Title:
An ontological approach to the representational lexicon in Functional Discourse Grammar

Author:
Christopher S. Butler
Honorary Professor, College of Arts and Humanities, Swansea University
Visiting Professor, School of Music, Humanities and Media, University of Huddersfield
Visiting Fellow, Centre for Translation Studies, University of Leeds

Abstract:
The aim of the article is to propose an outline of a model of the representational lexicon in Functional Discourse Grammar. Some problems with the mainstream account of the lexicon are discussed in the context of prior work by García Velasco (2007). It is proposed that structure building, including the choice of an appropriate predication frame, is mediated through the selection of lexical items to realise the concepts the language user wishes to express, and that although decompositional accounts of the lexicon need enrichment to account for lexical competence, they are still required, though at the conceptual level. It is then proposed that the link between lexical items and the Conceptual Component of the model should take the form of pointers from lexical items to a conceptual ontology, and that this ontology should be language-neutral across a range of culturally-related languages. It is shown that the FunGramKB system of ontology and lexicons, developed within the framework of the Lexical Constructional Model, can be adapted for use in Functional Discourse Grammar. The article then discusses the handling of alternations, collocations and lexical variation with situational content. Suggestions are also made for the treatment of prefabricated expressions which combine grammatical and lexical characteristics.

Keywords:
Functional Discourse Grammar
lexicon
concept
ontology

Highlights:

- The work of García Velasco (2007) is discussed and extended.
- The construction of structure, including that of the predication frame, is mediated through the selection of lexical material.
- Decomposition is still needed, but at the conceptual level.
- The FunGramKB ontological system can be adapted for use in FDG.
- Alternations, collocations, situational variation, prefabricated expressions can be handled.

1. Introduction

Without items carrying lexical meanings, there would be no language as we know it. Meanings are also conveyed by morphosyntax and prosodic phonology, but without lexical items or something equivalent to them linguistic communication would not exist. Functional linguistic theories are predicated on the claim that language is first and foremost a means for communication between human beings, and that this fact has a deep and all-pervasive influence on the forms that languages take. Logically, then, one would expect that a functional theory would have a great deal to say about lexical matters. Nevertheless, there are theories which are labelled as functional by their proponents but which as yet have paid scant attention to this area of linguistics. Functional Discourse Grammar (henceforth FDG, Hengeveld and Mackenzie, 2008) is one of these, as has been documented in Butler (2009a), which compares the treatment of lexical phenomena in FDG and in Systemic Functional Linguistics, which has a rich account of lexis. In particular, that article discusses the relationship between lexicon and syntax, paradigmatic lexical relations, syntagmatic lexical relations (collocation), lexis and sociocultural context, and the acquisition of lexis.

There are some important exceptions to the general neglect of lexical matters in FDG. García Velasco (2007), which acts as the point of departure for the present paper, argues for the replacement of the notion of word meaning by that of lexical competence (Marconi, 1997), and discusses the implications for FDG. Keizer (2007) argues that the distinction between lexical and grammatical phenomena, though still useful, is a non-discrete one. She shows that so-called 'grammatical items' can be classified, by means of particular tests, as either truly grammatical (primary grammatical words or grammatical operators/functions) or as lexical (secondary grammatical words, relabelled as lexical operators/functions). Keizer (2009) examines a range of verb-preposition combinations in English and proposes that their semantic and syntactic properties can be accounted for within FDG if lexemes are separated from predication frames, and if the latter include the 'composite predicate frame', the 'resultative frame' and the 'verb + prepositional complement frame'. Honselaar and Keizer (2009) present a detailed account of three Dutch lexemes, *bekend zijn* 'to be familiar, well known', *behandelen* 'to treat' and *trouwen* 'to marry', and discuss the relationship between lexemes and predication frames in the light of their findings.

The aim of the present paper is to propose some changes to García Velasco's account and to develop it into a more detailed model of the representational lexicon. The approach will borrow heavily from the FunGramKB model developed within the context of the Lexical Constructional Model (for detail and references see §4.4). I shall have nothing to say here about other aspects of the lexicon such as lexical items with interpersonal function.[1]

2. The need for a more highly developed lexicon in FDG

García Velasco points out that the lexicon in many contemporary grammatical theories has been viewed from a grammar-designing perspective, where one of the main aims is "to establish systematic links between the lexicon and syntax" (García Velasco, 2007, pp. 165-166). This, he states, has also been the case in Functional Grammar and its successor, FDG. In such an approach, substantive lexical differences, such as that between DOG and CAT, are irrelevant. García Velasco also observes that although such a view may be adequate for formal linguistics, in which language is seen as self-contained and autonomous, it is less so

[1] See O'Neill (2011) for a more wide-ranging account of the contents and structure of the lexicon which does include interpersonal lexemes, but stays more firmly within the boundaries of current FDG than the proposals made here.

from a functionalist perspective. García Velasco's aim is "to lay out the aspects of the lexicon component which are necessary in a functional characterization of communicative competence and to examine the implications for FDG" (p.166). In particular, he asserts that if we want to explain our ability to produce expressions complying with rules governing verbal interaction (Dik's pragmatic adequacy) we need to go beyond those properties of lexemes which are relevant to syntax, to study aspects of lexical competence as such. This viewpoint is very much in line with that expressed in Butler (2008a, 2009a, 2009b), where it is argued that a truly functional approach needs to go beyond the grammar itself, to include not only a rich account of lexical phenomena, but also accounts of conceptualisation and of the relationship between language and sociocultural context.

3. García Velasco's account of lexical competence and its relationship with FDG

In line with Dik's statement that communicative competence is the psychological correlate of a natural language, García Velasco defines lexical competence as "the ability to use words in appropriate and effective ways in verbal interaction", as part of communicative competence (p.166). His account is based on the work of Marconi (1997), who proposes that lexical competence consists of a referential component (the ability to map lexical items on to the world) and an inferential component (access to a network of connections between a word and other words and expressions). Such competence is flexible and open-ended, and so differs from one individual to another. In García Velasco's model meanings are sets of beliefs conventionally attached to lexical items by members of a linguistic community. He points out that a speaker does not need to have a perfect definition of a word in order to communicate efficiently, though there must be what he sees as a dynamic, online process of convergence between the beliefs of the speaker and hearer in relation to the lexical items used. Since the features associated with lexical concepts are open-ended, different aspects of meaning may be foregrounded in particular communicative contexts.

There is clearly a source of tension in the model of lexical competence advocated by García Velasco: lexical items are flexible and their specifications open-ended, but speakers and hearers must be able to converge enough for effective communication to take place. It does not seem reasonable to suggest that negotiation of meaning between interlocutors occurs totally afresh in each act of communication. Rather, as García Velasco observes, being conventional, or constitutive of 'normal' competence, is a crucial criterion for a feature to be regarded as central, as is agreed by Marconi and also within the Cognitive Grammar approach of Langacker, in which it is proposed that that part of the meaning of a lexical item which can be considered conventional is contextual knowledge which has been established through repeated occurrence (Langacker, 1987, p. 158). Thus, although "it is not possible to isolate the specific set of beliefs which are shared by all speakers for a given item", there is conventional knowledge which is shared by interlocutors, and "the semantic side of a lexical entry should thus be seen as an idealized representation of that partially common knowledge" (García Velasco, 2007, p. 179). This should not, however, be formulated in terms of the classical view of word meaning, based on necessary and sufficient conditions, which has been shown to be highly problematic, but rather in terms of sets of "pieces of information which could be assumed to be part of normal competence" (p. 179).

The espousal of lexical competence as the basis for lexical entries thus leads García Velasco to modify or replace the decompositional structures for lexical entries proposed in García Velasco and Hengeveld (2002) and subsequently adopted into mainstream FDG. He compares the two approaches using the example of transitive OPEN. The definition given by García Velasco and Hengeveld (2002, p. 114) is shown in 1:

1. *open* [V]
   [f_1: [CAUSE (x_1) [BECOME **open´** (x_2)]]]

This structure guides the selection of the predication frame in 2, which collapses information from the interpersonal and representational levels, and where T represents an ascriptive subact at the interpersonal level, R a referential subact and ♦ the position of the lexeme[2]:

2. (T_1: (f_1: ♦ (f_1)) (T_1)) ((R_1: (τ_1)$_{Ag}$ (R_1)) (R_2: (τ_2)$_{Pat}$ (R_2))

Shorn of its interpersonal level trappings, the matching of lexeme definition and predication frame leads to the simplified frame in 3 for transitive OPEN, given in García Velasco (p. 178), where 'π' represents any operators and 'e' a predication:

3. (π e_1: [(f_1: *open* (f_1)) (x_1) $_{Ag}$ (x_2) $_{Pat}$ (e_1))

García Velasco's new scheme maintains the link between lexical entry and predication frame, but substitutes the decompositional definition in 1 by the descriptive entry[3] in 4:

4. *Open*
   (a) Opening is an event.
   (b) By opening somebody allows entrance of something.
   (c) Thus, doors, etc. can be opened.
   (d) People often open doors to enter buildings.
   etc.

