THE DIFFERENT LEVELS OF LANGUAGE PATTERNING [1]

MARÍA JOSÉ LUZÓN MARCO
Universidad de Zaragoza

ABSTRACT

A lot of natural language is not freely generated, but it is formulaic or prepatterned language: there are prefabricated linguistic patterns that are resources for a more effective construction and interpretation of discourse. The purpose of this paper is to examine the different levels of language patterning and to show that patterning is inherent to language use. The user's co-selection of linguistic elements derives from the repeated co-occurrence of elements which has given way to typical uses/ patterns and usually to institutionalisation. Language patterning involves different levels of abstraction, including habitual co-selections of words, of phrases, of clauses, of semantic features and of functional units.

1. INTRODUCTION

Language patterning has been devoted a great deal of attention recently (Bolinger, 1976; Coulmas, 1981; Pawley and Syder, 1983; Fillmore *et al.*, 1988;

Tannen, 1989; Barkema, 1996), and the statement that «all discourse is more or less prepatterned» (Tannen, 1989: 42) has been explored from different perspectives. Language patterning has been discussed under different labels which refer in fact to the same phenomenon: patterning (Tannen, 1989; Biber, 1996), idiomaticity, phraseology and formulaic language (Fillmore *et al.*, 1988; Wong Fillmore, 1994), conventionalised or ritualised language (Coulmas, 1981; Atkinson, 1990). Wong Fillmore (1994: 230) defines formulaic language as «*conventionalised phraseology* including idioms, routines, *preplanned or prefabricated* phrasing that function as units or *patterns* for speakers*» (my own emphasis). This is only part of language patterning, since it seems to be restricted to prefabricated patterns at the lexical and phrasal level with no reference to discourse structure. We adopt the term «language patterning» here because we consider it a broader and more inclusive term, which may be used to give a proper account of how language works at all levels. Biber (1996: 173) uses this sense of patterning when describing *association patterns* as «the systematic ways in which linguistic features are used in association with other linguistic and nonlinguistic features». Language patterning refers to the co-occurrence or repeated co-selection of linguistic elements, that is, to the existence of prefabricated conventionalised linguistic patterns that are used as resources to construct and interpret discourse by users of language. The concept of pattern is related to the notion of *typicality* and *preference*. Hanks (1987: 121) makes an interesting point when discussing *selection preferences* at the lexical level.

> The basis of choice has its root in the notion of typicality. The words of English simply do not typically combine and recombine freely and randomly. Not only can typical grammatical structures and form classes be observed, but also typical collocates. The distinction between the possible and the typical is of greatest importance (...) But when we ask how a word is typically used, rather than how it might possibly be used, we can generally discover a relatively small number of distinct patterns.

Language patterning is concerned with the syntagmatic dimension of language relations: the combinations that linguistic elements enter into in the production of discourse. What happens with words also happens at other levels of language, since meaning cannot only be analysed at the lexical level. At all levels there are *selection preferences*, which means that although there are several alternatives, there are typical uses. Firth (1968: 176) considers that in order to make statements of meaning we have to «accept language events as integral in experience regarding them as wholes and as repetitive and interconnected» and then we have to deal with these events at various levels, beginning with the

context of situation. An important aspect of linguistic patterns is that they may consist of actual forms (e.g. *dark night*) or may include different levels of abstraction (e.g. adjective+ noun; Situation-Problem-Solution-Evaluation). The question to be answered is: at which levels is language patterned, or, to put it in another way, which units of language are patterned?

Atkinson (1990) distinguishes four different levels at which discourse conventions function: the macro-rhetorical, rhetorical, phrasal-clausal, and lexical level. Although this is a useful starting point we cannot establish such a clear-cut distinction for the study of language patterning. As we will see, it is not easy to determine the boundaries between a lexical and a phrasal pattern.

Our purpose here is to examine the different levels of language patterning [2] and describe patterning as systematic and inherent to language use and therefore pervasive. The assumptions on which this analysis is based have already been introduced: 1) To understand the norms that govern the users' language behaviour, linguistic analysis should deal with typical patterns of usage, rather than with possible linguistic forms; 2) There are different levels of meaning; 3) Language is patterned at all levels.

2. RESEARCH PERSPECTIVES IN LANGUAGE PATTERNING

Current research on language patterning has provided a wide range of accounts of this aspect of language. Three areas of linguistic research have taken a special interest in the patterning of language: corpus-based studies of lexis, studies concerned with language production and processing, studies of discourse structure and generic features.

2.1. Corpus-based studies of lexis

These studies are mainly concerned with phraseology, this being understood as the study of word combination. At this lexical level the most important concept is that of collocation, «the characteristic co-occurrence patterns of concept is that of collocation, «the characteristic co-occurrence patterns of words» (McEnery and Wilson, 1996: 71). Whether two words form a significant collocation or not is determined by statistical measurements and is measured in quantitative terms. In the compilation of the COBUILD dictionary collocates

---

[2] Although there may be recurrent co-selections at all levels, including co-selections of phonological and prosodic options, we will not deal with these selections here.

were considered to be «lexical items occurring within five words either way of the headwords with a greater frequency than the law of averages would lead you to expect» (Krisnamurthy, 1987: 70). Thus, collocation is seen as a probabilistic matter of occurrence. There is an interest in language as a product, in the meaning of words as a result of the items with which they collocate. The idea that underlies this type of research is Firth's (1957) statement that «you shall know a word by the company it keeps».

Traditionally, word combinations have been given different names (e.g. idioms, compounds, set phrases) and there have been numerous attempts to categorise them following criteria like fossilisation and flexibility. Collocation is part of what Bonelli (1996: 132) calls «co-selection» and Clear (1993: 272) refers to as «stereotyping»: habitual and distinct word combinations. Both Clear and Bonelli make a very important point regarding these combinations: they are not necessarily idiomatic. That is, they do not always have a meaning that is different from the combination of meanings of the words. As Bonelli (1996: 132) puts it, «there is a cline of co-selection ranging from words that are isolated to words that acquire a new idiomatic function by virtue of being co-selected with other words».

