An Introduction to the Study of Morphology

Vit Bubenik

An Introduction
to the Study of Morphology

Vit Bubenik

Printed in E.C.

Printed on chlorine-free paper

PREFACE

This introductory textbook to the study of linguistic morphology is based on four previous versions of a manuscript entitled *An Introduction to the Study of Morphology*. They were published in a mimeographed form by Memorial University of Newfoundland (St. John's, Canada) in 1978, 1982, 1986 and 1997, and were used at the third-year level in the Department of Linguistics.

Its current version is designed for use as a second- or third-year university level introductory textbook to linguistic morphology. Before taking this course, students should have previously completed one or two introductory courses to the whole discipline of linguistics at their first or second year at the university.

Its argumentation is built around the major turning points in the recent history of morphology linked with European and American scholars such as C. Hockett, P. H. Matthews, J. Bybee, W. Dressler, A. Spencer, A. Carstairs-McCarthy, M. Aronoff, and others. Its primary data are taken from representative Indo-European (English, German, Spanish, Latin, Greek, Russian, Sanskrit), Afro-Asiatic (Hebrew, Arabic, Berber) and several other languages (Turkish, Chinese, Algonkian and others).

The book consists of ten chapters explicating fundamental principles of morphology by means of (numbered) examples. All chapters (with the exception of the last one) are equipped with a number of pertinent exercises often arranged in the order of increasing difficulty. Its contents are as follows:

1. Introduction
2. Grammatical Units (words, morphemes, clitics)
3. Paradigmatic and Syntagmatic Relations
4. Inflectional and Derivational Morphology
5. Inflectional Categories Associated with Nominal Elements
6. Inflectional Categories Associated with Verbal Elements
7. Morphosyntactic Properties and their Exponents
8. Morpheme and Allomorph
9. Derivational Morphology (derivation and compounding)
10. Theoretical Models of Morphology

For pedagogical purposes it is necessary to deal with subject matters in individual chapters as consisting of several units (indicated by subheadings). Recommended Readings at the end of each chapter should provide further ammunition to both instructors and students of this course.

ii

During my twenty years of introducing the subject of linguistic morphology to third-year students of linguistics, languages, psychology, anthropology, sociology and other disciplines of Humanities and Social Sciences I benefitted enormously from various comments and suggestions made on the intermediate versions of the present textbook by my colleagues and students. At this point I want to acknowledge advice of and many helpful comments by the following scholars: Dr. A. Bartoněk (University of Brno), Dr. A. Erhart (University of Brno), Dr. J. Hewson (Memorial University of Newfoundland), Dr. B. Joseph (State University of Ohio), Dr. Stanislav Segert (University of California at Los Angeles), Dr. K. Strunk (University of Munich), Dr. H. Paddock (Memorial University of Newfoundland), Dr. H. Petersmann (University of Heidelberg), Dr. L. Zgusta (University of Illinois).

Many of my students during the 80's and 90's made a number of observations and suggestions on the style of the four previous versions, the clarity of their exposé and the level of difficulty of some of the exercises: Julie Brittain, Audrey Dawe, Barbara O'Dea, Kathy Francis, Margot French, Bernard Kavanagh, Angela Kotsopoulos, Dorothy Liberakis, Christa Lietz, Snezana Milovanovich, Sarah Rose, Donna Starks, Margot Stuart, and others. Many thanks for focusing my attention on the student point of view in composing this textbook.

And finally, I am grateful to three graduate students who formatted the fourth edition (1997) of the manuscript: Henry Muzale, Natasha Squires and Valeri Vassiliev. My special thanks are due to my research assistant Lawrence Greening who has been involved in editing, final text formatting, indexing and preparing a camera-ready copy for publication by Lincom Europa.

St. John's, April 1999

Vit Bubenik
Department of Linguistics
Memorial University of Newfoundland

PRELIMINARIES

Morphology in this book will be defined as that subdiscipline of linguistics whose subject matter is (i) grammatical **units** (morphemes and lexemes) and (ii) grammatical **categories**. The latter are traditionally divided into **primary** grammatical categories (i.e., 'parts of speech' such as nouns, verbs, pronouns, adjectives, adverbs) and **secondary** grammatical categories (such as nominal categories of gender, number and case, and verbal categories of person, number, tense, mood, aspect and voice). **Morphemes** are traditionally defined as the smallest meaningful elements in a language.