He assumes that the minimum required for competence in the use of this lexeme is (i) to specify it in the lexicon as event-denoting (rather than denoting a thing, property, etc.) and (ii) to specify the number of participants typically involved. In other words, parts (a) and (b) of the above entry are enough for the lexeme to be able to select an appropriate predication frame. Other specifications, regarded as pragmatic rather than semantic, are also needed to model full lexical competence for a lexeme, and may vary from one individual to another. García Velasco does not assume that the format of the entry should necessarily be propositional, as in 4(b), but allows the possibility of referential formats such as images or 3-dimensional representations (see Jackendoff, 1990, p. 34).

4. **Problems, revisions and expansions**

4.1 The order of selection of predication frames and lexemes

There is some confusion in the literature about the order of selection of predication frames and lexemes. García Velasco and Hengeveld (2002, pp. 113-114) propose that the meaning definitions of lexemes guide the selection of an appropriate predication frame, implying that the lexemes are chosen first, followed by the frames. This, however, is contradicted (without

[2] Note that the term 'lexeme' is being used here in a sense which is synonymous with 'lexical item'. However, see fn. 7 for a rather different usage.

[3] It is interesting to note that Honselaar and Keizer (2009, p. 1233), in their discussion of the meaning definitions necessary to account for the lexemes they analyse, say the following: "Even though this meaning definition will include only information that is linguistically relevant, the richness of the information needed to ensure that only acceptable combinations of lexemes and frames take place has led us to assume that this information will be highly descriptive in nature".

comment) in the summary of García Velasco and Hengeveld's paper in the editorial introduction to the book in which this article appeared, where it is stated that "a speaker selects a particular predication frame and then he/she chooses an appropriate lexeme, a process which is guided by the conceptual information in lexemes" (Mairal Usón and Pérez Quintero, 2002, p. ix). It is this latter interpretation of the facts which appears in Hengeveld (2005, p. 67), where it is stated that frames, at both the interpersonal and the representational levels, are selected first, and only then are lexemes inserted. This ordering is proposed in order to account for the choices speakers have in selecting lexemes with, for example, different connotations. The same ordering is proposed in Hengeveld and Mackenzie (2008, p. 19): "In the implementation of the grammar the frames are selected first, and only after that are lexemes inserted".

However, this proposal is unattractive from the viewpoint of cognitive adequacy, since it is surely more plausible cognitively to assume that the selection of content precedes, or at least is simultaneous with, the choice of an abstract meaning frame into which that content will be inserted (Butler 2009c)[4]. Indeed, Levelt (1989: 181), putting forward his influential psycholinguistically supported model of language production, says the following:

> A main thesis of this and the following chapters will be that formulation processes are lexically driven. This means that grammatical and phonological encoding are mediated by lexical entries.

In other words, the first step in formulation is to select at least one specific lexical item[5], and the lexical items then bring with them their grammatical characteristics. Similarly, Schönefeld (2001) concludes, if somewhat cautiously, that both production and comprehension processes are lexically driven.

García Velasco clearly implies that the selection of a lexeme is logically prior:

> Speakers will select a relevant frame on the basis of the specifications of beliefs associated to a lexical item, but the syntactically relevant information will only be present in the frame chosen. (García Velasco, 2007, p. 179)

It will also be recalled that García Velasco states that items 6a and 6b in the meaning definition given earlier are sufficient for the selection of the appropriate predication frame. The arguments presented in Honselaar and Keizer (2009) also point towards a lexically-driven process, though they remain open to the possibility that selection of lexeme and frame might be simultaneous[6].

---

[4] Proponents of FDG might wish to object at this point that the theory does not attempt to model language production by the speaker (see Hengeveld, 2004, pp. 366-367, Hengeveld and Mackenzie, 2008, p. 2). However, it does build in influences from certain aspects of the production model of Levelt (1989, 1999, notably its top-down organisation, the division into formulation and encoding and the acceptance of incrementality, in an explicit attempt to improve cognitive/psychological adequacy. Any attempt to make a linguistic model responsive to models of speech production is clearly a step in the direction of a speaker-based model, despite claims to the contrary.

[5] I say 'at least one' because it is generally agreed that production is incremental, so that once a lexical item has been selected, its associated structure can immediately start to be built.

[6] See, for example, Fernández and Cairns (2011, p. 141), who say that "the words and the structures are so closely related that the two processes take place practically simultaneously", though they also say that "it is most certainly not the case that the structure is constructed before the words are selected, nor are all the words selected before the structure is constructed", the last part being a reference to the incremental nature of processing.

Hengeveld and Mackenzie (2008, p. 19) give two reasons for the priority of frame selection over lexeme selection:

> This reflects the choice the Speaker often has in describing one and the same entity through a variety of lexemes with different connotations and/or denotation. It also provides a natural framework for understanding the phenomenon of coercion, through which lexemes that are strongly associated with a particular frame can be forced for expressive purposes into a frame that is usual coupled with lexemes of another meaning class.

However, neither of these arguments is persuasive. If by lexemes with different denotation Hengeveld and Mackenzie are referring to situations such as describing the same man as *my father* or as *my brother*, depending on who is speaking, this is surely a matter of the distinction between sense and reference, and the different lexemes refer to different concepts. In the case of the description of an entity by means of lexemes with different connotations (e.g. deciding (usually subconsciously) between *bustling* and *congested* to describe a town) it seems implausible to suggest that the speaker decides to say something about the town, so choosing an abstract modifying frame, and only then decides on the nature of the content to be conveyed ('crowded') and the connotation to be expressed (positive or negative). Rather, it seems sensible to postulate selection of the content and then of the appropriate frame, which in this particular case will be the same for both selections. As far as coercion is concerned, we shall see later that this can readily be accommodated within a model in which lexemes are chosen first.

Before we leave this area, one potential source of confusion needs to be cleared up. All that is being proposed here is that the development of structure occurs only as a consequence of lexical selection. The order in which different types of information about lexical items become available once these items have been selected is a quite separate matter. In Levelt's psycholinguistically-supported model activation is a two-stage process, structural information (which Levelt calls 'lemma information') being built before the phonological content of the lexical item is filled in (see Levelt 1999: 95ff). However, the crucial point is that the phonological content is that which corresponds to the already selected lexical material. It is not the case, as in the FDG proposals, that structure is constructed first, and lexical content decided only at the stage of inserting the phonological information[7]. This clearly has important implications for the design of a psychologically/cognitively adequate model.

4.2 We still need decomposition

García Velasco argues that decompositional meaning definitions are not rich enough to capture all the information relevant to the meaning of a lexeme. While agreeing that the decompositional definitions often proposed do need to be considerably enriched, I would argue that decomposition is still needed if we are to account for what Marconi calls the inferential competence of speakers and hearers. However, we shall see later that this decomposition is sited within the conceptual system to which lexical entries relate.

---

[7] Here, an unfortunate terminological confusion needs to be addressed. In some accounts, the level at which phonological information for a lexical item is specified is labelled the 'lexeme' level. This is clearly a more specific usage than the one in which 'lexeme' is used as an equivalent of 'lexical item'. It should also be noted that some researchers have doubted the need for a level of syntactically-specified lemmas: see the discussion in Harley (2008: 419-420).

Firstly, the ability to define words in terms of superordinates is still a part of the lexical competence of speakers. Any native English speaker without any language deficit could explain to a foreigner that 'a sparrow is a kind of bird'. The following corpus examples further illustrate the point:

5. "A jet is a kind of aircraft", said Angalo, the transport expert. (BNC HTH 110)

6. A prospectus is a sort of handbook which is supposed to tell you and your parents everything that you need to know about surviving at your school. (BNC C8N 847)

7. A nymph is a kind of fairy, I think. (BNC CK9 1636)

Note that the last example shows, through the modal expression *I think*, the contingent nature of lexical knowledge.

Secondly, speakers and writers can use superordinates to refer again to something which has been referred to more specifically:

8. The people who care for these snakes are very clever at milking them and getting the venom from the animal safely ... (BNC KRF 680)

9. A puppy which may have rabies is being held at a quarantine kennels in the region. The animal was seized after being brought ashore illegally from a foreign ship. (BNC K1J 2915-2916)

As we move towards superordinates which are further up the lexical hierarchy, this kind of reference gets less common, but is still possible:

10. Volunteers who sang to a sick dolphin for 36 hours to calm it returned the mammal to the sea at Gwithian, Cornwall, yesterday. (BNC CH2 815)

Proponents of FDG might, of course, wish to claim that the phenomena discussed in this section are matters of encyclopaedic rather than linguistic knowledge. However, as García Velasco (2007, p. 173) points out, citing Langacker (1987), the division between the two putative types of knowledge is problematic, and "lexical items, or linguistic expressions for that matter, are points of access to different bodies of knowledge against which we can make sense of them". The proposals in the present paper, in common with those of cognitive approaches such as Cognitive Grammar and Word Grammar, therefore take the view that a strict division between linguistic and encyclopaedic knowledge is neither possible nor desirable.

4.3 The link between the lexicon and the conceptual component

García Velasco describes the link between the lexicon and the Conceptual Component of FDG as follows:

> Basic specification, such as the fact that a given lexical item denotes an event or a thing should obviously be considered linguistic, but **the item should also be directly connected to the Conceptual Component, which provides specifications subject to revisions or validations**. To the extent that these specifications become

> conventional (in the sense of constitutive of normal competence) they will become part of the linguistic system. Thus, there are no lexical entries in the traditional sense, but rather **correspondences between conceptual information, i.e. specifications or properties we conventionally apply to concepts, and linguistic primitives used in the construction of linguistic expressions**. (García Velasco, 2007, p. 183, emphasis added)

The question which now arises is what exactly is in the Conceptual Component, and how it is linked to lexical items.

