Theorization and Representations in Linguistics

Theorization and Representations in Linguistics

Edited by

Viviane Arigne and Christiane Migette

Cambridge
Scholars
Publishing

Theorization and Representations in Linguistics

Edited by Viviane Arigne and Christiane Migette

This book first published 2018

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-0645-2
ISBN (13): 978-1-5275-0645-9

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Many colleagues have contributed in a variety of ways to bringing this publication about and we are grateful to each and every one of them. Our thanks are due to Nicolas Ballier, Philippe Bourdin, Evelio Cabrejo-Parra, Jean-François Chiantaretto, Pierre Cotte, Philippe De Brabanter, Ilse Depraetere, Jean-Pierre Desclés, Jean-Michel Fortis, Brendan Gillon, Geneviève Girard-Gillet, Jacqueline Guéron, Albert Hamm, James Hampton, Daniel Hirst, Jacques Jayez, Paul Larreya, Philip Miller, Sara Mondini, Eithne O'Neill, Jean Pamiès, Susan Pickford, Graham Ranger, Raphael Salkie, Michael Stubbs and Jennifer Vince. Our gratitude for their help and support also goes to the staff of the UFR LLSHS, notably, Rosa Abdesselam, Maud Loubet, Clara Picone, Houda Schiratti and Bérénice Waty, while the competence of our kind editorial assistant François-Xavier Mas was invaluable in bringing the book to completion. Lastly, we are indebted to the following members of the Pléiade research unit (EA 7338), Frédéric Alexandre, Maruszka Meinard and Françoise Palleau-Papin, as well as Anne-Valérie Dulac, Yann Fuchs, Mathilde Lévêque and Claire Parfait.

INTRODUCTION

VIVIANE ARIGNE
AND CHRISTIANE ROCQ-MIGETTE

The aim of this book is to question the representations related to theoretical frameworks in linguistics and at the same time reach a certain degree of insight into the way natural languages function. Access to theoretical representations can only be gained through an analysis of metalinguistic discourses, so that the volume can, in this respect, be seen as a sequel to our former publication *Metalinguistic Discourses* (2015). However, the contributions assembled here approach the topic of metalinguistic discourses from a new perspective. While still mainly addressing issues internal to linguistics, they also deal with the relationships between linguistics and other disciplines, studying how linguistics interfaces with related fields of study such as philosophy, logic, psychology and other cognitive sciences. The idea of this publication sprang from an international conference held on the subject of metalinguistic discourses in 2015 at Paris 13 University, in which interdisciplinary approaches were encouraged. While some of the chapters are authored by researchers who took part in the conference as speakers and then submitted a written version of their presentation for this publication, others are written by members of the scientific committee who volunteered to contribute their reflections on the conference themes.

The book is divided into two parts. The first part is entitled "Systems of Representation" and questions the very concept of metalanguage. The contributions gathered in the second part, "Linguistics and Cognition," are more specifically dedicated to the relations between linguistics, language and cognition from different angles. Needless to say, considering topics such as representations and cognition, some themes will be seen to recur at various points in the volume independently of the part they belong to. One might cite, for example, the question of categorization or the interdisciplinary approach to linguistics.

In the first chapter, entitled "The Relations between Object-Language and Metalanguage in Formalised and Natural Languages," Philippe de

Brabanter aims at clarifying the manner in which terms borrowed from logic can legitimately apply to natural languages. In the case of formalised languages, logicians assemble metalanguages to construct the former and describe their semantics. The metalanguage, which has its own lexicon, and the object-language therefore belong to two separate hierarchical levels. Things are far less obvious and the distinction blurred in the case of natural languages, as no clear distinction can be made between object-language and metalanguage, the object-language being the set of metalinguistic sentences of a natural language. As lexical items are apt to display various degrees of metalinguisticity, any neat hierarchy such as is found in logic is difficult to establish and this object-language can by no means be considered as a language in its own right. A better, more adequate theory of quotation seems to be a much more desirable objective than attempting to theorize about natural metalanguages. Nicolas Ballier agrees with Philip de Brabanter as regards the "impossibility of an abstract formal language (of the Fregean type)." In the second chapter, "What do Linguists Do when they Write Something: the Art of Stenography," he draws our attention to the role of notation in linguistic theories which, although not a metalanguage proper, is a vital component of linguistic metadiscourse. He develops the concept of stenography, examining the origins of the word understood as short hand and its use in linguistics and semiotics. He then goes on to show that a system of metalinguistic representations is entrenched in its notational procedures. This point is exemplified by the work of eighteenth century orthoepists and their reconceptualization of syllables and Nicolas Ballier compares the stenography of syllable division over a century of notational systems with the explicit metadiscourse found in the preface and entries of orthopeists' treaties. In his chapter "Metalinguistic Discourse in Chomskyan Theory and 'Representation'," Jean Pamiès also answers a need for clarification. After a definition of what is subsumed by metalinguistic discourse, he examines the protean use of the term 'representation' in Chomsky's work through the successive states of the theory. The word 'representation' is thus understood as mental representation, cerebral representation, theoretical representation, symbolic representation, semantic representation or intentional representation. For Jean Pamiès, these different uses of the term are a thread enabling one to follow the evolution of Chomsky's thought. Using and analysing quotations from the latter's work, he shows that, on the one hand, the interface between the different levels of representation cannot be scientifically defined and established and, on the other hand, that the link between mind and brain, absent in the theory, cannot be accounted for by the well-known metaphor of software and

hardware in a computer. In conclusion, it is impossible for Chomskyan theory to qualify as a science on a par with physics: in fact, the theory has become so cut off from observation and empirical data that it has drifted into pre-scientific speculation.