In the seventies the transformational-generative view of morphology as a section of syntax with its emphasis on relational aspects of language led to a neglect of the study of grammatical units and categories qua **forms**. However, it should be made clear that all the above mentioned grammatical units and categories can be studied most legitimately in three manners: morphological (or 'formal'), functional, and syntactic (or 'positional'). Any attempts to disregard formal aspects of language by overemphasizing functional or syntactic aspects are detrimental.

Inspection of various introductory books on linguistics will reveal another aspect of the current neglect of morphology. Given the fact that the English morphological system is rather poor compared with that of, say, Spanish or Latin, these books concentrate on the phonemic aspect of morphology (phonological conditioning of allomorphs). Of course, it is important to discuss such facts as the allomorphy of the 3rd Sg Pres /s/ ~ /z/ ~ /əz/ in English (in *he walks, loves* and *poaches*); this, however, should not detract our attention from the morphological aspects of the categories of person and number in Spanish, which display six different inflectional forms for three persons and two numbers (*amo, amas, ama, amamos, amais, aman*). Thus for Spanish, our task will be to account for accentual shift (*ámo ~ amámos*) in terms of morphological categories such as **stem** and **thematic vowel** (and phonological categories such as **penultimate syllable**). Furthermore, it is necessary to consider any linguistic structure as possessing two aspects, namely **syntagmatic** and **paradigmatic**. It is the latter aspect which was completely discarded by transformational-generative grammar, but which nevertheless is a proper domain of morphology. In the following chapters we will spend a lot of time on analyzing and constructing paradigmatic sets for the above mentioned grammatical units and categories. This approach to morphology is known as the **Word and Paradigm Model** (cf. Hockett 1954, Robins 1959) and this model is especially suitable for the analysis of inflectional languages which are morphologically complicated in that they do not always display a one-to-one relationship between morpheme and sememe (polysemy and polymorphy). The other morpheme-based approach, known as **Item and Arrangement Model**, is suitable to the analysis of agglutinating and polysynthetic
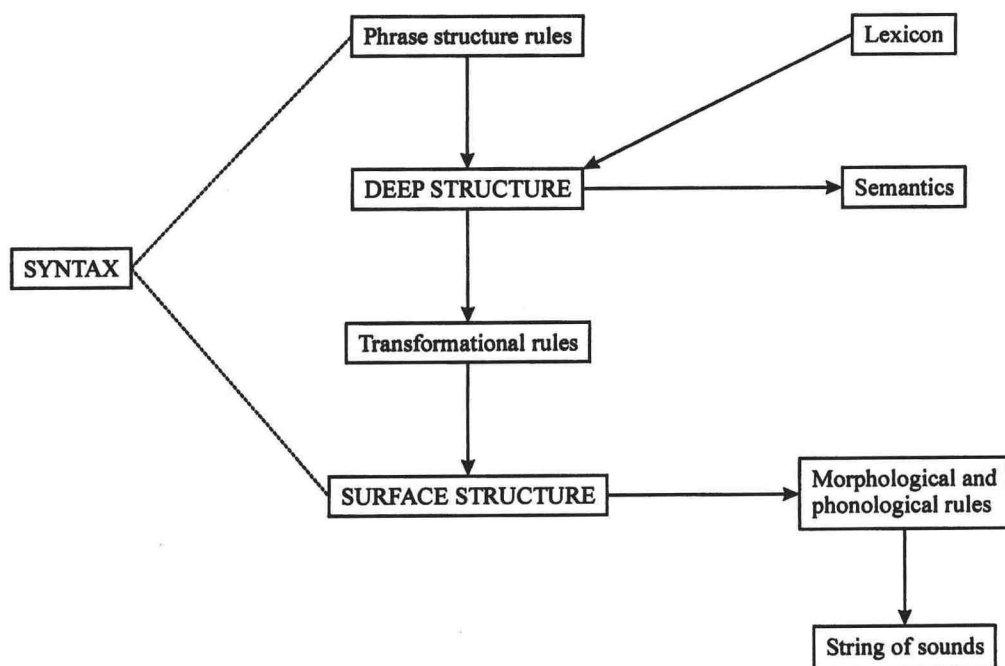
```
        ┌──────────────────────┐                    ┌──────────┐
        │ Phrase structure rules│                    │ Lexicon  │
        └──────────────────────┘                    └──────────┘
                    │                               ╱
                    ▼                             ╱
              ┌──────────────┐               ┌──────────┐
              │ DEEP STRUCTURE │──────────────▶│ Semantics │
              └──────────────┘               └──────────┘
  ┌────────┐         │
  │ SYNTAX │         ▼
  └────────┘   ┌────────────────────┐
              │ Transformational rules│
              └────────────────────┘
                    │
                    ▼
           ┌──────────────────┐        ┌─────────────────────┐
           │ SURFACE STRUCTURE │───────▶│ Morphological and   │
           └──────────────────┘        │ phonological rules  │
                                        └─────────────────────┘
                                                 │
                                                 ▼
                                        ┌─────────────────┐
                                        │ String of sounds│
                                        └─────────────────┘
```