For Hengeveld and Mackenzie, the Conceptual Component, although "the driving force behind the Grammatical Component as a whole" (2008, p. 7), is restricted, including only those aspects of cognition "that affect the immediate communicative intention". As an example, they cite Harder's example of what underlies the Spanish *Me temo que Juan está enfermo* ('I am afraid that John is ill'): the Speaker does his/her duty towards the Addressee by conveying relevant bad news ('John is ill'), mitigated by showing sympathy. However, this approach says nothing about the actual concepts involved. In order to implement García Velasco's ideas, we need the Conceptual Component to contain a set of concepts to which lexical items can be linked. We have seen that decomposition is still necessary, so we need a **hierarchically ordered ontology** of concepts.

In FDG as currently conceived, the degree of specificity of the Conceptual Component is somewhat unclear. In Hengeveld and Mackenzie (2008, p. 12), the conceptual component is described as prelinguistic, and it is claimed that "similar conceptual representations may receive different pragmatic and semantic representations in different languages", which strongly suggests that conceptual representations are not language-specific. Furthermore, it is explicitly stated (pp. 7-8) that what Slobin (1996) calls 'thinking for speaking', in other words the accommodation of conceptual information to the dictates of particular languages, would be part of the formulation process in FDG rather than of the Conceptual Component itself, and this again suggests that what is in the conceptual component lies at a deeper level than the modifications of conceptual structuring that Slobin says are necessary in order to account for how concepts get expressed through particular languages. However, Mackenzie (pers. comm.) tentatively suggests that the Conceptual Component of FDG may act as "an interface between the language-neutral, 'deeper' level and the specifics of the language chosen for use". Although it is clear that any attempt to formulate a truly universal ontology is doomed to failure, since very distant cultures may conceptualise the world in very different ways, the working assumption made here is that it may be worth trying to formulate general ontologies, at an appropriate level of granularity, for groups of languages which are closely related culturally.

In what follows, we shall be concerned only with what we might call the representational lexicon, referring to events, entities and representationally-based (rather than interpersonally-based) qualities. I have argued elsewhere that the Conceptual Component should be replaced by a **content component**, consisting of two sub-components, the conceptual content proper and the affective/interactional content, which would house not only the attitudes and emotions we express through language and paralinguistic codes, but also speech act forces and modal concepts (see Butler, 2008b, pp. 240-243, where possible neurological correlates of such a division are also discussed). Here, then, we are dealing with the conceptual content as such.

4.4 A possible model for adaptation in FDG: the FunGramKB ontology and lexicons

As a possible source of inspiration for how to model lexicons and the Conceptual Component in FDG, we shall examine FunGramKB, a computationally implemented system containing a language-independent ontology and language-dependent lexicons, designed for use with the Lexical Constructional Model. For overviews of the Lexical Constructional Model, see Ruiz de Mendoza and Mairal Usón (2008, 2011), Mairal Usón and Ruiz de Mendoza (2009), Butler (2009d), Ruiz de Mendoza and Gonzálvez-García (2011a, 2011b); for information on FunGramKB see Mairal Usón and Periñán-Pascual (2009), also the web site at www.fungramkb.com; and for a detailed account of the ontology and how it is constructed Periñán Pascual and Arcas Túnez (2010), Jiménez-Briones and Luzondo Oyón (2011).

The ontology component of FunGramKB is designed for a group of European languages which will eventually include English, Spanish, German, French, Italian, Bulgarian and Catalan, though most work has so far been done on the first two of these languages. So, for instance, the verbs OPEN (English), ABRIR (Spanish), OUVRIR (French), APRIRE (Italian), ÖFFNEN (German), etc., in sentences such as *John opened the door*, *Juan abrió la puerta*, *Jean a ouvert la porte*, would all be linked to a common concept in the ontology[8]. The ontology contains concepts which refer to entities and qualities as well as to events, and is linked to a separate lexicon for each language. Lexical entries are written in a COnceptual REpresentation Language (COREL), in which the elements make reference to the general ontology (see Periñán-Pascual and Mairal Usón 2010).

The ontology recognises three conceptual levels:

- **Metaconcepts**, marked with #: very general categories arrived at through comparison of standard ontologies, e.g. #ABSTRACT, #COLLECTION, #PSYCHOLOGICAL, #MOTION, #POSSESSION, etc.
- **Basic concepts**, marked with +: +HUMAN_00, +ON_00, +BE_00, +MACHINE_00, etc[9].
- **Terminal concepts**, marked with $: $GRASP, $BOW_00, $SUBTITLE_00, etc.

Basic and terminal concepts are stored with their ontological properties in the form of thematic frames and what are referred to as meaning postulates, though they are conceptual structures rather than implications as in the standard meaning postulates of logic.

For purposes of illustration, we shall continue with the concept of opening. For the event sense, the hierarchy leading to the relevant basic concept +OPEN_01 is shown in 11, where >> represents a move down the hierarchical structure:

11. #EVENT >> #MATERIAL >> +DO_00 >> +MOVE_00 >> +OPEN_01

The thematic frame shows the argument structure associated with the concept and the functions of the arguments[10]. For the event sense, this is as in 12[11]:

---

[8] Lexemes are given in small capitals, names for concepts in ordinary capitals.

[9] Suffixes _00, _01, etc, differentiate different though related concepts.

[10] The roles have the following glosses: AGENT: entity that makes another entity move, THEME: entity that changes its place or position, LOCATION: location in which an entity moves, ORIGIN: location from which an entity moves, GOAL: location to which an entity moves. It must be emphasised that these roles are conceptual in nature, being associated with other elements of a conceptual structure, and are not to be equated with semantic roles in any particular language. The thematic frame and meaning postulate are characterisations of the event of opening, not of the way in which this event is portrayed in any given language.

[11] The notational conventions employed in thematic frames and meaning postulates are as follows: ^ represents exclusion, | disjunction, * defeasibility,'e' an event, 'x' an argument (obligatory in the conceptual structure), 'f' a satellite (optional), 'pos' possibility.

12. THEMATIC FRAME:
(x1)Agent
(x2: +DOOR_00 ^ +WINDOW_00)Theme
(x3)Location
(x4)Origin
(x5)Goal

This shows that the event of opening something has an Agent, a Theme which is typically a door or window, a Location, an Origin state and a final Goal state.

The associated meaning postulate is as in 13:

13. MEANING POSTULATE:
*(e1: +MOVE_00 (x1)Agent (x2)Theme (x3)Location (x4)Origin (x5)Goal (f1: +HAND_00 ^+FOOT_00)Instrument (f2: (e2: +BE_01 (x2)Theme (x6: $OPEN_N_00)Attribute)) Condition (f3: (e3: +BE_01 (x2)Theme (x7: +OPEN_00)Attribute))Result)

This shows an 'exemplar concept' represented by a defeasible proposition (marked with *) – i.e. one which can be overridden if necessary (e.g. penguins are birds, and by default birds fly, but penguins do not). The meaning postulate states that an Agent moves a Theme in a Location from an Origin state to a Goal state, typically using hand or foot as an Instrument[12], and that the Theme thereby becomes +OPEN_00 as a result. The condition involving the attribute characterised as $OPEN_N_00 indicates that something must initially be closed ('not-open') if it is then to be opened.

The concept +OPEN_00 has the following thematic frame and meaning postulate:

14. THEMATIC FRAME:
(x1)Theme

15. MEANING POSTULATE:
*(e1: +BE_01 (x1)Theme (x2: +OPEN_00)Attribute (f1: (e2: pos +ENTER_00 (x3)Agent (x4)Theme (x1)Location (x5)Origin (x6)Goal))Purpose | (f2: (e3: pos +LEAVE_00 (x3)Agent (x4)Theme (x1)Location (x6)Origin (x5)Goal))Purpose)

This again shows a defeasible proposition in which a Theme has the Attribute of being open, for the purpose of allowing some Agent to enter or leave through the Theme.

The main part of the conceptual entry to which the verb OPEN is linked, taken from the FunGramKB website, is shown in Table 1[13]. As can be seen, it builds in the thematic frame and meaning postulate for the concept +OPEN 01.

Table 1: Conceptual entry to which the verb OPEN is linked

Associated with the conceptual entry is information on the properties of the lexeme itself, including the morphology of the verb (regular, in the case of the English verb OPEN), its Aktionsart (here, causative achievement, in the scheme used in Role and Reference Grammar: see Van Valin, 2005, pp. 32-42), any idiosyncratic features regarding the number

[12] The inclusion of the foot here as a typical instrument of opening is perhaps debatable.
[13] The conceptual entry in FunGramKB also includes an English paraphrase of the meaning of the concept, with illustrative examples. This adds nothing to the theoretical content and will be omitted here.

of general semantic roles (known in RRG as macroroles: Actor, Undergoer) it takes, what constructions it can occur in (OPEN takes part in the causative/inchoative alternation, resulting in sentences such as *The door opened*), any language-specific collocational information and whether it is associated with any particular style, dialect or domain.

