How to say things without words: Wordless instructions as texts

David Banks
How to say things without words: Wordless instructions as texts

David Banks

Introduction

[...] ‘and what is the use of a book,’ thought Alice, ‘without pictures’
[...] Lewis Carroll, Alice’s Adventures in Wonderland, 1865

1 In this article I wish to return to a subject which I treated briefly a short time ago (Banks 1998). It is common nowadays to find household equipment, particularly furniture, sold in kit form. The purchaser is required to assemble these items, and it is the document provided as assembly instructions which I would like to consider here. These documents contain few or no words, but it will be my hypothesis that they have numerous characteristics in common with written texts, to such an extent that it is reasonable to say that these wordless documents are texts themselves. I shall first look at some of the literature on the relationship between word and image. I will then present that part of the systemic functional model which I wish to use for purposes of analysis. The model will be used to analyse the semantic functioning of a wordless document, showing that its functions are those of a written text.

2 The relationship between verbal and non-verbal elements in written text has received considerable attention in recent years. Without attempting to be comprehensive or exhaustive, one might mention the work of Miller (1998), who compared research articles and their popularizations; he noted that images fulfil different functions in thematic structure in these two sub-genres. Sionis (1997a, 1997b, 1998) looked at the writing of native and non-native speakers in engineering and computer science research articles. He distinguished two types of PVM (partially verbal materials): those that are predominantly verbal, like equations, flow-charts, tree-diagrams, tables, and pie-charts; and those that are predominantly visual like photographs, drawings, graphs, and maps. Palmer and
Posteguillo (1998) studied graphs, tables, tree diagrams and maps in research articles finding differences between those that appeared in journals, collections and conference proceedings. Busch-Lauer (1998a, 1998b) analysed medical research articles, looking particularly at figures, bar charts, graphs and images and their pedagogical implications. Johns (1998) looked at the use of images in student writing in the area of macroeconomics. Myers (1995) considered the pedagogical implications of non-verbal elements in a biology textbook. There have also been at least two major contributions to this field. First, O’Toole’s (1994) study of the fine arts (painting, sculpture and architecture), treating them as semiotic systems parallel to texts, and Kress and van Leeuwen’s (1996) study of images and their use in texts, in fine and applied arts, including advertisements and newspaper layout. The latter, in particular, suggest a schema for visual layout which comprises a vertical axis with ‘ideal’ at the top and ‘real’ at the bottom, and a horizontal axis with ‘given’ to the left and ‘new’ to the right.

![Kress and van Leeuwen’s schema (1996)](image)

**Figure 1. Kress and van Leeuwen’s schema (1996)**

---

**Systemic Functional Linguistics**

The semantic component of the systemic functional model of language comprises three metafunctions: ideational, interpersonal and textual (Halliday 1994, Halliday & Matthiessen 1999, or one of the numerous introductions, Bloor & Bloor 1995, Downing & Locke 1992, Eggins 1994, Lock 1996, Thompson 1996). The ideational metafunction concerns that part of meaning which is representational. It is probably this metafunction which corresponds to the general public’s notion of “meaning”, in that it is the content of the message. This is expressed mainly through a process, with which one or more participants are usually associated, and in some cases circumstances. The interpersonal metafunction deals with the relationships associated with the speaker; these may be between the speaker and his hearer(s) and/or between the speaker and his message. The textual metafunction is that aspect of meaning which makes a string of words or clauses into a text. At the level of clause this includes phenomena such as thematic and information structure, and beyond the clause, cohesion.

In order to illustrate this I shall use a clause which occurred as a headline on the front page of the Sport supplement of *The Observer*, 30 Jan 2000:

> A million words after his arrival, Iron Mike beats Julius Francis in just four minutes.

In terms of its ideational content this clause can be analysed as in Figure 2.

![Figure 2. Ideational content](image)

---

**Figure 2. Ideational content**

<table>
<thead>
<tr>
<th>Circ: time</th>
<th>Actor</th>
<th>Pro: Mat</th>
<th>Affected</th>
<th>Circ: duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>A million ... arrival</td>
<td>Iron Mike</td>
<td>beats</td>
<td>Julius Francis</td>
<td>in ... four minutes</td>
</tr>
</tbody>
</table>
The process is encoded in the verb beats. As a process which takes place in the physical world, this constitutes a Material Process. It has two participants, Iron Mike, which encodes the conscious causer of the process, termed the Actor, and Julius Francis, which encodes the element which is modified by the process, and known as the Affected. This clause also comprises two circumstantial elements: a million words after his arrival, constitutes the circumstances of time, while in just four minutes gives the duration of the process.

Part of the interpersonal metafunction can be analysed in the following terms (see fig. 3).