Fig. 0.1 An earlier Transformational Model of language

languages. In these languages the segmentation of words does not present any major problems, since the morphemes and sememes are mostly in one-to-one relationship.

It should be mentioned that the earlier **Transformational Model** of language did not make any provision for the formal study of primary and secondary grammatical categories. These entities were taken for granted and the emphasis was laid on the study of transformational processes. Morphology was thus viewed only as a 'surface syntactic information', as shown in Figure 0.1.

In the eighties, with de-emphasis on the transformational component the place was made for meaning-based approaches to morphology. Linguists returned to a more traditional concept of morphology as a study whose domain is the relation between meaning (semantics) and form (morphology proper). Among the earlier studies along these lines, J. Bybee's *Morphology* (1985) has the lasting merit of freeing the morphological theorizing from genetic and areal biases (her hypotheses about inflectional morphology are based on a sample of fifty languages). In the eighties another approach to morphology gained prominence under the title of **Natural Morphology** in imitation of the title **Natural Phonology** (Hooper's *An Introduction to Natural Generative Grammar*, 1976). It was developed in Germany and Austria by W. Dressler and his co-workers, and is available in the collection of their articles entitled *Leitmotifs in Natural Morphology* (1987). Dressler operates with several explanatory principles (universals, typology,

system-dependency, paradigmatic structure and naturalness). The relationship between expression and meaning (Saussure's *signifiant* and *signifié*) remains the main concern. In addition, Dressler emphasizes the role of linguistic types as mediating between universal principles and language-particular behavior (universal principles of naturalness vs. system-dependent naturalness). One of the central concerns is the nature and organization of inflectional classes (the 'conjugations' and 'declensions' familiar from the traditional descriptions of many languages).

The influence of these ideas changed the study of formal syntax which in the eighties avoided the treatment of purely morphological phenomena and focused instead on the so-called interface questions such as the relation between morphology and syntax or that between morphology and phonology. To follow this change of mind one may consult Jensen (1990), Spencer (1991), Carstairs-McCarthy (1992), Aronoff (1993). Aronoff's pragmatic title, *Morphology by Itself,* marks the complete turn-about in the attitude of Generative Grammar towards morphology in that the latter is now considered not merely as an appendage of syntax and phonology; rather the author insists that linguistic theory must allow a separate and autonomous morphological component.

The reader of this manual might be surprised by the wealth of data included. This has been done on purpose, since I share Bybee's conviction (1985) that morphological universals cannot be fruitfully investigated unless we are willing to examine parallel areas of the grammars of individual languages. Morphology, of course, represents the biggest challenge to universalists' hypotheses since it is precisely here where languages differ most. Thus an important aspect of any course in morphology should be a practical and theoretical experience of analyzing phenomena which are foreign to English. Previous knowledge of the languages to be discussed is not presupposed, but the author hopes that this course will foster interest in their study.

Given the recent history of morphology, it is no surprise that there are only a few textbooks introducing linguistic morphology. The studies quoted above are not suitable for a second or a third year university course. Among earlier studies Matthews' *Morphology* (1974) has the merit of having been unique in pursuing word-based morphology independently of the generative concerns of the seventies. More recently, Bauer (1988) attempted a synthesis in the light of the influence of Natural Morphology on the field. Bauer's monograph provides both the general background to a number of morphological studies and various details of several theoretical approaches.

Neither Matthews (1974) nor Bauer (1988) contain any exercises which are essential to further progress in this field.

RECOMMENDED READINGS

Aronoff, Mark. 1993. *Morphology by Itself: Stems and Inflexional Classes*. Cambridge, Mass.: MIT Press.