2.2. Studies concerned with language production and processing

One of the most important contributions to the study of language patterning was Bolinger's (1976) suggestion that language users have at their disposal a repertoire of ready-made multi-word «chunks» that are used in the production of language [3]. That is, not all language is the result of using a series of generative/productive rules to combine minimal units. As Tannen (1989: 37) points out, «language is less freely generated, more prepatterned than most current linguistic theory acknowledges».

This idea has been supported by other linguists (Coulmas, 1979, 1981; Yorio, 1980), some of whom have described the role of formulaic language in first and second language acquisition (Peters, 1983; Nattinger and DeCarrico, 1992). An interesting study has been done by Pawley and Syder (1993), who explain the differences in fluency and novelty between native and non-native speakers by resorting to the concept of patterning.

---

[1] This was not a new idea, but was discarded by most linguists as a proper explanation for language production after Chomsky's (1965) claim that speakers produce new sequences of word, or «novel» sentences, and that language use is rule-governed.

A revealing approach to the study of language patterning is to describe it as a result of repetition and intertextuality. Tannen (1989: 44) considers that «all language is a repetition of previous language». There is a balance between repetition and novelty: «Language in discourse is not either prepatterned or novel, but more or less prepatterned» (Tannen, 1989: 38). Given that repetition occurs with all the units of language, Tannen considers that patterning can occur at any level: lexis, grammar, discourse structure, and even topic.

Even grammar, which has traditionally been considered the most rule-go-verned aspect of language, is described as a result of intertextuality. In Hopper's model of *emergent grammar*, grammar is «a set of...recurrent partials, whose status is constantly being renegotiated in speech» (1988: 118). From this pers-pective grammar also includes prior texts which are retrieved when using lan-guage. Fillmore *et al.* (1988) consider that part of a language user's competence cannot be explained without reference to his/her use of morphosyntactic patterns, which frequently have specific pragmatic functions.

2.3. Studies of discourse structure and generic features

A great number of studies have described the rhetorical patterns which or-ganise a text. Winter (1977) and Hoey (1983) describe discourse patterns such as *Problem-Solution* and *Hypothetical-Real*, and Meyer and Rice (1982) pro-pose several types of rhetorical organisation: *collection, causation, description, problem-solution, comparison.* They are ways of organising the topic which re-present abstract schemata (see Carrell, 1983).

Another important area of research in relation to patterning is that of gen-re analysis (Swales, 1990; Dudley-Evans, 1994). One of the most important fea-tures of a genre is its rhetorical structure. As Bhatia (1993: 13) puts it, «most often (a genre) is highly structured and conventionalised with constraints on allowable contributions in terms of their intent, positioning, form and functio-nal value». Highly structured and conventionalised genres tend to be associated with a specific rhetorical structure and with specific lexicosyntactic features.

3. WORD, PHRASE AND CLAUSE-LEVEL CO-SELECTION: COLLOCATION, COLLIGATION, AND COLLOCATIONAL PHRASES

One of the best studied aspects of language patterning is word combina-tion, given its interest for lexicography. As we have said above, traditional accounts have undertaken a classification of word combinations using criteria

such as fixity, collocability or idiosyncratic meaning. A good overview of the traditional terms for idioms and other lexicalised expressions can be found in Barkema (1996).

Barkema (1996: 127) points out that the traditional definition of *idiom* includes two aspects: «(a) idioms are expressions which contain at least two lexical items and (b) the meaning of an idiom is not the combinatorial result of the meanings of the lexical items in the expression». We are not interested in the distinction between idiomatic and non-idiomatic expressions, given that our purpose is not to categorise formulaic language. The study of idiomatic expressions as a special type of language leads to the view that language is only prefabricated at the lexical level and that patterns have a peripheral role in language.

Sinclair (1987) puts forward a more inclusive sense of the word «co-selection». He considers that actual language use follows two complementary principles: «the open-choice principle», which involves the potential to generate rule-governed sequences of words, and the «idiom principle». According to the «idiom principle» «a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments» (Sinclair, 1987: 320). It is from this point of view that the concept of co-selection is basic for understanding how discourse is produced. Sinclair does not mention that these single choices should have an idiosyncratic meaning. Word combinations with idiosyncratic meaning are only a part of (semi-) preconstructed phrases.

A useful concept to discuss language patterning at the lexical level is that of *collocation*. The term «collocation» was used by Firth (1957) to refer to a mode of meaning. Meaning by collocation is «an abstraction at the syntagmatic level and not directly connected with the conceptual or idea approach to the meaning of words» (Firth, 1957: 196). In his own words, «one of the meanings of *night* is its collocability with *dark*». In this connection, Biber (1996: 173) defines collocations as «characterizations of a word in terms of the other words that it typically co-occurs with». Following Firth the term collocation has been used to refer to any combination of words which co-occur frequently and regularly, e.g. *dark night*. The word which is taken as the focus of analysis is referred to as the «node» and the items which co-occur with it within a specified span (e.g. five words on either side of each occurrence) are its «collocates». Collocations can be more or less significant, depending on the strength of association between the words that collocate, which as we have said is measured with statistical methods. The concept of collocation seems to suggest the combination of two words, as Jones and Sinclair (1973: 19) indicate when defining collocation as «the co-occurrence of two items in a text with a specified environment». However, there are much longer collocational sequences, e.g. *from time to time*.

The existence of a syntagmatic co-selection is not only reflected in the concept of collocation, but also in the concept of *colligation* (Firth, 1968). Originally colligation was used to refer to «the interrelation of the syntactical categories within collocation», which constitutes the grammatical level of meaning (Firth, 1968: 28). While collocations are «*actual* words in habitual company» (Firth, 1968: 168) colligations cannot be defined in terms of words but at a more abstract level. Thus, Firth argues that the grammatical relations in «I watched» are not relations between the words *I* and *watched* but between the first person singular nominative personal pronoun and the past tense of a transitive verb. The term colligation has been extended by corpus linguists to refer to «the syntactic patterning found around nodes» (Barnbrook, 1996: 102), that is, the syntactic patterning of a specific word[4].

