Accommodating Multimodality in Functional Discourse Grammar

John H. Connolly, Loughborough University, Great Britain

Abstract

The process of human communication is essentially multimodal in character. The present paper is concerned with the question of how this multimodality, which is so evident in environments such as the World-Wide Web, should be addressed from the point of view of Functional Discourse Grammar (FDG).

Hengeveld & Mackenzie (2008: 1) echo the point of view expressed by Dik (1997a: 1-4) when he contended that a theory of grammar should form part of a broader "pragmatic theory of verbal interaction" and should be "devised in such a way that it can most easily and realistically be incorporated" into that broader theory. If we are to accomplish this aim, then we need to develop a multimodal, multi-level framework for the analysis and description of discourse, while at the same time seeking to ensure that FDG fits as seamlessly as possible into that framework,

This suggests that we should attempt to develop a semiotically-based approach to discourse analysis that includes, but is not confined to, human language, and which as far as possible employs the same descriptive apparatus for linguistic and non-linguistic modes of communication. Such an approach forms the basis of the present paper, which explores the idea further and assesses its viability.

1. Introduction

As befits a pragmatically oriented theory of human language, Functional Grammar (FG) has always been concerned with the process of interpersonal communication, and thus with discourse. This accommodation towards discourse has developed over the years, as manifested in the publication of books such as Dik (1997b: esp. ch. 18), Connolly et al. (1997: esp. chs. 1-5), Hannay & Bolkestein (1998: esp. chs. 1-5) and Steuten (1998); and it has now led to the proposal of a new version of FG known as Functional Discourse Grammar (FDG), in which Discourse Acts are actually the basic units of linguistic description; see Hengeveld & Mackenzie (2008: 4).

An important point about discourse which has been made over the past few years by a number of authors, notably Kress & van Leeuwen (1996: 39), Norris (2004: 9), O'Halloran (2004a: 1), Scollon & Levine (2004: 3), van Leeuwen (2004: 10) and Baldry & Thibault (2006: 19), and echoed in Connolly (2004: 94), is that it is not purely linguistic in character.[1] Rather, discourse is by nature multimodal. Spoken discourse incorporates non-verbal as well as verbal communication. Non-verbal communication may take the form of gestures, facial expressions, posture and various other phenomena. In written discourse, too, there are non-linguistic aspects to the communication; see for example Machin (2007). For instance, the way in which the text is laid out has

[1] O'Halloran considers the unimodal view of discourse to be "impoverished", while Scollon & Levine describe it as "distorting".

significance. Headings are usually placed on a line of their own, preceding the text that they introduce, and highlighted by some means, such as bold typeface and perhaps centring, while page numbers are set apart at the top or bottom of pages, and thus distinguished from any other numeric content; see further Connolly (2008: 278-282). Moreover, the use of features such as boldface, italicisation or underlining may contribute to the way in which texts are interpreted in relation to emphasis, for example.

Contemporary technology, of course, through the provision of digital multimedia, makes possible the creation of documents that are more highly multimodal than ever. Web pages are the obvious example. These can contain not only text but also still and/or moving images, speech, music, sound-effects and so forth, in more-or-less any combination.

Given, then, that discourse is in general multimodal, what are the implications for Functional *Discourse* Grammar? A possible, conventional answer to this question might run as follows. FDG is a *linguistic* framework and, as such, does not seek to incorporate other modes of communication. Now, there is nothing wrong with imposing such a restriction, and so those who wish to confine their attention to *linguistic* pragmatics are quite free to do so. Having said this, however, we should at least consider how talk and text relate to non-verbal modes of communication, as part of their communicative context. Pragmatics is concerned with the use of language in context, and FDG explicitly provides for a contextual component. For more on the issue of context within FDG, see Connolly (2004, 2007a, 2007b, 2008), Janssen, (2007), Butler (2008b), García Velasco (2008), Keizer (2008), Rijkhoff (2008) and Cornish (2009). Context has also been treated by many authors outside of FDG. Some useful discussions include Auer (1992, 2009), Goodwin & Duranti (1992), Gumperz (1992a, 1992b), Linell (1998: 127-158), Akman & Bazzanella (2003), Bach (2004, 2005), Fetzer (2004, 2007), Givón (2005), Fetzer & Fischer (2007) and Okada (2007).

Alternatively, can we be a little more adventurous? Can we develop a framework for the analysis and description of discourse which encompasses multimodality and which incorporates linguistic discourse as an integrated subcomponent? This would still allow for the purely linguistic treatment of discourse, but would not be limited to linguistic aspects. What are the implications of such a proposal?

The answer to this question depends centrally on the way in which discourse is to be treated within the overall theoretical framework. Consequently, we shall begin by giving some consideration to discourse and its analysis.

2. The treatment of discourse

As van Dijk (1997: 3-4) points out, the term "discourse" has several senses. In its most abstract sense it denotes "discourse in general", and in a slightly less abstract sense it can be applied to broad types of discourse, classified either by domain (e.g. "linguistic discourse") or outlook/ideology, (e.g. "the discourse of liberalism").[2] In a more concrete sense, it can be used to refer to an individual discourse, such as a particular conversation or a particular textbook. An individual discourse is termed a "microdiscourse" by Conley

[2] The term "genre" is often used to mean a discourse-type as classified by domain, though not usually a discourse-type as classified by outlook/ideology. For a recent discussion of genre, see Unger (2006).

& O'Barr (1998: 7), who contrast it with "macrodiscourse" in the more abstract sense found in expressions like "the discourse of liberalism".