Although the issue of cognition lies at the core of the analysis of representations as seen through linguists' metalinguistic discourses, it is more specifically addressed in the second part of the book. With "Cognitive Linguistics as One of the Cognitive Sciences: a Question of Terminology," Katarzyna Kwapisz-Osadnik presents a review of fundamental notions which are embedded in the history of cognitive linguistics and to be found in a variety of approaches and methodologies. Cognition not being the sole property of linguistics, she also studies the use of these notions in other disciplines, namely philosophy, psychology, mathematics and ethnology. The comparison reveals that identical terms in the lexicon may relate to different concepts, a state of affairs which is quite understandably conducive to a measure of confusion and misinterpretation. Focusing on Anglo-Saxon, French and Polish points of views, she insists on the fundamental holistic nature of the study of linguistic phenomena and the interdisciplinary nature of cognitive linguistics, a stand-point which may lead us to reassess the epistemological status of the various fields of research at work in the study of language. A specific example of a term being used in different scientific domains is that of 'prototype.' Jean-Michel Fortis, in his chapter entitled "Prototype Theory: the Origins of a Theoretical Patchwork," gives a detailed historical account of the origins and development of prototype theory, from its initial elaboration by Rosch in the 1970s to its transfer from psychology to linguistics by semanticists. The early days of the concept were contemporaneous with the categorization of colours, from Lenneberg's to Berlin and Kay's research. It was then that Rosch elaborated her own theory of categorization, extending its scope from colours and spatial shapes to 'semantic categories,' which she analysed in terms of discrete features. The transfer to linguistics occurred in the context of generative semantics, when Lakoff and Ross were working on their project of 'fuzzy grammar.' The version of the prototype theory implemented in this new lexical semantics was, however, drastically reduced to a distinction between basic meanings and, in the case of polysemous words, different and distinct meanings linked by family resemblance. To conclude, Jean-Michel Fortis notes that the concept of prototype has found other applications, in phonology and functional linguistics, as well as in attempts to isolate schematic units abstracted from usage patterns. Semantic categorization is also at the heart of "'That is Conjecture'. On English Assertive Shell-Nouns." Carla

Vergaro studies the metarepresentation of utterance acts, as named by illocutionary nouns in their function for shell-nouns. Basing her research on theoretical insights from cognitive linguistics and an empirical-conceptual approach to verbal communication, she contends that conceptualization can be studied through language, albeit in an indirect way only. In this respect, illocutionary shell-nouns are metalinguistic devices which are particularly appropriate for this kind of study. Nouns that name assertive speech acts are thus selected to conduct a corpus-based study and are then categorized on the basis of their complementation patterns and the potential collocations with determiners. This categorization is supported by descriptive and exploratory statistics. The study shows that both complementation and complement selection are linked to the semantics of the noun, exhibiting a correlation between semantic similarity and distributional similarity. In the final chapter, "Inference Processes Expressed by Languages: Deduction of a Probable Consequent vs. Abduction," Jean-Pierre Desclés and Zatkla Guentchéva also deal with categorization, which, in this instance, concerns mental operations, namely the different types of inference. Logical inference through deduction should not be confused with inference by abduction, the latter leading to the formulation of a more or less plausible hypothesis. As far as deduction is concerned, they underline the necessity to distinguish between a simple type of deduction and deduction that results in a probable consequent, contexts and inference processes being different in either case. The modal notion of uncertainty is shared by both abduction and deduction with a probable consequent, leading to the potential use of modalities such as possibility, probability, improbability, or to the explicit statement that what is expressed is a plausible hypothesis or a probable consequent. Inference is a cognitive process used in scientific reasoning as well as in the everyday use of language. Whereas scientific reasoning tries to make inferential processes explicit, they often remain implicit in the everyday use of language. As a matter of fact, the different types of inference may or may not be grammaticalised, depending on the language used. English lacks this possibility but Albanian, Bulgarian, Pomo or North-Amazonian languages, to name but a few, resort to specific formal devices to express abductive inference based on clues.

