New Directions in the Analysis of Multimodal Discourse

Edited by

Terry D. Royce • Wendy L. Bowcher

NEW DIRECTIONS
IN THE ANALYSIS OF
MULTIMODAL DISCOURSE

NEW DIRECTIONS
IN THE ANALYSIS OF
MULTIMODAL DISCOURSE

Edited by

Terry D. Royce
Teachers College, Columbia University

Wendy L. Bowcher
Tokyo Gakugei University

Cover design by Kathryn Houghtaling Lacey

Contents

Contributors

Anthony P. Baldry
University of Pavia

John Bateman
University of Bremen

Wendy L. Bowcher
Tokyo Gakugei University

Judy Delin
University of Leeds

Victor Lim Fei
National University of Singapore

Alfredo A. Ferreira
Hitotsubashi University

Renate Henschel
University of Bremen

Kangli Ji
Quinghua University

Emi Kobayashi
University of British Columbia

Masaki Kobayashi
University of British Columbia

Lynn Luo
University of British Columbia

J. R. Martin
University of Sydney

Christian M. I. M. Matthiessen
Macquarie University

Bernard Mohan
University of British Columbia

Terry D. Royce
Teachers College, Columbia University

Tammy Slater
University of British Columbia

Maree Stenglin
University of Sydney

Paul J. Thibault
Università di Venezia

Len Unsworth
University of New England

Preface

In recent decades there has been an increasing interest in analyzing modes of communication other than language. Much of this has emphasized the ways that specific modes realize meaning, with little emphasis on providing a set of principles that can be used to explain how these different modes interact intersemiotically. There has, therefore, long been a need for a volume that can open up the field of multimodal discourse analysis, address multimodality from both theoretical and practical perspectives, and indicate new directions in research and application.

The range and scope of the chapters in this volume offer groundbreaking insights into exploring and accounting for the various facets of multimodality in a range of texts and contexts. The predominant theoretical approach informing the chapters is Systemic Functional Linguistics (SFL), and in accordance with this theory the work presented takes a social semiotic view of multimodal communication.

This book has not been organized around any preconceived categories because each chapter not only uses a specific medium to illustrate certain issues but also extends theoretical boundaries of multimodal text analysis in new and different directions. The initial chapters specifically aim to tackle theoretical issues related to multimodal text analysis. The subsequent chapters focus on important research areas such as writing and graphology, genre, ideology, computational concordancing, and cross-cultural and cross-linguistic issues. A particularly new and interesting contribution that

this volume makes is the inclusion of an emphasis on the educational impli-
cations of multimodal discourse in first and second language contexts.

This volume represents an exciting contribution to the field of multi-
modal text analysis. It has wide appeal and is applicable to any researcher
and educator interested in multimodality and what this means in social in-
teraction.

Terry D. Royce
Wendy L. Bowcher

The Multimodal Page:
A Systemic Functional Exploration

Christian M. I. M. Matthiessen
Macquarie University

Multimodality—as it has come to be called—is an inherent feature of all aspects of our lives, as it has been, I believe, throughout human evolution. We can interpret this condition of pervasive multimodality "from above" in terms of the stratal organization of semiotic systems, by reference to the context of culture in which different semiotic systems operate, as suggested by Halliday (1977/2003):

Essentially, language expresses the meanings that inhere in and define the culture—the information that constitutes the social system.

> Language shares this function with other social semiotic systems: various forms of art, ritual décor and dress, and the like. Cultural meanings are realized through a great variety of symbolic modes, of which semantics is one; the semantic system is the linguistic mode of meaning. There is no need to insist that it is the "primary" one; I do not know what would be regarded as verifying such an assertion. But in important respects language is unique; particularly in its organization as a three-level coding system, with a lexicogrammar interposed between meaning and expression. It is this more than anything which enables language to serve both as a vehicle and as a metaphor, both maintaining and symbolizing the social system. (p. 83)

By viewing different semiotic systems "from above," from the vantage point of the context of culture in which they operate, we can see that these *semiotic systems complement one another in the creation of meaning.* The descrip-

1

tive and theoretical challenge is to make explicit how they complement one another—how they are coordinated in the process of making meaning and how their complementary contributions are integrated with one another (explicit in the way one would have to in modeling the generation of multimodal documents, as in Matthiessen et al., 1998). This challenge can only be met by taking account of both the semiotic systems themselves and the context in which they operate.

The view "from above" complements the stratal view "from below." This is the view of semiosis that is foregrounded when we adopt the term *multimodality*, drawing attention to the multiplicity of "modalities" within the expression plane through which meanings within the content plane are realized. The same is true of the more technological notion of "multimedia." Here "modality" has to be explored in terms of both the "channel" (e.g., graphic) and "medium" (e.g., written: printed; see Martin, 1992, pp. 508–516; Thibault, chap. 3, this volume). The expressive potential has been expanded through technological advances in both hardware (e.g., from analogue to digital) and software (e.g., new formats, techniques of compression, and standards of representation). The breakthroughs seem to have been driven "from below": they have, in the first instance, been concerned with the lowest level of the expression plane, making it possible to digitize audio-visual patterns of realization in different semiotic systems. These developments have opened up new possibilities within the content plane— possibilities that are now being taken up to different degrees and in different ways, as is shown by Len Unsworth's (chap. 11, this volume) study of electronically delivered books for children. But there is a sense in which there is, as yet, no widespread awareness outside expert research teams of what it would mean to technologize the content plane to complement the technologization of the expression plane.

Viewed from below, different semiotic systems thus operate in different realms—that is, in different modalities. But viewed from above, they all operate in the same realm—the realm of meaning. The assumption is thus that differences in modalities within the expression plane *decrease* as we move into the content plane toward the context, where different semiotic systems are integrated as complementary contributions to the making of meaning in context (see Fig. 1.1). As has become standard in diagrams of this kind, semiotic strata (levels) and planes are ordered from *low* to *high* along a dimension running from SE to NW (other orientations being used for dimensions other than that of stratification). The stratal or planar subsystems are represented by co-tangential circles, which increase in size with the stratal move upwards to show that stratal subsystems increase in size, in terms of both systemic potential and extent of units, as we move up the dimension of stratification. The convergence within the content plane of semiotic systems that are realized through different modalities of expres-
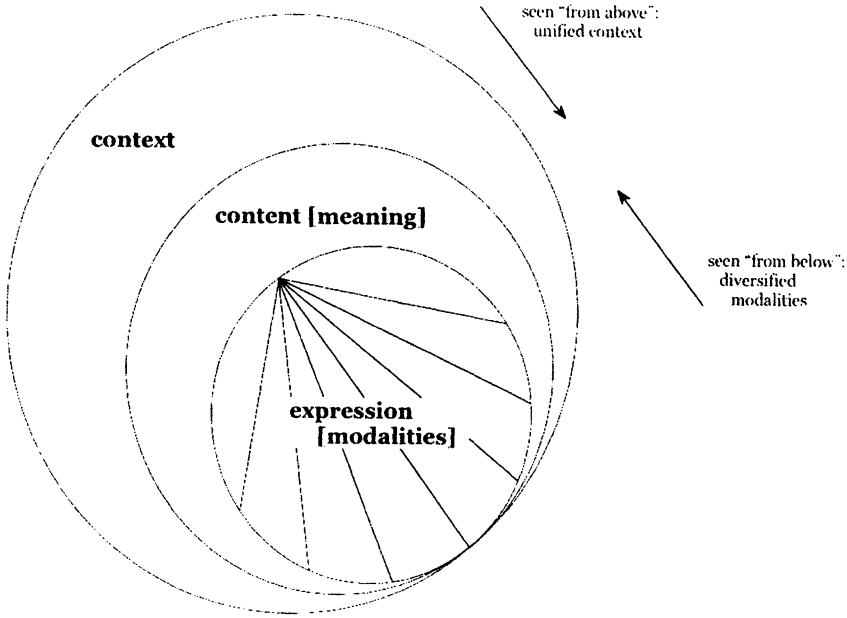
FIG. 1.1.   Multimodality: Differences in expression—differences in content?

sion would help explain why it is possible, up to a point, to translate instances (texts, drawings, paintings, ballets, pieces of music, and so on) within one semiotic system into those of another (cf. Matthiessen, 2002, and O'Toole's, 1994, translation of paintings, sculpture, and architecture into "the language of displayed art") or to generate instances from a common representation in meaning, as illustrated by the Multex multimodal presentation generation system described in Matthiessen et al. (1998), and why such translations are important in learning because they help learners achieve a more multifaceted and interconnected understanding of the relevant domain of meanings, as shown by Mohan (1986; chap. 9, this volume).

The semiotic dimension of stratification thus defines two views on multimodality—the view from above and the view from below; to these we need to add a third view—the view from within the content plane itself.

1. Viewed from above, from the vantage point of context, semiotic systems with different expressive modalities are coordinated and integrated in the creation of meanings in context. All multimodal presentations unfolding in context are like Richard Wagner's conception of an opera as a *Gesamtkunstwerk*, where the different contributions are woven together into one unified performance. The primary challenge is how to model this inte-

gration within the content plane. The integration can be modeled if this is done systemically (paradigmatically), as seen later in a section (see p. 24). Systemic distinctions shared across different modalities can then be realized in different ways in these different modalities—for example, they may be realized by function structures in language but "rendered" graphically.