The concept of collocation implies that the syntagmatic associations that a word has with its collocates determine the meaning of that word. Therefore, the word should not be considered the unit of meaning and of discourse production (Sinclair, 1996). However, it is not an easy task to determine the boundaries of the unit. For instance, the participle *indebted* is closely associated with the adverb *deeply* (i.e. the adverb that most significantly collocates with *indebted* is *deeply*), forming the collocation *deeply indebted*. However, *indebted* is also closely associated with the preposition *to*: *indebted to*. The question is: where is the boundary of the unit? There is an overlap of patterns which Hunston and Francis (1998: 69) call «flow of patterns»: «Whenever a word that is part of a pattern has a pattern of its own, this phenomenon of pattern flow occurs». Patterns flow into one another. Thus, *deeply indebted* flows into *indebted to*. This idea agrees with the proposal by Nattinger and DeCarrico (1992) that the use of language consists in the stringing together of prefabricated groups of words.

Another important point is that the co-selection is not only restricted to lexical and grammatical aspects but also to semantic aspects. McIntosh (1961) introduced the concept of *range* to refer to the words that co-occur with a specific item. For instance, the range of *molten* would be the list of words that may be qualified by this adjective (e.g. *metal, iron, lava*, etc.). The list of words which collocate with an item may belong to one or several lexical *sets* (Halliday, 1966). Thus, in some cases a repeated pattern may involve not only the

---

[4] The distinction between collocation and colligation is similar to the distinction other researchers have established between *lexical* and *grammatical collocation*. BENSON (1990: 61) defines lexical collocation as a recurrent combination of two lexical items, and grammatical collocation as a recurrent combination of a lexical word (verb, noun, adjective) and a grammatical word, e.g. *admiration for, abstain from*.

combination of two words but of more abstract elements: «word+ lexical set». The collocates of a word may be the realisation of a semantic feature that the word selects. If we take for instance the collocation *radical change*, this is in fact a realisation of a less obvious pattern: *radical*+ «noun expressing change». The most significant collocates of *radical* are *change, changes,* and *reform.* It also collocates strongly with other items indicating change such as *overhaul* and *departure*. Other lexical sets with which *radical* collocates strongly are words referring to a group (e.g. *party, groups, feminists*) or to ideas (e.g. *plan, idea*). Further examples are the collocations *I hate to admit* and *engrained prejudices. I hate to admit* is a realisation of a more abstract pattern: *I hate to*+ *discourse verb (usually concessive)* (e.g. *say, admit, recognise*). Similarly, *deeply engrained* illustrates another pattern: *engrained* tends to collocate with nouns expressing a negative attitude or feeling (e.g. *prejudices, hostility, attitudes, feelings, negativity, facets, sexism, racism, habits of dissent*). As we can see in this example, the semantic feature shared by the collocates of a word may include specific semantic prosodies (positive, negative, neutral).

A good example of the co-selection of lexical, grammatical and semantic elements is provided by Sinclair (1996) in his analysis of the collocation *naked eye*. The regularity with which the definite article *the* occurs at N-1 (one position before the node *naked eye*), and a preposition occurs at N-2 enables to consider them as part of the unit: *to/with the naked eye*. Additionally, the items that occur at N-3 and N-4 tend to belong to the field of «visibility», e.g. *visible, see (visible to the naked eye)*. This is what Sinclair calls «semantic preference».

A short list of randomly-selected concordance lines of the verb *budge* provides another example:

> given by the US for refusing to budge are falling away. Vietnam is now
> crisis. The Government is unlikely to budge on that requirement, but the
> Mrs Thatcher was not going to budge. The Italians still had a choice:
> the afternoon as the Bank refused to budge. However, despite the
> with two. The bad is that he will not budge until present or prospective
> The issue on which he will not budge - as he made clear in public in his
> Jamiat-e-Islami, which refuses to budge on the motoring laws and moreover
> chair. With the officials refusing to budge, Tarango suddenly exploded with

We can see here that the verb *budge* collocates strongly with preceding items that express the semantic idea of *unwillingness* (e.g. *refuse, will not, not going to*). The items that express this idea may be lexical or grammatical, which shows clearly that in fact *budge* is co-selected together with items with

a specific semantic feature, and its co-occurrence with specific words is a reflection of this co-selection.

See/visible to/with the naked eye is a phrase-long collocation. Phrase-long collocations with open slots have already been explored, but usually without paying attention to the fact that the slot may be filled by words sharing a semantic feature. Barkema (1996: 141-142) considers that the base form of an expression is «the simplest morpho-syntactic form that the expression can take». The base form may have one or several open function slots, e.g. *What the+* «noun» (*hell, devil, blazes*, etc.). Pawley and Syder (1983: 211) give examples of «sentence stems» with open slots such as *NP be-TENSE sorry to keep-TENSE you waiting*. This sentence stem is the result of a flow of patterns: *be sorry to, keep you waiting*. Nattinger (1980) categorises lexical phrases into six types following functional and structural criteria. One of the most important aspects of his discussion is that he includes within lexical phrases preassembled units of different lengths, up to sentence length (e.g. *not only X, but also Y; If I X, then I Y; I'll see you next week*).

Corpus researchers have emphasised that co-selection implies not only that lexical items tend to be associated with a limited range of syntactic structures but also that particular syntactic structures co-occur with particular lexical items (Francis, 1993: 143). The *Collins Cobuild Grammar* describes a structure that exemplifies clearly this point. The following prefacing structure can be used to comment on a fact, event or situation: *It+be+a+noun+that* (optional) (e.g. «It is a shame he didn't come»). The noun that functions as the complement in this preface is frequently one of the following: *disgrace, pity, wonder, marvel, shame, nuisance, surprise*. Similarly, the structure *It+ may+ sound+ adjective* can be used to reject a possible objection. The adjectives that most significantly function as complement here form a more or less restricted set. They tend to be negative evaluative adjectives, such as *ridiculous, crazy, corny, awful, silly, stupid, strange, mad, paradoxical*.

> (1) a. I know it sounds mawkish and unfeminist but I love being his girl.
> b. I know it sounds daft but I only found one Indian restaurant in Moscow and the food was awful.