However, there is an additional sense, which we can appreciate if we think, for example, of the current debate on global warming. If a debate is a kind of dialogue, and if a dialogue is a kind of discourse, then clearly the global warming debate is a kind of discourse. This is a concrete sense of the term "discourse", as the debate consists in concrete communicative events rather than abstract classificatory types (though of course the debate may be classified in terms of its domain (the interface between meteorology and ecology) and ideology (the conflict between environmentalism and consumerism)). However, it is made up of numerous individual discourses, such as political speeches, newspaper articles, and so on. We may therefore call it a "large-scale discourse" as opposed to an individual discourse.

These different senses of "discourse" form a four-level hierarchy:[3]

(1) (a) Hyper-level: discourse *in abstracto*.
(b) Macro-level: abstract discourse-type.
(c) Meso-level: large-scale discourse.
(d) Micro-level: individual discourse.

Discourse analysis, as developed and practised within linguistics, has been directed principally at micro-level discourse. Both FG and FDG have followed in this tradition. The higher levels in (1) supply (part of) the context of the micro-level discourse, and are therefore associated with what is termed its "discoursal context" in Connolly (2007a: 195, 2007b: 14).[4]

Dik (1997b: ch.18) proposed, in outline form, an eclectic but serviceable framework for the treatment of (micro-)discourse. His reasons for doing so are summarised in Dik (1997b: 409). He envisaged FG as a module within an integrated model of the natural language user (NLU), aimed at explaining how NLUs can communicate through verbal interaction; and he acknowledged that such interaction generally takes place not in isolated clauses but in larger and more complex, coherent units. Consequently (he stated), even if we were able to arrive at an optimal theory of the clause, this would still leave us well short of a theory that explained communicative competence. Hence, in order for FG to satisfy its avowed standards of adequacy, its ultimate goal had to be to provide an account of discourse.

He also pointed out that because clause structure is sensitive to discourse influences, no optimal theory of the isolated clause would be possible. This principle is very much at the heart of FDG as well.

Dik (1997b: 428) characterised the organisation of discourse, at the interpersonal level, in terms of the following hierarchy:

[3] The terminology employed here builds upon the nomenclature of Conley & O'Barr (1998).
[4] Note that the term "discoursal" rather than "discourse" context is being used here. The term "discoursal context" is explicitly intended to denote the *multimodal* environment of a text or fragment.

(2) (a) Discourse event
(b) Turn sequence
(c) Turn
(d) Speech act sequence
(e) Speech act

However, in FDG the two fundamental categories are:

(3) (a) Move
(b) Discourse Act

Thus, the Move is regarded as the highest-ranking unit, since it is believed to be the largest unit pertinent to grammatical analysis; see Hengeveld & Mackenzie (2008: 50). Put another way, FDG is concerned with those layers of the individual discourse that it makes sense to describe in terms of the whole set of linguistic levels of representation (the Interpersonal, the Representational, the Morphosyntactic and the Phonological); and the Move is considered the upper limit of the scope of morphosyntactic and phonological phenomena.

Of course, the validity of recognising higher-ranking units of discourse, such as the Exchange, is not denied; nor is the fact that the Interpersonal and Representational Levels may be seen at these higher-ranking layers. Such units are merely regarded as lying outside the scope of the actual grammar.

Although this is a principled stance to take on the coverage of a grammatical theory, nevertheless it does not negate Dik's view that what is ultimately needed is a theory that encompasses discourse as well.[5] In the present paper, therefore, current FDG will be regarded as a step along the way towards that broader theory, and discourse phenomena that operate above as well as within the Move will be recognised.

In what follows, then, we shall recognise discourse phenomena such as adjacency pairs and rhetorical relations, even though in current FDG adjacency pairs (within Exchanges) lie outside the scope of the grammar, given that they operate higher up the discourse hierarchy than the Move, whereas rhetorical relations (within Moves) do lie within its orbit. Given their pragmatic nature, these phenomena can be seen to pertain to the Interpersonal Level of discourse.

3. Multimodal discourse and semiotics

Given that linguistic discourse analysis is directed at the analysis of spoken and written text in terms of categories such as adjacency pairs and rhetorical relations, its analytical techniques have, obviously, been developed for the analysis of language-based discourse and are well adapted to that purpose. However, as soon as we step outside of language, we exceed the scope of linguistic discourse analysis and find ourselves in need of a more extensive framework than Linguistics can provide.

In Connolly & Phillips (2002) it is argued that a theoretical framework suitable to these more ramified needs is provided by Semiotics. This is because Semiotics is, in principle, broad enough to encompass any systematised method of communication and

[5] This view was echoed in Butler (2008a).

can, therefore, provide an appropriate framework for the handling of multimodal discourse.[6] If we want linguistic discourse, as analysed and described in accordance with the principles of FDG, to be regarded as an integrated component of such an overarching framework, then the implication is that FDG be treated as a branch of semiotic theory. This is consistent with de Saussure's view (1966: 16) that Linguistics is part of Semiotics,[7] and that the general principles of semiotics apply *a fortiori* to Linguistics. Furthermore, the idea of a semiotic approach to Functional Linguistics is, of course, already familiar from Systemic Functional Grammar; see, for instance Halliday (1978), who speaks of language as "social semiotic".[8]

Semiotics is based around the notion of "signs", which are characterised by the fact that they designate entities, processes or attributes other than themselves, a property which makes them very useful tools for the purposes of communication in contexts where they can be produced and interpreted by those who share a knowledge of the sign system involved. The communicative use of signs is a central concern in Semiotics, which thus has a clear functional character.