2. Viewed from below, from the vantage point of the expression plane, semiotic systems differ precisely because their expressive resources are drawn from different modalities, so when they are modeled explicitly they have to be specified in different ways, as when drawings are "rendered" and music is "synthesized" according to (what we can interpret as) systemic specifications (cf. Matthiessen et al., 1998; Winograd, 1968). The primary challenge here is to determine the extent to which different expressive resources construe qualitatively different content systems.

3. These two views are complemented by a third stratal view. This is the view from within the content plane itself. It is here that the main challenge lies: modeling the meaning-making resources of different semiotic systems in such a way as to provide a synthesis of the thesis that the systems are distinct (derived from below) and the antithesis that the systems function in a unified way (derived from above).

One place to start the exploration of multimodality is with the semiotic system of language since this semiotic system is an inherently multimodal one. This starting point also makes sense in the context of the present book where we are concerned with new directions in considering the multimodality of the page, computer screen etc.—with a kind of multimodality that typically involves language.

THE INHERENT MULTIMODALITY OF LANGUAGE

The inherent multimodality of language must have evolved out of multimodality in protolanguages as part of human evolution (Matthiessen, 2004) and it develops out of multimodality of protolanguages in the life of human children (Halliday, 1975).

Protolinguistic Multimodality

Protolanguages can, in principle, have the entire body as their expression plane (cf. Thibault's, 2004, notion of the "signifying body"): they are organized systemically within the content plane into microfunctional meaning

potentials (the core ones being regulatory, instrumental, personal, and interactional), and different modes of expression may be brought together within one microfunctional meaning potential or dispersed across more than one such potential. This key property of protolinguistic semiosis can be illustrated by reference to the microfunctional repertoire of chimpanzees (see Table 1.1): Expression modalities include gestures, facial expression, gaze, and vocalization. Interestingly, if we take account of the distinction of the two forms of consciousness identified by Halliday (1992)—action and reflection, we can see that there are strong correlations between meaning and expression as far as the examples given in the table go:

[form of consciousness:] action—regulatory, instrumental ↘ gesture

[form of consciousness:] reflection—personal, interactional ↘ various, but: face is involved in all except for 'excitement,' which is realized by vocalization; gesture combined with touch is used to express 'togetherness'

There is thus a strong tendency for a particular mode of expression to go with a particular mode of meaning: The active mode of meaning goes with the active mode of expression—gesture. We can also see that there is a tendency for a natural relationship between specific meanings and specific expressions: Gestures tend to resemble the physical acts that the meanings relate to; for example, the expression realizing 'I invite you' is like the act of pulling, and the expression realizing bonding is a mutual stare.

The same picture of expressions in protolanguage has emerged from the study of young human infants (e.g., Halliday, 1975, 1979; Painter, 1984): protolanguages are *multimodal* in the sense that they employ different mo-

TABLE 1.1
Protolinguistic Signs Used by Chimpanzees
Interpreted Microfunctionally

	Content	Expression
Regulatory	'I want to be groomed' [infant to mother]	[gesture:] hand raised in air
	'I refuse'	[gesture:] shaking head
	'I threaten: keep away(?)'	[gesture:] waving arm
Instrumental	'give me' (begging)	[gesture:] arm stretched out
	'I invite you'	[gesture:] arm stretched out (as in pulling)
Personal	[emotions]	[face, including eyes]
	excitement (+ identification (?))	[vocalization:] pant-hoot
Interactional	[togetherness]	[touch & gesture, face]
	Bonding	[face, eyes:] (mutual) stare

Based on Beaken (1996, p. 51); Hart (1996, pp. 115–117); cf. Kaplan & Rogers (1999, ch. 7) on primates in general and orangutans in particular; Marler (1998).

dalities within their expression planes; they are not monomodal.[1] This multimodality would in fact appear to be one of the keys to the evolution of protolanguage: *the multimodality increases the potential for signs that are natural and motivated, and iconic* because different modes of expression go with different modes of meaning. Protolinguistic expressions may, of course, be entirely arbitrary; but *early* expressions often appear to be derivable from a material context, as illustrated by Halliday (1998). For instance, the expression of the instrumental meaning of 'I want that' is related to the material act of grabbing the object; this may be compared with the Chimpanzee gesture presented in Table 1.1. This is iconic within the active mode of meaning. But a sigh as a form of expression is equally iconic—within the reflective mode of meaning.

Later Multimodality: Language and Other Semiotic Systems

As protolanguage evolved into language in the course of phylogenesis and as children make the move from protolanguage to language in the course of ontogenesis, the linguistic multimodality is retained and expanded; but it is also, as it were, distributed across language and other semiotic systems. Gestures and facial expressions become part of the expressive resources of what has been referred to as "body language." Vocalizations continue as the expressive resources of language, but vocal features on the periphery of these linguistic expressive resources—tamber, tempo, and loudness—become the expressive resources of what has been called "paralanguage." The boundaries between language and these other semiotic systems with aspects of the body as their expression plane are fuzzy rather than sharp. These semiotic systems shade into one another, and are coordinated in the unfolding of meaning within a given context when the linguistic system is instantiated as text with accompanying instances of body language and paralanguage. For example, certain interpersonal contrasts in language are realized vocally by contrasts in tone (pitch movement) accompanied by facial contrasts involving eyebrow movements;[2] textual contrasts in deicticity are often accompanied by pointing gestures; talking to babies may involve rounded, pouting lips—a feature that affects the sound but which is also vis-

---

[1]Various scholars have suggested that language first evolved with gestures as its mode of expression; but while gesture would have been significant all along, there is evidence for the long-term evolution of the vocal tract, and a monomodal early stage seems unlikely since it would have restricted the potential for iconicity in different modalities (see the following text): see Matthiessen (2004) for discussion and references.

[2]The face in general (including the eyes and the regions around the eyes and the mouth) is a domain for the expression of interpersonal meanings (from the point of view of the metafunctions of language). This is now reflected graphically in the form of the stylized facial expression of "smileys" used in email messages, messages to online fora and elsewhere.

ible; and, as detailed studies have shown, there is a complex relationship between addressing somebody in language and gazing at them. But language involves a higher degree of systemicization of its meaning potential—the highest degree of systemicization being grammaticalization.

Body language and paralanguage thus emerge as sets of distinct semiotic systems when protolanguage is transformed into language. But this does not mean that language becomes monomodal. On the contrary, it retains its potential for multimodal expression, as the existence of sign languages of deaf communities demonstrates; and the vocalizations of spoken languages can, in fact, be interpreted as multimodal: vocalizations include two "modalities"—articulation and prosody, the second of which may be both heard (intonation) and felt (rhythm) prenatally in the womb. These different modes of expression tend to be associated with distinct modes of meaning.

Multimodality and the Emergence of Writing

In human history, the inherent multimodal potential of language was expanded in a few cultures when writing emerged gradually as part of the evolution of city-based civilizations, first in Mesopotamia, Egypt, and China and later in Meso-America. As shown schematically in Fig. 1.2, language had existed in parallel with pictorial semiotic systems for hundreds, even thousands, of generations before writing was developed out of drawing (cf. Kress & van Leeuwen, 1996, p. 19). Clear evidence of this is provided by early cave paintings; but it seems very plausible that our ancestors had been
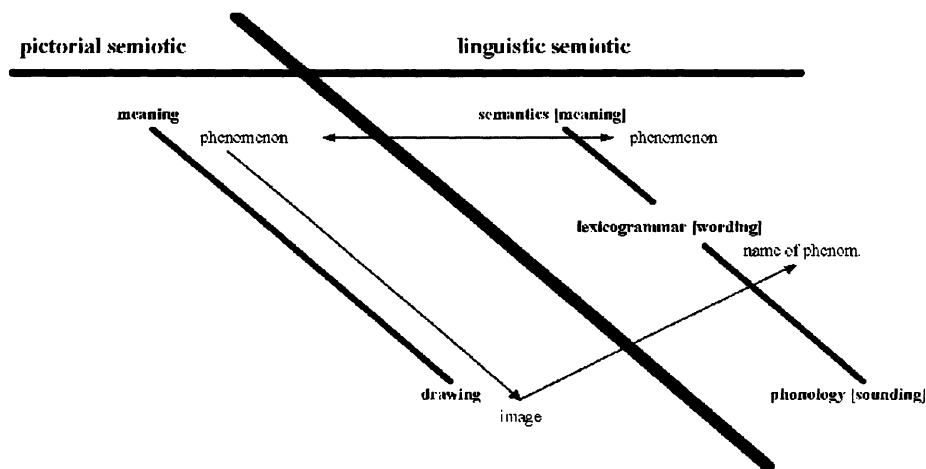

FIG. 1.2.   Conditions for development of graphology: Pictorial semiotic and linguistic semiotic as parallel systems.

drawing much earlier, using transient materials (like drawing in the sand, or in the air—where drawing and gesturing merge).

In the pictorial semiotic, the drawing system (expression plane) construes meanings (depiction within the content plane): that is, images created by the drawing system, realize phenomena construed within the pictorial system of meanings. To the extent that they are concrete, these phenomena are, of course, construals of objects in the material world; but these objects are construed semiotically by the human perceptual system. This means that the images that are drawn do not refer directly to objects in the material world. The relationship is *mediated* semiotically through the perceptual system, and drawing and meaning together constitute a semiotic system in its own right—the pictorial semiotic.