This book is an attempt to clarify some of the complex issues raised by theorization in linguistics. Metalinguistic discourses are intimately related to, and dependent on, theoretical frameworks and constructs, and one must also consider that the theory itself may also be seen as influenced by them. Metalinguistic discourses being held in natural languages, their lexicons

exhibit considerable instability, due to the symbolic and dynamic nature of language as well as to the theoretical framework. Taking into account other disciplines in an interdisciplinary perspective may lead to even more instability of the shared lexicon, as has been seen in the cases where cognition and cognitive sciences other than linguistics were concerned. Taking cognition into explicit consideration in the analysis of (meta)language reminds us that language is a symbolic system of representation processed by the human, cognitive, faculty of language. This is why the theme of representation is a fundamental concept in both the analysis of metalanguage and that of ordinary language, and in a number of cases closely associated to the issue of linguistic categorization. Language can probably be said to be a cognitive activity *par excellence* so that, as Katarzyna Kwapisz-Osadnik points out, "language […] proves to be the most direct and natural source of knowledge about human beings in all their intellectual, psycho-affective and behavioral complexity," and linguistics should be given pride of place among all the cognitive sciences.

PART I

SYSTEMS OF REPRESENTATION

CHAPTER ONE

THE RELATIONS BETWEEN OBJECT-LANGUAGE AND METALANGUAGE IN FORMALISED AND NATURAL LANGUAGES[1]

PHILIPPE DE BRABANTER

Introduction

The terms 'object-language' and 'metalanguage' form an inseparable pair,[2] so much so that there is no point in talking of one if there is no other. They were introduced into scientific parlance by logicians, in the early 1930s. Alfred Tarski is usually credited with coining, in 1931 or 1933, the term *metajezyk*, later to be translated into English as *metalanguage* (cf. Rey-Debove 1978, 7; Jakobson 1981, 25; Blackburn 1994, 239). However, in Schilpp (1963, 54), Carnap explains that he used the term *Metasprache* as early as 1931, in unpublished lectures on metalogic given in the Vienna Circle.

Be that as it may, it is clear that logicians of the early 1930s had a genuine need for terms like *metalanguage* and *object-language*. Just as they were striving to set up formalised languages that could do duty as 'languages of science', in the place of what they saw as flawed everyday languages, there had been a realisation that any description or definition of the semantics of formalised languages above a basic level of complexity had to be provided in a different language, on pain of generating semantic paradoxes such as the 'Liar paradox.' This different language had to be of a higher-order than the one whose semantic properties were being described or defined. For instance, for truth—a, perhaps *the*, central property in formal semantics—Tarski devised partial definitions which took the form of bi-conditional statements. These statements were couched not in the formalised language the truth of whose formulas was at stake, but in a higher-order metalanguage. (An illustration is provided in Section 2.1.)

Another motivation for developing hierarchies of languages is remedial in intent. The philosopher-logician Quine noted that logicians and mathematicians routinely failed to observe the distinction between producing sentences of the object-language and making claims about that language (Quine 1940, Chapter 4). This can lead to confusions where absolute clarity is required. These errors, however, can be corrected through strict adherence to the use/mention distinction. The mention of a term or formula is achieved by using a quotation (Quine 1940, 23). Thus, in

(1) **'Boston'** is disyllabic

the subject *'Boston'* is a quotation that names not the city, but the name of the city. Quotation is an unambiguous sign that the sentence that contains it is of a higher order than the term or formula mentioned. As we shall see, quotation is also of central importance in any discussion of the relations between object-language and metalanguage.

1. The relation between object-language (*Lo*) and metalanguage (*Lm*) in the theory of formal languages

Logicians *construct* object-languages and metalanguages. In so doing, they stipulate the features that these languages should display. Object-languages are formalised languages. Metalanguages are assembled in order to construct these formalised languages and/or describe their semantics. More often than not, at least in its initial stages, a metalanguage is informal, usually consisting of a (portion of a) natural language, suitably enriched with additional elements (Carnap 1947, 4; Church 1956, 47).[3]

1.1. The components of Lm

The views expressed in the literature as regards the building blocks of Lm vary slightly, but there is some consensus on the following elements. First, Lm must contain names of, and variables for, the symbols and the well-formed formulas of the object-language (Tarski 1983, 172–173;[4] Carnap 1947, 4; Church 1956, 60). Those quotations, conceived as names, are typically formed by enclosing the Lo symbols or formulas in quotation marks. Second, Lm must include semantic predicates, such as *true (sentence)*, *satisfy*, *denote*, *designate*, etc. (Tarski 1983 passim; 1944, 345; Carnap 1947, 4; Gupta 1998, 266). Third, it must contain either all the Lo symbols and formulas, or translations of these. Some logicians mention

just one of these alternatives. Thus, Reichenbach (1947, 10–11) favours a conception on which Lm also contains words from Lo. Others, like Tarski (1944, 350) or Prior (1967, 230), do not seem to regard the alternatives as significantly different.[5] A third group favour the use of translations as a "precaution against equivocation" (Church 1956, 63; cf. also Carnap 1947, 4).

The various components of Lm can be illustrated by an instance of Tarski's famous bi-conditional statements:

(2) **'Snow is white'** is a true sentence (in English) if and only if snow is white.

The subject is a quotation-name. *True sentence (in English)* is a semantic predicate. The clause that occurs after *if and only if* is either a sentence of Lo, or an extensionally equivalent translation of it. Whatever other words there are (*is*, *if and only if*) are also either Lo expressions or translations of these.