A sign, as has just been noted, involves the pairing of a form with a meaning. In Saussurean terms, the former is known as the "signifier" and the latter as the "signified".[9]

Following the tradition of Peirce (1931-1958), signs are generally classified into three categories:[10]

(4) (a) Symbolic signs.
(b) Iconic signs.
(c) Indexical signs.

Symbols involve an arbitrary, conventional relationship between signifier and signified. An example of a symbol is the word "river". On the other hand, a river could be denoted on a map by means of a blue line, drawn in such a way as to reflect or mimic the course of a river. This depiction of the river would be an icon, as it bears not an arbitrary relationship but a perceptible physical resemblance to what it signifies.[11] As for indices, these bear a natural (often causal) relation to what they signify. For instance, a puddle is an index if it is taken as a sign that there has recently been a shower of rain.

---

[6] The six-level semiotic framework proposed by Stamper (1991) is particularly attractive. See further Connolly (2002).

[7] Actually, de Saussure uses the term Semiology rather than Semiotics. He also envisages a somewhat narrower view of the field than is current today, restricting it to the study of what in Peirce's terminology would be called symbols.

[8] See in particular Halliday (1978: 123). Hodge & Kress (1988: 1-2) distinguish between what they term "mainstream semiotics", with its emphasis on structures and codes, and "social semiotics", which stresses function and interactivity in the communication process.

[9] Peirce uses an alternative nomenclature, but this need not concern us here.

[10] This classification is not absolute. It is possible for a sign to belong to more than one category at the same time. However, in practice it is usually possible to identify the primary or dominant category to which a particular sign belongs.

[11] Purchase (1998: 9, 1999: 249) subclassifies icons into two types: "concrete" and "abstract". A concrete icon is a relatively faithful representation of what it denotes. An example is a picture of a boat, since the visual representation looks very like the actual boat that it depicts. (Of course, it will probably not be *identical* to a view of the real boat, as it will probably be smaller.) An abstract icon bears less fidelity to the physical appearance of what it denotes. The representation of a river by means of a line on a map is an example of this type.

4. The analysis of multimodal discourse

Let us, then, explore the feasibility of carrying out semiotic discourse analysis on multimodal material in a manner compatible with the discourse-oriented approach embodied in FDG and cognisant of the scope envisaged in Dik (1997b: 409-441). We shall consider two scenarios:

(5) (a) Multimodal discourse realised by means of a combination of spoken language and non-verbal gestures.
(b) Multimodal discourse realised by means of a combination of written language and images (as, for instance, in a typical web page).

We shall see that both scenarios afford some encouragement in our enterprise.[12]

As a preliminary step, however, we need to make some brief remarks about the semiotics of non-verbal gestures and of (still) images. Gestures can be symbolic, iconic or indexical. For instance, in Northern Europe, nodding the head can symbolise agreement or affirmation; holding the flat palm of the hand downwards at a certain distance from the ground can indicate iconically the height of a child; and instinctively throwing up one's hands in horror can be an index of an unpleasant shock.

Images can function either as symbols or as icons. An example of an image serving as an icon has already been given. An example of an image acting as a symbol can be seen in the representation of a diode in an electrical circuit: see Figure 1. This is not an iconic representation, as diodes are not really that shape.[13]

Figure 1: Representation of a diode

As has been pointed out by Worth (1981: 174), pictures differ from language in that they do not admit of a propositional-type analysis; we do not normally say of a pictorial representation that it makes a factually true or false statement. On the other hand, we may say that it does or does not correspond with reality, as is the case with states-of-affairs represented in language.

[12] In order to keep the discussion within manageable bounds, we shall have nothing to say here about the semiotics of music, sound effects or moving images.
[13] In addition to their symbolic or iconic value, images can also contain indexical information. For example, discoloration can be a natural sign of age in a photograph.

4.1 Talk and non-verbal communication

To begin with, consider the following purely verbal Exchange:

(6) JIM*:* Where's the car?
KAREN*:* There.

A linguistic discourse analysis of this Exchange would reveal information such as the following. The Exchange consists of a Move by Jim followed by a Move by Karen. Jim's Move consists of a single, interrogative Discourse Act, while Karen's Move comprises a single declarative Discourse Act. These two Moves form an adjacency pair comprising a question followed by an answer.

Now suppose that the Exchange had, instead, taken the following form:

(7) JIM*:* Where's the car?
KAREN*:* *Points at the vehicle.*

This time, Karen's response is non-verbal. However, the meaning of her gestural response is the same as if it had been verbalised as *There*. As a Discourse Act, declarative in character, it similarly functions as the second Move in an Exchange, constituting the answer to a question within an adjacency pair.

This example illustrates the possibility of extending the methods of linguistic discourse analysis to a multimodal dialogue. However, because the dialogue combines two different sign-systems, the discourse analysis is, of course, semiotic rather than purely linguistic.

Let us exemplify this idea further. Consider the following Exchange:

(8) LOUISE*:* Are you comfortable?
MIKE*:* Yes. *(Spoken slowly, without tension, and with a low-falling tone.)*

This verbal Exchange would be analysed from the discourse point-of-view in the same way as (6), with the additional information that (as indicated by the prosodics) Mike shows a favourable, relaxed attitude.

Now imagine that the Exchange had, instead, been as follows:

(9) LOUISE*:* Are you comfortable?
MIKE*:* *Nods and smiles happily.*

Mike's gesture of nodding fulfils the same discourse function as the verbal response *Yes* would have done. In addition, the smile indicates a favourable, relaxed attitude. Hence, the multimodal Exchange may be given a semiotic discourse analysis similar to example (7), but again supplemented by the same attitudinal information as in (8).

