Lexical Plurals

A Morphosemantic Approach

PAOLO ACQUAVIVA

Lexical Plurals
A Morphosemantic Approach

OXFORD STUDIES IN THEORETICAL LINGUISTICS

GENERAL EDITORS: David Adger, Queen Mary, University of London; Hagit Borer, University of Southern California

ADVISORY EDITORS: Stephen Anderson, Yale University; Daniel Buring, University of California, Los Angeles; Nomi Erteschik Shir, Ben Gurion University; Donka Farkas, University of California, Santa Cruz; Angelika Kratzer, University of Massachusetts, Amherst; Andrew Nevins, Harvard University; Christopher Potts, University of Massachusetts, Amherst; Barry Schein, University of Southern California; Peter Svenonius, University of Tromsø; Moira Yip, University College London

Lexical Plurals
A Morphosemantic Approach

PAOLO ACQUAVIVA

OXFORD
UNIVERSITY PRESS

Contents

This page intentionally left blank

General Preface

The theoretical focus of this series is on the interfaces between subcomponents of the human grammatical system and the closely related area of the interfaces between the different subdisciplines of linguistics. The notion of 'interface' has become central in grammatical theory (for instance, in Chomsky's recent Minimalist Program) and in linguistic practice: work on the interfaces between syntax and semantics, syntax and morphology, phonology and phonetics, etc. has led to a deeper understanding of particular linguistic phenomena and of the architecture of the linguistic component of the mind/brain.

The series covers interfaces between core components of grammar, including syntax/morphology, syntax/semantics, syntax/phonology, syntax/pragmatics, morphology/phonology, phonology/phonetics, phonetics/speech processing, semantics/pragmatics, intonation/discourse structure as well as issues in the way that the systems of grammar involving these interface areas are acquired and deployed in use (including language acquisition, language dysfunction, and language processing). It demonstrates, we hope, that proper understandings of particular linguistic phenomena, languages, language groups, or inter-language variations all require reference to interfaces.

The series is open to work by linguists of all theoretical persuasions and schools of thought. A main requirement is that authors should write so as to be understood by colleagues in related subfields of linguistics and by scholars in cognate disciplines.

In this volume Paolo Acquaviva tackles the issue of the interaction between grammatical competence and lexical knowledge, focusing on the domain of number. He investigates cases where number is inherent to nouns (rather than being added by the grammatical systems) and argues that this kind of information is truly linguistic, rather than encyclopaedic. The ensuing picture of the interface between grammatical and lexical knowledge implies a certain set of expectations about the typological range of morphology/semantics connections in this domain, expectations which are argued to be met when a cross-linguistic perspective on this interface is taken.

David Adger
Hagit Borer

Acknowledgements

Abbreviations

#P	quantity phrase
1, 2, 3	first, second, third person
ABSTR	abstract nominalization morpheme
ACC	accusative
ADJ	adjective/adjectival affix
ASP	aspect
BP	broken plural (in Chapter 7)
CL	classifier
COLL	collective
COMPAR	comparative
CONJ	conjunctive particle
DEF	definiteness
DIM	diminutive
DP	determiner phrase
Det	determiner
FEM	feminine
FORMAL	formal register (in Korean)
FUT	future
GEN	genitive
IMPF	imperfective
INDEF	indefiniteness
INSTR	instrumental
MASC	masculine
N	noun
NEG	negation
NEUT	neuter
NOM	nominative
NP	noun phrase
PAST	past
PERF	perfective

PL	plural
PRES	present
PROGR	progressive
PRT	particle
REL	relational (of Arabic *nisba* adjectives)
SG	singular
SINGULAT	singulative
THAT	distal deictic
V	verb
VPT	verbal particle (in Breton)

1

Aims and assumptions

1.1 Lexical plurals as a morphosemantic concept

This book is a study in the relation between grammar and lexical competence. Its goal is to analyse how grammatical plurality can be an intrinsic component of certain nouns; or more concretely, to fully explain what it means to say that a noun is plural 'lexically'. The most obvious example and, I will argue, the least revealing, is represented by nouns with a fixed plural value, like *scissors.* Then there are lexically idiosyncratic plural forms, like *pence* from *penny.* Plurals that must be learned as whole word forms, like suppletive stems, also involve knowledge about certain words and not just about grammatical morphemes (affixal or otherwise). But the empirical domain of lexical plurality is much wider. It includes plural doublets and all instances of competing plural alternants, in so far as the choice between them is not automatically determined by grammar but involves choosing between distinct senses. For those who use *mice* for rodents and *mouses* for computer pointers, the choice between the two plurals is no more grammatically determined than that between *cat* and *dog.* Competing plural alternants often differ in form and grammatical diacritics beside meaning, but even when pluralization does not involve morphologically contrasting alternants, it may affect lexical semantics to such an extent that the question whether we are dealing with one noun or two becomes ineludible. It is not so clear that the plurals that appear in *she's got the brains in the family* or *the works of my watch were all gummed up* are inflectional forms of the same words that appear in the singular as *brain* and *work.* After all, if *brain* refers to an organ, *she's got the brains* does not mean that she has many cerebral organs. Or consider a plural like *waters* in *the river discharges its waters into the lake*; surely it does not refer to a set of waters in the same way as *books* refers to a set of books. Does that make it a lexical entry distinct from the singular *water*? Are *pence* (true of units of value) and *pennies* (true of coins) distinct lexical items? It depends on what we mean by lexical item. Once the question is brought into focus, it emerges in a surprisingly large variety of phenomena, many more and more diverse than a few English

examples would suggest. The immediate empirical goal of this investigation is to identify and categorize the mass of morphological and semantic phenomena characterizing plurality as lexical.

A major claim of this book is that the morphological and semantic viewpoint are both necessary to understand what these plurals can tell us about lexicality. Focusing on semantics alone would effectively reduce lexical plurals to those with an idiosyncratic interpretation, disregarding the very important fact that the non-canonical readings often correlate with a particular morphology. Focusing on morphology alone, symmetrically, would lead to a catalogue of idiosyncratic forms, missing the pervasive semantic generalizations within and especially across languages. Only by keeping track of morphology and semantics at the same time does a systematic connection emerge between certain conceptualizations in lexical semantics and certain morphological properties that do not reduce to contextual inflection. Lexical plurality is the linchpin connecting the two; it is the concept unifying all phenomena where plural is part of what it is to know a certain word.

1.2 Lexicality in morphology: stems and lexemes

The investigation deals with those properties of number which are peculiar to nouns. It is not a theory of plural or of grammatical number, but of its use as a constituent of words containing more than grammatical information. For this reason, I will not consider pronouns, even though they are of crucial importance for number systems, and the most thorough recent analyses of number have dealt with them in part (Noyer 1997; Corbett 2000; Harbour 2003; 2007) or exclusively (Harley and Ritter 2002; Moravcsik 2003). For the same reason I will not discuss adjectives either, whose number value is always determined syntactically. By concentrating on plural as an inherent specification, I intend to explore the role of grammar in making up substantive lexical words.

This latter notion is vague in the extreme, and another aim of this book is to make it clearer. When plurality is fused with a lexical base, it affects form and interpretation in ways that give us a window on the properties of that base. A comprehensive and detailed review of what happens when plural is not 'contextual' but 'inherent' inflection (to use the classic formulation of Booij 1994, 1996), allows us to see more clearly what exactly it is inherent in. Plurality may be embedded in that part of the word form which constitutes the input to context-triggered inflection. Calling this entity a *stem*, by itself, implies no theoretical choice; but the evidence will show that plurals of this sort often lack some typical traits of inflection, for instance by allowing competing alternants. When stem-internal plurals display such lexical traits,

they are part of a stem in the technical sense of Aronoff (1994): a form or set of forms referenced as a single entity by an autonomous morphological component, which does not necessarily have a particular meaning, and which spells out that form of a word that is input to syntactically driven affixation.

Beside being inherent in a stem, plurality may also be inherent in the abstract lexical base underlying all inflected realizations of a word, like the abstract BOOK underlies *book* and *books* in English. A lexical base in this sense is a *lexeme*, and many plural nouns, I will claim, are lexical because plurality is an integral part of the lexeme. This notion allows for a much more precise definition than one based on intuitive concepts such as noun, lexical entry, or semantic listeme. In tandem, the concepts of stem and lexeme are essential to clarify the 'lexicality' of a great amount of phenomena, and to trace back the empirical differences between, say, Italian and Arabic 'inherent' plurals to the characterization of plurality as, respectively, content of a lexeme and function of a stem.

1.3 Lexicality in semantics: conceptualization

The concept of lexeme brings in the semantic dimension of lexicality. Nouns are lexemes that refer to entities, and incorporate a conceptualization of the entities they refer to, primarily in terms of unity (whether they constitute discrete wholes) and identity (whether they are intrinsically identifiable). Through the number category, the grammar has a fixed way to affect the structure of the reference domain, making the noun's denotation range over collections instead of atoms (*books* ∼ *book*), or to collections of standard partitions instead of an undivided mass (*wines* ∼ *wine*). However, plurality often means more than this grammatically regimented reading, and affects the conceptualization inherent in the lexeme: *waters* does not mean 'many a water' in *the waters of the lake*, *funds* may mean 'funding' as opposed to 'many a fund', and *looks* may mean 'human physical features' rather than 'many a look'. Again, only a cross-linguistic perspective reveals the true extent of these lexicalization phenomena, bringing to light unexpected parallels and generalizations. My goal will be to categorize the attested readings and identify the fundamental properties of conceptualization underlying them. This task involves a characterization of the ontology defined by lexical plurals; not a description of some plural readings based on an assumed domain, but an analysis of the domain itself. Plurality, I will argue, affects lexical semantics when it brings about a conceptualization of the primitives of denotation as 'not-one', in ways that vary along the dimensions of unity and identity. These fundamental conceptual properties, along with others like cohesion and boundedness, allow for a coherent and revealing account of what might

otherwise appear as capricious irregularities, and highlight the connection between plurality and other grammatical expressions of part-structure conceptualization, like duals, 'collectives', and singulatives. Inherent in this hypothesis is the claim that the conceptualization encapsulated by lexical plurals is linguistic information, not unanalysable world knowledge. The observed gradience and variability are part and parcel of the phenomena to account for, and clearly contrast with the clear-cut oppositions of purely grammatical information, which are best expressed in terms of features. At the same time, the conceptual properties of nouns concern the characterization of lexemes as linguistic objects, and this directly translates into grammatical information, including plurality.