The conceptual entry to which adjectival OPEN is linked is shown in Table 2[14]:

Table 2: Conceptual entry to which adjectival OPEN is linked

What is particularly important about these representations is the fact that the ontology not only represents facts about how concepts are related to one another, but also builds in information of the pragmatic kind envisaged in García Velasco's account (e.g. typical things that are opened, typical purpose for which something is opened). It thus allows a formalisation of the kind of descriptive information given in García Velasco's lexical entries.

4.5 Adapting FunGramKB to FDG

4.5.1 Thematic frames and predication frames

FunGramKB thematic frames for events are very like FDG predication frames in that they indicate the number and semantic function of the arguments of predicates. Compare the thematic frame in 16 with the predication frame for transitive OPEN in 17:

16. THEMATIC FRAME:
(x1)Agent
(x2: +DOOR_00 ^ +WINDOW_00)Theme
(x3)Location
(x4)Origin
(x5)Goal

17. PREDICATION FRAME
(f_1: [(f_2) (v_1)$_A$ (v_2)$_U$] (f_1)) (Hengeveld and Mackenzie, 2008, p. 207)

The Agent of the thematic frame is the Actor of the predication frame, the Theme of the thematic frame is the Undergoer of the predication frame. These are the only two arguments needed for syntactic expression: the remaining arguments are in this sense supplementary, but needed for full interpretation. As FunGramKB thematic frames and meaning postulates represent the conceptual properties to which lexemes can be linked, they could be used as the conceptual correlates of lexical entries in FDG. Each lexical entry would have a pointer to the relevant conceptual entry in the ontology. For transitive OPEN, we would have a pointer to the conceptual entry in Table 1. If Agent and Theme were specified as the arguments needed for selection of the appropriate predication frame, this information would be enough for linking to the predication frame (through the thematic frame) as well as the statement of additional pragmatic information needed for full interpretation (through the meaning postulate). In addition, the lexical entry would contain information about the morphology of the lexeme, its

[14] 'Polar' and 'gradable' have their usual meanings here, in relation to qualities. A quality is dynamic if it can vary along the time axis (e.g. being hot), static if it cannot vary (e.g. being German). A quality is intersective if all individuals would perceive it in the same way (e.g. being naked), subsective if individuals might have different perceptions (e.g, being interesting) (see Periñán-Pascual and Arcas-Túnez 2008).

Aktionsart and constructional possibilities, register and dialect[15] characteristics, language-specific collocational patterns, etc. A provisional entry for OPEN is shown in 18[16]:

18. OPEN
→ OPEN_01
Syntactically profiled participants: Agent, Theme
Morphology: regular verb
Aktionsart: causative accomplishment
Constructions: Causative/inchoative, resultative, instrumental subject ...
Register: neutral
Dialect: standard
Language-specific collocations: ....

4.5.2 Dealing with alternations

We have seen that in FunGramKB the lexical entry contains a specification of the alternations in which (one sense of) a lexeme can appear. For instance, we have seen that OPEN participates in the causative/inchoative alternation; KICK can occur in the caused motion construction (*The player kicked the ball into the goal*, the resultative construction (*They kicked him black and blue*), and so on. Keizer (2009) shows, in the context of her analysis of verb + preposition combinations, that alternations such as the resultative correspond to different predication frames in FDG. Indeed, the causative/inchoative, caused motion and resultative constructions can be related to the list of predication frames given by Hengeveld and Mackenzie (2008, p. 207), as shown in Table 3.

Table 3: Predication frames for the causative/inchoative, caused motion and resultative constructions

The case of the resultative construction, as in *Polly carefully wiped the area dry*, is slightly more complicated, in that this, in FDG terms, involves secondary predication, and so two predication frames, as shown in simplified general form in the table, where the second predication frame represents 'the area' being 'dry'. Lexical entries would include a specification of the constructions they can enter into, each of these having a pointer to the corresponding predication frame.

Keizer (2009) also discusses cases of coercion, where a lexeme is used in a construction which it does not normally occur in, such as *Pat sneezed the foam off the cappuccino* (Goldberg, 2006, p. 73), where SNEEZE, normally used in the intransitive construction, is coerced into a caused motion interpretation by being placed in the frame Subj V Obj $Obl_{path/loc}$. Not every intransitive verb can occur in this construction (**Pat arrived/wept the foam off the cappuccino*), so that individual lexemes belonging to appropriate classes of verb would need to be marked as undergoing the alternation, i.e. as being combinable with the appropriate predication frame[17].

[15] In current FDG, each dialect is regarded as having its own grammar and lexicon. Given that dialects of a given language usually share many grammatical and lexical aspects, this approach seems to introduce an unwelcome amount of redundancy.

[16] Language-specific collocations would need to be determined by detailed corpus work on English and the other languages dealt with in the ontology: for discussion see §4.5.4.

[17] Note that all the phenomena discussed here would be dealt with in the 'grammaticon' of FunGramKB.

4.5.3 Worked examples

First let us consider the straightforward case of transitive OPEN, as in 19:

19. *Mr Milken opened the door*. (BNC ABF 2169)

Before we can deal with this example, a short digression on proper names is required. If we define a concept as "a nonlinguistic psychological representation of a class of entities in the world" (Murphy, 2002, p. 385), then the psychological correlate of a proper name cannot, strictly speaking, be a concept, since it does not represent a class of entities, but rather a single entity. This is what characterises the onomasticon as opposed to a general ontology. However, proposals have indeed been made to represent proper names as elements in an ontology. Krstev et al (2005) proposal a multilingual ontology of proper names in which the top level is labelled as conceptual and is constituted by a set of 'conceptual proper names'. For instance, for the city of Paris there would be a single 'conceptual' entity at this level, which would correspond to a set of different words, in different languages, at the linguistic level (Paris, Parijs, Parigi, Paryz, etc). Such entities are conceptual in the sense that they generalise over a number of different linguistic forms, just as do the items OPEN, ABRIR, ABRIRE, OUVRIR, etc, do. The difference is that the latter refer to a class of individual actions, while the former relate to a single entity. The usage of 'concept(ual)' here thus seems to be extension of its 'pure' use. However, proper names do clearly have a mental representation which is the counterpart of that for 'proper' concepts. I shall therefore use the term 'concept' in a sense which includes the mental entities corresponding to proper names[18].

The speaker or writer first selects the concepts MR MILKEN, OPEN and DOOR, together with the conceptual structure in which they occur, namely OPEN(MR MILKEN$_{Agent}$, DOOR$_{Theme}$)[19]. The concept of definiteness would also need to be selected in association with DOOR, though this lies beyond the scope of the present account. The concepts OPEN and DOOR are available in the general ontology, but the first must be sought in an onomasticon (repository of concepts relating to specific people, places, etc.), into which MR MILKEN has presumably been introduced earlier in the discourse. The selection of OPEN and DOOR activates the following conceptual entry, together with the associated morphological, Aktionsart, constructional and pragmatic information (Table 1 is repeated for convenience as Table 4):

Table 4: Conceptual entry to which the verb OPEN is linked

Table 5: Conceptual entry to which the noun DOOR is linked

Since the conceptual specification of the lexeme OPEN here contains both an Agent and a Theme, it corresponds perfectly to the conceptual configuration chosen by the speaker. Furthermore, the thematic frame mentions DOOR as a likely conceptual filler for the Theme slot. The thematic frame then activates the corresponding predication frame (the one given in 17 is repeated as 20 below):

20. $(f_1: [(f_2)\ (v_1)_A\ (v_2)_U]\ (f_1))$

[18] For information on how the onomasticon is populated in FunGramKB, see Periñán Pascual and Carrión Varela (2011).

[19] This corresponds roughly to what is called the Conceptual Logical Structure in FunGramKB.

This frame, filled in with values for its various components, is then mapped on to the appropriate structure at the morphosyntactic level. The mechanisms postulated here are in line with the production model of Levelt (Levelt, 1989, 1999; Bock and Levelt, 1994), with the important difference that a layer of language-specific semantics (the predication frames) is interposed between the conceptual and morphosyntactic structures, rather than lexemes being associated directly with syntactic frames.

Now let us consider the use of OPEN in 21:

21. *Slowly the door opened*. (BNC FPB 146)

The speaker or writer selects the concepts SLOW, DOOR and OPEN, all available in the general ontology, and the combination which we may represent as (SLOW(OPEN($DOOR_{Theme}$))), where the concept SLOW has scope over the whole event of the door opening. The conceptual parts of the entries for OPEN and DOOR are activated, together with that for SLOW shown in Table 6[20]:

Table 6: Conceptual entry to which SLOW is linked

Now, however, the thematic frame for the concept OPEN_01 does not match the conceptual configuration chosen by the speaker, which contains only one participant. The additional information in the constructional section of the lexical entry is therefore consulted, and it is found that OPEN takes part in the causative/inchoative alternation, the inchoative version having only a single Theme participant. This matches the chosen configuration, so the intransitive predicational frame shown in 22 is activated:

22. $(f_1: [(f_2)\ (v_1)_U]\ (f_1))$

Clearly, we have dealt only with the default, physical interpretation of OPEN, ignoring the many possible metaphorical uses, such as opening a meeting or a debate (though see §4.5.4 for further discussion). In order to cater for these, we would need to bring in theories of metaphor which lie beyond the scope of the present account – but is should be noted in passing that the Lexical Constructional Model within which FunGramKB is embedded incorporates detailed accounts of metaphorical and metonymic meaning extensions.