In the previous two examples, it has proved straightforward to extend the concepts of linguistic discourse analysis to multimodal material. Next, let us take a slightly different example:

(10) NEIL: I've found your pen.
Shows Olivia the pen.

Assume for the moment that Neil does not show Olivia the pen until he has finished his utterance. In that case, (10) can be described as a discourse Move consisting of two Acts, declarative in character, the first verbal and the second gestural. Moreover, there is a rhetorical relationship between the two Acts, such that the second constitutes "evidence" for the first.[14]

Suppose now that Neil showed Olivia the pen either before or at the same time as producing the utterance. Apart from the actual sequencing of the two communicative Acts, the discourse analysis would be unchanged. Neither the declarative nature of the Acts nor the rhetorical relationship between them would be altered.

This example illustrates an important point about the semiotic discourse analysis of multimodal material, namely that discourse relationships between communicative Acts are not necessarily sequential. However, this fact does not vitiate the analysis; it simply calls for an extra level of flexibility in the analysis compared with purely linguistic discourse analysis.[15]

Let us move on to another example:

(11) PAT: This is what the area looked like before they built the factory.
Shows Robert a picture of the area as it used to be.

The semiotic discourse analysis of this discourse Move is similar to that of (10), except that the rhetorical relationship between the two Acts is different. In (11) the second, gestural Act is an "elaboration" of the first, verbal Act. This elaboration is made possible (in the present example) because the gesture involves the presentation of an iconic sign conveying further information, namely the landscape that it depicts.

4.2 Text and images

Let us now move on to consider multimodal discourse conducted by means of a combination of written language and images. One issue that arises again here is the frequent absence of linearity in the organisation of the discourse. For instance, a single page may contain one or more pieces of writing and one or more images, all being visible at the same time.[16] Moreover, a single page may contain more than one thread of

[14] The treatment of rhetorical relationships is here based on Rhetorical Structure Theory; see Mann & Thompson (1987), Gulla (1997) and Verschueren (1999: 140-143).

[15] Relationships that are not of a rigid sequential nature are, of course, by no means a novel concept in Linguistics. We are familiar, for instance, with prosodic phenomena, such as intonation, having a structure that is, in effect, superimposed upon the syntax of an utterance. With particular reference to the statement of a claim and the statement of evidence to support that claim, we may note that if both these statements are expressed in language, then there is no fixed order that the two statements must follow. For instance, we could say either "The grass is glistening white. There has been a frost overnight." (with the statement of the evidence preceding that of the claim which it supports) or "There has been a frost overnight. The grass is glistening white." (with the reverse ordering).

[16] In some kinds of discourse, notably those which use web sites as their platform, there is scope for additional structural complications. The obvious example is the non-linearity attributable to the use of hyperlinks, which often result in the absence of any one predictable path through a document. (This can

discourse. For example, a news page may feature several stories, each with its own pictorial illustrations.

In order to help us analyse such a page, it may help if we make the following distinctions. Firstly, let us recognise the conventional dichotomy:

(12) (a) Foreground material, which can be composed of text, images or both.
(b) Background, against which the foreground material is set.

In the present context, the difference between the two types is, in essence, that foreground material contributes to the content of the page, whereas background does not. For instance, in a web page consisting of written text appearing against a backdrop of mottled "wallpaper", the text constitutes the foreground material, whereas the "wallpaper" forms the background. Secondly, let us subclassify foreground images in terms of the following categories:

(13) (a) Consolidated.
(b) Perfunctory.

A "consolidated" image is one that is integrated into at least one thread of discourse, which in some way it serves to illustrate. Any other foreground image is "perfunctory". Perfunctory images may be included in cases where the sheer presence of pictures or other content-bearing graphical material is felt by the designer to be important to the aesthetic appearance of a page, even though the images are not integrated into the text. Another example of a perfunctory image is a logo at the head of a page, which serves to lend it official status, but is not fully integrated into any thread of discourse. However, for the remainder of this paper we shall be concerned almost exclusively with consolidated images.

Fiske (1990: 103-6) points out that images can sometimes signify more than what they literally depict. For instance, a picture of a group of police officers may (in an appropriate context) stand for, or represent, the police in general. Let us develop this point a little, adding some examples of our own.

As a starting point, consider a full-length picture of Martin Luther King, accompanied by the text "Martin Luther King". In this example, the linguistic text and the picture share the same denotation without the interposition of any figurative interpretation. However, if the text "Martin Luther King" were illustrated by a picture of just King's face, then we would have a case of synecdoche (a part representing the whole), with the literal depiction of the image (the face) figuratively representing the entire man denoted by the accompanying text.

A slightly different example may be found in a book sales-catalogue in which the details of each book are accompanied by a picture of the book concerned. In this case, we have a type-token relationship. The image literally portrays an individual instance (or token) of the book concerned, whereas the accompanying text generalises across all the individual physical copies to the type, denoted by the book title, which the individual

---

happen whether the content is expressed by means of language text alone or by means of multimodal communication.) Another possibility is the embedding of one discourse within another, as for instance when a video forms part of a web page. However, complications such as these will not be addressed in the present paper.

tokens instantiate. For a more complex example, imagine now that the picture of the book appears on a page headed "gifts", which also contains pictures of a CD, a game and a video. Here, as well as the type-token relationship between the individual book depicted and the generic book, we have a case of class inclusion, since a book is (in this example) included within the broader category of gift, denoted by the page heading.

It is interesting to pause for a moment and note the kinds of relationships involved in these examples; they include sameness of denotation, the part-whole relationship and class inclusion. These are among the most important relationships recognised within the quite separate field of lexical semantics, where they pertain respectively to synonymy, meronymy and hyponymy; cf. Cruse (2000: 150-160) and Royce (2007: 70).