1.4 Lexicality in morphosyntactic structure

This approach has a structural side as well. Lexemes and stems only appear in a concrete syntactic context, which must be clarified in order to understand how lexical information is represented in the morphosyntactic representation. I assume that nouns are the innermost elements in a complex phrase headed by the determiner (DP), with several grammatical heads defining intervening projections between the determiner and the noun. I will take a syntactic number head (the inner one; see Section 6.6.2 in Chapter 6) to host the grammatical features that number-inflected nouns pick up in a syntactic derivation. This head encapsulates the grammatical, non-lexical element of number inflection, and so helps define lexical plurality by opposition. Nouns that are inflectionally plural, but do not spell out the number head, realize plurality lexically, through their stem and not through a grammatical morpheme. This happens when the noun is a monomorphemic stem without a discrete marker for plurality, but also when it contains a plural affix embedded inside a complex stem and bearing no relation to the syntactic context. A closer view of the structure of nouns also sharpens the intuition that plurality can be intrinsically associated with a noun. The various types of lexical plurals justify a constructional approach to what it means to be a noun. In this approach, the properties of nominality are distributed across distinct loci in a syntactic configuration, rather than concentrated on an unanalysable N head. Borer (2005) has argued that the number head contributes to determine the conceptual properties of a noun, in particular the granularity of its reference domain (cf. also Déprez 2005). Following and developing this insight, I will view the number head as the external, context-determined locus of part-structure conceptualization; another one is inside the noun, encoding a basic conceptualization that does not vary across syntactic contexts. I will identify

this with a head in its own right, [n], which according to recent proposals in Distributed Morphology (Marantz 2003; Arad 2003) combines with a category-free root to make up the syntactic construct we call a noun. To say that plurality is an integral part of a noun means, in structural terms, that it is encoded in [n], as part of the complex [n [ROOT]]. Plurality can thus be inherent in nouns as syntactic complexes, but not in roots, which by definition carry no grammatical information. The same root may be paired with another [n], perhaps with a different choice of gender or class diacritics, which is not intrinsically plural. Thanks to this decomposition, we can say that a noun is plural lexically, independently of the grammatical context around it, even though plurality is still a grammatical property with its own locus, separable from the noun's unanalysable core. Syntax thus helps us understand lexical plurality as a complex phenomenon, not reducible to a fixed marking on a 'lexical item'.

1.5 Inflection and derivation

A question that is central to this study concerns the distinction between inflection and derivation. In so far as the plurals here investigated differ from other plurals, they justify a distinction between morphology that is 'lexical' or word-creating, and morphology that is 'grammatical' or word-inflecting. In several languages, number has all the hallmarks of a derivational category, and carries information that shades into lexical semantics (a fundamental trait of word-formation formatives, as shown by Bybee 1985). While interesting in themselves, however, these cases do not shed much light on the relation between inflection and word formation. The really instructive cases, and those which I will focus on, involve languages where number is definitely inflectional as a morphosyntactic category, but has clear derivational properties on some words. This is the original insight encapsulated in Booij's (1994, 1996) concept of 'inherent inflection'. Following up on Booij's pioneering work, I will argue that plural has indeed different morphosemantic properties as an ingredient of a lexical base (lexeme or stem), or of the grammatical context for such a base; however, and this is crucial, the same plural forms may fulfil either function. This contrasts with the idea that morphology itself consists of two components, a lexical and a grammatical one (Split Morphology). The split concerns the uses to which morphological means are put: 'inflection and derivation are not two types, but two uses of morphology' (Aronoff 1994: 126; cf. also Stump 1998: 18–19; Borer 2005: 51–8). In turn, this conception presupposes a realizational approach to morphology, viewed as the translation of abstract linguistic information into a system of exponence

organized by its own principles (cf. the discussion in Anderson 1992). My position thus supports the autonomy of morphology and the separation, or lack of isomorphism, between morphological spell-out and the abstract information that it expresses (Aronoff 1994; Beard 1995).

Lexical bases are therefore legitimate, indeed necessary concepts, qualitatively different from affixes and from grammatically defined elements like pronouns or auxiliaries. In this respect, my stance is necessarily 'lexicalist' and focuses precisely on the difference between the lexical and non-lexical use of the same grammatical category. However, the lexical bases I have in mind are not atoms in a concrete morphosyntactic representation. I claim that lexemes and stems are part of linguistic knowledge; the model of Distributed Morphology of Halle and Marantz (1993), Marantz (1997), and Embick and Halle (2004) excludes both of them, so it is not compatible with my proposal. At the same time, I will follow that model in some crucial respects: the decomposition of nouns into roots and [n], the view that 'lexical items' are brought about by syntactic derivation, and the realizational view of morphology as a post-syntactic spell-out of an abstract input. Of these assumptions, only the first is specific to Distributed Morphology. Essentially, the reason for this inconsistency is simply that the evidence requires notions such as lexemes and stems, but the same evidence also shows that number as a lexicalized category is not an irreducible component of the atomic root; it falls, like gender, between this minimal root and contextual inflection—something that can be expressed naturally by adopting the Distributed-Morphological decomposition of nouns. To my mind, this suggests not so much a theoretical inconsistency, as the need to sharpen the reflection on lexicality within Distributed Morphology (a development foreshadowed in Noyer 1997). I do not argue for a separate lexicon of substantive words in addition to the atoms of morphosyntactic representation, but for a qualitative difference between lexical and non-lexical uses of these atoms. Stems and lexemes are the concepts that explicate this difference.

1.6 Structure of the book

Apart from this introductory chapter and a final conclusion on lexical and grammatical knowledge, the chapters of this book are grouped into two parts.

Part I, with Chapters 2–4, presents a typology of lexical plurality, which maps the range of phenomena traceable back to plurality being lexicalized. The data are categorized according to morphological and semantic concepts, which provide a unified framework for analysing phenomena that are typically discussed from partial perspectives. This typology, which may be read as a

self-contained introduction to lexical plurality, pursues a line of inquiry on the linguistic conceptualization of reality (Seiler and Lehmann 1982; Rijkhoff 2002; and, more specifically, Biermann 1982 and Tamm 2004), with the aid of analytical tools from theoretical morphology and from ontology and the philosophy of being.

Chapter 2 identifies the empirical domain by opposition. It shows that many plurals are 'lexical' in a sense that does not reduce to listedness, idiosyncrasy, or having a fixed number value. It then reviews the prototypical properties of inflection and illustrates plurals that fail to show them. The result is a structured sample of lexical plurals, arranged by the inflectional properties they fail to display.

The next two chapters turn to the morphological and semantic properties that make plurals lexical. Chapter 3 clarifies in what sense number is always inherent in nouns, and then catalogues the morphological phenomena where the expression of plurality coincides with, or is part of, the expression of a lexical word (plural words, suppletive forms, plural inside derivation, inherently plural noun classes).

Chapter 4 shows first of all that plurality is a self-standing semantic property, not reducible to a function from singulars. As a part of lexical semantics, plurality conceptualizes the primitives of the denotation as 'not-one', lacking *unity*, or whole-properties, and/or *identity*, or criteria of identifiability. Along with cohesion and boundedness, these concepts underlie a semantic typology that includes masses, measures, cohesive collections, indistinguishable objects, entities that instantiate kinds, and *tropes*, or property instances that are not entities.

Part II, Chapters 5–8, deploys the analytic tools explicated in Part I for the in-depth analysis of lexical plurals in Italian, Irish, Arabic, and Breton. Each study highlights different aspects of lexical plurality, and reaches the level of detail necessary to appreciate the true value of lexical plurals in the context of their respective language, beyond second-hand isolated examples and often misleading glosses. The choice of the four languages represents the best compromise between three factors: the intrinsic interest of the data; my familiarity with them; and the availability of (accessible) literature. These are therefore case studies, rather than an exhaustive encyclopaedia of lexical plurality.[1]

Chapter 5 focuses on a class of Italian plurals usually seen as an irregular inflectional class. My conclusion is instead that these are not the plurals of their respective singulars, but distinct, inherently plural lexemes, the output

---

[1] My own linguistic limitations have prevented me from discussing in any depth the evidence from Russian and Slavonic, which is interesting enough to have been the subject of a detailed monograph (Ljaševskaja 2004).

of word formation and not of inflection. Semantically, they share the property of denoting weakly differentiated entities. They are lexical because they are derived lexemes.

Chapter 6 examines a class of irregular Irish plurals used after numerals. Their interpretation as unit counters, their restriction to numeral contexts and their non-affixal morphology, taken together, suggest that these are inherently plural stems fulfilling a grammatical function, that of classifiers. They are lexical stems expressing non-individual units in a numberless context.

Chapter 7 analyses Arabic broken (non-affixal) plurals, which clearly display the properties of derived lexical stems. However, they are also productive and (usually) semantically transparent like inflectional plurals. To resolve the contradiction, I analyse them as Aronovian stems used as exponents of inflectional plurality. As the output of word formation put to inflectional use, they are lexical in form but grammatical in function.

Chapter 8 addresses the Breton plural system, which appears to neutralize any difference between inflection and word formation. The ambiguity is real, and is due to the radical separation of number exponents from their function in this language. The frequent use of number in a lexeme-forming function is one of the ways to express individuation and part-structure conceptualization, which is prominent in Breton noun morphology. It is the whole plural category that can be lexical in this case.

Finally, Chapter 9 recapitulates the results from Part I and Part II into a concluding discussion of plurality between grammatical and lexical competence. The latter involves knowledge of lexemes, stems, and a part-structure conceptualization. Plurality may be part of all three, as an ingredient of nominality. That is, I think, what it means to say that a noun is plural 'lexically'.

Part I
A Typology of Lexical Plurals

This page intentionally left blank

2

Varieties of non-inflectional plurals

2.1 Introduction

Since 'lexical' can mean so many different things, we cannot take lexical plurality to be a self-explanatory concept which could support an unequivocal definition. It seems much more practical to begin by the simple observation that not all plurals are alike, and that we are more willing to describe as lexicalized some of them rather than others. Pre-theoretically, the most straightforward sense in which a plural is lexical consists in not being something else; namely, in not being the regular outcome of a deterministic grammatical rule. To start our survey, therefore, this chapter will describe and categorize the many different ways in which plurals can look non-inflectional because of listedness and lack of regularity.

Sections 2.2, 2.3, and 2.4 open the survey by arguing against a simplistic reduction of lexical plurality to, respectively, idiosyncratic form, idiosyncratic meaning, or fixed value (as in *pluralia tantum*). The rest of the chapter characterizes lexical plurals by opposition to the prototypical properties of inflectional morphology. These are: in Section 2.5, the obligatory nature of inflectional rules or processes; in Section 2.6, the generality of application; in Section 2.7, the determinism of inflectional plural, which ensures that each input lexical item receives a unique plural match; and in Section 2.8, the semantic transparency of regular pluralization.

2.2 Lexical plurals ≠ irregular plurals

An influential tradition, from Bloomfield (1933: 274) to Chomsky (1995), through Chomsky (1965: 87), defines the lexicon as the repository of all the idiosyncratic forms of the language. Whether or not this says all that there is to say about the lexicon, and Aronoff (1994: 16–22) argues quite cogently that it does not, it certainly could provide a very neat way to delimit the empirical domain of lexical plurality. In fact, defining lexical plurals as those which are idiosyncratic is not incoherent; what I claim is that it is unrevealing.

To see this, consider more closely the suggestion that lexical plurals are all and only those which are idiosyncratic. Then *pence* or *depths*, but also *scissors*, *men*, *sheep*, *mice*, and *brethren*, should all be lexical. And so should be verbal forms *is*, *has*, *does*, *says*, as well as all strong verb forms if they cannot be reduced to a rule. And in Romance languages, widespread stem allomorphy would mean that most verbs have at least some lexical forms. In short, anything which is not the output of regular, general, deterministic inflectional operations would then be lexical. But this obscures some important distinctions. *Men* is irregular as a form, but its paradigmatic relation to *man* is exactly the same as that of any other plural: it is the only plural form and it is semantically transparent. By contrast, *brethren* and *mice*, for instance, have regular counterparts (*brothers* and, for those who accept it, *mouses* for 'computer pointers') and their meaning is not in the same way predictable from the singular. A noun like *scissors* must be lexically listed as lacking a singular altogether, quite a different property from having irregular exponence or being semantically specialized. And *depths* or *waters* are neither formally irregular nor *pluralia tantum*, but they are semantically non-transparent: whatever *depths* and *waters* mean, they do not mean 'many a depth' or 'many a water'. This is very different indeed from the purely formal idiosyncrasy of verbal forms like *has*, as well as from cases like *men* and other high-frequency irregular nouns.