1.2. Sentences and levels of language

For many logicians, Lm ⊃ Lo (e.g. Tarksi 1944, 350). Strictly speaking, the validity of this claim depends on whether one assumes that Lm includes translations of the symbols and formulas of Lo, or these symbols and formulas themselves. At any rate, logicians agree that Lm must be expressively richer than its Lo: it is made impossible to name an Lo element within Lo itself. Any sentence that contains a quotation-name is unambiguously a sentence of Lm. Likewise with sentences including a semantic predicate. What of those that include neither, e.g. *Snow is white*? If one assumes that Lm includes translations of Lo symbols and formulas, then there is no ambiguity as to the level of language: the form of the sentence will always reveal if it belongs to Lo or to Lm. If, on the other hand, one assumes that Lm ⊃ Lo in the strict sense, then there are formulas that could be regarded as being either 'object-level' or 'meta-level.' Those strings are "ambiguous as to level of language," as Reichenbach (1947, 10) puts it. This has two consequences. Thinking in sentential terms, a string like *Snow is white* can be both a sentence of Lo or of Lm. The lexical-level consequence is that one type of 'building block' of Lm sentences, such as *snow* or *is*, is also potentially an Lo element.

Does this ambiguity as to level of language pose problems for the hierarchy of languages? It does not, because there is never any point for the logician to produce an object-level sentence and intend it to be

understood as a sentence of Lm. Recall that Lm is set up to construct or describe Lo, so that any (relevant) sentence of Lm will contain a quotation-name and/or a semantic (or syntactic)[6] predicate.

1.3. Summary of the main properties of the Lo-Lm relation in the theory of formal languages

(i) Both Lm and Lo are languages in their own right;
(ii) Lm is set up to describe/construct *one* Lo;
(iii) Lm has its own 'lexicon.' The members of this set are listed as 'components' in Section 2.1.
(iv) For practical purposes, it is always possible to determine if a component of a sentence belongs to Lo or Lm;
(v) Lo is contained within its own Lm. (However, on a strict interpretation, the relation does not obtain if Lm is conceived of as containing translations of all the expressions of Lo.);
(vi) (In spite of (v)) Lm and Lo occur at distinct levels in a hierarchy of languages;
(vii) For practical purposes, it is always possible to determine if a sentence belongs to Lo or Lm.[7]

The reason why the boundaries and relations are so definite is that formalised languages and their metalanguages are designed in such a way as to display the desirable features just recapitulated.

2. In the theory of natural languages

Logicians, among them those who elaborated the relations between formalised object-languages and their metalanguages, are very much aware that these relations cannot be replicated exactly when it comes to natural languages. To begin with, they point out the flaws of natural languages, such as vagueness, ambiguity, irregularities and their unsuitability as languages for scientific discourse (Carnap 1937, 2; Tarksi 1944, 347; Reichenbach 1947, 6; Church 1956, 47). They also question the existence of a clear boundary between object-language and metalanguage (Reichenbach 1947, 16), readily assuming that the inclusion relation is inverted, with the natural object-language containing its metalanguage (Grelling 1936, 486). Relatedly, they also emphasise what Tarski (1983, 164) called the 'universality' of natural languages, i.e. their ability to talk about anything, which prevents them from being organised along a hierarchy of languages and, accordingly, lays them open to the semantic

paradoxes (Tarski 1983, 164). Tarksi consequently questions the very possibility of constructing an exact definition of truth for natural languages. The only remedy would consist in imposing on everyday languages a division into a sequence of object languages and metalanguages. Such a 'rationalization,' however, would be pointless because it would likely deprive a natural language of the features that make it a natural language. (Tarski, 1983, 267, as discussed by Simmons 2009, 559).

In the rest of the quite substantial Section 2, I look into the applicability of 'talk of metalanguage and object-language' in the context of natural languages.

2.1. Can the relations be explained using two terms?

Let us start with the extreme view that the natural Lo is the same language as its Lm. Several writers seem to hold that view. Thus, Carnap writes that "[o]bject language and metalanguage may also be identical, e.g. when we are speaking in English about English grammar, literature, etc." (1948, 4; also Gupta in Honderich 1995, 555). Initially, I will interpret this view literally, as a genuine identity statement, and examine its consequences.

Like Carnap, let us choose English as an example. The notion that needs to be captured here is 'the metalanguage that uses English to describe the natural language English.' The phrase is cumbersome, and will therefore be shortened to 'the English metalanguage.' Most authors conceive of the English metalanguage not as a language proper, but as a special use or function of English (e.g. Jakobson 1985, 117; Rey-Debove 1978, 7, 9; Gamut 1991, 27), or a particular discourse in English (Rey-Debove 1978, 42). In this context, *language* is used loosely. However, the English metalanguage can also be understood more strictly as a language, provided a set of lexical items is identified that, combined with a finite number of 'formation' rules, can be used to generate the set of all the sentences of the English metalanguage. (Section 2.2.3. is devoted to the 'metalinguistic lexicon,' or *metalexicon*, for short). For the moment, I will be content with pointing out that, based on the current state of our knowledge of English, no one can state a system of rules that, given the English lexicon, would generate all and only the sentences of the English metalanguage. Nor, as we will see in Section 2.4., is there any hope of ever achieving this goal. As a result, no substantial difference remains between the stricter and the looser interpretations, i.e. between 'metalanguage as lexicon + formation rules,' and 'metalanguage as

particular use.' For the sake of simplicity, in what follows, I will continue to talk of the English metalanguage as a language, i.e. as a set of sentences, all the time assuming that whatever point is made in terms of the set of sentences can also be made in 'use' terms.