Another commonality between images and language text is that, according to Kress & van Leeuwen (1996: 127-9), images can, at least to some extent, embody communicative Acts; see also O'Toole (1994). Kress & van Leeuwen refer to Halliday's (1994: 69) classification of speech functions in relation to their expected responses:

(14) (a) Giving information.
(b) Giving goods/services.
(c) Demanding information.
(d) Demanding goods/services.

Most images provide information of some kind, such as the physical appearance of a person or object. However, Kress & van Leeuwen suggest that some images demand a particular kind of (imaginary) service, namely that the viewer should agree to engage in a make-believe, temporary social relationship, for instance with an seductive-looking lady depicted as smiling and beckoning.

As well as embodying communicative Acts, images may fulfil rhetorical functions, as demonstrated by Kong (2006) and Matthiessen (2007: 33-36). For example, a portrait (whole or part) alongside a mention of the name of the person concerned may well function as an elaboration, providing information which is additional to that in the text and which is not readily conveyed textually.[17] Alternatively, a picture of a book in a page headed "gifts" serves as an exemplification of that term (rather than an elaboration with the aim of showing us what a book looks like). Both elaboration and exemplification are textbook examples of rhetorical relations, whose purpose is to link constituents of a text together in a coherent, comprehensible and identifiable manner, in order to progress the discourse; see Verschueren (1999: 140-142).

Further examples may be found on the website of the UK Natural History Museum (NHM). On a page entitled "Butterflies", NHM (2010b), the written text indicates that there are a number of common types of butterfly, and three examples are presented by means of images, labelled "clouded yellow", "red admiral" and "speckled wood", respectively. On another page, NHM (2010a), the written text offers guidance on the recognition of different types of British bumblebee, and makes reference to various shades of yellow or orange hair-colours that such bees may exhibit. These shades are illustrated with the help of an image, which elaborates upon the text giving a more exact idea of the shades in question than could be achieved through verbal means alone.

[17] Admittedly, the appearance might be so familiar that the depiction is in practice superfluous to the reader, but this is by no means always the case.

5. Relating linguistic and non-linguistic modes

5.1 Multimodal discourse and context

In order to accommodate multimodality within the approach to discourse adopted in FDG, it is necessary to postulate a broader discourse framework, into which FDG can fit as seamlessly as possible. This is what we have attempted to do in the present paper. In pursuance of the seamless fit, we have employed discourse-analytic constructs, in particular adjacency pairs and rhetorical relations, to bridge the divide between linguistic and non-linguistic modes.

In a multimodal individual discourse, each contributory mode supplies (part of) the discoursal context of the others. Thus, non-verbal modes belong to the context of the verbal, and vice versa. Of course, a full multimodal discourse analysis would need to include a semiotic analysis of all the relevant modes, both verbal and non-verbal. From a purely linguistic point of view, this kind of analysis would have the potential to provide semiotically-based descriptions of the discoursal context that were highly compatible with the description of language itself. Let us now give some further consideration to such a possibility.

5.2 Analysing discoursal context

In Connolly (2007: 14), a distinction is drawn between the following:

(15) (a) Discoursal context.
(b) Situational context.

If T represents a text or text-fragment that constitutes the focus of a discourse analysis, then the situational context lies outside of T or any other text, whereas the discoursal context consists in one or both of the following:

(16) (a) Co-text.
(b) Inter-text.

Co-text is supplied from within T, whereas inter-text is supplied from other discourse(s) than T.

Discoursal context is made up of signs drawn from systems such as the following, some of which were mentioned earlier:

(17) (a) Natural (human) languages.
(b) Formal languages, for instance computer programming languages such as Java.
(c) Nonverbal systems such as gesture,[18] facial expression and posture.
(d) Non-linguistic vocal effects.
(e) Layout and typography of written texts.
(f) Images, including pictures and line-drawings such as diagrams.[19]
(g) Music.
(h) Sound effects.[20]

Signs may, of course, be composites, drawing on more than one of these systems, for example labelled diagrams.

The analysis of semiotic systems other than natural language, in a manner that is nevertheless compatible with the analysis of natural language, has been pioneered by various authors in the Systemic Functional Linguistic (SFL) tradition, including O'Toole (1994, 2004), Kress & van Leeuwen (1996), Lemke (1998), O'Halloran (2004b,c), Baldry & Thibault (2006), Machin (2007) and Matthiessen (2007). A major thrust of this work has been the generalisation of Halliday's (1994: 179) metafunctional categories of "ideational", "interpersonal" and "textual" meaning into semiotic modes other than natural language. Terminology varies for the generalised categories, but this need not concern us here. The key point is that the SFL categories "ideational" and "interpersonal" correspond sufficiently closely to the FDG categories of "representational" and "interpersonal" to make it worth exploring the possibility that a generalisation of these FDG terms to non-linguistic modes would be equally possible.

Let us consider again some examples given earlier. Firstly, a gesture of pointing to indicate the position of a vehicle may be analysed both from a representational point of view, as shown in (18a), and from an interpersonal point of view, as shown in (18b):

(18) (a) The gesture serves to signify (or represent) the location within the immediate physical situational context where the vehicle is situated at the time.
(b) The gesture constitutes a declarative (or informative) Act, having the communicative content just described and involving the gesturer and the addressee as discourse-participants. The Act serves to make reference to the location concerned, but does not ascribe any specific property or attribute to that location.

Secondly, a diode symbol in a circuit diagram may likewise be analysed from a representational point of view, as shown in (19a), and from an interpersonal point of view, as shown in (19b):

[18] See further Lascarides & Stone (2009).
[19] Cf. Bretz (1971: 68).
[20] Cf. O'Halloran (2004c: 118).

