical connectivity, it would be useful to integrate the different levels of description in a single comprehensive network model.

An integrative network model needs to portray the diversity of associations that hold between words, distinguishing different types of semantic and formal connections. In order to be transparent and psychologically realistic, such a model has to portray lexical items as unified cognitive entities. The lexical information structure model provides an integrative perspective on lexical items, which satisfies these demands. It could be used to represent words in single-level networks. It could, however, be slightly confusing in a graphic illustration, if each item were represented by its lexical frame. An alternative display, which renders a network model more transparent, would be to illustrate the component properties of lexical knowledge as a set of layers centred around the semantic quality as their core attribute (cf. figure 3).
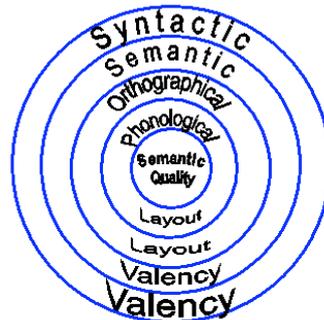
Figure 3 Lexical information structure: Layer Model

This display has the additional advantage of being able to rank the relative significance of the attributes in descending order from centre to periphery. The two inner circles, semantic quality and formal structure, can be said to constitute the 'hardware' of lexical knowledge, the linguistic sign. This portrayal is seen as representing the character of a lexical item particularly well: a conceptual content is encapsulated in a linguistic form. It shows the inseparability of the two dimensions; a form without a meaning would be an empty cover, a non-word, while a conceptual content without a formal frame would be a non-lexicalised

concept. The two outer layers, then, represent the item's combinatorial potential: information about putting lexical items to use, or 'lexical software'. Again, the semantic component is contained in, i.e., expressed through a formal shape.

The specific nature of the relations that hold between words in a comprehensive network could now be illustrated as associative connections at different levels of description, by linking the respective attribute layers. In this way, the specific character of the relationship of an item like horse to semantically or formally related items (e.g., pony, foal, or hoarse, house, respectively) and to collocationally or colligationally associated ones (e.g., stable, riding, or a, the, and an inflectional suffix like plural –s, respectively) can be integrated in a relatively transparent way. A further advantage of this approach is that multiple connections between lexical items become visible. Associations at several linguistic levels, as opposed to single connections only, could be regarded as a factor of strength of lexical links.

The above manner of illustrating lexical connectivity is thus seen as being useful for both making visible and disentangling the confusing complexity of lexical networks. Perhaps most importantly, it helps define an item's specific position in the lexical network.

5. Conclusion

The proposed frame model depicting the information structure of lexical items, and its underlying perspective on lexical information structure, has been shown useful for different purposes. It elucidates the interrelation of lexical and grammatical knowledge, providing a comprehensive and psychologically realistic view of the various dimensions of linguistic knowledge merging. This integrative view renders a description of the mechanisms involved in grammatical processing and phrasal construction more transparent, and can also be used for detecting the foundations of production errors. It has further been found suitable to provide a framework for illustrating the multidimensional structure of lexical networks.

6. References

Aitchison, J. (1989) *The Articulate Mammal*. London and New York: Routledge.
Aitchison, J. (1994) *Words in the Mind*. Oxford: Blackwell.
Barsalou, L.W. (1992a) *Cognitive Psychology*. Hillsdale, NJ.:Lawrence Erlbaum.
Bates, E. and MacWhinney, B. (1989) Functionalism and the Competition Model. In B. MacWhinney and E. Bates (eds.) *The Crosslinguistic Study of Sentence Processing*. CUP. 3-73.
Bybee, J. (1988) Morphology as Lexical Organisation. In M. Hammond and M. Noonan (eds.) *Theoretical Morphology*. London: Academic Press. 119-45.
Cohen, G. (1990) Schemata. In M.W. Eysenck, A.W. Ellis, E.B. Hunt, and P.N. Johnson-Laird (eds.) *The Blackwell Dictionary of Cognitive Psychology*. Oxford and Cambridge, MA: Blackwell. 316-7.

Ellis, N. (1997) The Epigenesis of Language: Acquisition as a Sequence Learning Problem. In A. Ryan and A. Wray (eds.) *Evolving Models of Language.* British Studies in Applied Linguistics. Clevedon: Multilingual Matters. 41-57.

Garrett, M.F. (1988) Processes in Language Production. In F.J. Niemeyer (ed.) *Linguistics: The Cambridge Survey* Vol. III. CUP. 69-96.

Green, D.W. (1993) Towards a Model of L2 Comprehension and Production. In R. Schreuder and B. Weltens (eds.) *The Bilingual Lexicon*. Amsterdam: John Benjamins. 249-78.

Herwig, A. (2004) *Aspects of Linguistic Organisation. Evidence from Lexical Processing in L1-L2 Translation*. Vasa: Åbo Akademi.

Lakoff, G. and M. Johnson (1999) *Philosophy in the Flesh*. New York: Basic Books.

Langacker, R. (1987) *Foundations of Cognitive Grammar 1. Theoretical Prerequisites.* Stanford, CA: Stanford University Press.

Langacker, R. (1990) *Concept, Image, and Symbol: The Cognitive Base of Grammar.* Berlin, New York: Mouton de Gruyter.

Lutjeharms, M. (1994) Worterkennen beim Lesen einer Fremdsprache. In W. Börner and K. Vogel (eds.) *Kognitive Linguistik und Fremdsprachenerwerb*. Tübingen: Narr. 149-68.

Mervis, C. and E. Rosch (1981) Categorisation of Natural Objects. In *Annual Review of Psychology* 32. 89-115.

Minsky, M. (1975) A Framework for Representing Knowledge. In P.H. Winston (ed.) *The Psychology of Computer Vision*. New York: McGraw-Hill. 211-77.

Perkins, M.R. and A. Wray (2000) The Functions of Formulaic Language: An Integrated Model. In *Language and Communication* 20. 1-28.

Singleton, D.M. (1999) *Exploring the Second Language Mental Lexicon*. CUP.