An important point here is the relation between collocation and the context of situation. Patterned language is used to phrase particular meanings more economically and efficiently; thus, collocations tend to have a specific pragmatic function. Formulae are usually associated with a specific situation and have become institutionalised. For instance, the formula above (*It may+*

sound+ adjective) has a rhetorical function: it helps to counter an opposing argument, by anticipating the hearer's evaluation. As Wong Fillmore (1994: 256) claims, these phrases:

> provide ready-made, handy ways of structuring parts of the arguments that the speaker is trying to lay out. This kind of phraseology is an important formulaic resource: it provides the speaker with convenient ways of structuring the arguments into more-or-less coherent pieces of discourse when needed. Such speech can be described as practiced; what is said is novel, but the way it is put is not.

Another question is how far from each other words can occur in the text and still be considered to collocate. In some accounts of co-occurrence or collocational relationship, collocates are regarded as adjacent words, as uninterrupted sequences of words. In Sinclair's (1966) and Halliday's (1966) seminal papers co-occurrence does not imply the occurrence of two words as an adjacent pair. There may be several words between the items that collocate. As Greenbaum's (1970: 11) example shows, they may even occur in different sentences:

(2) a. They *collect stamps*.
b. They *collect* foreign *stamps* only.
c. They *collect* many things, but chiefly *stamps*.
d. They *collect* many things, though their chief interest is in collecting coins. We, however, are only interested in *stamps*.

Jones and Sinclair (1973) tried to determine the length of the syntagmatic environment or span into which the influence of a word extends and concluded that the association between the node and words that occur at distances greater than four orthographic words on either side is not strong enough to be considered as providing information about the node. The program used to find collocates at COBUILD uses a span of a specified number of words either side of the node. Thus, if we want to find the collocates of *home* all words within a span of four words to the left and the right would considered collocates. Usually the statistical significance of the association between two items diminishes as they are more far away. That is, if we have:

$$\{N4\ N3\ N2\ N1\ \textbf{NODE}\ N1\ N2\ N3\ N4\}$$

the association between the node and N1 is usually stronger than the association between the node and N4. However, in some cases a word that is far away from the node may collocate strongly with this node. This is the case of some

conjunctions. When two items that are far apart in the text are co-selected one of them may be a grammar item. Specific lexical items show a high degree of co-occurrence with specific conjunctions, which provides a clue about the meaning of these items. For instance, the t-score[5] of the association between *fail* and *but* is {8.236131} because of the frequent occurrence of *fail* in a clause beginning with *but* (e.g. «but he failed to»), which reflects that *fail* is used to express counterexpectation; the t-score of the association between *happen to* and *if* is {13.967330}, which points to the pattern: *if+ subject+ happen to*.

If we search for the pattern *HAVE+ object+ infinitive*, we will see that a frequent collocating infinitive is *believe*, usually in the collocation *have us believe*. This collocation is almost always modified by the modal *would*: *would have us believe*. This pattern is used to claim that a situation or event is untrue or doubtful, in contrast to what other people say. This explains that the pattern *would have us believe* collocates strongly with the expression *or so* and with expressions of comparison, as the following randomly-selected concordance lines show:

> *or so* Fay Weldon and co would have us believe, any kind of hero. On The Late
>   is *as* rosy as the DTI would have us believe. By PAUL WALLACE
>   ago. *Or so* they would have us believe. By week's end, President (and
>   exhibition would have us believe, they still remain unanswered.
> more common *than* mythology would have us believe. He knows, however, of a Somerset
>   Imminent, *or so* he would have us believe. LLOYD HONEY GHAN Career
>   *as* vicious rumour would have you believe, the Sheriff of Nottingham.
> *Or so* the scientists would have us believe when they say that the cranium was
> Ball. *Or so* USA Today would have us believe. Hard to imagine a red-blooded

Similarly, *fail to be* is usually preceded by the modals *can* or *could* and collocates most significantly with participial forms which express an emotional reaction, e.g. *impressed, moved, struck, touched, convinced, affected, interested, amused, alarmed, horrified, seduced, aroused, fascinated, excited, delighted*. This can be seen in the following concordance lines:

> nerve-centres. **Only** a fool could fail to be *alarmed* by this gathering of
> **No** man, whatever his creed, could fail to be *thrilled* by the sight of the
> **Only** the hardest of hearts could fail to be *touched* by the story of nine-
> horrified and shocked. **No** one could fail to be *affected* by the people walking

---

[5] T-score is a statistical measurement which indicates «the degree of confidence with which we can claim that there is some association» between two items (CLEAR, 1993: 281). Although it is a relative measurement, a t-score higher than 2 can be considered to indicate a significant collocation.

> Family in London. **No** one could fail to be *impressed* by devoted royal fan
> perspective, and **only** a dullard could fail to be *intrigued* by his genius for
> and **only** a blind man could fail to be *moved* by such a sophisticated,

These concordance lines also reveal that the pattern *can/could+ fail to be+ participle expressing an emotional reaction* collocates with negative and restrictive items *(no-one, nobody, only)*, as a result of the fact that the subject is always negative or restrictive, e.g. *nobody/no-one+ relative clause, only+ indefinite noun phrase*. The following are examples of long-distance collocations:

> (3) a. *Nobody who* loves racing cars and admires the heroic history of the grands prix *could fail to be moved* by the sight and sound of Ron Dennis of McLaren driving the W165.
> b. *Nobody who* saw the diminutive figure of infant clarinettist, Julian Bliss, being taught recently on BBC1's Newsround by the great classical player Michael Collins *could fail to be impressed.*

Phrase-long and clause-long collocations result from the fact that some items collocate with specific types of syntactic elements (e.g. specific types of subjects) or of clauses (e.g. conditional clauses). These long-distance collocations usually have a specific pragmatic function. It is this pragmatic function that determines their collocation with other structures. If we take *end up,* this is a unit with a negative prosody (i.e. an item that occurs very often with negative items). The expression *you'll end up* is used with a clear pragmatic function: to express a warning by pointing to a negative consequence. This is the reason why *you'll end up* collocates so frequently with causal and conditional conjunctions, specially *otherwise*. This co-selection is a realisation of the semantic clause relation *cause-result.*

> right message across. *As a result, you'll end up* attempting the tricky job of
> know that *if you don't* have a list *you'll end up* with six vases. When I got
> likely, *if you* dress like Agassi *you'll end up* playing like the way he looks.
> dinner, *be* imaginative, *otherwise you'll end up* eating a meal just as
> Aargh, *don't* sit next to girls *cos you'll end up* with girl-germs. That sort

Thus, collocations such as *if...you'll end up* are realisations of more abstract patterns, such as clause relations.