(19) (a) (i) The diagram signifies the electrical circuit,
(ii) The symbol signifies a diode that constitutes a component of the circuit.
(b) (i) The diagram constitutes a declarative Act, having the communicative content described in (19a.i), and involving the author and the addressees as discourse-participants.
(ii) The symbol serves to make reference to the component concerned, and to ascribe to it the property of being a diode as opposed to any other type of component or connector.

Thirdly, an image showing a range of bumblebee hair-colours may similarly be analysed from a representational point of view, as shown in (20a), and from an interpersonal point of view, as shown in (20b):

(20) (a) The image signifies the bumblebee hair-colours.
(b) The image constitutes a declarative Act, having the communicative content just described and involving the web-page author and the addressees as discourse-participants. The Act serves to make reference to the bumblebee hair-colours concerned, and to ascribe to them a range of hues in the orange-yellow spectrum, as opposed to any other hues.

The analyses in (18-20) are phrased somewhat informally, in the hope of bringing out the parallels with natural language, while not imposing an unwarrantedly detailed linguistic description upon non-linguistic communication. Nevertheless, there is scope for a more formal notation, if and when agreement is reached as to what exactly it should contain.

As for the lower levels of structure, natural language is unique in having a both Morphosyntactic Level and a Phonological level. Nevertheless, other modes also have a level of form, of some kind. For instance, generally speaking, gestures may be characterised in terms of shape and movement; diagram-components in terms of shape and tone or colour, as well as graphical devices such as lines or arrows connecting them to other components; and pictures in terms of their composition.[21] If non-linguistic modes are amenable to semiotic analysis at the level of form, on the basis of their internal structure(s) and constituents, then some degree of compatibility with the linguistic description of natural language may well be achievable.

Needless to say, much further work is needed on the semiotic analysis of discoursal context, in order to provide a clearer idea of what is involved, and of the extent to which the idea is feasible. However, such research will bring a twofold benefit. It will not only enhance our ability to describe the discoursal context of natural language, but it will also, at the same time, develop our ability to carry out multimodal (rather than just linguistic) discourse analysis, if such is our desire.

[21] See further Kress & van Leeuwen (1996: 181-229).

6. Conclusion

The conclusions that may be drawn from this paper are as follows. First of all, a multimodal approach to discourse analysis that is compatible with FDG seems feasible, though it will need further exploration and development. This multimodal discourse analysis is applicable both to combinations of speech and gestural communication and to combinations of written text and images, at least. Furthermore, some concepts normally associated with lexical semantics, namely synonymy, meronymy and hyponymy, have turned out to have parallels in multimodal discourse.

Of course, there is no implication that linguists should be forced into multimodal discourse analysis. However, the option is certainly available for those who are interested, while on the other hand, multimodal discourse analysis also offers a handle on the description of discoursal context for the benefit of those who wish to confine themselves to linguistics but recognise the fact (embodied in the FDG framework) that language function cannot be fully accounted for in isolation from contextual considerations. In fact, the four-level discourse framework proposed in section 2 above may be seen as one contribution to the analysis of discoursal context.

Hengeveld & Mackenzie (2008: 1) echo the view expressed by Dik (1997a: 4) that a theory of grammar should form part of a broader "pragmatic theory of verbal interaction". Dik maintained that although a grammar, like F(D)G, should be distinguished from a theory of verbal interaction, nevertheless it should be constructed and formulated in such a manner as to permit it to be incorporated into the broader pragmatic theory. Given that human discourse is typically multimodal, it may be seen that the present paper represents an attempt to move towards such an integrated framework.

Acknowledgements

Thanks are due to fellow-participants in the 13th International Conference on FG (University of Westminster, London, UK), to Paul Wymer, and to an anonymous referee, for comments and discussion that have helped to improve this paper.

References

Akman, Varol & Carla Bazzanella (eds.). 2003. The complexity of context. Special issue of *Journal of Pragmatics* 35 (3).

Auer, Peter. 1992. Introduction: John Gumperz' approach to contextualization. In Peter Auer & Aldo di Luzio (eds.), *The contextualization of language*, 1-37. Amsterdam: Benjamins.

Auer, Peter. 2009. Context and contextualisation. In Jef Vershueren & Jan-Ola Östman (eds.), *Key notions for pragmatics*, 86-101. Amsterdam: Benjamins**.**

Bach, Kent. 2004. Minding the gap. In Claudia Bianchi (ed.), *The semantics/pragmatics distinction*, 27-43. Stanford: CSLI.

Bach, Kent. 2005. Context *ex machina*. In Zoltán Gendler Szabó (ed.), *Semantics versus pragmatics*, 15-44. Oxford: Clarendon Press.

Baldry, Anthony & Paul J. Thibault. 2006. *Multimodal transcription and text analysis: A multimedia toolkit and coursebook with associated on-line course.* London: Equinox.

Bretz, Rudolf. 1971. *A taxonomy of communication media*. Englewood cliffs, NJ: Educational Technology Publications.

Butler, Christopher S. 2008a. Back to basics: A reappraisal of the functional enterprise with particular reference to Functional Discourse Grammar. Paper presented at the 13th International Conference on Functional Grammar, University of Westminster, London, UK, 3-6 September, 2008.

Butler, Christopher S. 2008b. Interpersonal meaning in the noun phrase. In Daniel García Velasco & Jan Rijkhoff (eds.), *The noun phrase in Functional Discourse Grammar*, 221-261. Berlin: Mouton de Gruyter.