Finally, let us examine the example in 23:

23. *The man kicked the dog into a corner …* (BNC FRK 664)

All the concepts needed are available in the ontology, as shown in Tables 7-10[21]:

Table 7: Conceptual entry to which the verb KICK is linked

Table 8: Conceptual entry to which the noun MAN, in the sense of male human being, is linked

Table 9: Conceptual entry to which the noun DOG is linked

[20] 'Serial' indicates that the absence of the quality does not imply the presence of the opposing quality. Thus size and speed are serial, but openness is polar. The negation operator is represented by 'n'.

[21] Note that there is no concept +KICK_00 in the ontology. When the evidence does not support the creation of a more specific concept, according to the guidelines set out for the construction of the ontology, then a lexical item will be related to a more general concept, in this case that of hitting.

Table 10: Conceptual entry to which the noun CORNER is linked

The conceptual configuration HIT(MAN_{Agent}, DOG_{Theme}, $CORNER_{Goal}$) corresponds perfectly to the thematic frame to which KICK is linked, and the predication frame shown in 24 is activated:

24. $(f_1: [(f_2)\ (v_1)_A\ (v_2)_U\ (v_3)_L]\ (f_1))$

4.5.4 Collocation

We have seen that in FunGramKB, collocational information can be derived from two sources. Firstly, the information on characteristic conceptual values for arguments associated with a particular concept predicts that certain collocates will occur with the lexemes linked to this concept, across the various languages for which the ontology is considered to be valid. Secondly, there will also be language-specific collocations. We will deal with each of these in turn.

The fact that e.g. DOOR/WINDOW appear as typical arguments in the thematic frame for +OPEN_01, i.e. the fact that doors and windows are characteristically subject to the process of opening, leads us to expect that the corresponding lexical items will be frequent collocates of the verbs which indicate the process of opening something in all the various languages we are dealing with. As a further example, MEW is linked to the conceptual information shown in Table 11 (again taken from the English lexicon on the FunGramKB website):

Table 11: Conceptual entry to which MEW is linked

This shows that the characteristic Theme is an entity corresponding to the concept +CAT_00. A similar way of dealing with collocation has also been proposed in Systemic Functional Linguistics by Martin (1992) and Tucker (2006, 2009, pp. 421-422), though not within a cross-linguistic conceptual framework, the idea being that collocates can be modelled as frequent fillers of the various participant roles associated with particular lexical verbs, such as OPEN and MEW – for further discussion see Butler (2009a, p. 61).

In point of fact, the situation is more complex than the simple scenario incorporated into FunGramKB would suggest. In Table 12 are presented, in descending order of frequency, the top 20 nominal lemmas the forms of which collocate with forms of the lemmas OPEN, OUVRIR, ÖFFNEN, ABRIR and APRIRE in the Leeds Internet corpora of English, French, German, Spanish and Italian[22], within a span of 4 words on the right of the headword, as judged by the log likelihood statistic for estimating collocational strength. Table 13 shows the items for which equivalents occur in all five languages.

Table 12: Top 20 nominal lemmas the forms of which collocate with forms of the lemmas OPEN, OUVRIR, ÖFFNEN, ABRIR and APRIRE in the Leeds internet corpora

Table 13: Items with equivalents across the five languages

It could be argued that the process of opening associated with all of these items corresponds to the concept +OPEN_01 in the FunGramKB ontology, with some

[22] These corpora are available through the Intellitext interface at corpus.leeds.ac.uk/it.

modifications of the entry for that concept. The opening of gates is clearly very similar to that of doors and windows. Opening one's eyes or mouth is also a process of movement whose main purpose is to facilitate the passage of something (light; food, drink, air) through the aperture created. If we allow metaphorical rather than literal interpretations, opening one's heart to someone or something is also a way of "letting that someone or something in". With opening the way to something we are creating a space for that something to pass through. +GATE_00, +EYE_00, +MOUTH_00 and +HEART_00 are available in the FunGramKB ontology, and the English lexical item WAY is linked to the concept +PATH_00, so that all of these could be added to the specification of typical Themes in the thematic frame of +OPEN_01. However, there should be a notational convention indicating that these are only typical Themes, not the only possible ones. There are a number of items in the collocation lists which are not common to the top 20 in all languages but would also fit into the specification of Themes for +OPEN_01, such as BOX/BOÎTE/CAJA, FILE/DATEI/FILA, or with non-concrete entities POSSIBILITY/POSIBILIDAD, DÉBAT/DIBATTITO. Furthermore, the specification of a (typical) instrument of opening, currently hand or foot, should be removed, since this is clearly not appropriate in cases such as the opening of eyes or mouths, let alone hearts and ways, and opening with the foot is hardly typical even for doors and gates, and certainly not for windows.

Clearly, there are other types of opening which are implied by the lists of frequent collocates for OPEN and its equivalents in other languages. For example, a collocation found in the English, French and Italian top 20 lists is OPEN + FIRE and its equivalents, and it is hard to see how this could be accommodated within the concept +OPEN_01, even with metaphorical extension. It is likely that in such cases a new concept would need to be created within the OPEN family, or a single concept OPEN FIRE could be allowed if this were treated as an idiom (cf. TAKE A LEAK, mentioned in §4.5.5 below).

We also need to recognise that there are collocates in one language which do not have direct equivalents in the others, and so are not due to general cognitive properties of concepts, but rather to the lexical peculiarities of particular languages. As an example, consider the Spanish verb PRESTAR, which often corresponds to English LEND: for instance, both can be used in the context of loaning money or personal possessions. Collocational analysis of the Leeds Internet Corpora for the two languages shows that in both we also find the use of these verbs with an object noun phrase such as SUPPORT/APOYO or HELP/AYUDA. However, the most frequent collocate in Spanish is ATENCIÓN ('attention'), while the collocation LEND ATTENTION is not generally found in English, the verb normally used being PAY. A further clear example is the most normal and frequent way of expressing the idea of a very large change, increase or advance. In the Leeds Internet Corpus of English, the most frequent item at the first position to the left of the lexeme LEAP is QUANTUM, whereas in the corresponding internet corpus of Spanish the most frequent item one word to the right[23] of SALTO (the equivalent of LEAP) is CUALITATIVO.

4.5.5 Lexis in relation to situational context

Referring to my own work on lexical choice, which suggests the introduction into FDG of ways of dealing with the effect of contextual parameters such as degree of formality on the choice of lexical items, Hengeveld and Mackenzie explicitly rule out, in their own version of FDG, such a view of the Contextual Component and its relationship with lexical phenomena:

[23] To the right, because of noun + adjective ordering in Spanish.

> Many of the factors which he [Butler: CSB] himself includes in such a Component, like the factors that would induce selection of the informal lexeme *kid* rather than *child* in English to designate a child, would not find their way into an FDG Contextual Component. In an informal context, after all, a child may indeed be evoked by means of *kid*, but nothing prevents the choice of *child*. For this reason, factors relating to matters of genre, register, style, etc, will be included only where these can be shown to have a systematic effect upon grammatical choices in formulation … (Hengeveld and Mackenzie, 2008, p. 10).

Hengeveld and Mackenzie do not mention, however, that in a formal situation there would indeed be stylistic pressure to reject the informal form *kid* in favour of the neutral form, unless the speaker or writer wished to create a humorous effect or temporarily suspend or alter the formal tone. In later work, Mackenzie recognises the possibility not only that the Contextual Component "should be seen as also holding longer-term information about the ongoing situation", but also that it "should perhaps be ascribed even broader functions, encompassing such matters as the nature and purpose of the current interaction as well as the socio-cultural environment in which it is taking place" (Mackenzie, 2011, p. 1), though his paper continues to espouse a more conservative view, and he himself remains agnostic on the issue (Mackenzie, pers. comm.). This is a welcome development, since contextual factors such as the degree of formality of the situation as perceived or manipulated by the speaker, the genre of the interaction being engaged in, the dialect being used, and many more, do indeed influence grammatical and phonological as well as lexical choices. As a brief, random selection of such influences, consider the more frequent use of passive verbs in technical texts, the effect of formality on phonological realisation, or the use of verb forms in standard and non-standard dialects of English. Hengeveld and Mackenzie would no doubt reply that such influences are probabilistic in nature, and so fall outside the scope of FDG, which is concerned only with what is obligatory in the grammar, rather than with what is more frequent in certain contextual circumstances. However, as noted earlier, functional linguistics is claimed to be motivated first and foremost by the claim that language is primarily a form of communication, and that this fact is crucial to an understanding of what languages are like. If we are to take this claim seriously, we must surely investigate language as it is actually used by speakers and writers, and there is ample evidence from functional-cognitive usage-based models (see e.g. Bybee and Hopper, 2001; Bybee 2010) for the inherently probabilistic nature of language and the crucial role of frequency in how we process it and how it is learned. In the interests of psychological/cognitive and acquisitional adequacy, it is important for FDG to be responsive to such facts, though it is hard to see how this can be done within the model as currently conceived..