When the phenomena construed as meanings by the pictorial semiotic began to be construed also as phenomena in the semantic system of a language (such as Sumerian, Egyptian, Chinese, or Mayan), intertranslatability between the two semiotic systems came into existence. This meant (1) that the linguistic name for the phenomenon—that is, the lexicogrammatical word realizing the phenomenon—could be drawn and (2) that the image representing the phenomenon could be named or labeled, in speech. This intertranslatability provided the conditions for transforming images in the pictorial semiotics into representations of words in the linguistic semiotic— that is, into glyphs or characters in a graphology.

When images from the drawing system are gradually transformed into characters in the emergent graphology of a language, they are systemicized within this emergent graphology (writing system), as Fig. 1.3 illustrates. This is what happened as Sumerian writing evolved, a process characterized by Coulmas (2003) as follows:

> The layout of cases on early tablets was not fixed, and the signs [what I call "images" here, CMIMM], while more or less conventionalized in form, displayed considerable variation. The primary referents of the signs are physical objects. The line drawing of a bull's head refers to a bull or, perhaps, generically to cattle. The general form of the sign is significant, but the composition of the line drawing is not standardized. It is still the pictorial value that counts. But the more the scribes write the more they develop routines to produce the pictograms and in the process turn to impressing instead of scratching lines into the clay. Drawn lines are replaced by stylus impressions resulting in the characteristic wedge shapes that gave the Sumerian script its modern name: cuneiform, from Latin *cuneus* 'wedge'. . . . *Recognition of the signs is no longer based on similarity but on discrimination, as pictorial likeness is gradually replaced by the necessity to distinguish one sign from another.* Differentiation thus becomes the principal feature of the signs. For example, that the sign of a bull resembles a bull is now less important than that it differs from the sign of a cow. Hence the number and direction of wedges of which a given sign is composed are standardized. Signs come to be characterized as configurations of
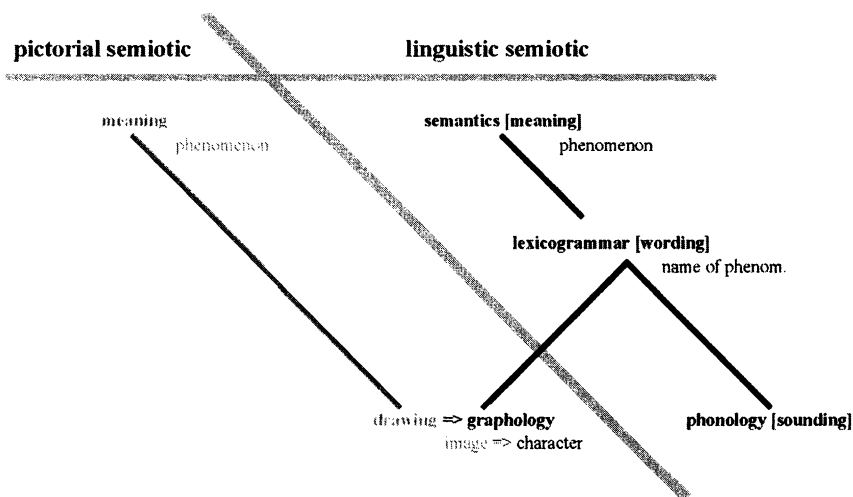
FIG. 1.3.   Emergence of graphology.

fixed numbers of strokes arranged in a fixed order. . . . The relationship be-
tween signs and objects is superseded by multiple relationships between signs
and other signs as the scribes' chief concern. *The signs thus become part of a
graphic* system *characterized by negative differentiation.* (p. 43; italics added)

As long as they are part of the expression system of the pictorial semi-
otic—that is, of the drawing system, images stand in a natural relationship
to the phenomena that they represent and their "pictorial" value is what
counts. But as they are transformed into characters in the emergent gra-
phology of a language, the relationship to what they represent—words or
morphemes in the lexicogrammar of that language—begins to change
from natural to conventional and this means that conventional grapho-
logical patterns based on systemic discrimination (valeur) rather than on
representation (signification) begin to develop.

Writing emerged in contexts that were distinct from those where spoken
language was used; there was thus a functional complementarity between
writing and speaking from the start (cf. Halliday, 1991): they occupied dif-
ferent registerial regions of instantiation along the cline of instantiation.
The contexts in which writing and the early registers of writing emerged
were ones in which the pictorial semiotic rather than spoken language was
being used—contexts of bookkeeping, relating to trade and administra-
tion: as Fischer (2001) put it, "Complete writing's crucible was accoun-
tancy" (p. 22). Referring to Olson (1989), Halliday (1991) noted that spo-
ken language is typically "communicative" whereas written language is
typically "archival." As Halliday (1991) pointed out, "writing was associated
from the start with non-propositional (and hence non-clausal) registers: for

example tabulation of goods for trading purposes, lists of names (kings, heroes, genealogies), inventories of property and the like"—these non-propositional meanings being ones that had previously been depicted. Written language thus replaced the pictorial semiotics in these registers of bookkeeping; for example, instead of drawing items that needed to be recorded scribes would now represent their names in writing by means of logograms. However, just as spoken language had remained in close contact with body language and paralanguage when it gradually evolved out of protolanguage hundreds of thousands of years before writing emerged, so written language remained in close contact with the pictorial semiotic: Among the written registers that evolved, many were multimodal from the start, with written text and image woven together within the same layout, as in many Egyptian displays (see further below).

The Multimodality of Language in Relation to Other Semiotic Systems

To sum up the discussion so far: Language is inherently multimodal, having the potential for (at least!) three kinds of expression system—phonology, graphology, and sign. Graphology and sign are (typically) manifested visually (with braille as an example of tactile graphology), whereas phonology is manifested aurally; but whereas graphology is (typically) manifested as traces of bodily movements in some relatively permanent material that is extraneous to the body (clay, papyrus, paper, wax, bone, stone), phonology and sign are manifested as transient bodily movements that are perceived aurally or visually. These different modalities of the linguistic expression plane have evolved together with (and, in the case of graphology, been taken from) other semiotic systems. The situation is summarized diagrammatically in Fig. 1.4. Here the region of the expression plane that can be manifested on the printed page includes graphology and visual "paralanguage."

MULTIMODALITY, MEDIA OF EXPRESSION, MODES OF EXPRESSION, AND MODES OF MEANING

Modes of Expression Manifested in Media of Expression in Language

Figure 1.4 shows how language makes contact with other semiotic systems within the expression plane because it is itself multimodally diversified in expression. This is thus how multimodality across semiotic systems appears to us when we view it "from below," from the vantage point of the expres-
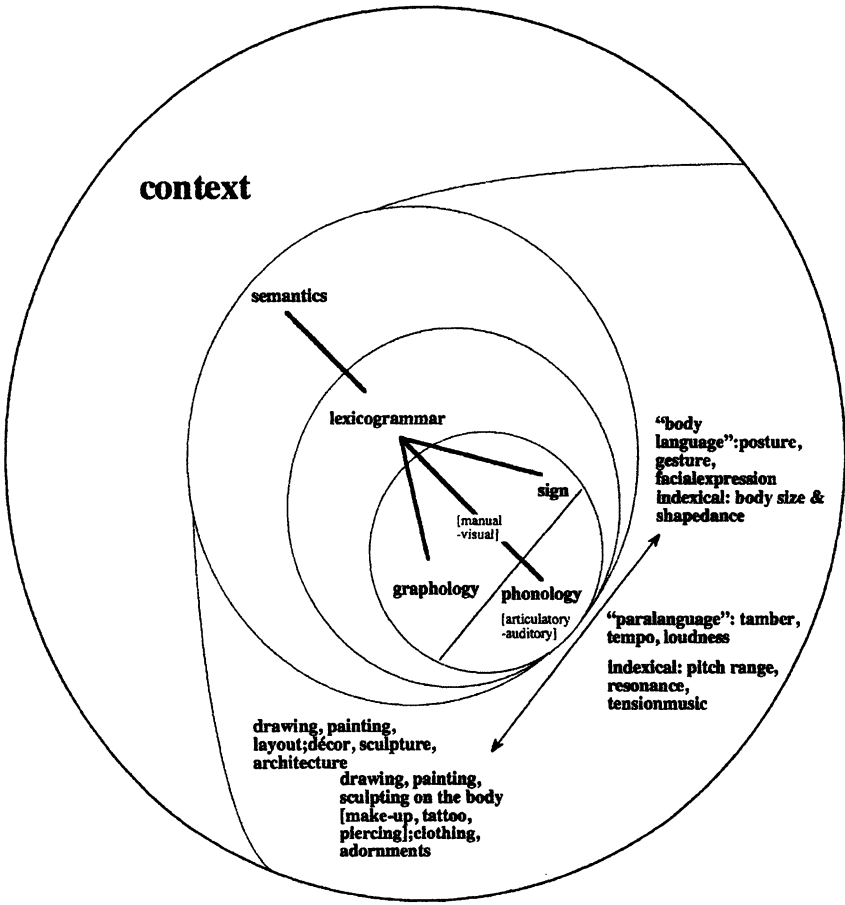
FIG. 1.4. The inherent multimodality of language, and links to other semiotic systems.

sion plane. To complement this expression-plane view, we also need to explore how language makes contact with these other semiotic systems within the content plane; but first I examine multimodality within language a little further.