The clause *if you'll excuse me* is a formulaic way to realise an apology. It is what Nattinger (1980) calls a «situational utterance», a sentence that has a function in a particular social interaction. This situational utterance is used when the speaker wants to terminate an interaction, leaving the place or asking

the other(s) to leave. This formula co-occurs very frequently with the structures *I have to* and *I have+ object+ to infinitive*, as the following random concordance lines show:

> we'll know.» He rose. «If you'll excuse me, I have some work to do.»
> in Stalin's time. «Now, if you'll excuse me, I have other people to see
> here comes Mr Schrader. If you'll excuse me, I have a lot to attend to. I'm
> she answered simply. «Now, if you'll excuse me, I have work to do.» Billy
> kind of information. Now if you'll excuse me, I have other things to do.» She
> and got to her feet. «If you'll excuse me, I have to reserve a ticket for
> said Conder briskly, «if you'll excuse me, I've rather a lot of work to
> doesn't feel like it! Now if you'll excuse me, I've got work to do. I hope you
> told her firmly. «Now, if you'll excuse me, I've got customers to serve.»

Since the function of *If you'll excuse me* is to put an end to the interaction, this cannot be done politely without providing a reason or justification for it. The clause including *have to* provides a justification in the form of an external obligation. We can see, therefore, that the co-selection of linguistic forms can be accounted for by considering the function of these forms in the interaction.

Similarly, if we take the following example:

> (4) I could lecture you like a mum until I'm blue in the face, but the
> only people that can prevent this happening again are yourselves

it does not make sense to say that *blue in the face* is a formulaic expression. The pattern is in fact: *can/ could+ discourse verb+ until+ be+ blue in the face+ (but)...* and it is the whole pattern that has a meaning: there is no point in doing something. The pattern may admit variants, such as another type of verb.

> the ARL scream and sue until it is blue in the face. As it is now, Super
> argued until they were all shades of blue in the face and then capitulated.
> spans. But they can talk till they're blue in the face with the Brendas,
> can issue decrees until he is blue in the face, but they are ignored or
> So you can admit it till you're blue in the face but you get nothing back
> down he can criticise it until he is blue in the face, but it will require the
> You can shout until you are blue in the face, but psychiatric beds
> be true. I've soundproofed till I'm blue in the face, but the noise of a
> talked to Gooch about it until I was blue in the face but it was all falling

Similarly the pattern *it+ sounds+ adjective* frequently collocates with two types of hedges: *may* (i.e. «it may sound») or *I know* (i.e. «I know it sounds»), and with the conjunction *but*: *It may sound/ I know it sounds + adjective+ but*.

This can be explained if we consider the function of the phrase: to counter an opposing argument.

The co-occurrence of two formulaic clauses (e.g. «if you'll excuse me, I have some work to do»), or collocations which involve conjunctions (e.g. «if...you'll end up») can therefore be explained at a more abstract level: they are typical realisations of a clausal semantic relation or of two acts which co-occur, e.g. excuse/ justification in the case of *If you'll excuse me, I have something to do*. This is, thus, the following level of patterning.

4. TEXT LEVEL CO-SELECTION

Given that words, clauses, and their connections can only be interpreted within the higher level of discourse, it is necessary to study the structure of discourse sequences to round off our discussion of patterning.

At the level of text the existence of patterns beyond the sentence has been a topic of research from several perspectives. Some studies have revealed that both spoken and written discourse are organised and that there exist different levels of discourse structure. Other pieces of research have focused on the patterns of frequency which some linguistic features exhibit in specific text types. We are not going to present a detailed discussion of patterning at this level, since that would imply an overview of a broad area of discourse studies. To discuss all these studies in detail is unproductive, since the results are already well known. However, a brief review is necessary in order to prove that habitual co-selection takes place at all levels and that in some cases co-selection at a more concrete level results from co-selection at a more abstract or higher level.

We should consider two types of patterns: those involving the co-selection of functional units, which result in particular rhetorical structures; and, those involving the co-selection of lexico-syntactic elements at the textual level.

4.1. Patterns involving co-selection of functional units

The idea of co-selection is a basic assumption in the studies of conversational structure. Conversations are considered to consist of sequences of more basic units, such as reciprocal openings or closings, or exchange clusters (Hoey, 1993). At the simplest level, the concept of *adjacency pair* is based on the principle of co-selection, although in this case the co-selection takes place across turns (Sacks, 1967). Adjacency pairs are two *related* utterances which

are ordered in such a way that the first utterance must belong to the class of first pair parts and the second to the class of second pair parts. Thus, a Greeting predicts a Greeting, a Question is followed by an Answer, and a Complaint by an Apology or by a Justification. An interesting point is that although some first pair parts may have different seconds there are preferred and dispreferred seconds. For instance, the preferred second of Invitation is Acceptance, and the dispreferred second is Rejection.

Sinclair and Coulthard (1975) propose a structural approach to the description of classroom interaction using the idea of rank-scale with which Halliday explains how grammar is organised. The scale has the following ranks: Transaction, Exchange, Move, Act. Each rank consists of an ordered combination of units from the rank below. They propose a three-move structure for Exchange- Initiation, Response and Follow Up, which is not obligatory in all contexts. For example:

Teacher:	INITIATION	What do we do with a saw? Marvelette?
Pupil:	RESPONSE	Cut wood.
Teacher:	FOLLOW UP	We cut wood.