There is already work in FDG which puts forward this view. Connolly (2007a), in a paper on context in FDG, begins as follows:

> Any approach to language that merits the epithet 'functional' has to take into account not only the lexical, morphosyntactic and semantic resources afforded by the language system, but also the ways in which those resources may be deployed for the purposes of communication. An important fact about communication is that it always takes place in a context, and such contextually-situated use of language constitutes the essential concern of pragmatics. (Connolly, 2007a, p. 11)

Connolly puts forward a model of context as a structured but dynamic construct, changing to some degree as the discourse proceeds. Four broad distinctions are made within context as a whole (Connolly, 2007b): discoursal vs. situational, physical vs. socio-cultural, narrower vs.

broader, mental vs. extra-mental. Discourse context is constituted by "the surrounding (relevant) multimodal discourse, including both the linguistic and non-verbal aspects", while the situational context comprises "the part of the context that falls outside of the current (or any other) discourse", and is divided into its physical aspect, including factors such as time and space, and socio-cultural aspects, "notably social organisation and norms of thought and behaviour" (Connolly, 2007a: 14-15). Both discoursal and situational context can be seen in narrower or broader terms. The narrower discoursal context is the rest of the discourse being described, and is what Hymes (1972) referred to as the 'setting'; it includes the animate and inanimate entities present, and location in time and space. The broader discoursal context includes other discourses which may be relevant ('intertextuality'). The narrower situational context is provided by the immediate surroundings, and constitutes what Hymes called the 'scene', including (i) the discourse participants, their social and psychological properties, including the social relationship between them, (ii) the type of occasion involved, characterised by, for instance, its degree of formality, and (iii) the purposes and outcomes of the discourse. The broader situational context is what is provided, physically and socioculturally, outside the immediate context. Finally, the mental context is "that part of the context that resides in the minds of the producers and the interpreters (including analysts) of a discourse or fragment", and the extra-mental context is anything outside the mental sphere. Context is thus characterised as a multidimensional, hierarchically-structured part of a model of interacting participants.

Connolly (2011) offers revisions and enhancements of what he now calls his extended model of context (EMC). In particular, he presents an architecture for the discoursal and situational context components of the EMC which is intended to support a dynamic implementation of FDG. He gives an account of the ways in which both of these components influence the Grammatical Component. Crucially for our purposes here, Connolly supports the incorporation of probabilistic information into FDG, as advocated earlier in the present section:

> Hengeveld and Mackenzie's restriction of contextual relevance to factors with a systematic influence on grammatical choices also calls for comment. They interpret this criterion very narrowly, confining it to cases of 'mandatory' influence [...]. However, in principle, systematic *quantitative* relationships are also possible, and if FDG is to offer an attractive face to sociolinguists and psycholinguists, for example, then, it is possible that statistical systematic relationships may, indeed, need to be countenanced. (Connolly, 2011, 20, emphasis in original)

Connolly also discusses the interaction of the Grammatical, Contextual and Conceptual Components, pointing out that since context must be mentally represented in order for the information in it to be used, the Conceptual Component obligatorily acts as an intermediary in handling this information.

We have already seen that FunGramKB provides, in lexical entries, information about style, register and domain, which would fall into the narrower situational context (subclass: type of occasion) in Connolly's model. For instance, the entry in the English lexicon for TAKE A LEAK, which has the underlying concept +EXPEL_00 and a thematic frame indicating that what is characteristically expelled is urine and that the location of expulsion is a body part, is also marked as slang in the style section, while the lexical item MASTICATE is marked as formal. Once again, this would be a good model for FDG to follow.

4.6 Between grammar and lexis: prefabricated expressions

Hengeveld and Mackenzie (2008) adopt a strict division between grammar and lexicon in FDG, a position which is also held in RRG (see Van Valin, 2005, pp. 158-161)[24]. We have seen that lexemes in FDG are independent of the predication frames into which they are slotted once the frames themselves have been selected. The division between the grammatical and the lexical is also evident in the distinction, carried over from Functional Grammar, between operators, which (apart from those having their effect at the phonological level) are realised morphosyntactically, and modifiers, which are lexical in their realisation. This approach contrasts strongly with that in Systemic Functional Linguistics and in cognitive models, where grammar and lexis are seen as on a continuum, with no sharp dividing line between them.

The approach taken in mainstream FDG faces several problems. The distinction between lexical (content) words and grammatical (form) words is a notoriously slippery one, and this, as has been pointed out by critics in the literature, casts some doubt on the validity of the operator/modifier distinction. Furthermore, it is necessary, at some point in the model, to relate the meanings of operators such as those for tense, and their modifier counterparts, temporal adverbials. A productive test bed for the grammatical/lexical distinction is prepositions (or, more generally, adpositions), and here we have two contemporaneous but rather different proposals: Hengeveld and Mackenzie (2008, pp. 251, 259-260) make a distinction between lexical adpositions, such as English *during*, *before*, *after*, *prior to*, and grammatical ones, such as *at*, *from*, *for*, *until*/*till*/*to* used in time expressions, which express semantic functions; Keizer (2008), on the other hand, claims that almost all English prepositions are lexical, the exceptions being *of* and *by* when used to introduce arguments. Keizer (2007), in her detailed study of the criteria which can be used to support the lexical/grammatical dichotomy in FDG, concludes that although this distinction is useful, as is that between operators and predicates, there is no one-to-one relationship between the two classifications.

However, perhaps the most serious problem with a sharp distinction between the grammatical and the lexical is that is makes it difficult for FDG to accommodate certain kinds of 'prefabricated' expression, which combine obligatory grammatical structures with partially-constrained filling of slots by lexical items. A clear instance is the family of expressions which we may represent schematically as in 25, which underlies examples such as 26-28 (for an account based on Systemic Functional Grammar, see Tucker, 1996):

25. NP HAVE + *n't the* {faintest/slightest/foggiest/least/remotest} {idea/notion/concept/clue}

26. *I haven't the faintest idea who they are*. (BNC KCX 2029)

27. *...when or where that will be we haven't the faintest idea*. (BNC JYE 4237)

28. *I haven't the foggiest idea*. (BNC KCX 6431)

There is overwhelming evidence from corpus studies that partially filled expressions of this type are extremely important, both quantitatively and qualitatively, in the language we routinely use. It is much easier to account for such expressions in a constructional approach

[24] The following brief discussion summarises that in Butler (2009a, pp. 56-59).

to language which accepts that there is a continuum between words, partially filled expressions with slots, and abstract constructions such as the ditransitive, resultative, etc. (see e.g. Goldberg, 2006, p. 5).

One possible way of approaching this problem is suggested by the observation in Hengeveld and Smit (2009, p. 1120) that "predication frames are stored as primitives in the fund, just as lexemes", the fund being taken over from Functional Grammar as "[t]he component that contains all the predicates and terms from which predications can be construed" (Dik, 1997, p. 58). It would therefore be possible to propose that the fund consists of a continuum, with simple lexemes at one end, predication frames at the other, and schemata for partially prefabricated expressions between the two ends, an idea which is left as a possibility for further research.

5. Conclusion

In this article, I have taken inspiration from ideas put forward by García Velasco (2007), in order to sketch a model for the lexicon in FDG and its interaction with the Conceptual and Contextual Components. While totally agreeing with García Velasco that an account of the lexicon in a functional theory needs to go well beyond the specification of properties relevant to the morphosyntax, to embrace the lexical competence of language users in all its richness, I have proposed a number of developments and modifications of his account which, I believe, may prove useful in the elaboration of a model for lexical phenomena in FDG.

Firstly, I argue for something which is implicit in García Velasco's proposal, namely that the construction of structural frames is mediated through lexical selection, rather than frames being selected first as proposed by Hengeveld and Mackenzie (2008). Secondly, I show that despite García Velasco's proposal to minimise the decompositional aspects of many current approaches to the lexicon, decomposition is indeed still needed. Thirdly, I attempt to put some flesh on the bones of García Velasco's claim that the lexicon needs to be connected to the Conceptual Component, proposing that what is needed is a hierarchically structured ontology which is language-neutral, within the scope of a set of culturally-related languages, together with a set of lexicons, one for each language. I suggest that the FunGramKB model already elaborated within the Lexical Constructional Model provides a suitable basis for adaptation in FDG, and demonstrate the parallels between the thematic frames of the FunGramKB model and the predication frames of FDG. I also show how constructional alternations, such as the causative/inchoative, caused motion and resultative, could be handled within the model I have proposed. Three worked examples are provided. Finally, the paper demonstrates how collocation, lexical variation with register and dialect, and partially prefabricated expressions can be dealt with in the model.

Acknowledgements

The research for this article was partially financed by the research project INCITE09 204155 PR (Autonomous Government of Galicia) and FFI2010-19380 (Spanish Ministry of Science and Innovation). I am grateful to Daniel García Velasco, Lachlan Mackenzie, Ricardo Mairal Usón and Carlos Periñán Pascual for their valuable comments on earlier drafts, which have led to substantial improvements.

References

Bock, K., Levelt, W.J.M., 1994. Language production: grammatical encoding. In: Gernsbacher, M.A. (Ed.), Handbook of Psycholinguistics. Academic Press, New York, pp. 945-84.

Butler, C.S., 2008a. Cognitive adequacy in structural-functional theories of language. Language Sciences 30, 1-30.