In Fig. 1.4, multimodality in language is represented as being confined to the expression plane: There are three different expression systems (graphology, phonology, sign), but there is only one system within each stratum of the content plane—one lexicogrammatical system and one semantic system. This is of course partly based on a theoretical decision linguists have made about how to model language; but extensive descriptive work and discourse analysis in systemic functional linguistics over the last 40 years or so

have both shown that within the content plane the "modal" difference between speaking and writing is simply part of the more general phenomenon of registerial variation in language—that is, variation in the instantiation of semantic and lexicogrammatical options within the total meaning potential of language according to context of use (situation type). There are certain lexicogrammatical systems and their semantic correlates that are present in the content system of speaking but not in the content system of writing because they depend on the expressive potential of intonation. These are the interpersonal system of KEY and the textual system of INFORMATION. But these systems can be located as registerial partitions available in the spoken mode within the overall lexicogrammatical system common to spoken and written registers (cf. Matthiessen, 1993, on registerial partitions; see Bateman, Matthiessen, Nanri, & Zeng, 1991, on this form of representation). With respect to all other systems, the lexicogrammatical and semantic differences between speaking and writing can be modeled in probabilistic terms as registerial settings of probabilities of instantiation (cf. Halliday, 1978; Matthiessen, 1993, 2002, in press). This is significant for the description and modeling of language in relation to other semiotic systems: Generalizations about the division of labor among these different semiotic systems often seem to be of the order of generality of registers (such as recipes) or families of registers (such as procedures), as is illustrated by Martinec's (2003) study of recipes and other procedural texts accompanied by images in English and Japanese.

In other words, while protolanguage shows a tendency to correlate different microfunctional modes of meaning with different modes of expression, as shown by Halliday (1992) for the protolanguage of one young child and as illustrated for chimpanzees in Table 1.1, (adult) language is not characterized by such a simple correlation between modalities of expression and modes of meaning. There is, however, a relationship between metafunctional modes of meaning and modes of expression. This relationship, first identified and theorized by Halliday (1979) and subsequently explored further in a number of publications (e.g., Martin, 1996; Matthiessen, 1988), is more abstract and powerful than the kind of relationship we find in protolanguage between microfunctional modes of meaning (action: instrumental, regulatory; reflection: interactional, personal) and modalities of expression (gestural/vocal). It is based on modes of expression rather than on modalities of expression. Modes of expression are abstract patterns of organization that generalize across the modalities of speaking and writing, and they are highly relevant to our understanding of the inherent multimodality of language.

While modalities of expression are closely tied to low-level channels within the material realm such as manual–visual or articulatory–auditory, modes of expression are more abstract principles of patterning in expres-

sion that are manifested within these channels. Since they are more abstract, they can be manifested in different modalities, as we shall see presently. These abstract principles of patterning are set out in the leftmost column of Table 1.2: The experiential mode of expression is configurational (particulate), the interpersonal mode of expression is prosodic, and the textual mode of expression is promotional (culminative, or wavelike). (In addition, we need to take account of the logical model of meaning, which is serial, or chain-like.) These different modes of expression are manifested in different media of expression such as sequence, segments, and intonation. These media of expression are likely to vary from one modality of expression to another, although there are clearly identifiable correspondences.

In spoken language, the media of expression are segments (realized by articulatory sequences), sequence (realized by temporal unfolding) and intonation (realized by pitch movement). In written language, the media of expression are segments (realized by graphemic sequences) and sequence (realized by graphological unfolding horizontally [left to right or right to left] or vertically). Intonation is of course not present in written language (cf. Halliday, 1985); and features that might serve as graphological equiva-

TABLE 1.2
Metafunctions and Modes of Expression in Language (Spoken) and Image

Metafunction: Mode of Expression	Language (Spoken)		Pictorial System (Image)	
	Primary Medium	Secondary Medium	Primary Medium	Secondary Medium
Experiential: Configurational	configurations of linguistic segments (realized conventionally by articulatory sequences)	—	configuring of figures (realized naturally by "picture elements")	. . . coloring
Interpersonal: Prosodic	prosody through intonation (tone)	juncture prosodies (segments placed at junctures with the whole unit as their domain)	suffusing through hue/ tone (color)	[relation to viewer:] gaze, viewing distance & projection
Textual: Promoting (wave, pulse)	promotion through intonation (tonicity)	promoting particles (segments marking highlighted elements)	highlighting through brightness (color)	highlighting through composition

lents (e.g., font styles—italics, bolding, and so on) are very varied in use.[3] They are thus not systemic in the way intonation is, but are rather more paralinguistic in character. These different media of expression are deployed in different ways by the different modes of expression, as indicated in Table 1.2. Let us consider the different media of expression one by one, noting how they are deployed by the metafunctional modes of expression within the grammatical zone of lexicogrammar.[4]

1. As a medium of expression, intonation is complex; it can be deconstructed into both tone contour (continuous pitch movement) and tonic (location of major pitch movement). And this deconstruction of the expressive resource of intonation into two modes has been taken up and deployed by the grammars of many languages: One mode is deployed interpersonally as a prosody (the direction of the pitch movement),[5] and the other textually as a wave with a peak of intonational prominence (the location of the major pitch movement).

2. As a medium of expression, sequence is complex; it can be deconstructed into both relative sequence of paired elements, such as Subject and Finite (realizing mood type) and Subject and Complement (possibly with Predicator as an intervening internal juncture, realizing modal responsibility), and linear, absolute sequence, from initial to final position in the clause and other grammatical units. And this deconstruction of sequence

---

[3]For example, double and single quotes, small capitals, bold, italics, and underlining are used in different ways by different authors, by different publishing houses (reflected in their style sheets) and they also vary across registers. The patterns of use are thus fairly local and have not been standardized for the overall system of written English.

[4]Of the three media of expression, only segments are deployed within the lexical zone of lexicogrammar in the form of lexical items because it is only segments that can expand sufficiently in number to realize the very large number of the open-ended and ever-changing set of lexical features of a language.

[5]It is important to distinguish between tone extending over the domain of a tone group (intonation unit) and tone extending over the domain of a syllable. Tone within the domain of tone groups is deployed by the grammar as a medium for the prosodic mode of realization within the interpersonal metafunction, and the realizational relationship is a natural one (as opposed to a conventional one). In contrast, tone within the domain of syllables is not deployed prosodically by the grammar. Instead, it is deployed by either grammar or lexis in the same way as articulatory patterns are, as part of the phonological shape of a grammatical or lexical item (such tones may in fact derive historically from the effect on pitch exercised by phonemes that have disappeared; cf. Yip, 2002, pp. 35–38, on "tonogenesis" in general, and Mithun, 1999, p. 25, on the development of tonal contrasts in certain languages in North America): here the realizational relationship is a conventional ("arbitrary") one, and it is not associated with any particular metafunction. This is familiar from the discussion of lexical distinctions that are realized tonally in "tone languages." Grammatical distinctions may also be realized tonally in this way, as happens with "tone cases" in some Western Bantu languages (Creissels, 2000, p. 234).

into two modes has been taken up and deployed by the grammars of many languages: One mode is deployed interpersonally, as a prosody of movement either from Subject to Finite or from Finite to Subject and as juncture prosodies where interpersonal particles are placed either at the interpersonal "overture" of a clause as an interactive event or at the "interpersonal" finale, and the other mode is deployed textually, as a wave with a peak of sequential prominence (typically at either the beginning or the end of the clause).

3. As medium of expression, segmental marking might be have been thought to be the medium of expression for the experiential configurational mode par excellence. However, segmental markers—that is, grammatical items at various ranks—are in fact deployed according to all the different modes of expression. Segments can highlight the peak of a textual wave as prominent over other elements,[6] they can serve as juncture prosodies, marking the value of an interpersonal prosody that suffuses the whole clause, or they can differentiate the component parts of an experiential configuration. These different modes of deploying segmental markers are also shown in Table 1.3 with one or two languages mentioned as examples within each cell.

The different metafunctional modes of deploying segmental marking can be illustrated by reference to Japanese. Figure 1.5 shows an analysis (in three displays, [a], [b] and [c]) of an example of a Japanese clause, due to Kazuhiro Teruya (cf. Teruya, in press, forthcoming).

1. Textually, one element is singled out as a peak of thematic prominence in the clause as message wave or "pulse"—*anata wa*. It is marked by *wa*; and the remainder of the clause is thematically nonprominent. That is, textually, the *wa*-marked element is not part of a configuration of other elements. The fact that *anata* is present rather than absent gives it some degree of prominence.

2. Interpersonally, the whole clause is enacted as a move giving information with a check for confirmation directed to the listener. This interpersonal selection is realized by the interpersonal particle *ne*, which serves as a juncture prosody at the end of the clause when the speaker is about to hand over to the listener. This *ne* is not configured with any of the other elements

---

[6]Related to this use of "promotional" items, there is an alternative textual deployment of segments: the degree of phonological (graphological) "presence" or "weight" of an item may serve to realize the degree of textual prominence. Nonprominence is realized by nonsalience and phonological reduction, as happens with pronominal items that realize recoverable (identifiable) and given information. The limiting case of this reduced presence is of course absence—that is, ellipsis used to indicate continuous information.