Calling lexical all these varieties of idiosyncrasy is not just vague, but vague in a misleading way. The meaning of some plurals, like *brethren* (confrères) or *mouses* (computer pointers), differ enough from that of the singular to suggest that the latter is ambiguous, or perhaps even that there are two homophonous singulars:

(2.1)   a.   mice   b. mouse ——— mice

         mouse

                                    mouses   mouse ——— mouses
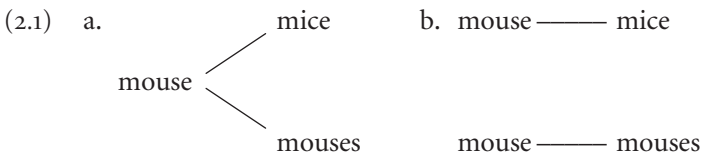
To describe this situation, we need a distinction between forms (*mouse*, *mice*, *mouses*) and abstract 'bases', so that we can say that one plural form does not block the other because they are not really alternative forms of the same base. It is not so obvious to decide in what sense *mice* and *mouses* correspond to different bases; we could be talking about two readings licensed by the same lexical entry ('sense' for Pustejovsky 1995); or about two semantic listemes associated with the same lexical entry; or, assuming homophonous

singulars, of two quite distinct lexical entries. In any case, the choice between *mouses* and *mice* is related to lexical semantics, and speakers choose between them as they select one 'sense' over another, or one word over another. By contrast, *men* is just an irregular form, the obligatory realization of an inflectional cell, without any connection with the choice between 'senses', or words. But this important difference disappears if being lexical means being listed or idiosyncratic, because the forms *men*, *mice*, and *mouses*, as well as the hypothetical bases behind the latter two, are all equally listed. The forms are not the same entities as these bases, yet this is what we say if we conflate the two notions under the same label of lexical, identified with idiosyncratic.

Besides, a reduction of lexicality to idiosyncrasy would be unsatisfactory because entries listed in the mental lexicon can, and often do, include elements regularly formed by perfectly regular operations. Speakers apparently do not synthesize online all regularly inflected forms, but access the most frequent ones as pre-packaged units (Stemberger and McWhinney 1988; Clahsen 1999; Booij 1999). And, as Jackendoff emphasizes (1997: 153–78, 2002: 152–95), the range of what is listed in a language extends far beyond the basic morphemes and their irregularities, to include idioms in the strict sense, collocations and syntactically complex structures of various size. At least, the concept of listedness should distinguish between a morphological dimension, in which the forms are listed that play the role of underived building blocks for morphology, and a conceptual-mental dimension, which stores all idioms. This corresponds, as far as I can see, to the distinction between Vocabulary and Encyclopaedia outlined by Marantz (1997) and sharpened by Harley and Noyer (2000). To clarify the concept of lexical base necessary to account for the behaviour of lexicalized plural nouns, later chapters will justify reference to stems and lexemes, which are neither basic morphological building blocks nor entries in the mental lexicon. But for the moment, the point is more generally that lexical plurality cannot be reduced to idiosyncrasy, or listedness, or lack of regularity (cf. Di Sciullo and Williams 1987). Lexical plurals derive their interest from the fact that they are peculiar words, and although listedness is important, reducing wordhood to listedness is wrong.

2.3 Lexical plurals ≠ semantically irregular plurals

The foregoing observations might suggest that what I have been calling lexical plurals correspond to a particular subtype of idiosyncratic plurals, those with idiosyncratic meaning. It is true that semantic opacity is an important aspect of most plurals that involve lexical knowledge; but it cannot be a definitional property.

To begin with, specifying what counts as semantically idiosyncratic is far from straightforward. Maybe a plural is lexical if its meaning does not arise through the composition of the meaning of the stem and of the plural affix? This may certainly work for *depths* and other such cases. But take *louses* or *oxes*. In so far as speakers accept these forms at all, their interpretation is based on a metaphorical extension of the core meaning of the nouns: *louse* as 'contemptible person' and *ox* as 'large and uncouth man'. But, as plurals of these singulars, *louses* and *oxes* are absolutely transparent and compositional. *Louses* is semantically listed as a word, but it is compositional as a plural. Similar, if not identical, considerations apply to all those cases where a plural doublet disambiguates two senses conflated in one singular form, as *brothers* ∼ *brethren*, or the Italian *membri* 'members' ∼ *membra* 'limbs', mirrored by the Dutch doublet *leden* 'members' ∼ *ledematen* 'limbs' (Donaldson 1987: 37). The existence of two senses with separate expressions in the plural is a matter of lexical knowledge. Yet, both plurals are totally compositional, each with respect to its sense.

Taking semantic unpredictability as the definitional criterion would prove too limiting in other important regards. Notice that no *pluralia tantum* could then be regarded as lexical plurals: strictly speaking, there can be no sense in which the meaning of plural is related non-compositionally to that of the singular, if there is no singular. And yet, if there are words for which plural must be listed as a lexical property, these are undoubtedly *pluralia tantum*. Identifying lexical plurality with semantic irregularity would also exclude a whole class of phenomena where plural is sensitive to lexical semantics. For instance, Donaldson (1987: 38) lists for Dutch 'a small group of neuter nouns that preserve an old plural ending in -*eren* (compare Eng. *children*)'. The list is as follows (omitting alternative plurals for *hoen* and *volk*, and for *been* and *blad* in the sense of 'leg' and 'page, paper leaf'):[1]

(2.2)	SINGULAR	GLOSS	PLURAL	(Dutch)
	been	'bone'	beenderen	
	blad	'leaf'	bladeren	
	ei	'egg'	eieren	
	gelid	'joint'	gelederen	
	gemoed	'mind'	gemoederen	
	goed	'goods, wares'	goederen	
	hoen	'fowl'	hoenderen	
	kalf	'calf'	kalveren	

---

[1] On the connection between lexical semantic interpretation, relative markedness of singular and plural, and morphological exponence in these and other cases, see Tiersma (1982).

kind	'child'	kinderen
lam	'lamb'	lammeren
lied	'song'	liederen
rad	'wheel'	raderen
rund	'cow, ox'	runderen
volk	'nation, people'	volkeren

The survival of an ancient plural form precisely for these words has a semantic motivation. The referents of these plurals are all weakly individual notions: physical objects naturally occurring in cohesive aggregates (bones, joints, wheels), entities liable to being experienced as interchangeable units (leaves, songs, eggs, children, animals), notions conceptualized as a mass (mind, goods), and the intrinsically collective 'people', whose plural is as likely to express a multitude of persons ('nations') as of peoples. As we will see in the following chapters, the same categories occur in language after language as the semantic common denominator of various types of irregularity in plural formation. This kind of semantic motivation for certain morphological classes is a very important and widespread phenomenon, and it has definitely to do with lexical knowledge. But this pluralization pattern is not semantically irregular. What is lexicalized here is not the semantic relation between singular and plural, as for instance in *depth* ∼ *depths*. If that was the definitional property of lexical plurals, the impressive consistency with which irregular plurals correlate with certain semantic categories across languages would have nothing to do with lexical plurality—a paradoxical and unsatisfactory result.

2.4 Lexical plurals ≠ *pluralia tantum*

Some nouns only exist in the plural, like *clothes* (or in the singular, like *fun*). What better example is there of plurality being part of lexical competence? Actually, we have already considered some cases where number is a lexicalized property on nouns that are not *pluralia* or *singularia tantum*; still, *pluralia tantum* might appear as a core case of lexical plurals.

In fact, *pluralia* (and *singularia*) *tantum* are convenient descriptive labels, but no more. Being plural or singular is a grammatical property; but *pluralia* and *singularia tantum*, despite common assumptions, do not really make up a class of lexical items defined by the grammatical property of having a number specification built in. By this I do not mean that the notion is incoherent, nor do I intend to deny the unassailable fact that being plural or singular is part of the intrinsic specification of certain nouns: I chose *clothes* and *fun* as examples

because *a clothe* or *funs* are not part of English in any of its current varieties. What I claim is rather that everything interesting (linguistically significant) that can be said about fixed-number nouns can also be said about non-fixed-number nouns—specifically, about all those which instantiate number as a lexical property. *Pluralia tantum* are not a particular way to instantiate lexical plurality, but simply lexical plurals whose only distinctive property is that they have no singular. One could group them together on the basis of this property, but then one could also group together all lexical plurals with irregular morphology, or indeed all those beginning with *f*. Whether these or other groupings systematically correlate with other properties is an empirical matter, and lacking a singular apparently does not. We will now consider in turn the evidence for denying *pluralia tantum* the status of a grammatical natural class.

2.4.1 *Intrinsic fuzziness of the concept*

If having a fixed plural value was a clear-cut property, it should be possible to list the *pluralia tantum* in a language, allowing for some fuzziness because not all speakers have the same vocabulary, but overall with a certain precision. English, with several nouns like *clothes* or *scissors*, would appear to be a language with such a fixed-number noun class. Yet this class turns out to have much more blurred boundaries than, say, the class of strong verbs.

   One source of variation derives from a peculiar morphology–semantics mismatch connected with the -*s* ending. Names of diseases, games, and disciplines like *mumps*, *measles*, *billiards*, *tactics*, or *semantics*, predominantly trigger singular agreement, and others like *blues*, *means*, and *news* always do so. If being a *plurale tantum* is viewed exclusively as a morphological property, then all these nouns are plural, because they derive from bases without the -*s* and cannot be further suffixed (*newses*, *meanses*, unlike for instance *summonses*). But it seems wrong to treat as plural a noun like *news*, which is always syntactically singular (*the news is/*are bad*); other nouns in this group have variable agreement, like *means*. In these cases, then, the relation between morphological and syntactic plurality is too loose to allow a clear decision about *pluralia tantum* status. This is not a matter of variation in the individual use of single words.[2]

   Something similar applies to plurals that denote single objects. Few native speakers would say *a scissors*, but many such 'summation plurals' can take the

---

[2] Notice that the lack of a singular in *pants*, *scissors*, or *trousers* cannot be a matter of morphological well formedness, witness the grammaticality of *pant leg*, *trouser leg*, or *scissor blade* (and, for those who accept it, *many a trouser* or *many a scissor*).

singular indefinite article with premodification: *a garden shears*, *a curling-tongs*, etc. (Quirk *et al.* 1985; cf. Allan 1980 for discussion). Again, while the *-s* ending is definitely a plural morpheme (*many garden shears*, not **shearses*), agreement may be in the singular. If the plurality of these nouns is ambiguous, so is their status as *pluralia tantum*.

Another factor preventing a neat delimitation of the class has to do with listedness. *Amends* is certainly a plural noun wherever it occurs; but it only occurs in the fixed collocation *make amends*. Is the noun a *plurale tantum*, or is, rather, the idiom that enforces plurality on a noun not used elsewhere? The same applies to other nouns restricted to idioms, like *creeps* or *jitters*, or to fixed collocations, like *throes*, but also to plurals in name-like definite descriptions, like *the antipodes*. These are nouns and they are only plural; but they only exist as parts of a listed phrase.