Conley, John M. & William M. O'Barr. 1998. *Just words: Law, language and power*. Chicago: University of Chicago Press.

Connolly, John H. 2002. Accommodating natural language within the Organisational Semiotic framework. In Kecheng Liu, Rodney J. Clarke, Peter B. Andersen & Ronald K. Stamper (eds.), *Coordination and communication using signs*, 3-21. Boston: Kluwer.

Connolly, John H. 2004. The question of discourse representation in Functional Discourse Grammar. In J. Lachlan Mackenzie & María de los Ángeles Gómez-González (eds.), *A new architecture for Functional Grammar*, 89-116. Berlin: Mouton de Gruyter.

Connolly, John H. 2007a. Mental context and the expression of terms within the English clause: An approach based on Functional Discourse Grammar. In Mike Hannay & Gerard J. Steen (eds.), *Structural-functional studies in English grammar*, 193-208. Amsterdam: Benjamins.

Connolly, John H. 2007b. Context in Functional Discourse Grammar. *Alfa: Revista de Lingüística* 51 (2). 11-33.

Connolly, John H. 2008. Freestanding noun phrases within documents: A pragmatic approach based on Functional Discourse Grammar. In Daniel García Velasco & Jan Rijkhoff (eds.), *The noun phrase in Functional Discourse Grammar*, 263-285. Berlin: Mouton de Gruyter.

Connolly, John H. & Iain W. Phillips. 2002. Semiotics and the theoretical foundations of multimedia. *Semiotica* 141. 169-184.

Connolly, John H., Roel M. Vismans, Christopher S. Butler & Richard A. Gatward (eds.). 1997. *Discourse and pragmatics in Functional Grammar.* Berlin: Mouton de Gruyter.

Cornish, Francis. 2009. *Text* and *discourse* as *context*: discourse anaphora and the FDG contextual component. In Evelien Keizer & Gerry Wanders (eds.), *Web Papers in Functional Discourse Grammar* 82, 97-115.

Cruse, D. Alan. 2000. *Meaning in language: An introduction to semantics and pragmatics*. Oxford: Oxford University Press.

Dijk, Teun A. van. 1997. The study of discourse. In Teun A van Dijk (ed.), *Discourse as structure and process*, 1-34. London: Sage.

Dik, Simon C., ed. Kees Hengeveld. 1997a. *The theory of Functional Grammar, part 1: The structure of the clause.* Berlin: Mouton de Gruyter.

Dik, Simon C., ed. Kees Hengeveld. 1997b. *The theory of Functional Grammar, part 2: Complex and derived constructions.* Berlin: Mouton de Gruyter.

Fetzer, Anita. 2004. *Recontextualising context: Grammaticality meets appropriateness*. Amsterdam: Benjamins.
Fetzer, Anita. 2007. Context, contexts and appropriateness. In Anita Fetzer (ed.), *Context and appropriateness: Micro meets macro*, 3-27. Amsterdam: Benjamins.
Fetzer, Anita & Kerstin Fischer. 2007. Introduction. In Anita Fetzer & Kerstin Fischer (eds.), *Lexical markers of common grounds*, 1-13. Amsterdam: Benjamins.
Fiske, John. 1990. *Introduction to communication studies.* 2nd edn. London: Routledge.
García Velasco, Daniel. 2008. Functional Discourse Grammar and extraction from (complex) noun phrases. In Daniel García Velasco & Jan Rijkhoff (eds.), *The noun phrase in Functional Discourse Grammar*, 321-363. Berlin: Mouton de Gruyter.
Givón, T. 2005. *Context as other minds: The pragmatics of sociality, cognition and communication*. Amsterdam: Benjamins,
Goodwin, Charles & Alessandro Duranti. 1992. Rethinking context: An introduction. In Alessandro Duranti & Charles Goodwin (eds.), *Rethinking context: Language as an interactive phenomenon*, 1-42. Cambridge: Cambridge University Press.
Gulla, John Atle. 1997. Combining Functional Grammar and Rhetorical Structure Theory for discourse representation. In John H., Connolly, Roel M. Vismans, Christopher S. Butler &, Richard A. Gatward (eds.), *Discourse and pragmatics in Functional Grammar,* 75-89. Berlin: Mouton de Gruyter.
Gumperz, John J. 1992a. Contextualisation and understanding. In Alessandro Duranti & Charles Goodwin (eds.), *Rethinking context: Language as an interactive phenomenon*, 229-254. Cambridge: Cambridge University Press.
Gumperz, John J. 1992b. Contextualisation revisited. In Peter Auer & Aldo di Luzio (eds.), *The contextualization of language*, 39-53. Amsterdam: Benjamins.
Halliday, Michael A.K. 1978. *Language as social semiotic: The social interpretation of language and meaning.* London: Arnold.
Halliday, Michael A.K. 1994. *An introduction to Functional Grammar,* 2nd edn. London: Arnold.
Hannay, Mike & A. Machtelt Bolkestein (eds.). 1998. *Functional Grammar and verbal interaction*. Amsterdam: Benjamins.
Hengeveld, Kees & J. Lachlan Mackenzie. 2008. *Functional Discourse Grammar: A typologically-based theory of language structure*. Oxford: Oxford University Press.
Hodge, Robert & Gunther Kress. 1988. *Social semiotics.* Cambridge: Polity Press.
Janssen, Theo A.J.M. 2007. A speaker/hearer-based grammar: the case of possessives and compounds. In Mike Hannay & Gerard J. Steen (eds.), *Structural-functional studies in English grammar,* 353-387. Amsterdam: Benjamins.
Keizer, Evelien. 2008. Reference and ascription in Functional Discourse Grammar: An inventory of problems and some possible solutions. In Daniel García Velasco & Jan Rijkhoff (eds.), *The noun phrase in Functional Discourse Grammar*, 181-219. Berlin: Mouton de Gruyter.
Kong, Kenneth. 2006. A taxonomy of the discourse relations between words and visual. *Information Design Journal* 14. 207-230.
Kress, Gunther & Theo van Leeuwen. 1996. *Reading images: The grammar of visual design.* London: Routledge.
Lascarides, Alex L. & Stone, Matthew. 2009. A formal semantic analysis of gesture. *Journal of Semantics* 26, 393-449.