TABLE 1.3

Segmental Marking as Medium of Expression

Class of Item	Textual ↗ Wave	Interpersonal ↗ Prosody	Experiential ↗ Configuration
Clause	Particles	"topic" or "focus" markers (before or after relevant element)	interpersonal particles (typically at junctures) marking mood & key, affect, "politeness", modal assessment, speaker status Mandarin: *ma/ ne/ ba* etc. and Japanese: *ka/ ne/ yo* etc. clause-final juncture as Negotiator	
Group/ Phrase	Adpositions	Mandarin: *a* etc. after the topical Theme ["pause particles": Li & Thompson, 1981, p. 86] adpositional markers highlighting (topical) Theme or (Focus of) New	adpositional marking of arguability status, but: *verbal: auxiliaries marking modality, polarity* German: Subject ↘ nominative; Complement, Adjunct ↘ oblique	nominal: multivariate configurations of adpositional cases, marking participants/circumstances Japanese: configuration of participants and circumstances *ga* + *o* + *ni* + *kara* + . . .
Word	Affixes	Japanese: *wa*; Korean: *un/nun*; Tagalog: *ang*	*verbal: affixes marking mood, modality, polarity* Japanese (modality, polarity); Greenlandic (mood)	nominal: multivariate configurations of affixal cases marking participants/circumstances Finnish (affixal cases marking participant and circumstance roles)

16

[a]

	Anata **wa**	kinoo no gogo goji **ni**	kooen **de**	Suzuki-san **o**	mimashita	ne?
	you	yesterday 'of' afternoon 5	park	Suzuki-san	see: formal + past	
wave	- **wa**					
prosody					(Predicator)	**ne**
configuration	Senser [0]	Time **ni**	Place **de**	Phenomenon **o**	Process	

'You saw Suzuki yesterday afternoon at 5 in the park, didn't you?'

[b]

Anata **wa**   kinoo no gogo goji   **ni**   kooen **de**   Suzuki-san   **o**   mimashita   n<u>e</u>?

| Senser | Time | | Place | Phenom. | Process |

[c]

Anata **wa**   kinoo no gogo goji   **ni**   kooen **de**   Suzuki-san   **o**   mimashita   ne?
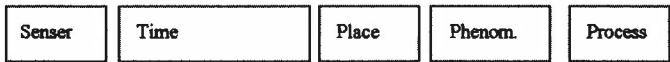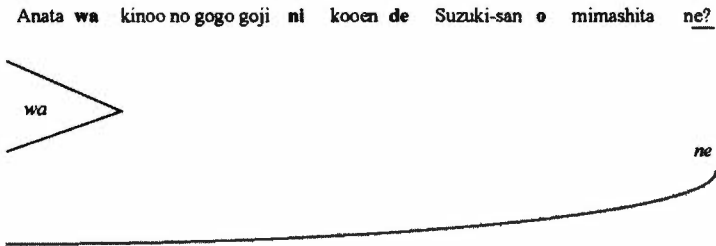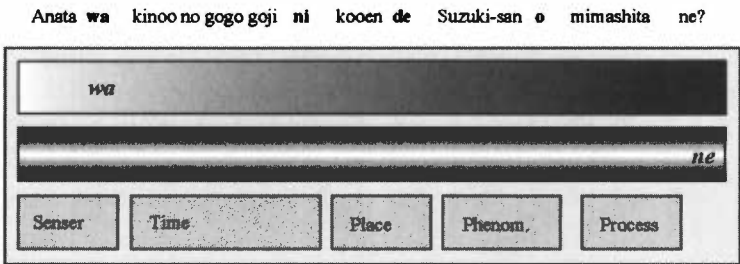
FIG. 1.5.   Japanese clause as wave, prosody, and configuration.

of the clause; it is "suprasegmental" and marks a feature that colors or suffuses the whole clause.

3. Experientially, the clause is a configuration of elements, each with its own distinctive contribution to the whole: a process, participants involved in the process (Senser, Phenomenon) and attendant circumstances (Time, Place). The participants and circumstances are marked by "postpositions,"

each indicating a distinct role in the configuration: Senser-□ + Time-ni + Place-de + Phenomenon-o + Process. (The exception is the Senser; it would have been marked by *ga* if it was nonthematic and it is helpful to think of it as being "*gawa*" (as opposed to "*owa*"): Like thematically highlighted participants in general, it only carries the thematic marker *wa*, whereas thematically highlighted circumstances are marked by postposition + *wa*.) Here the segmentally marked elements do form a configuration of component parts with distinct roles in the configurational whole.

The three media of expression of intonation, sequence and segments can thus serve as "carriers" of the three modes of expression, the modes of expression constituting different modes of deploying the media of expression (see Table 1.4). This principle has two implications: (1) there may be

TABLE 1.4
Medium and Mode of Expression Within
the Grammatical Zone of Lexicogrammar

Medium of Expression	Mode of Expression		
	Textual ↘ Wave	Interpersonal ↘ Prosody	Experiential ↘ Prosody
Intonation	intonational prominence: location of major pitch movement	intonational suffusion: extension of pitch movement	—
	[realization of INFORMATION]	[realization of KEY, MOOD]	
Sequence	sequential prominence: location of poles on a cline from early (initial) to late (final)	sequential suffusion: relative sequence as marker for whole clause [mood type: Subject ° Finite; modal responsibility: Subject ° Complement, with Predicator as potential internal juncture]	sequential phases (stages): sequence as default iconic representation of the unfolding of a figure [but easily overridden by textual/interpersonal factors]
	[realization of THEME]	[realization of MOOD]	[realization of TRANSITIVITY]
Segment	segmental prominence: particles used to "highlight" element of structure	segmental suffusion: initial or final juncture particles with the whole clause as their domain	segmental configuration: distinct "particles" for multivariate roles
	[realization of THEME (thematic particles), or INFORMATION (focus particles)]	[realization of MOOD]	[realization of TRANSITIVITY]

other media that can act as "carriers" of waves, prosodies and configura-
tions; and (2) while these three media are often used to realize comple-
mentary systems (as in the case of THEME and INFORMATION, realized in
many languages by sequence and intonation, respectively), it is also possi-
ble that they could serve as alternative realizations within the same system
in one language or, typologically, across languages.

The second point means that *alternations in medium of expression operate
within one and the same metafunctional mode of expression.* That is, textual waves
may be alternatively manifested sequentially, intonationally, and segmen-
tally; and interpersonal prosodies may be alternatively manifested sequen-
tially, intonationally, and segmentally. As far as I know, no language deploys
intonation experientially as a configurational mode of expression; but the
configurational mode of expression can be manifested either sequentially
or segmentally. We may find such alternations in the deployment of media
of expression within a single language; and we certainly also find them as a
range of alternatives across languages.

Media of Expression in Language and Image

In representing the different modes of expression diagrammatically as in
Fig. 1.5, I have shown that the principle of alternative media of expression
for the same mode of expression can in fact be generalized across semiotic
systems: I have shown how wave, prosody, and configuration can be mani-
fested through visualization in a pictorial semiotics, as in Halliday (1979),
Matthiessen (1988), and Martin (1996).

The general principle is that, as abstract kinds of patterning, the differ-
ent modes of expression (wave, prosody, and configuration) can be mani-
fested in different media of expression and that the range of media of ex-
pression available will depend on the "modality" of a given semiotic
system. For example, while prosodies and waves can be manifested
through intonation in spoken language, they can be manifested through
color in images: Prosody is manifested through the hue or tone that suf-
fuses an image or part of an image, and wave is manifested through bright-
ness, with high degree of brightness serving to highlight textually promi-
nent regions. At the same time, color is also representational—that is, it
represents the construal of human perception of color in the world
around us, and it is part of the expressive resources of pictorial systems for
realizing experiential configurations. Similarly, composition is used in dif-
ferent ways by the different metafunctional modes of expression. The ex-
periential metafunction deploys composition to construe configurations
within an image—configurations of elements, of figures, and of whole
scenes, while the textual metafunction deploys composition to promote

and demote what is represented in an image—centrality in a composition being a typical way of textual promotion. Further, perspective provides alternative strategies for construing our experience of a scene; but at the same time these alternative strategies also enact different relationships with the viewer of an image since perspective will place him or her relative to the scene represented in the image.

Examples of the metafunctional deployment of these and other media of expression in spoken language and pictorial systems (image) are set out systematically in Table 1.2. This table allows us to see the potential for translation between language and pictorial systems: Compare the discussion in Matthiessen (2001) of Ruskin's "translation" of Turner's painting *The Slave Ship*.

One interesting difference between language and image with respect to the media of expression has to do with degree of conventionality. In language, intonation and sequence tend to stand in a natural relationship to the meanings they realize; for instance, the expressive contrast between rising and falling pitch is a natural realization of the contrast in certainty between uncertain and certain—perhaps even with a physiological basis (rise and uncertainty being manifestations of tension, and fall and certainty of relaxation—cf. Bolinger, 1978).

However, with certain exceptions (like onomatopoeia and sound symbolism), segments stand in a conventional relationship to the meanings they realize, which is one reason why learning vocabulary in a language from a different genetic family from that of one's mother tongue is so hard! This relates to the fact that the rank scale of the lexicogrammar of a language is not congruent with the rank scale of its phonology.