A deeper source of uncertainty is that in all too many cases it is far from clear whether a plural noun is an inherently plural lexical item or the plural of an existing singular. I am referring to the numerous singular–plural pairs, pointed out as problematic by Corbett (2000: 176), whose semantic relation is not transparent. For example, the singular *look* denotes a looking event (or gesture); *looks* can mean a plurality of such events, or it may have the idiomatic reading 'physical attractiveness'. If the idiomatic plural is a distinct lexical item, it is a *plurale tantum*. But how semantically distant must singular and plural be, in order to count as different lexical items? Corbett (2000: 176) states that the different meanings of *fund* and *funds* 'do not match up completely as the singular and plural of the same lexical item', a formulation that seems to imply that singular and plural must be semantically transparent to qualify as forms of the same lexical item. Indeed, his discussion takes place in the context of *pluralia* and *singularia tantum*, the point being that an English pair like *fund* ~ *funds* conceals two number-defective lexical items. However, Corbett himself points out earlier on (2000: 84–7) that switching number can change the countability status of a noun, as in *three coffees*. So, one and the same noun can have distinct countability status depending on number; a difference in countability is not enough to show that two forms belong to two distinct lexical items. Yet it is hard to see what else differentiates *fund* and *funds*, which the *Shorter Oxford Dictionary* (14th edn, 1990) glosses as 'a stock of money, esp. one set aside for a purpose; (in *pl.*) money resources' (notice *resources* as a plural mass noun). Once it is agreed that singular and plural can express distinct countability readings for the same lexical item, and there are very good reasons for doing so, the rationale for deciding when word forms belong to the same lexical item can become very dim indeed. Some cases are easy: *looks* 'physical attractiveness' and *look* 'act of looking', or *works*

'mechanism', 'factory' and *work* 'activity', are semantically different words, because they are true of different things; and the totally transparent *car* and *cars* must be forms of the same word. But between these extreme cases lies a continuum of surprisingly many plurals, semantically non-compositional in very different ways and to different degrees: *bearings, brains, crops, depths, dimensions, directions, foundations, gates, heavens, heights, holidays, intricacies, manners, mists, plans, preparations, proofs, resources, results, skies, snows, suspicions, thoughts, times, views, waters, winds*... The list is long and, more importantly, open-ended. How is one to tell if *sorrows* is the compositional plural of *sorrow*? Can one count one's insecurities or fears? Given that *I saw you in my dreams* can be true when a single dream is involved, is *dreams* a *plurale tantum*, even though it means the same as the singular? Does hurting someone's feelings amount to hurting a collection of abstract things, each of which is a feeling, or is *feelings* yet another *plurale tantum*? If a cottage is in the woods, is it among many distinct woods?[3] In all of these cases, like in so many others, the semantic relation between singular and plural differs from the canonical one–many opposition of regular count nouns, which is basic and cuts across nouns and pronouns ([± group] of Harley and Ritter 2002). Cases like *snows* or *rains* are particularly important, because they differ from their singular neither in countability (*rain* is mass just like *rains*) nor in what they are true of, which is always a meteorological occurrence or a substance. Yet the plural does indeed have some kind of plural interpretation, referring to a multiplicity of raining events, as in *the Autumn rains*; it is just that a single event is not *a rain*. Is *rains* a distinct, plural-only lexical item, just because of this different conceptualization? This is far from self-evident (I will take up the issue in detail in Chapter 4). Since the same uncertainty surrounds mass plurals of count singulars (*dream* ∼ *dreams*), the plurals for which the status of *plurale tantum* is uncertain make up a sizeable group. Hence, the boundaries of the class of *pluralia tantum* are significantly blurred, and for principled reasons.

Finally, it is often unclear whether the lack of singular is a grammatical fact at all. From a theoretical viewpoint there is a sharp difference between a noun being inherently plural as a fact of internalized grammar, and a noun being perfectly regular but having a negligible token frequency as singular. Even without statistical evidence, I think that a word like *nit* (in its proper, not metaphorical sense) would exemplify this latter state of affairs. It is not a *plurale tantum*, in so far as speakers could in principle use its singular. But what can be done 'in principle' is not always clear, even to individual speakers;

---

[3] This particular example brings out the word dependent character of the use of plural as a massifier. *In the woods* can mean something like 'in a wooded environment', but *in the forests* suggests more strongly a plurality of distinct forests.

so the membership of the class of *pluralia tantum* inevitably varies and depends on the vagaries of idiolectal usage. In a number of cases there is no clear distinction between nouns whose lexical representation lacks the singular form and nouns which, simply, tend not to be used in the singular because of their meaning. English examples like *embers*, *reins*, *traces* (of a draught horse) and *tropics* illustrate this category of quasi-*pluralia tantum*, for which a singular form may be admissible in some context. But membership in a grammatical class is not a matter of degree for a given lexical item. The existence of such cases suggests that proper *pluralia tantum* like *clothes* represent the end-point of a spectrum, ranging from total admissibility of both singular and plural to total exclusion of singular. This would make them, as stated, a useful descriptive label, but nothing more than that.

2.4.2 *Lack of a specific semantic correlate*

The intrinsic fuzziness of the concept is not the only reason why *pluralia tantum* cannot be a grammatically defined class of lexical plurals. As is well known, the meaning of *pluralia tantum* is not random. Koptjevskaja-Tamm and Wälchli (2001: 630) identify the following main categories, which essentially confirm from a wider typological perspective those proposed by Delbrück (1893: 147–65) for Indo-European (cf. also Corbett 2000: 175–6): substances (Lithuanian *putos* 'foam'), complex artifacts (Russian *časy* 'clock, watch'), environment types (Russian *džungli* 'jungle'), diseases (English *measles*), periods of time (Russian *sutki* 'day-night cycle'), festivities (Finnish *häat* 'wedding'), activities with multiple participants (Russian *prjatki* 'hide-and-seek'). Let us briefly review the content of these categories.[4]

Substance *pluralia tantum* denote masses along a continuum from aggregates of discrete objects like *clothes* to entities with a not clearly defined but still granular part structure, like *oats* or *entrails*, to granular substances like *suds*, to totally homogeneous masses. It is not clear if English has any *pluralia tantum* for this last class, but the noun for 'water' is plural in several languages: Sanskrit *apas* ('almost always plural'; Delbrück 1893: 149), Hebrew *mayim* (formally a *duale tantum*; Schwarzwald 1991: 593), Akkadian *mû* (von Soden 1969: 76), Swahili *maji* (Contini-Morava 1999: 6), Turkana *ā-kipì* (Dimmendaal 1983: 211), Rendille *bicé* and Somali *biyyó* (Oomen 1981: 50).

---

[4] Wierzbicka (1988) deserves a separate mention. Her in depth study of plurality and mass has a much more ambitious goal than categorizing *pluralia tantum*, namely, to derive the use of number for mass nouns from the conceptualization of the referents. This deeper level of analysis will be relevant in Chapter 4. What I must make clear here, however, is that Wierzbicka's conclusions seem to me to establish too direct a link between grammatical number and world knowledge. While I agree that the different number values of *grass* and *gravel* on one hand and *oats* and *groceries* on the other are not arbitrary, her direct linking of singular with large mass and plural with small composite mass seems too deterministic, and at odds with cross linguistic variation ('grass' is plural in Turkana, 'gravel' in Breton).

What Quirk *et al.* (1985) call 'summation plurals' refer to single, discrete entities having a perceptually salient internal articulation, objects like *spectacles* or functionally related sets of unattached parts like *pyjamas*. Different languages obviously vary in the conceptualization of such complex objects: scissors are plural in English, French (*ciseaux*), Italian (*forbici*), and Russian (*nožnitsy*) but not in German (*Schere*), Swedish (*sax*), or Hungarian (*olló*); among *pluralia tantum* that would strike English speakers as unusual, Corbett (2000: 174) mentions *niicugnissuutet* 'radio' in Central Alaska Yup'ik, and we may add the Polish *skrzypce* 'violin' (Wierzbicka 1988: 536) and the Russian *sani* 'sledge'.

Two categories that display a remarkable cross-linguistic consistency are names for diseases and for culturally salient event-types articulated into sub-events: festivities, ritual occasions, or names for games. Some Russian examples are *imeniny* 'name's day', *zamorozki* 'light frosts', *pokhorony* 'funerals'. In Latin, we have festival names like *saturnalia*, the fixed calendar dates *idus*, *calendae*, *nonae*, and occasions like *feriae* 'holidays' and *nuptiae* 'nuptials' (traditionally a sequence of events).

Finally, in English and in many other languages plural is inherent on nouns that straddle the divide between mass and abstract, like *goods* or *arrears*, abstracts like *auspices*, *manners*, or *thanks*, and others that basically refer to manifold instantiations of a property, specifying the property but not the instances: these are often nominalizations like *belongings*, *riches*, or *valuables*, whose 'singular' counterpart is not a noun. The shift from manifold to abstract has a particular significance, because it explains some puzzling *pluralia tantum* that seem to have no connection with semantic plurality. The Latin *liberi* 'children', for instance, can only be plural (and masculine); it is in fact an abstract designation, something like 'offspring', and it can refer to a single individual in the appropriate context (see Wackernagel 1926: 95).

Having briefly considered the semantic range of *pluralia tantum*, we are now in a position to establish an important fact: these broad semantic categories are relevant whenever plurality is lexicalized, not just with *pluralia tantum*. It simply does not matter whether or not the noun in question is invariably plural. Plurality as manifold part structure within a single discrete object appears on plurals that, like the classical Greek *hármata* 'chariots' or *tóksa* 'bows', can refer to single implements (like *scissors*) but have a singular (Brugmann 1900: 170). The various shades of mass interpretation can also be expressed by plurality on nouns that are not *pluralia tantum*: the relation of *funds* to *fund*, or *holidays* to *holiday*, for instance, is that of a complex mass to a discrete or bounded entity (in the sense of Jackendoff 1991; cf. Chapter 4 below). Along with the plural-only nouns for 'water', we also find Latin *aqua* ∼ *aquae*, with a mass reading of the plural paralleled by many European languages (for water scattered in space or as

a name for the cosmic element), including the English *water* ∼ *waters*. Many abstract nouns occur both in singular and in plural, sometimes with the semantic effect of manifold instantiation that is also relevant for *pluralia tantum* (these matters will be taken up in the next chapters). Also, cases like *furnishings* are *pluralia tantum* because their base is not a noun (*a furnishing); but the same function associated with these plurals, that is, manifold unspecified instantiations of a property, appears where the base is a noun (singular), as in *depth* ∼ *depths*. The lack of a singular form for the noun is really a fortuitous accident; what matters is the use of plurality.