Butler, C.S., 2008b. Interpersonal meaning in the noun phrase. In: García Velasco, D., Rijkhoff, J. (Eds.), The Noun Phrase in Functional Discourse Grammar. (Trends in Linguistics, Studies and Monographs, 195). Mouton de Gruyter, Berlin and New York, pp. 221-261.

Butler, C.S., 2009a. Lexical phenomena in Functional Discourse Grammar and Systemic Functional Linguistics. In: Slembrouck, S., Taverniers, M., van Herreweghe, M. (Eds.), From *will* to *well*: Studies in Linguistics Offered to Anne-Marie Simon-Vandenbergen. Academia Press, Gent, pp. 55-67.

Butler, C.S., 2009b. Criteria of adequacy in functional linguistics. Folia Linguistica 43 (1), 1-66.

Butler, C.S., 2009c. From conceptualisation to morphosyntax in Functional Discourse Grammar: the role of predication frames. Paper given at workshop on the Morphosyntactic Level in Functional Discourse Grammar, Gijón, Spain, 17-18 September 2009.

Butler, C.S., 2009d. The Lexical Constructional Model: genesis, strengths and challenges. In: Butler, C.S., Martín Arista, J. (Eds.), Deconstructing Constructions. John Benjamins, Amsterdam and Philadelphia, pp. 117-152.

Bybee, J., 2010. Language, Usage and Cognition. Cambridge University Press, Cambridge.

Bybee, J, Hopper, P. (Eds.), 2001. Frequency and the Emergence of Linguistic Structure. (Typological Studies in Language 45.) John Benjamins: Amsterdam and Philadelphia.

Connolly, J.H., 2007a. Context in Functional Discourse Grammar. ALFA Revista de Lingüística 51 (2), 11-33.

Connolly, J.H., 2007b. Mental context and the expression of terms within the English clause: an approach based on Functional Discourse Grammar. In: Hannay, M., Steen, G.J. (Eds.), Structural-Functional Studies in English Grammar: In Honour of Lachlan Mackenzie. John Benjamins, Amsterdam and Philadelphia, pp. 193-208.

Connolly, J.H., 2011. The contextual component within a dynamic implementation of the FDG model: structure and interaction. Written version of a paper presented at the Workshop on the Contextual Component in Functional Discourse Grammar, University of Barcelona, Spain, 8-9 September 2011.

Dik, S.C., 1977. The Theory of Functional Grammar. Part 1: The Structure of the Clause. Ed. Kees Hengeveld. Mouton de Gruyter, Berlin and New York.

Fernández, E.M., Cairns, H.S., 2011. Fundamentals of Psycholinguistics. John Wiley, Chichester.

García Velasco, D., 2007. Lexical competence and Functional Discourse Grammar. ALFA Revista de Lingüística 51 (2), 165-187.

García Velasco, D.. Hengeveld, K., 2002. Do we need predicate frames? In: Mairal Usón, R., Pérez Quintero, M.J. (Eds.), New Perspectives on Argument Structure in Functional Grammar. (Functional Grammar Series 25.) Mouton de Gruyter, Berlin and New York, pp. 95-123.

Goldberg, A.E., 2006. Constructions at Work: The Nature of Generalization in Language. Oxford University Press, Oxford and New York.

Harley, T.A., 2008. The Psychology of Language: From Data to Theory. 3rd ed. Psychology Press, Hove and New York.

Hengeveld, K., 2004. Epilogue. In: Mackenzie, J.L., Gómez-González, M.L.A. (Eds.), A New Architecture for Functional Grammar. (Functional Grammar Series 24.) Mouton de Gruyter, Berlin and New York, pp. 365-378.
Hengeveld, K., 2005. Dynamic expression in Functional Discourse Grammar. In: de Groot, C., Hengeveld, K. (Eds.), Morphosyntactic Expression in Functional Grammar. (Functional Grammar Series 27.) Mouton de Gruyter, Berlin and New York, pp. 53-86.
Hengeveld, K., Mackenzie, J.L., 2008. Functional Discourse Grammar: A Typologically-Based Theory of Language Structure. Oxford University Press, Oxford and New York.
Hengeveld, K., Smit, N., 2009. Dynamic formulation in Functional Discourse Grammar. Lingua 119 (8), 1118-1130.
Honselaar, W., Keizer, E., 2009. Lexicon and frames in FDG: a treatment of Dutch *bekend zijn* 'to be familiar, well known', *behandelen* 'to treat' and *trouwen* 'to marry'. Lingua 119 (8), 1212-1241.
Hymes, D., 1972. Models of the interaction of language and social life. In: Gumperz, J.J., Hymes, D. (Eds.), Directions in Sociolinguistics: The Ethnography of Communication. Holt, Rinehart and Winston, New York, pp. 35-71.
Jackendoff, R., 1990. Semantic Structures. MIT Press, Cambridge, MA.
Jiménez Briones, R., Luzondo Oyón, A., 2011. Building ontological meaning in a lexico-conceptual knowledge base. ONOMÁZEIN 23, 11-40.
Keizer, E., 2007. The lexical-grammatical dichotomy in Functional Discourse Grammar. ALFA Revista de Lingüística 51 (2), 35-56.
Keizer, E., 2008. English prepositions in Functional Discourse Grammar. Functions of Language 15 (2), 216-256.
Keizer, E., 2009. Verb-preposition constructions in FDG. Lingua 119 (8), 1186-1211.
Krstev, S., Vitas, D., Maurel, D., Tran. M., 2005. Multilingual ontology of proper names. In: Vetulani, Z. (Ed.), Proceedings of Second Language and Technology Conference, Poznań, pp. 116-119.
Langacker, R.W., 1987. Foundations of Cognitive Grammar. Volume I: Theoretical Prerequisites. Stanford University Press, Stanford, CA.
Levelt, W.J.M., 1989. Speaking: From Intention to Articulation., MIT Press, Cambridge, MA.
Levelt, W.J.M., 1999. Producing spoken language: a blueprint of the speaker. In: Brown, C.M., Hagoort, P. (Eds.), The Neurocognition of Language. Oxford University Press, Oxford, pp. 83-122.
Mackenzie, J.L., 2011. The Contextual Component in a dialogic FDG. Written version of a paper presented at the Workshop on the Contextual Component in Functional Discourse Grammar, University of Barcelona, Spain, 8-9 September 2011.
Mairal Usón, R., Pérez Quintero, M.J., 2002. Introduction. In: Mairal Usón, R., Pérez Quintero, M.J. (Eds.), New Perspectives on Argument Structure in Functional Grammar. (Functional Grammar Series 25.) Mouton de Gruyter, Berlin and New York, pp. vii-xii.
Mairal Usón, R., Periñán-Pascual, C., 2009. The anatomy of the lexicon component within the framework of a conceptual knowledge base. Revista Española de Lingüística Aplicada 22, 217-244.
Mairal Usón, R., Ruiz de Mendoza Ibáñez, F.J., 2009. Levels of description and explanation in meaning construction. In: Butler, C.S., Martín Arista, J. (Eds.), Deconstructing Constructions. John Benjamins, Amsterdam and Philadelphia, pp. 153-198.
Marconi, D., 1997. Lexical Competence. MIT Press, Cambridge, MA.
Martin, J.R., 1992. English Text: System and Structure. John Benjamins, Amsterdam and Philadelphia.
Murphy, G.L., 2002. The Big Book of Concepts. MIT Press, Cambridge, MA.

O'Neill, G., 2011. Humming, Whistling, Singing, and Yelling in Pirahã. Context and channels of communication in FDG. Written version of paper presented at the Workshop on the Contextual Component in Functional Discourse Grammar, University of Barcelona, Spain, 8-9 September 2011.
Periñán-Pascual, C., Arcas-Túnez, F., 2008. A cognitive approach to qualities in NLP. Procesamiento del Lenguaje Natural 41, 137-144.
Periñán-Pascual, C., Arcas–Túnez, F., 2010. Ontological commitments in FunGramKB. Procesamiento del Lenguaje Natural 44, 27-34.
Periñán Pascual, C., Carrión Varela, M.L., 2011. FunGramKB y el conocimiento cultural. ANGLOGERMANICA ONLINE 2011, 87-105.
Periñán-Pascual, C., Mairal Usón, R., 2010. La Gramática de COREL: un lenguaje de representación conceptual. ONOMÁZEIN 21 (1), 11-45.
Ruiz de Mendoza Ibáñez, F.J., Gonzálvez-García, F., 2011a. Illocutionary meaning revisited: Subjective-transitive constructions in the Lexical Constructional Model. In: Stalmaszczyk, P. (Ed.), Cognitive and phenomenological turns in philosophy of language and linguistics (Łódź Studies in Language). Peter Lang, Frankfurt am Main, pp. 65-77.
Ruiz de Mendoza Ibáñez, F.J., Gonzálvez-García, F., 2011b. The illocutionary value of subjective transitive constructions: A view from the Lexical Constructional Model. Paper given at the 44th Annual Meeting of the Societas Linguistica Europaea, University of La Rioja, Spain, 8 September 2011.
Ruiz de Mendoza, F.J., Mairal Usón, R., 2008. Levels of description and constraining factors in meaning construction: an introduction to the Lexical Constructional Model. Folia Linguistica 42 (2), 355-400.
Ruiz de Mendoza, F.J., Mairal Usón, R., 2011. Constraints on syntactic alternation: lexical-constructional subsumption in the Lexical-Constructional Model. In: Guerrero Medina, P. (Ed.), Morphosyntactic Alternations in English. Functional and Cognitive Perspectives. Equinox, London and Oakville, CT, pp. 62-82.
Schönefeld, D., 2001. Where Lexicon and Syntax Meet. Mouton de Gruyter, Berlin and New York.
Slobin, D.I., 1996. From "thought and language" to "thinking for speaking". In: Gumperz, J.J., Levinson, S.C. (Eds.) Rethinking Linguistic Relativity. Cambridge University Press, Cambridge, pp. 70-96.
Tucker, G.H., 1996. So grammarians haven't the faintest idea: reconciling lexis-oriented and grammar-oriented approaches to language. In: Hasan, R., Cloran, C., Butt, D.G. (Eds.), Functional Descriptions: Theory in Practice. John Benjamins, Amsterdam and Philadelphia, pp.145-178.
Tucker , G.H., 2006. Systemic In*corpora*tion: On the relationship between corpus and Systemic Functional Grammar. In: Thompson, G., Hunston, S. (Eds.), System and Corpus: Exploring Connections. Equinox, London and Oakville, CT, pp. 81-102.
Tucker, G.H. 2009. In search of the grammarian's dream: locating lexis in a systemic functional grammar. In: Slembrouck, S., Taverniers, M., van Herreweghe, M. (Eds.), From *will* to *well*: Studies in Linguistics Offered to Anne-Marie Simon-Vandenbergen. Academia Press, Gent, pp. 415-426.
Van Valin, R.D., Jr., 2005. Exploring the Syntax-Semantics Interface. Cambridge University Press, Cambridge.