Figure 3. Interpersonal function

<table>
<thead>
<tr>
<th>A million ... arrival</th>
<th>Iron Mike</th>
<th>beats</th>
<th>Julius Francis</th>
<th>in ... four minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residue</td>
<td>Mood</td>
<td>Residue</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mood choices form part of the interpersonal metafunction because they impose roles on the speaker and hearer. For example, the selection of interrogative mood imposes the role of questioner on the speaker, and that of potential answerer on the hearer. If the example had contained the words Iron Mike has beaten ..., then the order Subject (Iron Mike) followed by Finite (has) would indicate the choice of declarative mood; reversing this order would indicate the choice of interrogative mood. The Subject and Finite form the Mood Element; the rest of the clause is known as the Residue. In the clause to be analysed the Finite and Predicator have been fused in a single word, beat. Only that part of this word which constitutes the Finite is included in the Mood element, the rest forming part of the Residue. The illustration given here does not exhaust the semantic content of the metafunctions in question. For example in this clause it could be argued that the word just encodes some interpersonal content.

The textual metafunction includes both thematic and information structure. Thematic structure distinguishes between Theme, the speaker’s starting point, and Rheme, what he says about that starting point. In this clause the thematic structure can be analysed as follows in figure 4.

Figure 4. Textual function

<table>
<thead>
<tr>
<th>A million ... arrival</th>
<th>Iron Mike beats Julius Francis in ... four minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theme</td>
<td>Rheme</td>
</tr>
</tbody>
</table>

The time circumstance serves as the speaker’s starting point; that is he sets up a circumstantial framework as his starting point, the rest of the clause then follows against that background; thus a million words after his arrival functions as Theme. He then elaborates the rest of his clause against this, and consequently Iron Mike beats Julius Francis in just four minutes functions as Rheme.
The textual metafunction also comprises information structure which distinguishes between the Given, that which the speaker presents as what he wishes to be taken for granted, and the New, which is the information he wishes to focalize, either as being fresh (the unmarked case) or contrastive (marked). In spoken language the tonic accent marks the culmination of the New. Since the position of the tonic accent has to be inferred by the reader in written text, there is a tendency for it to occupy the unmarked position, which is the head of the final constituent of the clause. This gives an analysis in figure 5.

Thus, minutes is the presumed site of the tonic accent, and the final constituent, in just four minutes, is New; however, in this unmarked case, we do not know if the New extends further to the left in this clause. In this particular case, it would be possible to consider the possibility of a marked reading. It happens that the result of this boxing match was a foregone conclusion; no-one doubted that Mike Tyson would win, and probably in record time; the only question was how long it would take. In these circumstances, we might envisage a contrastive stress on four, in which case that word would constitute the only new information, and the whole of the rest of the clause would be Given.

Since the three metafunctions operate simultaneously, the three analyses can be conflated to give a (still partial) analysis of the semantics of this clause (see fig. 6).

**Wordless Documents**

Much furniture is today sold in kit form. The hapless buyer returns home to find that his purchase consists of a box full of hardware, accompanied by a sheet of paper, ostensibly of instructions for assembling the said hardware into a piece of furniture. At best this document will have summary wording accompanying a set of pictures (cf. Appendix 1, which show the instructions for assembling a convertible settee); the wording may be of a very general nature and not actually give any of the instructions (cf. Appendix 2, which shows the instructions for a CD rack); and in some cases wording is virtually non-existent (cf. Appendix 3, which shows the instructions for a newspaper rack, and Appendix 4, for an office chair). What follows will be mainly an analysis of the document shown in Appendix 4.
The Ideational Component

In this document (Appendix 4), the processes are systematically expressed as arrows. These are all material processes, that is they are all concerned with change of some sort in the physical world. They all concern movement: moving different items and bringing them into contact in various ways. The items illustrated are the participants. In this case they are all items which undergo the movements indicated by the arrows. Hence they constitute the Affecteds of the different processes. The Actor of these processes is not illustrated in the document. However, it is clearly understood from the context that the purchaser is the supposed Actor. Although it should not be expected that wordless documents present a mirror-image of verbal texts, it might be pointed out here that in written instructions, the unmarked form uses imperative mood, where the Actor is also unexpressed.

The Interpersonal Component

The illocutionary force of this document is obviously directive. (I presume my (only slightly) oblique reference to Austin (1975) in the title of this paper will have escaped no-one). The document’s objective is to alter the state of affairs in the world, though in this case perhaps one should talk of fit of world to document, rather than world to word (Searle 1969, 1979). It is interesting to notice the importance of the context in distinguishing the interpersonal component of this document. We expect kits to be accompanied by assembly instructions, and many people probably look for this type of document when they unpack a product of this type. However, it can be noted that the absence of an encoded Actor is also related to the reading of the document as having directive illocutionary force. The interpersonal component’s directive nature is imperative in the unmarked verbal case, and as pointed out above, this involves unencoded Actor too.