In short, the categories that are relevant for *pluralia tantum* are a subset of those relevant for lexical plurality. Koptjevskaja-Tamm and Wälchli (2001: 629–30) make the point explicitly, noting that some nouns 'behave like prototypical *pluralia tantum*, but happen to have a singular form as well' (cf. 2.4.1 above). This would be a contradiction in terms, were it not for the fact that there is a shared semantic basis that groups together these plurals over and above the lack of a singular. Since their study centres on circum-Baltic languages, one of whose features is the great frequency of plural-only nouns, they retain the traditional label of *pluralia tantum* while adding that 'a more appropriate term would be lexical plurals'. For present purposes, the conclusion must be formulated differently. There is no property, morphological or semantic, that isolates plural-only nouns from among the rest of phenomena characterizable as lexical plurality. The definitional property of having no singular turns out to be shallow and sometimes accidental, often (as in English) practically impossible to define and circumscribe. This state of affairs resembles the status of the mass–count distinction. With some cross-linguistic variation, mass and count are often no more than tendencies rather than strict grammatical determinations of lexical items; different contextual tests give different results, judgements are often a matter of degree and there is strong lexical and idiolectal variation (see Allan 1980 and especially Pelletier and Schubert 1989). While they remain necessary as descriptive concepts, mass and count cannot be defined as grammatical properties of lexical items outside of a context, as Borer (2005) cogently shows. In the same way, I think, *pluralia* and *singularia tantum* are indispensable descriptive concepts, but they are not genuine linguistic classes. Therefore, we cannot build a notion of lexical plurals around that of *pluralia tantum*.

2.5 Lexical vs. inflectional plurals: lack of obligatoriness

The intuition that plurality may be lexical means, above all, that it may display properties not usually associated with regular inflectional morphology. Unfortunately, defining what it means to be inflectional is a matter of continuous

debate. Like the philosophical question of the reality of universals, the distinction between inflection and derivation is too close to the theoretical heart of the discipline to be settled conclusively, as if it were no more than an empirical question. Still, the conflicting views expressed by the many alternative theoretical orientations have considerably clarified the issues involved. Among the vast literature that directly or indirectly addresses the question, Aronoff (1994: 126) spells out the fundamental insight on which my analysis is based: 'inflection is the morphological realization of syntax, while derivation is the morphological realization of lexeme formation'. Even after making this clear, a precise definition of inflectional morphology can prove difficult and controversial. However, its prototypical traits are readily listed, and that is what matters for the present descriptive purposes. The ways in which plurals can fail to exhibit these characteristics serve as guidelines for a first understanding of lexical plurals.

To begin with, inflection is obligatory. The plural of *book* is *books*, and failure to use that form in a morphosyntactically plural context (like *these are   *) results in ungrammaticality. Regardless of its exponence, one of the values defined by the number opposition is mandatory in any one grammatical context.

As is well known, there are many languages in which plural has an altogether different status. Many languages lack an obligatory number opposition on nouns, and use instead a single form that can notionally correspond to our singular or plural. General number, as Corbett (2000: 10–12) refers to this (after Andrzejewski 1960), is illustrated by the following examples:

(2.3) a. otokonoko-ga asonde-iru (Japanese; Nakanishi and Tomioka 2004: 113)
           boy-NOM   play-PROGR
           'a boy is playing/boys are playing'

      b. tôi dã   mua bao   (Vietnamese; Löbel 2000: 268)
           I   PAST buy bag
           'I bought a bag/the bag/bags/the bags'

The notion of a nominal form neutral between singular and plural appears straightforward; but this impression quickly dissolves on closer scrutiny. In order to avoid being bogged down in this complex issue, I will just make three general points to clarify the issue of optionality.

First, one must carefully distinguish between a morphological and a semantic sense. English certainly has an inflectional number opposition, which partitions its noun forms in two mutually exclusive and jointly exhaustive sets, singular and plural. Some nouns may only have one form, like *clothes*;

others may fail to formally express the opposition, like *a barracks/several barracks*; but every noun has one and only one value in any given context. Notice that the British and Irish construction *my family live here* does not blur the singular–plural contrast, because the head of the NP which triggers plural agreement is morphosyntactically singular (*this/*these family live here*); the number value of a noun phrase may be ambiguous in terms of agreement, but even in constructions like *twenty faculty* the head noun itself is morphologically singular, as contrasting with *faculties*. And yet, the morphological categories of singular and plural can lend themselves to an interpretation that is neither singular nor plural. This happens in what Carlson and Pelletier (1995: 85) call kind-oriented talk, 'when we do not care about the object-level identity of the objects, as in *we filmed the grizzly in Alaska*' (cf. Section 4.5 in Chapter 4). English has no general number as a separate category beside singular and plural, but the number values made available by the morphology of the language can be used in transnumeral sense, to use the term originating from the German-speaking tradition (see Biermann 1982). Link (1998: 221) makes the point very clearly in connection with bare plurals:

(2.4) a. tigers are a subspecies of wild cats

  b. Otto motors were invented by the German engineer Nikolaus August Otto

Since it is obvious that no single tiger can be a (sub) species, and Otto motors cannot have been invented over and over again, we see that the number distinction *singular vs. plural* is neutralized in these contexts. Bare plural have the *form* of a plural, but their reference is *transnumeral*. (Link 1998: 221)

The same applies when the kind-referring expression is grammatically singular (examples from Carlson and Pelletier 1995):

(2.5) a. the lion has a bushy tail
  b. the tiger is widespread in southern Asia
  c. the American family has 2.3 children

Single lions have tails, but the set of all lions does not have a tail; conversely, the set of tigers but not individual tigers may cover a geographical area; and neither the set of families nor single families have fractionary offspring. The first point, then, is that semantic transnumerality can be encoded through grammatical number.

  The second point is that the details of this encoding vary enormously across languages. As Corbett (2000: 9–19) makes clear, some languages are like English and every word form of a noun must fall in one of the grammatical numbers; others oppose a category that can express both singular or plural

(like those in (2.3) above) to an unambiguous plural, or to a singular, often to plural *and* to singular. Notice that a language may lack an obligatory number opposition in nouns while having it in pronouns, or verbs, as is often the case in North American languages (see Mithun 1999: 79–87). The singular, especially when it is formally unmarked, often serves simultaneously as a morphological category contrasting with plural (or other number values) and as general number; cf. the following remarks from Greenberg (1974: 30):

A considerable number of classifier languages (e.g. many Iranian and Turkic lan guages, Korean) have what are generally described as plural affixes. However, closer examination seems to show that in almost every instance the 'unmarked' singular is in fact a form which like the collective in languages with a compulsory plural, is non committal in regard to number.

If a language has a general form for semantic singular and plural, and a specifically plural form just for semantic plurality, the morphological plural is strictly speaking redundant. Plurality markers are then said to be optional, or used mainly where the context does not otherwise disambiguate.

However, things are never so simple, and this leads to the third point. In some cases, the idea that plural markers are optional is really a misperception, caused by the mistaken assumption that the primary function of such markers is to express plurality. What pass for optional plural markers often encode a variety of meanings only indirectly related to a reading as 'more than one x': distribution in space or time, distinctness of type (as opposed to token), material or functional cohesion, collectiveness, association; Cheng and Sybesma (1999: 536–8), for example, make exactly this point about the Chinese suffix -*men*.[5] I will discuss these and other semantic categories at different stages in this and in later chapters; for the moment, the point is that this kind of information is in principle distinct from plurality proper, so much so that there are systems where the two coexist and are expressed by distinct morphemes.[6] For this reason, the lack of number as an inflectional category in a language does not imply that plural is optional, not because it is obligatory but because it is not really plural. In fact, once the semantic and

[5] It is often stated that classifiers and number are in complementary distribution, being two alternative ways to syntactically encode countability. This should be taken to apply to inflectional plural, not to the 'optional' markers under consideration. Unterbeck (1993: 183 5) shows that the Korean *tyr* can pluralize a noun in a classifier construction, and cites similar examples from Yucatec, Ojibway, Tarasco, and Jacaltec; see also Rijkhoff (2002: 43) on Imbabura Quechua. Other exceptions are mentioned by Aikhenvald (2003: 249). Borer (2005: 93 5) states that 'plural inflection *is* classifier inflection', but then goes on to illustrate the complementarity of plural and classifiers by means of an *optional* plural suffix in Armenian, which blurs the matter somewhat.

[6] See Mithun (1999: 91 2) for the co occurrence of collective and singular/plural affixes in Yana and Zuni. Besides, Georgian features a type of distributive reduplication (Gil 1988) alongside a grammatical opposition of singular and plural.

often pragmatic characterization of such putative 'plural' morphology is made clear, it may well turn out that these markers, while redundant for plurality, are indispensable for certain readings. For Korean, for example, Unterbeck (1993) has shown that the suffix *-tyr* (*-tul* in other transcriptions) cannot be omitted when referring to a plural referent that is also specific, non-generic, and animate (fully confirming the claim of Song 1975).

The contrast between obligatory and optional number, then, is misguided; but it is a misguided way to describe a real opposition. What distinguishes languages like English or Arabic from languages like Chinese or Oromo is the obligatoriness of number as a morphological and syntactic determination of nouns. Every noun must count as singular or plural (or dual) in the former group, because being one or the other is part of what it is to be a noun (in a context), and number underlies syntactic agreement. By contrast, languages with a well-established category of general number do not require a specific choice of number morphology as a grammatical determinant of nouns. 'Plural' morphology can then be optional in the sense of being determined by the intended interpretation of the head noun, independently of the grammatical context. It is a matter of choosing the right sense for the right word, not the right grammatical form of a word. It is, in sum, lexical—like choosing between *brain* and *brains*.

2.6 Lexical vs. inflectional plurals: lack of generality

Another prototypical trait of inflectional morphology is its generality. If a language has number inflection for certain lexical categories (nouns, adjectives, verbs), the relevant oppositions apply to all lexical items in those categories. A noun with defective paradigm like the singular-only *fun* is no exception, because it too has a number value, and so does not neutralize the singular–plural opposition. In many instances, however, plurality does not apply in this blind fashion typical of inflectional processes. There are three types of possible restrictions on pluralization: syntactic, categorial, and semantic. Of these, the last one most clearly points to lexicality, because it makes pluralization dependent on the choice of a specific noun and its semantic content. However, there are phenomena where lexical restrictions only surface in certain syntactic contexts, or combine with a categorial restriction to nouns, as opposed to adjectives or pronouns. We will now consider how these types of restriction relate grammatical plurality to lexical knowledge.

2.6.1 *Contextual restrictions*

Many languages neutralize the number opposition in the context of numerical modification: especially in agglutinating languages where plural is a discrete

suffix, nouns governed by a numeral must take the unsuffixed form, equivalent to the singular. This happens across the Turkic and Uralic languages, but also, for instance, in Georgian (many more examples in Rijkhoff 2002: 38–41):[7]

(2.6)	Gloss	Singular	Plural	'Three N'	
a.	'day'	kun	kun-lar	uč kun	
				(Uzbek; Bodrogligeti 2002: 120)	
b.	'house'	ház	ház-ak	három ház	(Hungarian)
c.	'man'	k'ac-i	k'ac-eb-i	sam-i	k'ac-i (Georgian)
		man-NOM	man-PL-NOM	three-NOM	man-NOM

There is nothing lexical in this, in so far as all nouns are affected. Sometimes, however, numerical modification brings out lexical restrictions in the use of number. In several languages that require the plural after numerals above 1, the singular is exceptionally required or permitted on nouns that denote units of measurement, or objects likely to serve as measures of quantity. English, especially in its non-American dialects, is such a language:

(2.7) a. three   cent/dozen/Euro/fathom/foot/hundred/hundredweight/
              million/pound/score/stone/thousand

        b. three   quid/bob

The expressions in (2.7a) can all appear, in different contexts, as nouns with a regular -*s* plural; *quid* and *bob* are instead invariable (see Section 6.3.2).[8] In these cases, the neutralization of the number opposition is determined both by the context and by the interpretation of the few nouns involved. Since not every unit noun with a suitable semantics follows this pattern (*three metre, *three mile), the irregularity is lexically encoded and is not automatically brought about by the noun's unit interpretation. Something similar happens in Irish, where a set of nouns likewise centring on measures is irregular in being instead *plural* after the numbers 3–10 (while singular would be the norm). Since these counting plurals will be mentioned again in Section 2.7.2 and analysed in detail in Chapter 6, I will not exemplify them here nor enter into the complications, of which there are many.