In contrast, in pictorial systems, all media of expression stand in a natural relationship to the meanings they represent (unless of course they represent symbols from a different kind of semiotic system such as mathematical symbols). This relates to the fact that the rank scale of the content plane of a pictorial system is congruent with the rank scale of its expression plane. Since the relation between content and expression is natural rather than conventional, systems within the different metafunctions tend to vary together rather than independently of one another. For example, if a ray of sunshine breaks through the clouds to illuminate a group of people in a traditional landscape painting, this will at the same time serve to promote them textually against the background of the landscape. Similarly, it is not possible to change perspective or gaze in an image without changing both what is represented in the image and the way in which it interacts with the viewer. Perspective and gaze work interpersonally in pictorial systems because they imply a location of the viewer in an extension of what is represented.

Construal in Language and Multimodality

The potential for "translating" linguistic models into other semiotic systems as in Fig. 1.5 is of course built into our linguistic resources for construing them within the content plane by means of the linguistic metaphors of wave, prosody, and configuration. That is, the ideational resources of language provide us with a range of strategies for modeling ("conceptualizing") all modes and media of expression; and we can then use nonlinguistic semiotic systems such as drawing, painting, animation, sculpture, or music for representing these models construed within the content plane of language (compare the discussion of the evolution of writing in Figs. 1.2 and 1.3).

Other semiotic systems can thus serve to "concretize" abstract semantic models in language based on the concrete grounding of these models. One particularly important area here is abstract space: Languages construe many concrete domains of experience based on the model of space. One of these is degree (as noted by Whorf, 1956): Degree is construed as vertical location in space and change degree as vertical movement in this abstract space. Thus in construing the degree of temperature, we say that the temperature is very high, that it is falling, and so on. Similarly, when intonation is visualized in phonetic analysis, it is represented as a vertical movement extending over time. (A related alternative is to construe degree as volume in three-dimensional space.) This whole metaphorical system for construing degree on the model of space is central to discourses of economy and finance, and this abstract space can be concretized through visualization in graphs of various kinds (cf. Halliday & Matthiessen, 1999).

The nonlinguistic semiotic systems may provide highly schematic or stylized representations (as in the representation of a wave in Fig. 1.5 [b], or the representation of a pulse in Fig. 1.5 [c]); but they still concretize the linguistic metaphors by visualizing them or making them perceptually accessible in some other modality (one can, for example, play a text score of the kind illustrated in Matthiessen, 1995, by mapping features onto notes, thus making it possible to "hear" the analysis of a text). The visual mode is of course privileged—in human perception, in the linguistic construal of what we perceive,[7] and in the potential for creating representations in the form of images. This translation across semiotic systems is very important in science, engineering, and in education. Mohan (1986) shows how it plays a significant role in facilitating learning.

---

[7]This is illustrated by Viberg's (1984) typological study of verbs of perception: The ones for visual perception are always more highly elaborated than the ones for other modes of perception.
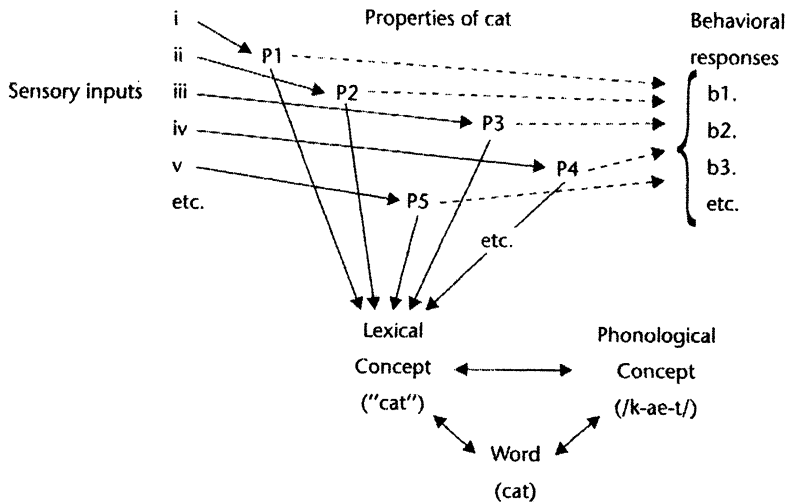
FIG. 1.6. The "lexical concept" of 'cat' as an integrative node in a network
of relationships. From Bickerton (1995, Fig. 1.1).

The potential for intersemiotic translation is thus created by language:
Linguistic meanings constitute the nodes in networks of relationships link-
ing them to meanings in other semiotic systems and in the "biosemiotic"
systems of perception (cf. reflection in protolanguage) and motor action
(cf. action in protolanguage). This view is supported by the model pre-
sented by Bickerton (1995), as illustrated by reference to the "lexical con-
cept" of 'cat' in Fig. 1.6.

This drawing could be further extended by additions of concepts of cat
from other semiotic systems linked to the lexical concept of cat. These
would include pictorial concepts linked to sensory concepts and (in the
case of people who can draw) to the motor actions for drawing a cat.

From a neurological point of view, the potential for intersemiotic trans-
lation based on language means that language is the one semiotic system
that involves virtually all regions of the brain, integrating these regions into
a vast network of relations (see Fig. 1.7). It would appear that the principles
of organization are different within the content plane and within the ex-
pression plane.[8] Within the expression plane, the determining factors in lo-
calization are modality (written vs. spoken) and direction of instantiation
(production vs. analysis). Within the content plane, the determining factor
is the metafunctional dispersal of meaning. Within the ideational meta-

---

[8]This is my interpretation as a linguist of a range of sources (e.g., Deacon, 1997; Eccles,
1989; Edelman, 1992; Edelman & Tononi, 2000; Ellegård, 1982; Greenfield, 1997; Lamb,
1999).

localization according to
metafunctional mode of
meaning

ideational –
experiential: distribution
according to
phenomenal realm
(e.g. parietal, temporal)
logical: ~ frontal lobe
reasoning?

Interpersonal –
distribution according to
realm of enactment:
emotion – frontal lobe
intentional action – frontal
lobe
[intonation: right
hemisphere]

higher degree of integrative organization

semantics

lexicogrammar

phonology

phonetics

textual –
distribution according to
realm of attention:
[ordinary] focus of
attention –
thalamus

STRATIFICATION
[metarundant patterns over
patterns]

higher degree of localization in
different perceptual /
motor areas

writing phase:
motor cortex

reading phase:
eyes => thalamus
=> occipital lobe –
primary visual
area => Wernicke's
area

articulatory phase:
Broca's area =>
motor cortex :
mouth & tongue
regions (precentral
gyrus)

evolutionary specialization:
articulatory & auditory

auditory phase:
ears -> thalamus
=> temporal lobe :
primary auditory
cortex =>
Wernicke's area

distribution according to direction
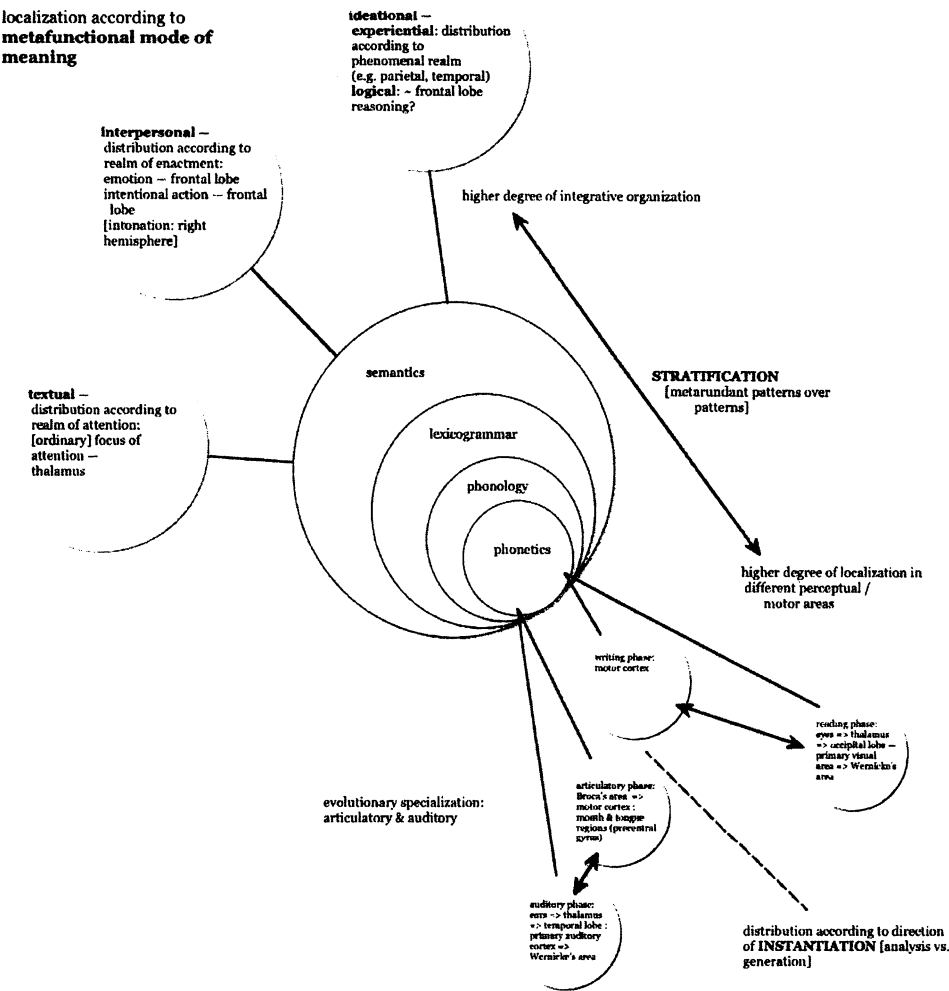of **INSTANTIATION** [analysis vs.
generation]