(Hoey, 1993)

Sinclair and Coulthard consider the Exchange as the primary unit of interaction. Other researchers have also claimed that the basic unit of analysis has three parts. The two basic units of this type that Hinds (1982) identifies are question-answer-(acknowledgement) and remark-reply-(acknowledgement). Thus, conversational exchanges follow a patterned structure. The components of these patterns are no longer lexico-grammatical or semantic elements, but the functions performed in every turn.

The fact that interaction is structured has also been revealed by the study of speech events. Some speech events can be analysed in terms of sequences of speech acts and have a fixed structure which has to be conformed to by the speakers. As McCarthy and Carter (1994: 117) put it:

> the kinds of patterns thus observed are sets of elements in sequence, the presence and ordering of which represent an idealized version of a particular culture's requirements for the realization of an activity such as inviting, apologizing, or whatever (...) The whole emphasis is on a sequence and an ordering of acts which together perform an activity recognized by members of the speech community.

An example of a patterned speech event is the compliment, which Hatch (1992) regards as having a structure with obligatory and optional elements:

(Compliment solicit) Compliment act+ Acknowledgement
(agree/ deny/ redirect focus)+ bridge.

In the following example the pattern is not conformed to and B does not produce an acknowledgement, which gives rise to some problems in the communication:

(5) A. I really like your scarf
B. Ohnn, no. It's nothing.
A. No, I really like it.
B. It is not new.
A. I still like it anyway.
B. (.4) ((Smiles))
A. Uhhh, well, are you uh going to class? (Hatch, 1992: 138).

Other studies have focused on the semantic relations between parts of the text and on the overall logical and rhetorical organisation of the text, which helps reader's comprehension. Co-selections at this level are reflected either in the organisation of the whole text or in the relation between textual segments. The basic idea is that the organisation of the text reproduces mental schemata in the user's mind. For instance, Mann and Thompson's (1987) rhetorical structure theory is concerned with the semantic relations that may hold between two portions of the text. Examples of these relations would be cause-consequence, instrument-achievement, generalisation-instance.

Several patterns of text organisation have been described by Hoey (1983), e.g. *problem-solution, general-particular* and *hypothetical-real.* The Problem-Solution pattern, for instance, consists of two basic parts narrowly related: *problem* and *solution.* They may be preceded by a *situation* part, which establishes the basis for the problem, and followed by an *evaluation* part, which provides the assessment of the solution. Hoey (1994: 28) argues that the sentences in (6) can be combined in twenty-four different sequences, but this is the preferred or unmarked one, «the only one that can be read without special intonation and make perfect sense».

(6) I was on sentry duty. I saw the enemy approaching. I opened fire. I beat off the enemy attack.

| Situation | Problem | Solution | Evaluation |

These patterns are usually associated with specific text types. Van Dijk (1977) states that the underlying structure of narratives is Setting-Complication-Resolution-Evaluation-Moral and that of scientific discourse is Introduction-

Problem-Solution-Conclusion. Adams Smith (1987) claims that the Unexplained-Explanation pattern (Situation- Unexplained- Procedural- Findings- Interpretation- (Evaluation)- Explanation) is very useful to describe the structure of a biomedical research paper. This leads to the concept of genre.

A genre is a communicative event with a conventionalised structure and features which constrain the linguistic elements that the writer can use and their function. This implies that only a restricted set of elements of the whole resources of language will be used and these elements will fulfil specific functions and have specific meanings within the genre. A concept that foregrounds the patterned nature of genre is the concept of *generic structure potential*, introduced by Hasan (1978). This is an abstract schema which specifies the total possible range of patterns which can be selected within a genre. This suggests that the structure of a genre allows for variation: a genre includes both obligatory and optional elements. The sequence of obligatory elements defines the limits of the genre.

In genre analysis the overall text organisation is described as being composed of moves, functional units based on purpose, and all the syntactic and lexical choices are explained by relating them to this higher level organisation. This rhetorical organisation is illustrated by Bhatia's (1993) analysis of a sales promotion letter:

1. Establishing credentials.
2. Introducing the offer: offering the product or service, essential detailing of the offer, indicating value of the offer.
3. Offering incentives.
4. Including documents.
5. Soliciting response.
6. Using pressure tactics.
7. Ending politely.

4.2. Patterns involving the co-selection of lexico-syntactic elements at the textual level

In the previous section we have seen that there are different rhetorical patterns to organise discourse, i.e. textual patterns (e.g. Problem-Solution), generic patterns. One of the aspects studied in relation to rhetorical patterns is the signalling of these patterns or the association of some linguistic items with the pattern. For instance, Winter (1977) and Hoey (1983) have studied the lexis associated with the Problem-Solution pattern and with the Hypothetical-Real

pattern. The following fragment illustrates how the pattern Problem-Solution is lexically signalled:

Problem	Testing may be for both teachers and students one of the most *unpleasant* aspects involved in the teaching learning process. It is *difficult* to choose the right way to do it. Teachers often find that after working long hours to prepare the text they *fail* to get the best of their students, who, at the same time, feel that the exam was *unfair* in some way.
Solution	How could I change this? This is one of the questions I put to myself, and I *found the answer* in a training course for teachers (...). We were encouraged to put the students in the role of protagonists in the whole training process, which, of course, included testing (...)
Evaluation	Finding themselves in the position of teacher made them aware of the difficulties of producing a test, and they came to a *more thorough understanding* of my role as a teacher. (...) By the end of the experiment we were *all pleased* with the results: the students because they had done something *new, creative and meaningful*; and for me, it was very *rewarding* in terms of motivation.

In a text organised following this pattern we are likely to find three types of items: items expressing need or problem (e.g. *unpleasant, difficult, fail, unfair*), items expressing solution (e.g. *solution, answer, way*), positive evaluative items (e.g. *pleased, new, creative, meaningful, rewarding*). Thus, there is a co-selection of these types of words, motivated by the pattern that structures the text.

The fact that the concept of genre implies the use of conventionalised linguistic elements which are most appropriate to achieve a specific purpose also explains that there are sets of linguistic features that tend to co-occur in a genre. This has been extensively investigated by Biber (1988), who uses the term «dimension» to refer to these sets of features. For instance the co-occurring features associated with dimension 2 (*Narrative vs. Non-Narrative*) are the following: past tense verbs, third person pronouns, perfect aspect verbs, public verbs, synthetic negation, present participial clauses.