Bauer, Laurie. 1988. *Introducing Linguistic Morphology*. Edinburgh: Edinburgh University Press.

Bybee, Joan. 1985. *Morphology: A Study of the Relation Between Meaning and Form*. Amsterdam: Benjamins.

Carstairs-McCarthy, Andrew. 1992. *Current Morphology*. London: Routledge.

Dressler, Wolfgang U., W. Mayerthaler, O. Panagl & W. U. Wurzel, eds. 1987. *Leitmotifs in Natural Morphology*. Amsterdam: Benjamins.

Hockett, Charles F. 1954. "Two models of grammatical description". *Word* 10.210–231.

Hooper, Joan B. 1976. *An Introduction to Natural Generative Phonology*. New York: Academic Press.

Jensen, J. 1990. *Morphology*. Amsterdam: Benjamins.

Matthews, Peter H. 1974. *Morphology: An Introduction to the Theory of Word-Structure*. Cambridge: Cambridge University Press.

Robins, Robert H. 1959. "In defense of WP". *Transactions of the Philological Society* 57.116–144.

Spencer, Andrew. 1991. *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Oxford: Blackwell.

CONTENTS

CHAPTER ONE
INTRODUCTION

1.1 *Language and its Units*

Human language is a particular kind of sign system which bridges two areas of the nonlinguistic universe: non-linguistic real (or imagined) world, i.e. the things we talk about, on the one side, and physical speech sounds produced by human speech organs, on the other. Put differently, language is a mechanism that connects meaning with sound.

Various linguistic schools differ in the number of language levels (subsystems) they posit. Even the number of units assigned by various linguistic schools to each linguistic level is far from being agreed upon. Since the purpose of this book is not to argue for any particular linguistic school, we will simply enumerate and briefly characterize the concepts which appear in most European and American writings. Most linguists, no matter of what persuasion, recognize the following units: **distinctive features, (allo)phones, phonemes, morphophonemes, (allo)morphs, morphemes, lexemes** (words), **(allo)semes** and **sememes**. The first three may be called phonological units; morphs, morphemes and lexemes may be called grammatical units; sememes represent 'semological' or commonly semantic units.

(1)   Language Levels (Subsystems) Units
    (i)   phonology   distinctive features, phones, phonemes
    (ii)   morphology   morphs, morphemes
    (iii) lexicology   lexemes
    (iv)   semantics ('semology')   sememes

The phoneme has been defined as a family (class) of sounds in a given language that function as one and to which the speakers react as one sound. The members of this class are (allo)phones, which occur in mutually exclusive phonetic environments, and which share at least one phonetic feature. Phonetic features are building blocks of phones (e.g., /g/ is a 'bundle' of closure, velarity and voice). Two phones are said to be in contrast if they occupy analogous slots in two different morphemes or lexemes, i.e., if they occur in paradigmatic distribution, such as *fine* vs. *vine*. On the other hand, this opposition does not necessarily hold on the morphophonemic level, e.g., *knife* vs. *knive-s*. Here the allomorphs /najf/ ~ /najv/ belong to the same morpheme {najf} and the same lexeme *knife* whereas /fajn/ and /vajn/ are two different morphemes {fajn} and {vajn} and two different lexemes *fine* and *vine*. Thus allomorphs are not only held together by **morphophonemes**, implemented by phonemes, but they are also linked to the same semantic unit: **sememe**. Morphemes are the universal units of grammatical analysis and they are established on