FIG. 1.7.   Language and the brain: the distributed and integrative nature of language.

function, meanings appear to be linked to the phenomenal realms that
they construe; and this takes us back to Bickerton's (1995) linguistic cat.
The picture that is emerging would thus help explain the multimodal
power of language: Its expression plane is multimodal, and is processed as
such; but its content plane transcends this multimodality, while at the same
time providing the resources for construing phenomena from different
perceptual realms, integrating them seamlessly into a unified experiential
system, and mapping them onto the interpersonal and textual domains.
And language might in turn help explain the multisemiotic power of the
brain. There is increasing evidence that language and the brain evolved to-

gether toward ever greater complexity (e.g., Deacon, 1992, 1997; Edelman, 1992) and this semiotic-biological coevolution may have paved the way for music and visual arts, as suggested by Deacon (1992):

> Brain-language co-evolution has made us all linguistic savants or prodigies in some sense. Like the savant, we effortlessly accomplish the miraculous mental feats of language calculation because we have some peculiarly overdeveloped and highly specific cognitive abilities. But also like the savant these special abilities tend to organize many of our thought processes and bias what we attend to and seek out in the world. Perhaps even the uniquely human fascination with combinatorially structured sounds (music) and sights (visual arts) can be explained as a secondary consequence of this cognitive reorganization. (p. 70)

TEXT AND IMAGE ON THE PAGE

Language is, as discussed earlier, inherently multimodal (see Fig. 1.4); and in each modality, it operates together with a range of other semiotic systems: These semiotic systems draw on different aspects of the resources of the expression plane, but the boundaries are fluid and indeterminate. The modal differences within the expression plane are, I suggested, mediated by the metafunctional principle of organization and realization: The different metafunctional modes of meaning tend to be realized by different metafunctional modes of expression, and these different modes of expression are manifested in different media of expression. The media of expression will vary according to the modalities of the expression plane; but the modes of expression are more abstract patterns of organization that have the power to generalize across semiotic systems of different modalities. For example, prosody can be manifested aurally as pitch movement and visually as color, the constraint being that the medium has to be able to "carry" a pattern of suffusion with polar contrasts (such as falling vs. rising pitch or warm vs. cold hue).

Semiotic Systems of the Printed Page

Let us now narrow the focus to those semiotic systems that are realized "on the printed page." These systems include:

- Language, written (with the potential for being read aloud in spoken language)
- Visual paralanguage: font family, type face ("style"), layout (graphic design)

• Visual (pictorial) semiotic systems defining images of different kinds: drawings, paintings, photographs, maps, graphs, charts, and so on

In the creation of printed pages, the semiotic labor is divided among these different semiotic systems: they are coordinated in the process of making meanings, and integrated (more or less tightly) into a *Gesamtkunstwerk*.[9] The nature of the division of semiotic labor depends on the context: Different recurrent settings of the field, tenor, and mode parameters of context (i.e., different situation types) are associated with different registers, each of which is characterized by some particular division of labor among the semiotic systems listed earlier.

It will not be possible to exemplify the full potential of the printed page here, so instead I will draw on texts from one register—health reports issued regularly by the WHO, their *Weekly Epidemiological Record.* An example of part of one page from such a report is given in Fig. 1.8. As illustrated by this page, the main semiotic labor in this register is carried out by language; reports are in English and French and may be accompanied by elements from other semiotic systems—graphs (as in this case), maps, and tables. These displays have captions in English and French, and include numbers and labels in both languages. They are thus designed to be read in English or French, with support from displays.

The overall organization of the page is paralinguistic rather than linguistic: the layout frames the texts and the images and includes a header in English and French outside the borderlines indicating the organization of the framing. The English and French texts are placed in parallel columns with the same font, font size, and type face so that they are on equal footing (unless the horizontal ordering from left to right is seen as significant). The image is placed below these two parallel paragraphs and the border lines indicate that it accompanies both texts. It seems that the placement of images on the pages of WHO reports is based primarily on constraints "from below"—that is, considerations of where they will fit on the page and of proximity to the relevant passage of text. Printed pages are very varied in this respect. Different conventions are in operation in different registers; for example, the composition of pages with advertisements tends to be under control "from above" by the textual metafunction (cf. Kress & van Leeuwen, 1996). As we will see shortly, images play a supportive role in WHO re-

---

[9]Kusukawa (2000) discusses the consequences of the distinct technologies and professions involved in the production of text and image in the first centuries of printing, noting the discussion of the use, value, and role of images in the 16th century. Since different technologies and professions were involved, multimodal pages would sometimes be produced "from below" on the expression plane with little regard for the view "from above," from the content plane. Books where closer relationships between text and image were explored within the content plane included those dealing with anatomy.

FIG. 1.8.   Example of WHO health report from the *Weekly Epidemiological Record*, 9th June, 1995, No. 23, excerpt from page 163.

ports. They are introduced, labeled, and interpreted linguistically. Printed pages vary considerably in this respect too, ranging from pages relying entirely on semiotic systems other than language to ones constituted entirely in language. The variation in the division of labor among the semiotic systems contributing to a printed page is, as already noted, registerial in nature, which means that the variation correlates with different contextual settings.

Context

Printed pages appear in a wide range of registers that occur in different situation types characterized by different combinations of field, tenor, and mode settings. These situation types in which the registers of printed pages occur obviously fall within the overall context of culture of a society, with certain general constraints placed on field, tenor, and mode, as indicated in Table 1.5. The situation type in which the WHO *Weekly Epidemiological Record* occurs is set out in Table 1.6.

TABLE 1.5
Context Values of "the Printed Page"

	Field	*Tenor*	*Mode*
First order	social process: typically reflection (rather than action); dissemination or archiving of meanings, under conditions of sufficient financial support	social (institutional) roles: writer to (segment of) general public; range of institutional roles for writers appearing in print	—
Second order	field of experience (topic domain): not constrained, except by tenor-based values—moral (e.g., decency), legal (censorship, security laws), ideological (e.g., objectivity, privacy) & "self-censorship"	speech roles: not constrained (except as a reflection of social roles) affect: often "tempered" due to public and lasting nature of printed page valuation (of field): neutral or charged, but subject constraints on value systems (moral, legal, ideological)	role of semiotic systems in context: rhetorical mode: (typically) constitutive (rather than ancillary); cline from language only to visual semiotic only with various intermediate combinations mode: written medium: graphology channel: page, printed

TABLE 1.6
Field, Tenor, and Mode Values
of WHO's *Weekly Epidemiological Record*

	Field	Tenor	Mode
First order	social process: reflection on state of communicable diseases; dissemination of current info about these and archiving	social roles: medical expert(s) to other medical experts, including health officers in various institutions and members of the general public with special interests	—
Second order	field of experience (topic domain): medical conditions in the field: changes in state of communicable diseases—from single cases to epidemics; applications of countermeasures institutional (rather than personal) view of outbreaks of diseases and institutional responses	speech roles: expert giving information (to be accepted), with assessments of information varying from reported facts to inferences affect: neutral valuation (of field): neutral	role of semiotic systems in context: rhetorical mode: multimodal document constitutive; text in English & French with accompanying images (maps, graphs) and tables elaborating information evaluated in the text mode: written medium: graphology channel: page, printed

We can interpret this as a continuation of the development that started with the emergence of writing (as discussed before). As noted earlier, writing emerged in contexts complementing those in which spoken language was used, supplementing or replacing an earlier pictorial semiotic system—contexts of administration and book keeping in which meanings needed to be "archived." The settings of values within field, tenor, and mode were thus initially highly constrained to quite a limited range. However, since then these constraints have gradually weakened, and the field, tenor, and mode potentials have themselves expanded as writing moved into an ever-wider range of new contexts, new written registers emerged and the number and nature of the members of a society having access to these registers increased as literacy spread with new developments in education and as printed documents became increasingly important as a commodity of commerce and as an instrument of control by state and church (cf. Burke, 2000).

Written texts and images have existed together in many registers since the emergence of writing, and inscriptions are an integral feature of many sculpted objects and architectural artifacts (e.g., stelas erected by rulers in Mesopotamia recording their achievements). In Sumer, records of transactions and other types of bookkeeping changed from drawn records to written records; but in Egypt, writing was more multiregisterial from an early stage, and there were many displays that were fully multimodal with writing and drawing as integral parts of the display (see e.g., Jean, 1992, pp. 31–37). The boundaries between writing and drawing may be fuzzy; but the semiotic distinction between them is clear—both in terms of what they represent (see Fig. 1.3) and in terms of their inherent potential for being read (writing can be read aloud, but drawing cannot—it must be "translated" into language first).

This kind of multimodality with image and written text seems to have a feature of "literate" cultures in general. However, there have been periods when the division of labor between text and image has changed in a significant way. Let me just give two examples from western history, both taken from Olson (1994). The first concerns the change in the division of labor between text and image in Bibles and Books of Hours, with text gradually taking on a more primary role.