Now, if English and its metalanguage are identical, then every English sentence must be about English. In other words, English is the only topic that can be talked about in English. This is an absurd consequence that flies in the face of the oft-recognised universality of natural languages, i.e. their ability to talk about 'anything' (e.g. Tarski 1983, 164; Hjelmslev 1966, 175; Droste 1989b, 931).

The view just discussed is unlikely to be seriously entertained by any scholar. Most probably it results from a loose use of *be* in the relevant passages. More promising as a characterisation of the relation between Lo and Lm in the theory of natural languages is a view according to which Lo *contains* Lm. This view, which was already attributed to Grelling above, is widely accepted by linguists (Jakobson 1981, 1985; Harris 1968, 17, 152, 1988, 34–35, 1991, 274–78; Greimas and Courtés 1979, 226; Droste 1983, *passim*, 1989a, 21, 1989b, 930), but does it stand up to scrutiny?

Before we can begin to answer this question, it must be pointed out that this view that Lo ⊃ Lm potentially conceals two claims, one about the English metalanguage *stricto sensu*, the other about English metalinguistic sentences at large.

It seems trivially true that English contains the English metalanguage in the narrow sense. Consider examples (2) above and (3).

(3) ***With*** is a preposition.

Both are unquestionably English sentences, and, using quotations, they make points about an English sentence and an English word, respectively. All the other words involved are themselves English. So, yes, *Lo ⊃Lm* is true on that conception of the English metalanguage. However, that conception is too narrow to be interesting, because the English metalanguage is a mere subset of English metalinguistic sentences. There are a great many such sentences that are about other languages, real or imagined, as in (4) and (5), about pseudo-words (6) or non-words (7), or even about no particular language at all, as with many statements about languages in general or about language (in the sense of 'language faculty'), as in (8) and (9):

(4) Did Galileo really say **"Eppur si muove"**?
(5) **"Grildrig"** was the name the Brobdingnagians gave Gulliver.

(6) This is because ***wug*** [wʌg] was one such pseudoword used by Jean Berko Gleason in her wug test 1958 experiments. (Wikipedia)
(7) It is nonsense to say, **'Wbnjnmrtk.'** (Droste 1983, 687)
(8) Nouns are typically used to denote objects and people.
(9) The V-features of an inflectional element disappear when they check V, the NP-features when they check NP. (Chomsky 1993, 30)

All of the above are English sentences, never mind the non-English elements occurring in (4) to (7). Hence they belong to the set of English metalinguistic sentences. And yet none of them talks about English (or English specifically).

This means that the two terms *Lo* and *Lm* do not suffice to characterise the relations (of inclusion or others) we are interested in. Sentences (4) to (9) are part of the natural language English. But English is not their object-language. A dissociation needs to be made between Lo and natural language. If we take Lm to be the set of English metalinguistic sentences, we see that it comprises English sentences that talk about (i) English, (ii) other (possibly imaginary, or even maybe impossible) languages, (iii) no particular language. Irrespective of whether Lm sentences are about (i), (ii), or (iii), they are all sentences of the natural language English. Hence our discussion cannot dispense with a third term, 'natural language,' or *Ln* for short.

2.1.1. Droste's proposal

In the 1980s, the linguist Flip G. Droste made a proposal that built on these three terms. Unsurprisingly, his account looks upon Lm as a particular use of Ln. What of Lo? Just as I did above, Droste warns against equating Lo with Ln. Instead, he provides a definition of Lo via a set of elements of Lm, the set of proper names that refer to some linguistic entity (cf. the quotation-names in Section 1.1). Of this set of names, Droste says the following: "the proper names constitute a lexicon which is part of Lm or, rather, is essential to the creation of the special use Lm" (1983, 696). To which he adds: "And the items of this lexicon refer to entities which, taken together, define the set Lo exhaustively. The Lo element [e.g. the referent of *'Boston'* in (1)] is an autonym of the Ln element but under no condition whatsoever should it be identified with it" (1983, 696–97).