---

[7] Thanks are due to Alice Harris for help with the Georgian data, and specifically for providing this example. Rijkhoff (2002: 39) quotes a Georgian example with 'two', but in general the status of plural with numerals is better gauged with numbers from three up, because some languages (not including Georgian) show the effects of an old dual after two; such is the case for Irish, as we will see in Chapter 6.

[8] The irregular singular under discussion is different from the pre nominal or pre adjectival modification exemplified by *three foot tall* or *a three page document*, which are not likewise lexically restricted.

2.6.2 *Categorial restrictions*

The second type of restriction on number refers to the grammatical category of the base. Again, there is nothing lexical in the fact that a certain exponent for plurality only appears on nouns and not on other categories that inflect for number. The English plural formation in *-s* is as inflectional as one could wish, but it applies to nouns and never to pronouns (a distinction typical of inflecting languages, as opposed to agglutinating ones). Conversely, the Korean marker *-tyr* has the characteristic distribution of an 'optional' marker on nouns (in fact, it is governed by semantic and discourse factors rather than by grammar), but it is a necessary component of third-person plural pronouns, just like the Chinese *-men* (Unterbeck 1993: 190–2). Much more significant are cases where both nouns and adjectives inflect for number, but a certain pattern of plural exponence can only appear on nouns. The Italian irregular plurals in *-a*, which I will extensively discuss in Chapter 5, provide an example. The final *-a* of forms like *miglia* 'miles' does not express plurality anywhere else in the language; what is more, this formal irregularity is compounded by a change in the gender. Many, perhaps most, adjectives fall in the same inflectional class as *miglio* when they are masculine; but none ends in *-a* in the plural, regardless of gender. If the irregularity concerned the form alone, there would be no reason why it should not apply to some of the thousands of adjectives whose masculine singular ends in *-o*. In this case, special plurals are in fact plural-only derived nouns; adjectives, whose number value is instead determined by agreement, cannot display this type of inherent plurality. To consider a historically related example, Romanian features a large class of nouns which are masculine when singular but feminine when plural; not a few of them have the plural ending *-uri*, developed from a Latin neuter plural ending. Since *-uri* is now productive enough to have been applied to loanwords like *inputuri* 'inputs', one would think that it could spread to the feminine plural form of adjectives as well, which inflect for gender and number just like nouns and in fact largely show the same ending *-e* as many feminine plural nouns (e.g. *cas-e bun-e* 'good houses'). Instead, *-uri* remains exclusively confined to nouns, as a marker associated with those nouns lexically characterized as being ambigeneric (it also appears, exceptionally, on a few feminine nouns like *carne* ∼ *cărnuri*; see Chapter 5 below).

As these examples show, the fact that a certain plural marker is restricted to nouns does not in itself show that plurality has been lexicalized; but such a categorial restriction often accompanies lexical restrictions, and shows that plurality has become something more than the pure inflectional property that adjectives agree in.

2.6.3 *Lexical semantic restrictions: animacy and plurality split*

Finally, and most importantly, plurality can be restricted by the semantics of individual nouns. Lexical semantics can restrict the application of number in two main ways: by motivating a distinction between a set of nouns to which the number opposition applies and another set to which it does not; and by motivating subgroupings among nouns to which the opposition applies.

The first type of restriction leads to the phenomenon known (since Smith-Stark 1974) as plurality split, whereby languages exclude certain categories of lexical items from the singular–plural opposition on the basis of their meaning. In many languages, only items referring to humans and/or animates can be singular or plural, while the opposition is neutralized for inanimates. More generally, as Corbett (2000) shows, all number oppositions available in any given language tend to pattern in such a way that neutralizations, if there are any, concern nouns with lowest animacy that make up a continuous segment of the so-called animacy hierarchy (Corbett 2000: 90):

(2.8)   *The animacy hierarchy*
         speaker > addressee > 3rd person > kin > human > animate > inanimate

In other words, if a language distinguishes singular and plural on humans and inanimates, we expect it on non-human animates too. The categories 'speaker' and 'addressee' reflect the inclusion of pronouns, alongside nouns, in the lexical items affected by the plurality (and more generally number) split. For nominal number, the main factor in determining which classes of nouns take part in grammatical oppositions is certainly animacy. In a great many languages, only nouns with animate or human referents have a singular and a plural; other nouns either are not marked for number at all, or only optionally (keeping in mind that this usually means that pluralization is determined by discourse factors as opposed to grammar). Since a representative exemplification would take up too much space without adding anything to the main point, I will just mention a few cases for illustrative purposes (see Corbett 2000, chs 3–4, for a fully worked-out typology, as well as Rijkhoff 2002: 34–8).

Palauan maintains a clear distinction between human nouns, on which plurality is marked by the prefix *rę-*, and non-humans, which are not marked (Josephs 1975: 43). By contrast, in languages as different as Miya (Chadic) and Korean (Altaic), any noun can be pluralized in isolation (that is, preserving morphological well-formedness), but certain contexts enforce the pluralization of animate nouns alone. The animate/inanimate divide is more clearly grammaticalized in Miya: 'In phrases which are semantically plural, [+animate] nouns must be marked for plurality; for [−animate] nouns, plural

marking is optional' (Schuh 1989: 175). Another possibility is to mark both animates and inanimates for plurality, but with two distinct sets of affixes, as Rijkhoff (2002: 34) reports for Abkhaz. Finally, there are languages where plural marking is the exception rather than the rule. In Igbo, no noun is marked for number except for 'person' and 'child', which have suppletive plurals (Creissells 2000: 246). In other languages, only nouns for animates and humans have a plural form, and it is optional. This happens in most North American languages (Mithun 1999: 82–3), in Kilivila, spoken in the Trobriand Islands (Senft 1986: 45–6), and in Oromo (Owens 1985: 94).

Number clearly cannot be an inflectional category in the nominal system of a language if it only concerns a few nouns; such exceptional plural formations as Kilivila *tau* ∼ *tauwau* 'man ∼ men' are best seen as lexical derivations (or simply as distinct words in case of suppletion; cf. Section 3.3.3 in the next chapter). In this scenario, the language expresses plurality by means of its vocabulary, not of its grammar.

A clarification is needed to prevent a possible misunderstanding. Of course, even the English *-s* pluralization, inflectional as it is, is sensitive to lexical semantics: most English mass nouns are singular and cannot be pluralized without shifting their interpretation (cf. *wine* as a kind of beverage versus *wines* as sorts or concrete portions of wine), and some cannot be pluralized at all. **Funs* sounds odd because the meaning of the word as an abstract property generalization sits uncomfortably with the plural reading expressed by *-s*. Because of this perfectly natural semantic filtering, **funs* is ruled out and *fun* is ruled in; but it is ruled in as a singular, not as a general form. There is no plurality split in English, because the number opposition applies across the whole semantic range of nouns. By contrast, the restrictions I have been discussing in this subsection limit the application of the number opposition to semantically defined subsets of the vocabulary.[9]

---

[9] To highlight the difference between semantically motivated *singularia* (or *pluralia*) *tantum* and nouns to which number simply does not apply, consider the following example from Harar Oromo:

(i) eerúu c'úfá k'ote   (Owens 1985: 89)
     field   all   dug
     'he cultivated all the fields'

Owens explains that *eerúu c'úfá* does not mean 'all the farm', but specifically 'all the fields': 'If the noun it modifies is countable, *c'úfá* refers to a set of individuals, not to a spatial extent'. Then, the general form *eerúu* is not really ambiguous between singular and plural, because a singular interpretation would have licensed the reading 'all the field', 'the whole field'. Apparently the determiner *c'úfá* enforces a plural reading 'fields', but this has no bearing on the morphology of *eerúu*, which remains non plural (Owens does not say if a plural form would be acceptable; if not, a plural marker would actually be impossible in a context requiring semantic plurality). Exactly the opposite happens with the English *singulare tantum* **funs*, which is unacceptable because semantic and morphological plurality go hand in hand.

2.6.4 *Lexical semantic restrictions: other semantic categories*

A different type of lexical restriction obtains when the number opposition applies across the board, but groups of semantically related nouns form morphological subregularities. English zero-plurals like *sheep* or *deer* provides a convenient example, in this case based on nouns of animals. Being semantically homogeneous is not a clear-cut formalizable notion, so it is entirely expected that the status of some nouns may be dubious and that the semantic motivation should be less clear in some cases than in others. For example, the precise use of the plurals *fishes* and *fish* would seem to vary, and nouns like *elephant* or *lion* can have a zero plural (still countable, exactly like *sheep*) in the language of hunters and conservationists, where animals are referred to as tokens (cf. the discussion in Corbett 2000: 68).[10] Humans and, to a lesser extent, animates make obvious candidates for semantically homogeneous categories, and it is not hard to find examples where plural marking is sensitive to them. In Early Italoromance, for instance, the plural of certain kinship terms featured a stem extended by the suffix *-an-* (for example, *barba* 'uncle' ∼ *barbani* 'uncles'), which in some northern dialects became the form of feminine plural human nouns (Rohlfs 1968: 41–2, §371; notice the generalization kin > humans, in conformity with the animacy hierarchy). The Russian human nouns like *tatarin* 'Tatar' ∼ *tatary* 'Tatars', where it is the singular that displays an irregular stem extension in *-in-*, illustrate another morphological irregularity tied to a semantically restricted class (nationalities and other humans like *barin* 'landowner' or *gospodin* 'lord'). The restriction of certain plural markers to humans in Breton will be discussed in Chapter 8 below.

So-called 'collective numerals' like the Irish *triúr, ceathrar . . . deichniúr* 'three, four . . . ten people' are a special case of restriction to humans. This construction resembles phenomena that are attested in certain classifier languages, where numerals can be fused with different sorts of classifiers giving rise to counting forms that typically include a series for counting people. In Salish, in particular, people and animals have a certain prominence, because numerals for counting animates do not incorporate a particular classifier, but are simply reduplicated (see Anderson 1999). Although Irish is not a classifier language and the number opposition is firmly grounded in its grammatical system, this parallel will come as no surprise since, as I will

---

[10] The zero plural on *air craft (landing , space craft)*, which has nothing to do with animals, is truly exceptional. It resembles zero plurals like *a fleet of 20 sail*, *244 cannon*, *300 shell*, admissible in a military context that emphasizes the reading as indistinguishable tokens; but *aircraft* is different from these cases because it has no suffixed plural **aircrafts* (like **sheeps*, but unlike *lions*). Perhaps *craft* is really a singular abstract noun used in a concrete function, as in *20 personnel* (cf. note 8 in Chapter 3).

argue, Irish numerals govern a classifier head. Other languages too express plurality on numerals when they quantify over humans: Hungarian suffixes *-en/-an* to numerals as well as to quantifiers like *hány* 'how many?' and *sok* 'many' (Rounds 2001: 244), and Russian employs a special set of numerals with plural endings before human nouns of masculine grammatical gender; Wade (1992: 209) explains that 'the use of the collective numeral emphasizes the cohesiveness of the group, by contrast with the individualizing nature of the cardinals'. It must be added that Russian 'collective numerals' are also used for inanimate referents (typically with plural-only nouns), and so do not fully parallel the Irish personal numerals.