Leeuwen, Theo van. 2004. Ten reasons why linguists should pay attention to visual communication. In Philip Levine & Ronald Scollon (eds.), *Discourse and technology: Multimodal discourse analysis,* 7-19. Washington, DC: Georgetown University Press.

Lemke, Jay. 1998. Multiplying media: visual and verbal semiotics in scientific text. In James R. Martin & Robert Veel (eds.), *Reading science: Critical and functional perspectives on discourses of science,* 87-113. London: Routledge.

Linell, Per. 1998. *Approaching dialogue: Talk, interaction and context in dialogical perspectives*. Amsterdam: Benjamins.

Machin, David. 2007. *Introduction to multimodal analysis.* London: Hodder Arnold.

Mann, William C. & Sandra A. Thompson. 1987. Rhetorical Structure Theory: A framework for the analysis of texts. *Papers in Pragmatics* 1. 79-105.

Matthiessen, Christian M.I.M. 2007. The multimodal page: A systemic functional exploration. In Terry D. Royce & Wendy D. Bowcher (eds.), *New directions in the analysis of multimodal discourse,* 1-62. Mahwah, NJ: Erlbaum.

NHM. 2010a (accessed February 15). *British bumblebee identification guide*. Available at: http://www.nhm.ac.uk/nature-online/life/insects-spiders/identification-guides-and-keys/bumblebees/index.html/.

NHM. 2010b (accessed February 15). *Butterflies*. Available at: http://www.nhm.ac.uk/nature-online/life/insects-spiders/bug-forum/?q=image/tid/26/.

Norris, Sigrid. 2004. *Analysing multimodal interaction: A methodological framework.* New York: Routledge.

O'Halloran, Kay L. 2004a. Introduction. In Kay L. O'Halloran (ed.), *Multimodal discourse analysis: Systemic Functional perspectives,* 1-7. London: Continuum.

O'Halloran, Kay L. 2004b. On the effectiveness of mathematics. In Eija Ventola, Cassily Charles and Martin Kaltenbacher (eds.), *Perspectives on Multimodality,* 91-117. London: Continuum.

O'Halloran, Kay L. 2004c. Visual semiosis in film. In Kay L. O'Halloran (ed.), *Multimodal discourse analysis: Systemic Functional perspectives,* 109-130. London: Continuum.

Okada, Moeko. 2007. Whose common ground? A misunderstanding caused by incorrect interpretations of the lexical markers of common ground. In Anita Fetzer & Kerstin Fischer (eds.), *Lexical markers of common grounds,*183-194. Amsterdam: Benjamins.

O'Toole, L. Michael. 1994. *The language of displayed art.* London: Leicester University Press.

O'Toole, L. Michael. 2004. Opera ludentes: The Sydney opera House at Work and Play. In Kay L. O'Halloran (ed.), *Multimodal discourse analysis: Systemic Functional perspectives,* 11-27. London: Continuum.

Peirce, Charles S., ed. Charles Hartshorne & Paul Weiss. 1931-1958. *Collected papers,* vols. 1-8. Cambridge, MA: Harvard University Press.

Purchase, Helen C. 1998. Defining multimedia. *IEEE Multimedia* 5 (1). 58-65.

Purchase, Helen C. 1999. A semiotic definition of multimedia communication. *Semiotica* 123. 247-259.

Rijkhoff, Jan. 2008. Layers, levels and contexts in Functional Discourse Grammar. In Daniel García Velasco & Jan Rijkhoff (eds.), *The noun phrase in Functional Discourse Grammar,* 63-115. Berlin: Mouton de Gruyter.

Royce, Terry D. 2007. Intersemiotic complementarity: A framework for multimodal discourse analysis. In Terry D. Royce & Wendy D. Bowcher (eds.), *New directions in the analysis of multimodal discourse*, 63-109. Mahwah, NJ: Erlbaum.

Saussure, Ferdinand de, tr. Wade Baskin. 1966. *Course in general linguistics*. New York: McGraw-Hill.

Scollon, Ronald & Philip Levine. 2004. Multimodal discourse analysis as the confluence of discourse and technology. In Philip Levine & Ronald Scollon (eds.), *Discourse and technology: Multimodal discourse analysis*, 1-6. Washington, DC: Georgetown University Press.

Stamper, Ronald K. 1991. The semiotic framework for Information Systems research. In Hans-Erik Nissen, Heinz K. Klein & Rudy Hirschheim (eds.), *Information systems research: Contemporary approaches and emergent traditions*, 515-527. Amsterdam: North-Holland.

Steuten, Ans A.G. 1998. *A contribution to the linguistic analysis of business conversations within the Language/Action Perspective*. Delft: Steuten.

Verschueren, Jef. 1999. *Understanding pragmatics.* London: Arnold.

Unger, Christoph. 2006. *Genre, relevance and global coherence: The pragmatics of discourse type*. Basingstoke: Palgrave Macmillan.

Worth, Sol, ed. Larry Gross. 1981. *Studying visual communication.* Philadelphia: University of Pennsylvania Press.