The Textual Component

In terms of thematic structure, Theme is the speaker’s starting point. In this document it is reasonable to take the items to be assembled as being the starting point of the message. These are illustrated in the top right sector of the document. The Rheme, what is said about the Theme, is then the procedure for assembling the items, and these are illustrated in the three icons which follow the order, top left, bottom left, bottom right. The document seems therefore to present a distinct Theme-Rheme structure. In parallel with this is information structure which distinguishes between Given and New. It seems reasonable here to take the icons illustrating the items to be assembled as Given; they are indeed what the buyer finds in front of him before he attempts to carry out the processes. The icons illustrating the assembly procedure will then constitute the New. Thus in this document Theme coincides with Given, and Rheme coincides with New, just as they would in an unmarked written document.

One might ask whether this document is in fact a well constructed (“well-written”) document. According to Kress and van Leeuwen (1996), given information typically appears to the left, at least in cultures based on Western European languages. This
document places the given information top right, and then the development follows a circular pattern moving through top left and bottom left to end in the bottom right position. I would concur with Kress and van Leeuwen to the extent that I feel that this is not the optimal layout, and would have felt more at ease with the given information in top left position, which is where Kress and van Leeuwen’s hypothesis would have placed it. The document in Appendix 3 does have this optimal layout.

It will be noted that when discussing the ideational component, it was possible to attribute an ideational structure to each individual icon, whereas here, textual structure has been attributed only to the document as a whole. Indeed it is not immediately obvious how one could attribute textual structure to the individual icons in this document, though perhaps this might be possible where individual icons have a more complex structure. Nevertheless, in the document under analysis, it would, I think, be unjustified to go beyond the macro-textual structure that has been suggested here.

**Cohesion**

It is usual in studies of cohesion to distinguish between exophoric reference, which links items referred to in a document to items in the extralinguistic world, and endophoric reference, which is intralinguistic and links items within the document itself. In this document the items illustrated top right represent the hardware supplied by the vendor, and hence can be construed as exophoric. These illustrations are then taken up in the procedure icons, so these can be construed as endophoric. Though it is not the case in this document, this type of endophoric reference can be reinforced by indexation; this is the case in the document shown in Appendix 3, where some of the items to be used in the assembly process are indexed $F_1$, $F_2$, $F_3$, and these are used in some of the following icons. Consequently, it can be claimed that this type of document contains both exophoric and endophoric reference.

**Grammatical Metaphor**

The concept of grammatical metaphor has been developed by Halliday (Halliday 1988, 1994a, 1994b, Halliday & Martin 1993, Halliday & Matthiessen 1999) to account for non-congruent realizations of semantic features. Thus a quality is congruently realized by the word class of adjective, e.g. *beautiful*, but it can be non-congruently realized as a noun, e.g. *beauty*. It is this non-congruent realization that Halliday terms grammatical metaphor. The commonest form of grammatical metaphor is the nominalization of processes, that is the realization of processes, which congruently are expressed as verbs, in the form of nouns. Diachronic studies (Halliday 1988, 1994b, Banks 1996, 1997a, 1997b, Perrot 1998) show that there has been increasing use of grammatical metaphor, particularly nominalized processes, notably in scientific and technical writing. One of the effects of using nominalized processes is that the resulting noun groups can then be used as participants in further processes. In many cases these new processes may turn out to be realized by relational verbs, frequently of low semantic content. For example the following is the first sentence of a scientific research article (Lindsay *et al.* 1996):

> Atomic and molecular collisions are of fundamental physical interest and play a central role in a wide range of environments.
If the first clause of this sentence is unpacked, i.e., expressed in congruent, non-
metaphorical form, we get something like the following:

Atoms and molecules collide and this fundamentally interests us from a physical
point of view ...

Thus we see that in the metaphorical form preferred by the scientific writers, the process
\textit{collide} is expressed in its nominalized form \textit{collision}, and the process \textit{interest} is expressed
by its homographic noun. These are then used as participants linked by the simple
relational process \textit{are}.

In the document we are analysing, all of the processes have been reduced to the same
simple iconographic form, that of the arrow. The semantic content of the message seems
to be concentrated in the participants of the processes. It is true that the content of the
processes remains material, and has not been reduced to a relational process, but even in
written documents, the reduction to relational process is the limiting case and not
necessary in all cases of grammatical metaphor. It simply illustrates what can ultimately
be done when grammatical metaphor is used. Nevertheless, it is my hypothesis that this
iconographical simplicity in terms of the processes is akin to what happens in the
phenomenon of grammatical metaphor. I would like to call this “iconographic
grammatical metaphor”.

\textbf{Other Processes}

It has been noted that all of the processes in the document under study are material
processes, and in a sense, this does not seem unreasonable. In a written document of
assembly instructions material processes would certainly predominate. However, it is of
interest to ask whether such documents are restricted to material processes. It does seem
that while the vast majority of processes in documents of this type are naturally going to
be material, other types of process may occasionally occur.