The lexico-grammatical realisations of the different moves of a genre also form patterns of co-occurrence. For instance, in the discussion of a research paper which includes moves such as *Location of results, Observation, Statement of results, Reference to previous research,* or *Claims*, there is a co-occurrence of semi-patterned units which signal or realise these moves: *Table 1 shows/lists/illustrates that; We found that; These findings/results show/indicate that; Our results confirm/are in agreement with; This study has demonstrated/ We have shown/ This study suggests that.* Similarly, in a formal letter *Dear Sir* is likely to co-occur with *Yours faithfully*: they are formulaic expressions associated with a specific genre and therefore likely to co-occur.

Textual and generic patterns are patterns at the metadiscursive level or at the level of superstructure. There are also patterns at the conceptual level, the level of macrostructure. These patterns can be related to the concept of register, defined by Halliday as «the *semantic patterning* that is characteristically associated with the 'context of situation' of a text» (Halliday, 1978: 14). Paying attention to the «aboutness» of texts, Phillips (1989) discovered the existence of meaningful syntagmatic lexical sets in texts (i.e. sets of lexical items which could be grouped in terms of their collocational patterning, that is, in terms of their frequency of syntagmatic association). They are meaningful because there is a clear relation between «the syntagmatic organisation of words into lexical sets and the conceptual concerns of the text» (Phillips, 1989: 53). Some examples of sets are the following: (1) *charge, density, symmetry, uniform, distribution, total*; (2) *angular, precess, constant, direction, swing, given, clearly*. The networks created by these sets recur over different parts of the text and in this way establish links between these parts, revealing the macrostructure of the text. This lexical patterning is dependent on register, which supports Firth's (1957: 14) claim that the key words of any restricted language are likely to exhibit characteristic collocations, which «will help justify the restriction of the field». We can see, therefore, that patterning at the lexico-syntactic level is determined by higher level constraints.

5. CONCLUSION

The present paper provides an overview of the concept of language patterning and shows that patterning is inherent to language use. Patterning is related to typicality of use (i.e. what is typically said, rather than what can be said), but does not impose any limits on creativity: variation is always possible.

We have seen that patterns at all levels are meaningful and that co-selection has a cognitive basis: as research on language production suggests, the language stored in the mind does not only consist of individual morphemes and words, but also of longer chunks, which are retrieved from memory as such. These chunks are associated with specific meanings. The co-selection of linguistic elements should be explained within a functional framework: the context of situation and the meaning that the user tries to convey determines the selection and co-selection of items. Meaning is not only conveyed by words but also by patterns.

This paper has highlighted some implications for language teaching. The teaching of language should pay attention to the syntagmatic aspect, to the habitual co-selections of native speakers, which will provide naturalness to language use. Conventionalised syntagms or patterns at all levels are very useful in

the production of language by non-native speakers because they contribute to economy and effectiveness: they are an effective way to construct and interpret meaning. Students do not need to generate language from scratch each time they want to say something, but can use pre-assembled blocks or follow pre-established patterns that are kept in their mind as schemata.

6. REFERENCES

ADAMS SMITH, DIANA E. (1987): «Variation in field-related genre», *ELR Journal,* 1:10-32.

ATKINSON, DWIGHT (1990): «Discourse analysis and written discourse conventions», *Annual Review of Applied Linguistics,* 11: 57-76.

BAKER, MONA, FRANCIS, GILL and TOGNINI-BONELLI, ELENA (eds.) (1993): *Text and Technology,* Amsterdam, Benjamins.

BARKEMA, HENK (1996): «Idiomaticity and terminology: a multi-dimensional descriptive model», *Studia Linguistica,* 50/2: 125-160.

BARNBROOK, GEOFF (1996): *Language and Computers. A Practical Introduction to the Computer Analysis of Language,* Edinburgh University Press.

BAZELL, C. E., CATFORD, J. C., HALLIDAY, M. A. K. and ROBINS R. H. (eds.) (1966): *In Memory of J. R. Firth.* London: Longman.

BENSON, MORTON (1990): «Collocations and general-purpose dictionaries», *International Journal of Lexicography,* 3/1: 23-35.

BHATIA, VIJAY K. (1993): *Analysing Genre: Language Use in Professional Settings,* London, Longman.

BIBER, DOUGLAS (1988): *Variation Across Speech and Writing,* Cambridge, Cambridge University Press.

BIBER, DOUGLAS (1996): «Investigating language use through corpus-based analysis of association patterns», *International Journal of Corpus Linguistics,* 1/2: 171-197.

BOLINGER, DWIGHT (1976): «Meaning and Memory», *Forum Linguisticum,.* 1/1: 1-14.

CARRELL, PATRICIA (1983): «Some issues in studying the roles of schemata, or background knowledge, in second language comprehension», *Reading in a Foreign Language,* 1: 81-92.

CHOMSKY, NOAM (1965): *Aspects of the Theory of Syntax,* Cambridge, MA,: MIT Press.

CLEAR, JEREMY (1993): «From Firth Principles. Computational Tools for the Study of Collocation» in Baker, M., Francis, G. & Tognini-Bonelli, E. (eds.), pp. 271-292.

COULMAS, FLORIAN (1979): «On the sociolinguistic relevance of routine formulae», *Journal of Pragmatics,* 3: 239-66.

COULMAS, FLORIAN (1981): «Introduction: conversational routine» in Coulmas, F. (ed.): *Conversational Routines: Explorations in Standardised Communication Situation and Prepatterned Speech.* The Hague, Mouton, pp. 1-18.

COULTHARD, MALCOLM (ed.) (1994): *Advances in Written Text Analysis,* London and New York, Routledge.

DIJK, TEUN, VAN (1977): *Text and Context. Explorations in the Semantics and Pragmatics of Discourse,* London, Longman.

DUDLEY-EVANS, TONY (1994): «Genre analysis: an approach to text analysis for ESP» in Coulthard, M. (ed.), pp. 219-228.