Several remarks are in order. First, in order to dispel confusion, it must be made clear that Droste uses the term *autonym* differently from Carnap, who coined the term, and Rey-Debove, who wrote extensively about it. Whereas they both regard an autonym essentially as a quotation—Rey-Debove's autonym is understood as a 'metahomonym' of its referent,[8]

Carnap's as a name for itself—Droste's autonym is the referent of that name. I will therefore not adopt Droste's confusing use of the term, and will stick to the phrase *Lo element* in the discussion below. Second, it is not obvious why Droste should warn against confusing Lo elements with Ln elements. Granted, not all the referents of quotations are elements of Ln. Quotations may also, as Droste points out, designate 'quasi-expressions' (cf examples (6) and (7) above). It is therefore correct to say that Lo cannot be the same thing as Ln. However, on Droste's account (and on anybody else's), Lo also contains "linguistic elements such as sentences, words, morphemes, etc." (1983, 697). When those are English sentences, words, morphemes, are they not Ln elements? No, says Droste: "*The Lo elements*, finally, *refer to Ln expressions* or to things which are in the complement set of Ln (where Ln is defined as a set of sentences)" (1983, 697; italics mine). One ends up with the following picture: quotation-names refer to Lo elements, (some of) which in turn refer to Ln elements (while the rest refer to elements from other languages, or to 'quasi-expressions'). Applied to examples (1) and (2) above, this means that the quotations in those sentences refer not to Ln but to Lo elements, say BOSTON and SNOW IS WHITE, which elements in turn refer to the English word *Boston* and the English sentence *Snow is white*.

The postulation of a double layer of reference appears unwarranted. For one thing, it is not economical. For another, it leads to something of a paradox: in using quotation, one does not in fact speak about what one could be expecting and expected to speak about: whenever a quoter intends to say something about an Ln element, and thinks she is, she is in fact talking about an Lo element, which, Droste insists, cannot be the same thing. Access to the intended referents is only gained indirectly, secondarily, as can be seen in the short analysis of (1) and (2) above.

2.2. An amended three-term account

Although Droste's account does not ultimately make the grade, I want to retain three of his crucial insights:

(i) a third term is needed if one is to be able to define Lo and Lm in the context of natural languages;

(ii) the so-called 'object-language' is not a language (cf. Droste 1983, 690), but a set of 'quotables' instead. The first claim is uncontroversial. The second—Lo is a set of quotables—is rather more so. As is shown in Section 2.2.2., this assumption results in a substantial mismatch between Lm and Lo;

(iii) Lm is a particular use of Ln, which I interpret as equivalent to the claim that Lm ⊂ Ln. This claim faces some severe challenges at the lexical level: in particular, it is not evident that quotations, all of which have a reasonable claim to membership of Lm, can be said to belong to Ln too. I propose a novel solution to this problem, which allows me to conclude that Lm does form part of Ln.

The discussion will be broken down into three main sections, each adopting a different angle. In Section 2.2.1., I focus on Ln ∩ Lo, and explain which elements fall inside or outside this intersection. In Section 2.2.2., I try to state the exact relations between Lo and Lm. Though that question may seem to call for a trivial answer—"Lo is the set of referents for Lm expressions… by definition!," things turn out to be more complex after all. Finally, in Section 2.2.3., I pick up all the previous threads and make a proposal compatible with maintaining the claim that Lm ⊂ Ln.

2.2.1. Ln ∩ Lo

I have said above that I endorsed Droste's view that Lo is the set of referents of the quotations that can occur in Lm sentences.[9] I also accept the consequence that Lo is not a language, but a heterogeneous set of suitable referents. I do not, however, agree with Droste that the intersection between Lo and Ln is empty. I argue that Ln ∩ Lo contains all the sentences of Ln, and also the whole Ln lexicon, to be understood in a broad sense, as containing words, but also morphemes, and possibly phrases and constructions, depending on one's theory of the lexicon. Evidence of this is provided by examples like (10) to (13), which quote an English sentence, English words, an English derivational suffix, and an English phrase, respectively.

(10) **"This is yellow"** is a sentence epitomizing these three categories. (*The Corpus of Historical American English (COHA)*, Davies 2010-)
(11) **"This"** represents the uncharacterized particular; **"yellow,"** the unparticularized character. (COHA)
(12) *Vasectomy. It's the* ***-ectomy*** *that puts me off*. (Barnes 2001, 109)
(13) **"Control freak"** is the phrase often used to describe Messier. (*The Corpus of Contemporary American English*, Davies 2008-)

Perhaps more difficult to classify are arbitrary fragments of an Ln string. To give but a few examples (with Ln = English): sub-morphemic or non-morphemic strings like *-chine* (as in *machine*) or *-ervation* (as in

preservation, *conservation*, etc.) (cf. (14)); incomplete supra-morphemic units (*quiet and reserved is seen*), as in (15); strings that are judged unacceptable, as in (16):

(14) Ever noticed that the ***-chine*** of *machine* is pronounced just like *sheen*?
(15) In the sentence *In some cultures being quiet and reserved is seen as a sign of politeness*, the string ***quiet and reserved is seen*** is not a syntactic constituent.
(16) It is incorrect to say, **'Him saw her'** (Droste 1983, 687)

Though (14) to (16) are Lm sentences containing quotations of strings made up of 'bits of Ln,' these strings are none the less not part of Ln, and therefore fall outside Ln ∩ Lo. In the case of (15) and (16), this conclusion rests on the sensible assumption that the referents of *quiet and reserved is seen* and *Him saw her* are these strings taken as a whole, and not each word taken separately. One who adopted the latter view of quotation—Geach (1957) is a rare example—would say the referents of the quotations in (15) and (16) *do* fall within Ln ∩ Lo.