a semantic and distributional basis. For instance, *go* and *wen-(t)* are usually grouped together into one morpheme {go} because both mean "go", and distributionally they behave in exactly the same way as *sleep* and *slep-(t)*. However, there is no regular morphophonemic tie between the former pair whereas there is one in the latter case in the sense that there are more examples of the alternation /i/ ~ /ɛ/ as in *weep* and *wep-(t)*; consequently, /go/ and /wen/ should not belong to the same morpheme {go}. Here we witness that two different morphemes {go} and {wen} can represent the same semantic unit. This fairly well-known phenomenon, neglected by earlier theoretical treatments of morphology, is called **suppletion** or **polymorphy**. The opposite phenomenon is called **polysemy**. These phenomena are shown in Figure 1.1. For instance, in English the morpheme {s} (= /s/ ~ /z/ ~ /əz/) represents the 3rd Pers Sg of verbs, and the possessive and plural on nouns. In Arabic the same **discontinuous morpheme** /i–ā/ may represent the singular in *kitāb* "book" and plural in *kilāb* "dogs" (singular *kalb*). In other words, morphology and semantics are independent of each other even if they were collapsed in many introductory textbooks to linguistics. What is of particular interest in the study of morphology is the nature of the link-up between morpheme and sememe in the linguistic sign; it may be one-to-one but also two-to-one or one-to-two. This is illustrated in Figure 1.2. It should be emphasized that systematic confrontation of **morphemes** and **sememes** (the smallest elements of the semantic content of language) was done mostly by structuralist linguistic schools, whereas it was neglected by generativists, who concentrated more on relational aspects of language and tended to disregard units in favor of rules. Also it should be mentioned that the background for distinguishing morphology from semantics was provided a long time ago by the work of linguists dealing with typology of languages. In one type of language, commonly denoted as **agglutinating** (e.g., Turkish) each sememe is expressed by a separate morpheme, while in another type, called **inflectional** (e.g., Latin), one morpheme can express more than one sememe. Consider the inflectional forms of the word for "man" in Latin and Turkish given in (2):

(2)	Latin	Spelling	Turkish	Spelling
	vir "man"		adam "man"	
	vir - ī	virī	adam - Ø - ın	adamın
	Sg/Gen		Sg Gen	
	vir - ōrum	virōrum	adam - lar - ın	adamların
	Pl/Gen		Pl Gen	

In the Latin form *vir-ī*, *-ī* expresses two sememes (grammatical meanings) namely the singular and the genitive case; *-ōrum* expresses the plural and the genitive case. Here the relationship between morphology and 'semology' is one-to-two. On the other hand, in Turkish, each sememe is expressed by a separate morpheme: *-Ø* (zero) expresses the singular, *-lar* the plural and *-ın* the genitive case. The relationship between morphology and 'semology' is here one-to-one.

MORPHEME₁   SEMEME   MORPHEME   SEMEME₁

MORPHEME₂   SEMEME₂

Fig. 1.1 Polymorphy and polysemy

M   S   M — S   S   M
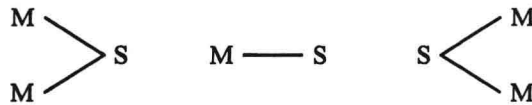
M   M

Fig. 1.2 Morpheme and sememe in linguistic sign

The distinctive features of sound have been studied extensively since Trubetzkoy's and Jakobson's pioneer work in the thirties. They are relatively easy to study because they are only a few (between twelve to seventeen in most languages). The **distinctive features** of meaning are parallel to phonetic distinctive features, but they are much more numerous and consequently much more difficult to study. Nevertheless, much has been accomplished at the level of semantics by so-called **componential analysis** limited to a few areas of lexicon, such as kinship terminology, animals, colors, etc.; there are many more semantic areas which are notoriously difficult to decompose into their semantic features. Consider, for instance, how the semantic features [+male], [+female], and [+young] combine with generic meanings of animal species, as shown in (3).

(3)	Generic Meaning	Male	Female	Young
	horse	stallion	mare	foal
	goose	gander	goose	gosling
	dog	dog	bitch	puppy
	cat	tom-cat	cat	kitten
	man	man	woman	child

The independence of morphology and semantics becomes quite clear in that the same form can represent two meanings: generic and male (*dog, man*), generic and female (*goose, cat*).

1.2 *Units and Rules*

It should be kept in mind that linguistic units cannot exist in language without rules governing their distribution. Both units and rules (or 'items' and their 'arrangements') are equally important in any serious attempt to describe the functioning of language. Their mutual

relationship is of a complementary nature, i.e., it is misleading and detrimental to try to order them hierarchically, or to over-inflate either the entitative component (unit) or process component (rule) in linguistic descriptions. The study of phonotactic rules (constraints on phonological sequences) is a domain of phonology; the study of syntactic rules (**lexotactics** or rules governing distribution of words in sentences) is a domain of syntax. In morphology we will be dealing with **morphotactics**, rules of word formation. **Derivational morphology** (derivation proper and compounding) is currently treated with a strong bias towards morphophonemics; it will be shown that the semantic aspects of word formation are equally important and interesting.

1.3 *Language and its Symbolic Aspect*

We may start this section by examining one of the many problem-ridden definitions of language (Wardhaugh 1972:3):

A language is a system of arbitrary vocal symbols used for human communication.