Plural morphology may be sensitive to semantic categories more specific than [human] or [animate]. Modern Hebrew, to take a clear example, has a binary gender opposition, which is reflected in the two plural endings *-im* for masculine nouns and *-ot* for feminines. The relation between these suffixes and gender, however, is loose enough to allow for some masculine plural nouns taking the 'feminine' ending and conversely (Aronoff 1994: 76 speaks of 'approximately 80' masculines taking *-ot* and 'thirty or so' feminines taking *-im*). In addition, feminine singular nouns take one of several suffixes, or no suffix at all, without any connection with the choice of plural suffix. The result, as Aronoff puts it, is that 'exceptional individual nouns are specified for individual singular and plural affixes *independently*' (p. 78; emphasis in the original). In other words, 'there are no noun paradigms in the language'.

Against this backdrop, it is easy to see subregularities. Two groups of feminine nouns with *-im* in the plural have an obvious semantic basis: the small animals *cfardea* 'frog', *tola'at* 'worm', *cipor* 'bird', *yona* 'dove', *nemala* 'ant', and *dvora* 'bee' and the plants *xita* 'wheat', *se'ora* barley', *te'ena* 'fig', *šošana* 'rose', *'afuna* 'pea' and *šibolet* 'ear (of grain)' (Schwarzwald 1991: 590–1). Other semantic fields emerge as relevant in the distribution of the suffix *-ayim*. This acts as a dual suffix, beside the plural *-im* and *-ot*, on a lexically restricted set of nouns. Schwarzwald (1991: 587–8) lists the following: *yom* 'day', *ša'a* 'hour', *šavua* 'week', *xodeš* 'month', *šana* 'year', *pa'am* 'time', *'elef* 'thousand', *me'a* 'hundred'. The lexical restriction clearly refers to units of time and of the counting system, not to notions categorized by their animacy (Fischer and Jastrow 1980: 89 report the same restriction for urban Arabic dialects in Morocco). If *-ayim* expressed dual as a category on a par with singular and plural, then (as Corbett 2000: 99 notes) Hebrew would have a number value that does not apply to a continuous segment starting from either end of the animacy hierarchy. Coupled with the fact that the dual is optional when two entities are referred to (Corbett 2000: 96), this strongly suggests that we are dealing with a lexically and not grammatically defined

class: -*ayim* can attach to a semantically homogeneous group of nouns to form derived nouns with the meaning 'two Ns'. This is very different from claiming that the inflectional system of Hebrew number has dual as a value alongside singular and plural.

The claim that -*ayim* is not an inflectional number affix receives strong empirical support by the detailed analysis of Tobin (1999), who views the affix as an exponent of dual but points out that dual has a much more restricted application that singular and plural, that it can co-occur with the plural on the same noun form, and that its use for referring to two-membered sets depends on the lexical item involved and not just on the quantity of referents. However, Tobin (1999: 93–5) also presents compelling evidence for the productivity of -*ayim*. Either the dual or the numeral 'two' plus plural can be used in many expressions, and the choice is 'more of a lexical rather than a purely grammatical choice in a way similar to choosing between two synonyms' (p. 93). What is more, 'speakers can idiosyncratically attach the dual to almost any noun for a comic, facetious, pejorative, metaphorical or rhetorical effect in a way similar to creating neologisms' (p. 94). This suggests that -*ayim* is a word-forming affix, always available in principle but lexically associated with unit nouns, on which it expresses duality. With another semantically defined group of nouns (this time body parts and clothes for them), -*ayim* has instead a plural meaning, referring to two or more than two instances, and contrasting on some occasions with regular plurals with a different meaning. Below are a few examples from the longer list in Schwarzwald (1991: 588–9):

(2.9)	SINGULAR	GLOSS	PL -*ayim*	GLOSS	PL -*im*/-*ot*	GLOSS
	regel	'foot, leg'	raglayim	'feet, legs'	regalim	'pilgrimages'
	yad	'hand'	yadayim	'hands'	yadot	'parts'
	ʕayin	'eye'	ʕenayim	'eyes'	ʕayanot	'springs'
	safa	'lip'	sfatayim	'lips'	safot	'languages'
	šen	'tooth'	šinayim	'teeth'		
	na'al	'shoe'	na'alayim	'shoes'		
	gerev	'sock'	garbayim	'socks'		

Rather than considering -*ayim* a dual ending, which would leave the cases in (2.9) as exceptions, it seems more appropriate to treat the suffix as a means to construct plural lexically rather than inflectionally. With one set of nouns, having a unit interpretation as a necessary but not sufficient condition, -*ayim* builds a plural whose special property consists of referring to two-membered collections. With another group, centred on body parts, the same suffix is instead interpreted as a simple plural (clearly because most are paired organs for which 'two' and 'all' coincide). This is a good example of lexical number.

2.7 Lexical vs. inflectional plurals: lack of determinism

Along with obligatoriness and generality, the main ingredient in what we intuitively perceive as inflectional regularity is determinism. By this I mean the property of entirely determining a unique form, automatically and necessarily selected by the grammar without room for choice. In Latin, for example, the ablative singular of *lupus* 'wolf' cannot be anything else but *lupo*, and in French the synchronically unpredictable form *yeux* /jø/ 'eyes' represents the only possible plural of *oeil* /œj/, regardless of the irregularity. Carstairs (1987: 31) calls 'inflectional parsimony' this very strong tendency of inflectional paradigms to realize each cell by no more than one word form, and in so doing he points out that things could stand otherwise. In fact, as we have seen with doublets like *mice ~ mouses*, grammatically equivalent plural alternants are exceptional but not impossible, and when they arise languages tend to semantically differentiate them. When a noun has more than one plural form, therefore, the alternants are no longer the grammar-driven outcome of an inflectional process, but involve a lexical choice. I will distinguish four types of differentiation: stylistic, contextual, semantic, and semantic-grammatical. These descriptive categories will guide us through a broad spectrum of lexicalized plurals, from sporadic exceptions in otherwise deterministic inflectional paradigms, to languages where pluralization is fundamentally non-deterministic and routinely involves lexical choices.

2.7.1 *Stylistic variants*

Carstairs himself (1987: 30) acknowledges that, rarely, inflectional doublets may be as equivalent semantically as they are grammatically. This happens in Finnish, where nouns in the genitive and inessive plural attach the case endings to a stem optionally enlarged by the extension *-oi-* (*-oitt-* or *-oid-* in the genitive plural):

(2.10)	Gloss	Nominative singular	Genitive plural	Inessive plural	(Finnish)
	'potato'	peruna	perunien	perunissa	
			perunoitten	perunoissa	
	'circle'	ympyrä	ympyrien	ympyrissä	
			ympyröitten	ympyröissä	

The alternative forms seem to be genuine equivalents; the few informants I could consult expressed at most a preference for one or the other alternant on isolated words (one informant, for example, judged the illative plural *mustikoissa* to be

much better than *mustikissa* for the noun *mustikka* 'blueberry'). There is no evidence here that plurality is lexicalized in any sense on either alternant. Nor, I think, should lexical plurality be invoked for doublets that arise from socio-linguistic variability in the degree to which non-native vocabulary is adapted: in English, the plurals of *formula* and *cherub* may be sounded *formulae* and *cherubím*, or *formulas* and *cherubs*, according to the cultural awareness and the intentions of a speaker.[11] I will therefore ignore them in what follows.

2.7.2 *Contextual and semantic-contextual variants*

The case of contextually differentiated variants is altogether different. In the clearest scenario, contexts determine a pattern of complementary distribution between alternants. Irish counting plurals, to be analysed in Chapter 6, illustrate this category: the special plural forms required by the numerals 3–10, here exemplified by *bliana* 'years', do not fit in any other context, and their regular counterparts (here *blianta*) are generally unacceptable with numerals:

(2.11) a. trí bliana   (Irish)

       b. *trí bhlianta
         'three years'

(2.12) a. *an iomarca bliana

       b. an iomarca blianta
         'too many years'

I will propose in Chapter 6 that doublets like *bliana* and *blianta* are not actually two lexical items in their own right, but two morphologically and syntactically distinct entities, one a fully-fledged noun (the regular plural), and the other a distinct stem form of the same noun, inserted in the syntactic slot of a classifier. This raises all sorts of questions, first among them whether this syntactic difference warrants viewing counting plurals as non-inflectional. I will argue that it does, in that these special forms also have a special meaning as indistinguishable token units. Irish counting plurals thus represent the forms assumed by certain nouns when employed in a classifier capacity; in this role, they are actually numberless, neither plural not singular.

  Irish counting plurals loosely resemble what Stankiewicz (1962) called quantifying plurals in several Slavonic languages, understood as special plural forms that occur after numerals (or after some numerals) and contrast with a

---

[11] Archaic forms also belong here, if they have no distinctive trait but their archaic flavour. I will briefly refer to the historical stratification in speakers' vocabulary in the context of Arabic broken plurals (Chapter 8).

non-count plural, or an affective one, or an unmarked one. The plural of Macedonian *list* 'leaf' has *listovi* as unmarked form, *lista* after numerals and *lisje* as non-count collective, for example (Stankiewicz 1962: 7). As Stankiewicz notes, such alternants are not in complementary distribution after numerals, which proves that they are lexical derivations and not grammatically determined allomorphs. Slavonic quantifying plurals differ from the Irish ones in one major respect: they do not isolate a semantically homogeneous small group of nouns used as standards for quantifying. Other languages, which have no morphological class of counting plurals, have isolated examples of nouns with a special classifier-like reading. In Dutch, for example, 'the plural of *stuk* (piece) is *stukken*, but *stuks* occurs too with the meaning of items, for example *Ik heb tien stuks gekocht* I bought ten (pencils, balls, etc.)' (Donaldson 1987: 34). The unit-value interpretation of *pence* parallels in this sense that of 'singulars' like *three fathom* (cf. Section 2.6.1 above and Chapter 6). These are all special classifier constructions, as already Greenberg (1974: 33) had seen, and the plurals that appear in this capacity are lexical, because inflectional plurality is incompatible with classifier constructions (see note 5 above).

Unlike Irish counting plurals, Germanic unit nouns are not all tied to numerals, as shown by phrases like *a few stone*, *a few bob* or the German *alle Mann am Deck* 'all hands on deck' or *drei gemischte Eis* 'three mixed ice-cream[s]' (where *Mann* and *Eis* are singular but *alle* and *gemischte* are morphologically plural). This greater freedom from a specific context makes Germanic counting forms more similar than their Irish counterparts to the next category to be examined, that of semantically differentiated irregular plurals.

2.7.3 *Semantic variants*

Semantically differentiated plural doublets are extremely common. In fact, it is a plausible suggestion that every language with the appropriate morphology (where number is an inflectional category with some allomorphy among plurals) has at least some. The previous subsection has considered plural doublets where the difference in meaning correlates with distinct contexts, and the next one will focus on those where it correlates with distinct gender and noun class values. This leaves cases like *brothers* ~ *brethren*, where the distinction concerns meaning alone, neither syntax nor morphological class.