There are many more elements of Lo that do not belong to Ln. A precise categorisation of these must be underpinned by a thorough discussion of what elements deserve the label 'linguistic.' Before I proceed to do just that, however, I will begin by introducing the various cases that I think should be distinguished:

– First, any morpheme, word, phrase, sentence, etc. of *another* natural language, actual or imaginary, is also part of Lo. In (17), the first quotation refers to a word in imaginary Brobdingnagian, the next two to Latin words.

(17) She gave me the name of ***Grildrig***, which the family took up, and afterwards the whole kingdom. The word imports what the Latins call ***nanunculus***, the Italians ***homunceletino***, and the English *mannikin*. (Swift 1960, 108)

– Second, Lo\Ln (i.e., Lo minus Ln) contains strings that comply with English phonotactic rules but are not attested in the productions of native speakers. Take *wug* in (6), or creations like *brillig*, *slithy*, *toves*, *gimble*, *outgrabe*, etc. from Lewis Carroll's *Jabberwocky*.
– Third, Lo\Ln contains any arbitrary sub- or supra-morphemic string from any natural language other than English—the non-English counterparts of the strings quoted in (14) to (16).

– Fourth, what might be called 'non-words' relative to English, such as *Wbnjnmrtk* in example (7) above. These are different from items like *wug* or *brillig* in the sense that they violate the phonotactic rules of English.
– Fifth, strings that are impossible in any natural language [if these can be devised and quoted].

The classification above depends on our ability to distinguish between what is linguistic, what is pseudo-linguistic, and what is non-linguistic.

2.2.1.1. *'linguistic'*

The term 'linguistic' should be defined independently of any particular language: a linguistic item is any string that exists in the lexicon or can be generated by any given grammar. Hence, the quotations of an Italian sentence in (4) and of Latin words in (17) indisputably refer to a linguistic Lo element. One might question whether strings in imaginary 'languages,' as are quoted in (5) and (17), also deserve to be called 'linguistic.' In the case of Brobdingnagian, the few words mentioned by Swift suggest that it works along very similar principles to English or other Indo-European languages. So, we have no trouble regarding *Grildrig* (or *Glumdalclitch*, *splacknuck*, *Lorbrulgrud*, or *relplum scalcath*) as linguistic entities. But one could imagine 'communication systems' that are so different that they do not self-evidently rate as 'linguistic.' A case in point is a version of (written) 'Martian' considered by Bennett, in which the symbols are not "arranged on a directional line" (Bennett 1988, 413). For example, vertical relations between (word-like) clusters of symbols might be relevant at the same time as horizontal arrangements are; or the order of symbols within a cluster might be a matter of aesthetics, with only the *number* of symbols being relevant to meaning. What Bennett is trying to do with this example is to warn the language scholar against parochialism: maybe languages are possible which have completely different organising principles from those that are known to us. Should we therefore dismiss these as not being languages? I certainly cannot answer this question here. But just the possibility of raising it shows that fixing the limits of what deserves to be called 'linguistic' is a tricky affair.[10]

2.2.1.2. *'pseudo-linguistic'*

Unlike 'linguistic,' 'pseudo-linguistic' can only be defined meaningfully relative to a particular language. For the purposes of this paper, the term

will be regarded as synonymous with *pseudo-English*. Psycholinguists call 'pseudo-words' those word-like strings that, although not recorded in dictionaries, comply with the phonotactic rules of the language spoken by the subjects of their experiments. As regards higher levels of grammatical organisation, pseudo-phrases, -clauses and -sentences should similarly comply with the morphological and syntactic rules of the language. On such a definition, Lewis Carroll's *'Twas brillig, and the slithy toves did gyre and gimble in the wabe* (from his *Jabberwocky*) is a pseudo-English sentence containing six pseudo-English words. By the same token, the poet Henri Michaux's *Il l'emparouille et l'endosque contre terre; Il le rague et le roupète jusqu'à son drâle* (from his *Qui je fus*) is pseudo-French.

Quotations of pseudo-English sequences are not infrequent:

(18) "Nothing much, to be honest. Nothing helpful. I was just wondering what **"Feminian"** means."
"…?"
"I wonder if it's a real geological term, or if Kipling just made it up." (Barnes 1982, 119)

(19) Beneath the window is a bilingual rubbish bin with a spelling mistake. The top line says PAPIERS (how official the French sounds: 'Driving licence! Identity card!' it seems to command). The English translation underneath reads **LITTERS**. What a difference a single consonant makes. (Barnes 1985, 82–83)

Neither *Feminian* nor *litters* are linguistic items in English: the first is not recorded in any of the major dictionaries I have consulted, and the second, at the present time, can only be used as a noncount noun.[11] However, given an expansion of the English lexicon and a not unusual change in grammatical status (a different exploitation of the grammatical potential), both pseudo-words would become actual linguistic items.