Even the remarkable illuminations framing medieval Bibles and Books of Hours can be seen as non-verbal devices for indicating how the accompanying texts were to be understood. Morrison (1990) has provided a fascinating account of the changing role that these illuminations played in the expression of meaning in the twelfth century. Earlier illuminations were a critical part of the text. The interpretive assumption was that the words together with the pictures presented modes or components of information which when pulled together by the reader provided an epiphany, a revelation, of the true meaning, which ultimately was God. Consequently, texts were written in a "nuclear" manner consisting of a series of points or parts which, while not related logically, led the viewer or reader to form a synthesis. In the eleventh and twelfth centuries illumination changed its character in that images became subordinated to the text which came, increasingly, to be seen as the primary conveyor of meaning. And with the emphasis on the text came, Morrison suggests, the greater concern with the logical form of a text, with single, clear lines of argument and universal unambiguous meanings of terms. This linearity was also reflected in the development of a single linear perspective in art. Thus advances in the visual arts reflected changes in ways of reading . . . (p. 112)

The second example comes from scientific writing about plants; Olson (1994, p. 225) quotes an observation by Boas (1962) regarding the shift from image to text in this register:

Illustrations delighted the eye and supplemented the text; but in botany and anatomy they did more, for they could convey what words, as yet insufficiently subordinated to technical needs, could not. There was as yet no technical language accurate in meaning and universally known, fit to explain in detail the necessary description of form; in fact botany dispensed with pictures when, in the eighteenth century, such a technical language was developed.

The development of scientific registers in the period of early modern science has been discussed by Bazerman (1988), Slaughter (1986), Halliday (1988) and Swales (1990).

According to Kress and van Leeuwen (1996, pp. 21ff), the last few decades have shown another change in the division of labor between text and image in a range of registers, including the "macro-registers" of news, school science textbooks, and books for young children. In reference to school science textbooks, they describe this as a "dramatic shift from the verbal to the visual." This change in the division of labor predates the Internet and the creation of the World Wide Web by several decades; but these more recent technological developments are, of course, now contributing to the change, aided by multimedia technologies. These changes have affected mode in the first instance.

Mode has expanded with the emergence of new semiotic technologies, printing technology being the critical one in our discussion. The nature and roles of the "printed page" are changing with the emergence of new semiotic technologies of printing available to small organizations or even individual persons—technologies such as mimeographing, photocopying, desktop publishing, and "virtual" publishing through the World Wide Web. For example, WHO's *Weekly Epidemiological Record* is now available in portable document format for free as a download, as well as in ordinary print format for an annual subscription. This illustrates a general trend for various publications. Electronic versions are made available free of charge or at the lower price than that charged for the printed version; and the user then decides whether to print out the document locally or not. But new semiotic technologies relating to the spoken mode are also changing the role of the written page, as are the technologies of modal cross-over—text-to-speech and speech recognition systems.

The "printed page" is thus a matter of mode in the first instance—written mode, graphological medium, and print channel, although as we have seen, there are clear implications for field and tenor as well. Mode is concerned with the role played in a context by language and other semiotic systems in relation to one another and in relation to social processes. As shown in Fig. 1.9, the subsystems of mode relate to the stratal subsystems of language and other semiotic systems: Rhetorical mode relates to context and the content systems (semantics in the first instance), medium to the expression systems of phonology and graphology (in the case of language),
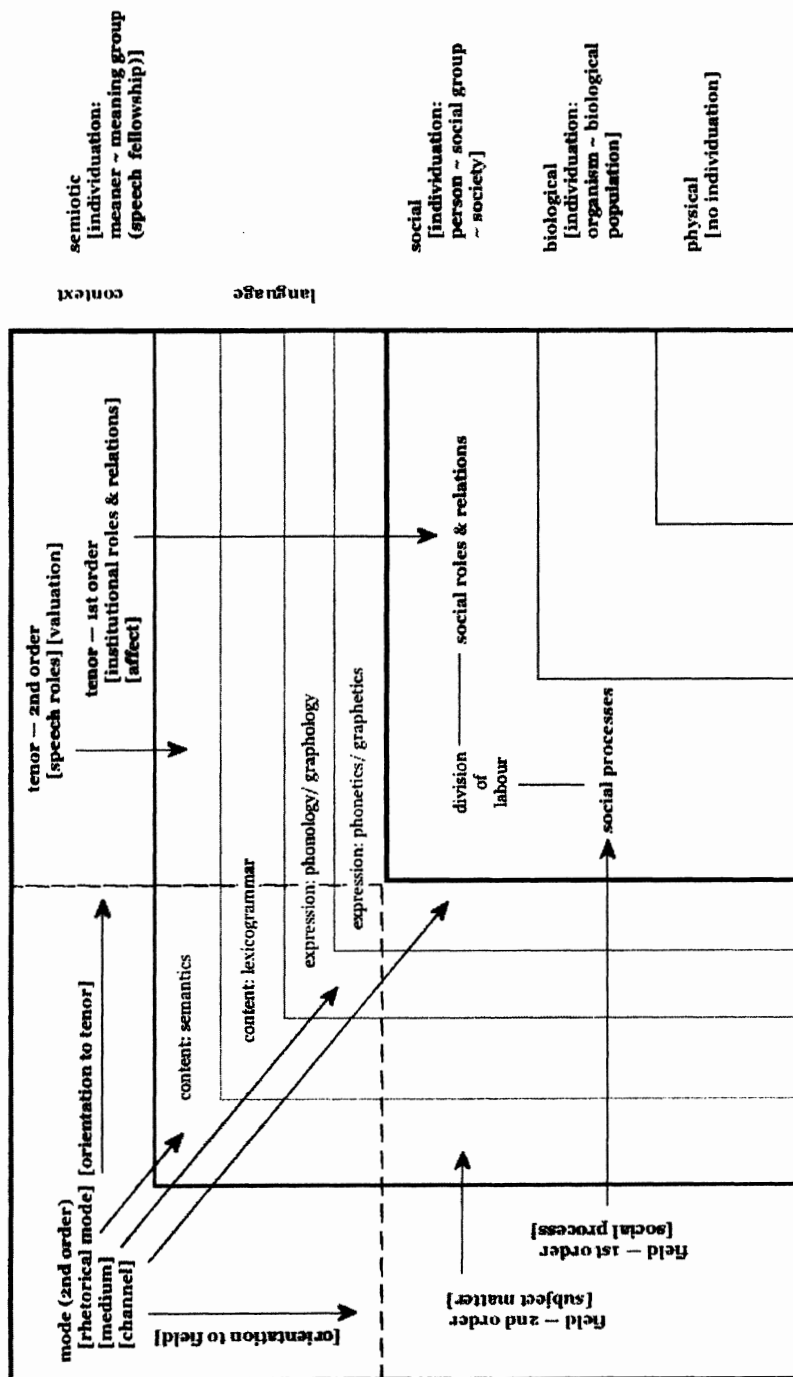
FIG. 1.9. Mode—the role of semiotic systems in context.

and channel to their manifestation in phonetics and graphetics (in the case of language).

Rhetorical mode is concerned with division of sociosemiotic labor—with the division of labor between semiotic systems and social systems within a given context, with the division of labor between field and tenor as the main source of organization, and with the division of semantic labor among the semiotic systems involved in the creation of a multimodal text and also with the relations that link these different semiotic contributions together in the text. Rhetorical mode thus serves to allocate the responsibility for content to different semiotic systems, just as medium and channel serve to allocate the responsibility for expression to different modalities. This allocation is represented diagrammatically in Fig. 1.10 and the primary variables are set out in Fig. 1.12 on page 37. As the diagram in Fig. 1.10 indicates, the resource for linking text and image are rhetorical relations—relations of projection and expansion. We now turn to this resource.
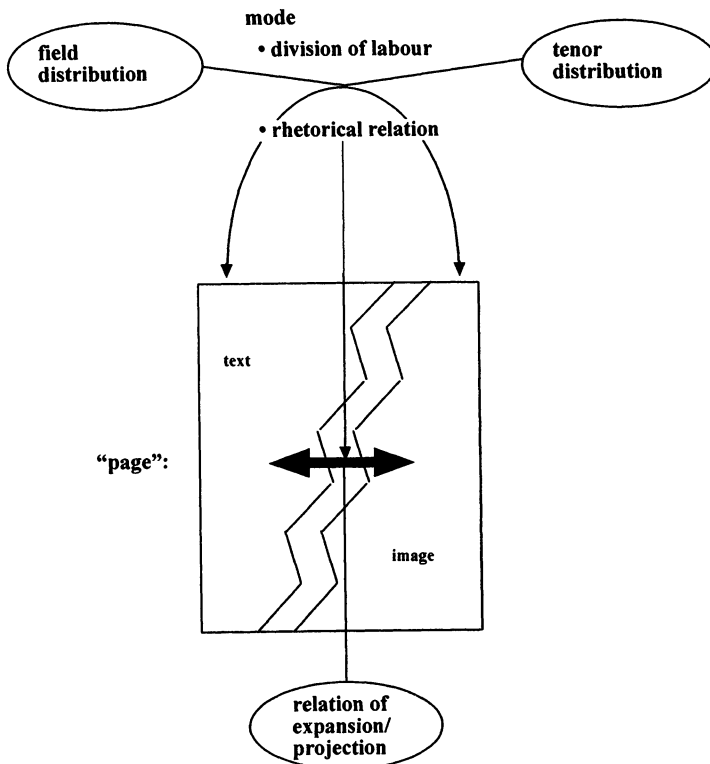
FIG. 1.10.   The "printed page"—context, division of labor, and rhetorical relations.