Table 1

CONCEPT:	+OPEN_01
THEMATIC FRAME:	(x1)Agent (x2: +DOOR_00 ^ +WINDOW_00)Theme (x3)Location (x4)Origin (x5)Goal
MEANING POSTULATE:	*(e1: +MOVE_00 (x1)Agent (x2)Theme (x3)Location (x4)Origin (x5)Goal (f1: +HAND_00 ^ +FOOT_00)Instrument (f2: (e2: +BE_01 (x2)Theme (x6: $OPEN_N_00)Attribute))Condition (f3: (e3: +BE_01 (x2)Theme (x7: +OPEN_00)Attribute))Result)

Table 2

CONCEPT:	+OPEN_00
SEMANTIC TYPE:	dynamic, polar, gradable, intersective
THEMATIC FRAME:	(x1)Theme
MEANING POSTULATE:	*(e1: +BE_01 (x1)Theme (x2: +OPEN_00)Attribute (f1: (e2: pos +ENTER_00 (x3)Agent (x4)Theme (x1)Location (x5)Origin (x6)Goal))Purpose | (f2: (e3: pos +LEAVE_00 (x3)Agent (x4)Theme (x1)Location (x6)Origin (x5)Goal))Purpose)

Table 3

Alternation	Example	Predication frame
Causative/inchoative	*Mr Milken opened the door*. (BNC ABF 2169)	$(f_1: [(f_2)\ (v_1)_A\ (v_2)_U]\ (f_1))$
	Slowly the door opened. (BNC FPB 146)	$(f_1: [(f_2)\ (v_1)_U]\ (f_1))$
Caused motion	*She crammed the meat into her mouth*. (BNC HTY 3825)	$(f_1: [(f_2)\ (v_1)_A\ (v_2)_U\ (v_3)_L]\ (f_1))$
Resultative	*Polly carefully wiped the area dry*. (BNC H7W 4264)	$(f_1: [(f_2)\ (v_1)_A\ (v_2)_U]\ (f_1))\ (f_3: [(f_4)\ (v_2)_U]\ (f_3))$

Table 4

CONCEPT:	+OPEN_01
THEMATIC FRAME:	(x1)Agent (x2: +DOOR_00 ^ +WINDOW_00)Theme (x3)Location (x4)Origin (x5)Goal
MEANING POSTULATE:	*(e1: +MOVE_00 (x1)Agent (x2)Theme (x3)Location (x4)Origin (x5)Goal (f1: +HAND_00 ^ +FOOT_00)Instrument (f2: (e2: +BE_01 (x2)Theme (x6: $OPEN_N_00)Attribute))Condition (f3: (e3: +BE_01 (x2)Theme (x7: +OPEN_00)Attribute))Result)

Table 5

CONCEPT:	+DOOR_00
MEANING POSTULATE:	+(e1: +BE_00 (x1: +DOOR_00)Theme (x2: +BARRIER_00)Referent)

Table 6

CONCEPT:	+SLOW_00
SEMANTIC TYPE:	dynamic, serial, gradable, subsective
THEMATIC FRAME:	(x1)Theme
MEANING POSTULATE:	*(e1: +BE_01 (x1)Theme (x2: +SLOW_00)Attribute) +(e2: +BE_00 (x2)Theme (x3: +SPEED_00)Referent) *(e3: n +BE_01 (x1)Theme (x4: +FAST_00)Attribute)

Table 7

CONCEPT:	+HIT_00
THEMATIC FRAME:	(x1: +HUMAN_00 ^ +ANIMAL_00)Agent (x2: +FOOT_00)Theme (x3)Origin (x4)Goal
MEANING POSTULATE:	+(e1: +PUT_00 (x1)Agent (x2)Theme (x3)Origin (x4)Goal (f1: +FAST_00)Speed (f2: +HARD_00)Manner)

Table 8

CONCEPT:	+MAN_00
SEMANTIC TYPE:	
MEANING POSTULATE:	+(e1: +BE_00 (x1: +MAN_00)Theme (x2: +ADULT_00)Referent) +(e2: +BE_01 (x1)Theme (x3: +MALE_00)Attribute)

Table 9

CONCEPT:	+DOG_00
SEMANTIC TYPE:	
MEANING POSTULATE:	+(e1: +BE_00 (x1: +DOG_00)Theme (x2: +MAMMAL_00)Referent) *(e2: pos +PROTECT_00 (x1)Theme (x3: +HOUSE_00)Referent) *(e3: pos +HUNT_00 (x1)Theme (x4: +ANIMAL_00)Referent)

Table 10

CONCEPT:	+CORNER_00
SEMANTIC TYPE:	+rigid, +dependent
MEANING POSTULATE:	+(e1: +BE_00 (x1: +CORNER_00)Theme (x2: +ANGLE_00)Referent) +(e2: +COMPRISE_00 (x3: 2 +WALL_00)Theme (x1)Referent)

Table 11

CONCEPT:	$SOUND_00
THEMATIC FRAME:	(x1: +CAT_00)Theme (x2: +SOUND_00)Referent
MEANING POSTULATE:	+(e1: +SOUND_01 (x1)Theme (x2)Referent (f1)Beneficiary (f2: +MOUTH_00)Instrument)

Table 12

English	French	German	Spanish	Italian
DOOR	PORTE	TÜR	PUERTA	DIBATTITO
EYE	ŒIL	AUGE	CAMINO	FORUM
WINDOW	VOIE	FENSTER	OJO	PORTA
MOUTH	BOUCHE	TOR	PASO	OCCHIO
FIRE	FENÊTRE	MUND	VENTANA	BOCCA
ACCOUNT	FEU	DATEI	POSIBILIDAD	FINESTRA
POSSIBILITY	PORTIER-PORTIÈRE	TÜRE	BOCA	STRADA
FILE	DÉBAT	PFORTE	ESPACIO	VARCO
BROWSER	COFFRE	SCHLIEßEN	HORIZONTE	FUOCO
GATE	ENQUÊTE	THÜR	BRECHA	INCHIESTA
MARKET	PUBLIC	HERZ	CORAZÓN	SPORTELLO
BOX	BRÈCHE	BLICK	PERSPECTIVA	FILA-FILE
MIND	LIVRE	SCHLEUSE	MERCADO	PARENTESI
OPPORTUNITY	MONDE	WEG	CAJA	DISCUSSIONE
STORE	TIROIR	ORDNER	ARCHIVO	DIALOGO
HEART	PERSPECTIVE	DECKEL	ABANICO	SPIRAGLIO
FLOODGATE	PASSAGE	KLAPPE	CUENTA	CUORE
SOURCE	BOÎTE	SCHRANK	MENTE	CANCELLO
WAY	CŒUR	FLASCHE	PAR	PERSPETTIVA
OFFICE	CHEMIN	UMSCHLAG	CAMPO	CASSETTO

Table 13

English	French	German	Spanish	Italian
DOOR	PORTE/ PORTIÈRE	TÜR/ TÜRE/THÜR	PUERTA	PORTA/SPORTELLO
WINDOW	FENÊTRE	FENSTER	VENTANA	FINESTRA
GATE	PORTE	TOR/PFORTE	PUERTA	CANCELLO
EYE	ŒIL	AUGE	OJO	OCCHIO
MOUTH	BOUCHE	MUND/KLAPPE	BOCA	BOCCA
HEART	CŒUR	HERZ	CORAZÓN	CUORE
WAY	VOIE/CHEMIN/PASSAGE	WEG	CAMINO/PASO	STRADA/VARCO