Note that it is not unreasonable to assume that all lexical and grammatical innovations begin as pseudo-elements, in the sense of "possible but not yet acceptable." Some of them retain that status, while others end up being accepted as part of the relevant language. This dynamic dimension of lexicons suggests that the boundary between pseudo-English and actual English is fuzzy. One source of fuzziness is the fact that the object 'English' is no more than a useful theoretical fiction. It might perhaps be more accurate to write that, at a certain moment in time, some of what I have called 'pseudo-English' strings are genuine members or products of the lexicons or grammars of some idiolects, but still fall

outside too many such lexicons or grammars to be regarded as English in the standard collective sense of the term.

The intersection between pseudo-English (or pseudo-Chinese, etc.) and linguistic items need not be empty. That is because a pseudo-linguistic item with respect to language L_1 might exist as a linguistic item in language L_2. In language contact, interference is rife: the creations of Belgian students of English (*constatate*, *factures*, *tenniswomen*) are pseudo-English strings with sometimes a linguistic existence elsewhere: *factures* and *tenniswomen* exist in the written system of French.

2.2.1.3. *'non-linguistic'*

What do I mean by 'non-linguistic' elements, and why do they matter at all in the current discussion? It is a fact that 'just about anything' can be quoted. In (20) and (21), strings of letters are used to quote noises and musical sounds, respectively. In (22), a transcription of a spoken (and gestured) example, some bodily motions (which are not to be confused with lexical items of a sign language) are used to 'quote' a particular feeling:

(20) M pushes the penis on a gilt cherub which is flying up the mirror frame, and the whole edifice slides back with a great **"gzzhhd."** (*New Statesman*, 20/12/99, 6)

(21) ***PA-PA-PA-PA PUM PUM PUM*** went Sir Jack as Woodie, cap under arm, opened the limo door, "***Pum pa-pa-pa-pa pumm pumm pumm***. Recognize it, Woodie?" (Barnes 1999)

(22) And I'm like [SIGHING AND ADOPTING FACIAL EXPRESSION DENOTING DISAPPOINTMENT].

The quoted elements are not linguistic entities, because none of them is part of a natural language. Here as elsewhere, it is important not to confuse the quotation (which in (20) and (21) resorts to linguistic means—letters, syllables) with the quoted entities. Now, is the quotability of the quoted elements reason enough to include them within Lo? The answer can only be a yes. That is because I have decided to take on board Droste's idea that Lo is a set of quotables. Restricting this set to just linguistic quotables would leave us with the question of what to do with the other quotables. Should they be included in a different set, one that would introduce a fourth term into the present discussion? This seems unnecessary, as, on the current picture, non-linguistic and linguistic quotables are already clearly distinguished: while the latter belong to $Ln_i \cap Lo$ (the intersections of Lo with some natural language), the former do not. Besides, as we saw above,

Lo is not a language. So there should be no problem with it also encompassing sounds, noises and body language. (I return to the consequences that this has on the Lm-Lo relation in Section 2.2.2.)

Just as tricky are occurrences such as *Him saw her* in example (16) or Zellig Harris's example of a meaningless string *slept the a the* (1991, 31). They are neither English nor pseudo-English, since they do not comply with the syntactic rules of English. Yet they appear less non-linguistic than the noises, sounds or body language quoted in (20) and (21). Likewise with *Wbnjnmrtk* in example (7). As initially intended by Droste, *Wbnjnmrtk* is not to be read as a transcription of only the consonants of a spoken expression that would also contain vowels, as occurs for example in Arabic script. *Wbnjnmrtk* is to be understood as complete. As such, it seems to defy the phonotactic rules of any natural language: is *Wbnjnmrtk*, for all that, a non-linguistic object? There seems to be no definitive answer to this question. Bennett's warning against theoretical parochialism (1988, 413–14) might suggest that a fourth category is needed, between pseudo-English and non-linguistic, that of 'pseudo-linguistic items with respect to *any* language.' The rationale would be that there is no theoretical restriction against these strings occurring in Martian, so maybe they are not genuinely non-linguistic. This, however, might put us on a slippery slope to acceptance of just anything as being pseudo-linguistic in this extended sense, since nothing proves that other similar noises, sounds, facial expressions, etc. could not be meaningful parts of a yet-to-be-discovered language. I will refrain from going this way here, but, as my reader can see, this is not a decision that is (or can be) grounded in empirical evidence.

2.2.1.4. *Intermediate recap*

All in all, Lo is a peculiar collection of objects whose core is a linguistic component but whose periphery also includes a variety of pseudo-linguistic and non-linguistic entities. The contents of Lo raise questions to do with the relations between Lm and Lo and with the question whether Lm can be said to be included within Ln. I turn to these issues in the next two sections.

2.2.2. Relations between Lm and Lo

We saw that, in the theory of formalised languages, there were two possible relations between Lm and Lo. Either Lm contains Lo as a part, or Lm contains (exact) translations of all the formulas of Lo. The situation is