In the document shown in Appendix 5, there seem to be two mental processes of
perception. In the part shown as 5, an eye is represented (mid-right) indicating that the
assembler has to see/notice certain features. In the part shown as 5b, the curved lines
round the mallet (mid-right) seem to suggest that the assembler should hear something
click into place. This is perhaps not totally obvious, but if it were not the case one would
expect the movement to be indicated by an arrow as are other types of movement in this
document. Another example of this appears in the document shown in Appendix 6. Here
the jagged lines in the central “magnifying” circle seem to indicate that the assembler
should again hear the piece click into place. Since these are motivated perception
processes some may prefer to think of them as behavioural processes. It is perhaps also
possible to see a representation of relational process in Appendix 5. The mathematical
sign of equality (\(=\)) appears between the image of a mallet and that of a hammer with a
wooden block, indicating that either of these may be used with equal efficiency. One is “as
good as” the other. Thus, although material processes dominate, as is to be expected in
this type of document, other types of process are not excluded and will occasionally
occur.
Conclusion

My hypothesis is that these documents are texts. That is, I am not claiming that they are similar to, parallel to, or in some way aping texts, but that they are texts. They function as texts, both externally and internally. They have characteristics which function semantically in the ideational, interpersonal and textual metafunctions. They have interclausal textual features like cohesion, and they have features which function in the same way as grammatical metaphor. Since they function as texts, there seems to be no reason why they should not be called texts, or quasi-verbal texts.

Perhaps this is not all that surprising, in that many languages from ancient Sumerian cuneiform and Egyptian hieroglyphs to contemporary Chinese have writing systems that are iconographic (Campbell 1997, Ouaknin 1997). It is perhaps only those of us more familiar with alphabetic languages and writing systems that are likely to forget this. Iconographic systems have the advantage of permitting communication between individuals who do not necessarily understand each other’s languages, and who cannot communicate on a verbal basis. It is perhaps for this reason that iconographic systems are being developed for this type of communication where the language of the addressees may be varied and unforeseeable.

“What is the use of a book without pictures?” was Alice’s question. Indeed, we might ask whether we really need anything else.

BIBLIOGRAPHY


Busch-Lauer, Ines. 1998a. “Non-verbal elements in German and English medical texts”. In Lundquist, Lita, Heribert Picht & Jacques Qvistgaard (eds.), LSP, Identity and Interface, Research,


Lindsay, B.G. et al. 1996. “Charge transfer of 0.5-, 1.5-, and 5-keV protons with atomic oxygen: Absolute differential and integral cross sections”. Physical Review A 53/1, 212-218.


**APPENDIXES**

**Appendix 1**

![Appendix 1 Image]

**Appendix 2**

---

*ASp, 27-30 | 2000*
ABSTRACTS

Many household items are today sold in kit form. The assembly instructions provided are mainly iconographic, some totally so, with no words at all. This obviously has the commercial advantage of being suitable, no matter what the native language of the purchaser. The question I address is whether such documents can be considered to be texts. The Systemic Functional model provides for three levels (metafunctions) of meaning: the ideational metafunction deals with the representational aspect of meaning; the interpersonal metafunction is concerned with the relationships of the speaker, both with his interlocutors and with his message; the textual metafunction deals with that which makes a string of words into a text. Wordless assembly instructions have features that correspond to all three metafunctions. Hence it can be argued that these documents, even when they have no words, function in the same way as a written text.

Beaucoup d'objets ménagers se vendent de nos jours sous forme de kit. Les instructions d'assemblage fournies tantôt mêlent texte et illustrations; tantôt même sont purement iconographiques. Ce choix présente l'avantage économique de convenir à tous les acheteurs potentiels, quelle que soit leur langue maternelle. Peut-on considérer néanmoins ces documents comme des textes et jusqu'à quel point? Le modèle systémique fonctionnel prévoit trois strates (métafonctions) sémantiques. La métafonction idéationnelle traite de l'aspect référentiel de la signification; la métafonction interpersonnelle concerne les relations qu'entretient le locuteur avec d'une part ses interlocuteurs et de l'autre son message; la métafonction textuelle traite de l'aspect de la signification qui fait qu'une chaîne de mots constitue un texte. Les instructions sans paroles possèdent des caractéristiques qui correspondent à chacune de ces trois métafonctions. Par conséquent, on peut émettre l'hypothèse que ces documents, bien que sans paroles, fonctionnent de la même façon qu'un texte écrit.
INDEX

**Mots-clés:** cohésion, image, métafonction sémantique, métaphore grammaticale, texte, thématisation  

**Keywords:** cohesion, image, metafunction (semantic), metaphor (grammatical), text, thematic structure

AUTHOR  

DAVID BANKS