In view of our discussion above it is preferable to view language as a 'system of (sub)systems' (with 'levels' such as phonology, morphology, lexology, semantics). The above definition makes no provision for the societal and cultural aspects of language. The term vocal in the definition over-emphasizes the fact that the primary medium of language is sound and that writing is only a secondary representation of the primary speech. Let us now examine the remaining term **arbitrary symbols** which brings us back to the Saussurean concept of the **linguistic sign**. According to Saussure the linguistic **sign** is made up of **signifier** and **signified**: *signe = signifiant + signifié*. It may be remarked that the Saussurean dichotomy continues a respectable tradition of semantics starting in Ancient Greece with the Stoics that had an identical dichotomy σημαῖνον /sēmaînon/ plus σημαινόμενον /sēmainómenon/ (σημαίνειν /sēmaínein/ "signify"). The basic assumption here is the *word* (i.e., the basic unit of syntax and semantics) as a linguistic sign composed of two parts: the *form* of the word (signifier) and what is meant (signified), or its *meaning* (concept). It will be shown in chapters dealing with inflectional morphology that the form of a word must be distinguished from its inflected ('accidental') forms which the word assumes when it functions in the sentence. It must also be mentioned that this terminology can be confusing since the 'form' of a word (signifier) could be taken to 'signify' both the 'concept' (mental image) and the 'thing' itself (referent). As is well-known, there exists extensive scholastic literature bearing on the relationship between 'concepts' and 'things', but all this is only of marginal interest to linguists. However, we have to keep in mind that the domain of linguistic meaning does not include the referent. Obviously, we can deal with 'things' themselves only by means of 'concepts', as expressed by the scholastic dictum *voces significant mediantibus conceptibus* "words signify by means of concepts". Hence the line between **form** (signifier) and **referent** (thing) in the famous 'semiotic' triangle reproduced in Figure 1.3 is only dotted.
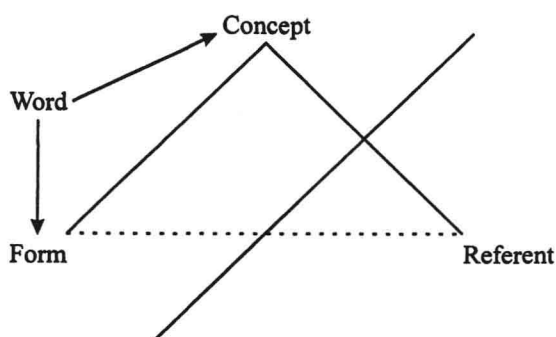
Concept

Word

Form   Referent

Fig. 1.3 Semiotic triangle

The relationship which holds between words (as units of linguistic meaning) and things (i.e., their referents) is the relationship of **reference**. Linguistically, words can be viewed as forms signifying **concepts**, extralinguistically (i.e., referentially) as linguistic signs referring to, or naming, extralinguistic things.

In explaining the nature of the sign, Saussure states that it is arbitrary in that one signified will have different signifiers in different languages, and almost all these signifiers were 'chosen' arbitrarily. Linguistic signs or symbols have to be learned when one acquires one's language, since they are based on a learned conventional relation; in most cases, the names we give to things are conventional, not of natural origin. However, there are two other types of linguistic signs (as defined by linguists working in semiotics), namely **icons** (literally 'pictures') and **indexes** which have to be defined referentially. **Icons** express mainly formal, factual similarity between the meaning and the form; in icons, there is physical resemblance between the shape of the sign and its referent (here, the line between form and referent is solid rather than dotted). Onomatopoeic words like *bang, thump, roar,* etc. are examples from English for this phenomenon of direct representational connection between a word and something in the 'real' world. As is well known, all languages possess highly iconic words by which speakers try to imitate the sounds of nature. **Indexes** express mainly factual, existential contiguity between meaning and form. The indexical features of language include relational concepts of place and time such as *here - there, now - then, I - you - he, this - that.* Their reference is multiple (e.g., *you* can theoretically refer to millions of addresses) and only other linguistic elements in discourse can disambiguate their meaning.

1.4 *Iconic Tendency of Language*

Onomatopoeic words are only one subcategory of icons, those sometimes called **images**. Linguists working in **semiotics** (the study of signs and sign systems) distinguish two more subclasses, namely **diagrams** and **metaphors**. **Diagrams** are characterized by a similarity between form and meaning that is constituted by the relations of their parts. A classical example