FILLMORE, CHARLES; KAY, PAUL and O' CONNOR, MARY (1988): «Regularity and idiomaticity in grammatical constructions: the case of *let alone» Language,* 64/3: 501-538.

FIRTH, J.R. (1957): *Papers in Linguistics,* London, O.U.P.

FIRTH, J.R. (1968): *Selected Papers of J.R. Firth. 1952-1959.* Ed. by F.R. Palmer. London, Longman.

FRANCIS, GILL (1993): «A corpus-driven approach to grammar- principles, methods and examples» in Baker, M., Francis, G. & Tognini-Bonelli, E. (eds.), pp. 137-156.

GREENBAUM, SYDNEY (1970): *Verb-Intensifier Collocations in English,* The Hague, Mouton.

HALLIDAY, M. A. K. (L966): «Lexis as a linguistic level» in Bazell, C.E., Catford, J.C., Halliday, M.A.K., and Robins, R.H. (eds.), pp. 148-162.

HALLIDAY, M. A. K. (1978): *Language as Social Semiotics: the Social Interpretation of Language and Meaning.* London, Edward Arnold.

HANKS, PATRICK (1987): «Definitions and Explanations» in Sinclair, J.M. (ed.): *Looking up,* London and Glasgow, Collins ELT, pp. 116-136.

HASAN, RUQUIYA (1978): «Text in the systemic-functional model» in Dressler, W.U. (ed.): *Current Trends in Text Linguistics.* Berlin, de Gruyter, pp. 228-246.

HATCH, EVELYN (1992): *Discourse and Language Education,* Cambridge, C.U.P.

HINDS, JOHN (1982): «Japanese conversational structure», *Lingua,* 57: 301-326.

HOEY, MICHAEL (1994): «Signalling in discourse: a functional analysis of a common discourse pattern in written and spoken English» in Coulthard, M. (ed.), pp. 26-46.

HOEY, MICHAEL (1993): «The Case for the Exchange Complex» in Hoey, M. (ed.), *Data, Description and Discourse,* London, Harper Collins Publishing, pp. 115-138.

HOEY, MICHAEL (1983): *On the Surface of Discourse,* London, Allen und Unwin.

HOPPER, PAUL (1988): «Emergent grammar and the *a priori* grammar postulate» in Tannen, D. (ed.): *Linguistics in Context: Connecting Observation and Understanding.* Norwood, NJ, Ablex, pp. 214-234.

JONES, SUSAN and SINCLAIR, JOHN (1973): «English Lexical collocations», *Cahiers de Lexicology,* 23/2.

KRISHNAMURTHY, RAMESH (1987): «The process of compilation» in Sinclair, J.M. (ed.): *Looking up,* .London and Glasgow, Collins ELT, pp. 62-85.

MANN, WILLIAM and THOMPSON, SANDRA (1986): «Related processes in discourse», *Discourse Processes,* 9/11: 57-98.

McCARTHY, MICHAEL and CARTER, RONALD (1994): *Language as Discourse: Pers-pectives for Language Teaching*, London and New York, Longman.

McENERY, TONY and WILSON, ANDREW (1996): *Corpus Linguistics*, Edinburgh Text-books in Empirical Linguistics.

McINTOSH, ANGUS (1961): «Patterns and ranges», *Language,* 37: 325-37.

MEYER, BONNIE and RICE, E.G. (1982): «The interaction of reader strategies and the organization of text», *TEXT,* 2/1-3: 155-192.

NATTINGER, JAMES (1980): «A lexical phrase-grammar for ESL», *TESOL Quarterly,* 14,3: 337-44.

NATTINGER, JAMES and DECARRICO, JEANETTE (1992): *Lexical Phrases and Langua-ge Teaching*, Oxford University Press.

PAWLEY, ANDREW and SYDER, FRANCES (1983): «Two puzzles for linguistic theory: native-like selections and native-like fluency» in Richards, J. and Schmidt, R. (eds.): *Language and Communication*, London, Longman.

PETERS, ANN (1983): *The Units of Language Acquisition*, Cambridge, C.U.P.

PHILLIPS, M. (1989): *Lexical Structure of Text,* Discourse Analysis Monographs, n. 12.

SACKS, H. (1967): Unpublished lectures notes.

SINCLAIR, JOHN (1987): «Collocation: a progress report», offprint from Ross Steele and T. Treadgold (ed.) *Language Topics. Essays in Honour of M. Halliday.* Amsterdam, John Benjamins.

SINCLAIR, JOHN (1966): «Beginning the study of lexis» in Bazell, C.E., Catford, J.C., Halliday, M.A.K., and Robins, R.H. (eds.) 1966: 148-162.

SINCLAIR, JOHN and COULTHARD, MALCOLM (1975): *Towards an Analysis of Dis-course. The English used by Teachers and Pupils*, Oxford, O.U.P.

SINCLAIR, JOHN (1991): *Corpus, Concordance, Collocation*, Oxford, O. U. P.

SINCLAIR, JOHN (1996): «The search for units of meaning», *TEXTUS IX.*

STUBBS, MICHAEL (1996): *Text and Corpus Analysis*, Oxford, Blackwell Publishers.

SWALES, JOHN (1990): *Genre Analysis: English in Academic and Research Settings,* Cambridge, England, C.U.P.

TANNEN, DEBORAH (1989): *Talking Voices*, Cambridge, New York, C.U.P.

TOGNINI BONELLI, ELENA (1996): *The Role of Corpus Evidence in Linguistic Theory and Description.* Unpublished PhD.Thesis. University of Birmingham.

WINTER, EUGENE (L977): «A clause relational approach to English texts: a study of some predictive items in written discourse», *Instructional Science,* 6/1: 1-92.

WONG FILLMORE, LILY (1994): «The role and function of formulaic speech in Con-versation» in Grimshaw, A.D. (ed.): *What's going on here? Complementary Studies of professional talk.* Vol. XLII. Ablex Publishing Corporation. Norwo-od, New Jersey.

YORIO, CARLOS (1980): «Conventionalised forms and the development of communi-cative competence», *TESOL Quarterly,* XIV/ 4: 433-42.