In English, the meaning of *brethren* (confrères) is specialized with respect to that of the regular *brothers*; the opposite is true of the pairs *oxen* ~ *oxes*, *lice* ~ *louses*, and *mice* ~ *mouses* (that is, it is that regular form that is associated with a particular meaning, here metaphorical). Dutch, in which -*en* and -*s* are both widespread plural suffixes, has more examples. Beside doublets that differ in

gender, Donaldson (1987: 33) provides the following eight doublets whose only formal difference is the suffix:

(2.13)   hemelen 'heavens' ~ hemels 'canopies'   (Dutch)
         hersenen 'brains' (as organ or food) ~ hersens 'brains' (as food)
         letteren 'literature' ~ letters 'letters' (of the alphabet)
         middelen 'means' ~ middels 'waists'
         tafelen 'tablets' (of the law) ~ tafels 'tables' (for food)
         vaderen 'forefathers' ~ vaders 'fathers'
         wateren 'waterways' ~ waters 'waters'
         wortelen 'carrots' ~ wortels 'roots'[12]

Phonologically, the suffix expected after the endings *-el*, *-en*, and *-er* is *-s*. Indeed, three of the forms in *-en* are semantically idiosyncratic: *letteren* as 'literature' (or 'letters'), 'forefathers', and 'waterways'. However, there is nothing intrinsically more idiomatic in 'carrots' than in 'roots' (both nouns are countable), although the former ends in *-en*; and the *-s* plural *hersens* has a mass interpretation that does not arise compositionally from the pluralization of 'brain'. Two conclusions can be drawn from these Dutch doublets, supported by the previous English and German examples. First, there is no one-to-one relation between irregular form and irregular meaning: *-en* marks a semantically regular plural in *oxen* and *wortelen*, but not in *brethren* and *wateren*. Second, the semantic distance between the two plural forms varies considerably, from 'means' ~ 'waists', to the much more closely related 'brains as organs' ~ 'brains as mass', all the way to 'words in context' ~ 'words in isolation'.

   Russian fully supports both conclusions. Semantic differentiation between the members of plural doublets is relatively common, if lexically restricted, in Slavonic languages, especially those which lost gender distinctions in the plural (Stankiewicz 1962: 6). Wade (1992: 59–60) lists 23 doublets among masculine plural forms, all involving the alternation between the nominative endings *-i* (or *-y*) and stressed *-a* (or *-ja*). They are divided in three groups: eleven show a clear semantic differentiation, two a specialized interpretation, and ten a mere stylistic difference. Here are some examples from all three classes, with accents added to mark stress:

(2.14)   a. konduktorá 'bus conductors' ~ kondúktory 'electrical conductors'
                                                              (Russian)

             toná 'colour shades' ~ tóny 'musical tones'
             tsvetá 'colours' ~ tsvéty 'flowers'

---

[12] I reproduce Donaldson's glosses; several Dutch speakers informed me that *wortels* may have either meaning.

b. lagerjá 'camps' ∼ lágeri 'political camps'
   učiteljá 'teachers' ∼ učíteli 'teachers of doctrine'

(2.15) STANDARD/WRITTEN COLLOQUIAL/JARGON GLOSS   (Russian)
        gódy   godá   'years'
        pékary   pekarjá   'bakers'
        slésari   slesará   'locksmiths'

Examples like (2.15) might suggest that *-a/-ja* cause a shift in register towards informality and/or professional jargon, but (2.14) shows that this cannot be true. Cases like (2.14b) and (2.15) together might be taken to indicate that the function of these alternative endings is to provide a plural with a specialized interpretation, but specialization is not what is involved in the differentiated pairs in (2.15a). Again, we must conclude that the alternative endings do not 'mean' one thing or another, but merely provide the morphological means for a semantic differentiation, which spans the whole range from variation in register to a decidedly lexical shift in the noun's interpretation.[13]

When neither contextual restrictions nor other grammatical categories play a role, then, plural doublets exemplify the separation of morphological form and function (in the sense of Beard 1995 and Aronoff 1994): the formal oppositions they provide are put to different semantic uses, and there is no one-to-one relation between alternative endings and lexical plurality, any more than between endings and particular interpretations.

This conclusion has been reached on the basis of a few typologically similar languages, in which plural doublets are a relatively marginal phenomenon. In other languages, the indeterminacy of nominal plural formation is firmly entrenched in the morphological system itself. Breton provides a striking example. The richness and complexity of its nominal number system justify devoting the entirety of Chapter 8 to it. Leaving the details to that chapter, I will only cite here some remarks from Trépos (1956), a whole monograph on the Breton plural written with the insight of a linguist who was also a native speaker. In the first chapter, Trépos took particular care to make plain the manifold variability of Breton plural nouns. Having illustrated the strong dialectal differentiation, he adds:

S'il n'est past trop surprenant, à la reflexion, de trouver des pluriels différents en des régions différentes, il l'est plus de les trouver dans la bouche d'une même personne, qui

---

[13] I don't mean to suggest that the interpretation of nouns ending in *á* is entirely unpredictable. John Bailyn has pointed out to me that recent computer related loanwords featuring the English *er* suffix systematically have *á* plurals (*userá*, *driverá*, *printerá*), and that *brat* 'brother' has an idiomatic plural *bratvá* 'brothers, members of a gang'. There is clearly a pattern, pointing to some productivity of *á* as a plural suffix with an informal characterization. My point is that this is not *the* meaning of *á* in plural doublets.

utilise instinctivement l'un de préférence à l'autre, selon le sens, ou simplement la nuance, qu'elle veut exprimer....Le breton possède t il donc, pour le nombre, des moyens d'expression que l'on ne peut faire rentrer dans le cadre d'une grammaire classique?

[If, on reflection, finding different plurals in different regions is not too surprising, it is more so to find them on the lips of one and the same person, who instinctively uses one form instead of the other according to the meaning, or simply the shade of meaning, that he or she wants to convey....Does Breton, then, possess for number some means of expression which cannot be made to fit the frame of a classical grammar?] (Trépos 1956: 23)

The short answer is 'oui', namely plurality as a lexical property, and Trépos says as much when, at the close of the same chapter, he warns against the temptation to isolate one form as *the* plural of a noun. To do so, he intimates, would fail to recognize that the choice among alternative plurals is one of the characteristic traits of the language. He explicitly contrasts (p. 32) the unacceptability of morphologically ill-formed French words like **chevals* and **oeils* for *chevaux* and *yeux*, 'horses' and 'eyes', with the effect of a novel and unusual plural on the ears of a native speaker, which would amount at most to an 'amused surprise' (*une sorte de surprise amusée*). In sum, as nonce forms make especially clear, plural in Breton can act as a word-forming category. As we will see in Chapter 8, its real interest is that the same morphology also spells out plural as an inflectional category.
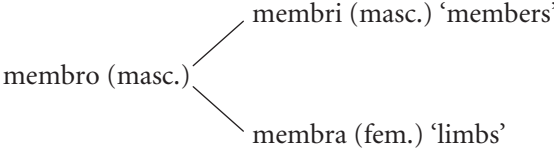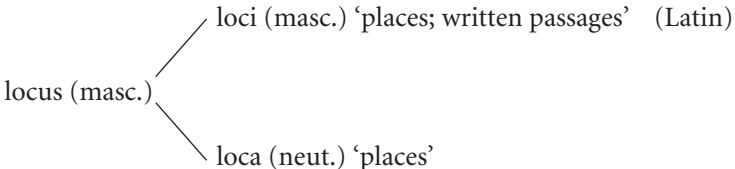
   Notice that a language may have non-unique plural forms on a large scale without necessarily displaying the peculiar richness and syntactic flexibility that are typical of Breton. Berber has a complex but overall fairly regular system of plural construction, in many ways parallel to the Arabic system (Idrissi 2000; see Chapter 7). However, speaking of the Tamazight dialect, Penchoen (1973: 14) specifies: 'Although there are a number of fairly regular rules, even Tamazight speakers accept, and may even use actively, different plural forms for the same noun'. This quote highlights the distinction between the array of plural exponents and the map between singular and plural. The former may range from small to complex, the latter from deterministic to unpredictable, independently of each other.

2.7.4 *Semantic-grammatical variants*

The variants we have considered up to now are grammatically equivalent. Others are differentiated not only semantically but also morphologically, typically by the use of different class diacritics like gender or noun classes. The tendency to differentiate grammatically equivalent forms has an obvious functional basis. When the nouns of a language are already divided up into

categories that have or can have some semantic motivation, it is only natural that plural doublets should be differentiated according to these categories.

A doublet such as the German *Bände* 'volumes' (masculine) ~ *Bänder* 'ribbons' (neuter) is not a good illustration, because the shared singular *Band* has one gender or the other according to the interpretation: *Band* ~ *Bände* is masculine and means 'volume(s)', *Band* ~ *Bänder* is neuter and means 'ribbon(s)', but apart from having homophonous singulars they are perfectly regular nouns. Gender differentiates two plurals only when a singular with one gender is paired to two plurals with distinct genders.[14] The following are cases in point:

(2.16) a.   membri (masc.) 'members'   (Italian)

          membro (masc.)

                         membra (fem.) 'limbs'

       b.   loci (masc.) 'places; written passages'   (Latin)

          locus (masc.)

                         loca (neut.) 'places'

Such a schematic presentation distorts somewhat the actual relations between the noun forms, because it suggests that the singular is perfectly ambiguous and that the gender opposition disambiguates two clearly distinct interpretations. The facts are more complex and more revealing. To keep to the two examples above, the singular *membro* only ever means 'limb' as an euphemistic term for 'penis', and is otherwise unambiguously the singular of *membri*. It is the feminine *membra* which stands out, in asymmetric relation to the regular pair *membro* ~ *membri*. In the Latin *locus*, on the other hand, the two plural forms largely overlap in meaning, but are semantically distinguished by the fact that only the regularly masculine *loci* could refer to written passages or rhetorical patterns.[15]

---

[14] In principle, gender may signal lexical plurality when a noun has only one plural form, differing from the singular in gender value. As far as I know, the gender mismatch between the French *orgue* 'organ' (masc.) ~ *orgues* 'organs' (fem.) (both referring to the musical instrument) has no semantic correlate.

[15] Lexical meanings cannot be precisely discriminated in a corpus language, especially in one as historically stratified as Latin. While the details vary (presumably for each dialect, if not for each attestation), the lexicographic and grammatical traditions agree in attributing distinct semantic values to *loci* and *loca*, as well as to several other doublets (cf. Schön 1971: 39 82).

As can be seen, just two doublets in immediately related languages suffice to show the intrinsic variability of this kind of opposition between plurals. Some singulars are ambiguous, others are not; sometimes the semantically compositional plural is the morphologically regular one, sometimes it is not; some plural doublets distinguish two relatively distinct senses, some others a general and a specialized sense. As we concluded for the doublets in the previous section, there is no direct, transparent relation between morphological irregularity and interpretive opacity. This theme will be pursued in much greater detail in Chapter 5. We will see there that some of these Italian doublets feature clear semantic distinctions, like 'limbs' ∼ 'members', or 'entrails' ∼ 'alleys'; but others display a much more elusive semantic differentiation, for example *ossa* 'bones (as body parts, although not necessarily in a single skeleton)' ∼ *ossi* 'bones (as unrelated pieces)'. Parallel cases from classical Greek show more clearly that the elements of plural doublets differ in conceptualization rather than in basic lexical meaning: the masculine *keleuthos* 'path', *kyklos* 'circle' and *meros* 'thigh' have regular masculine plurals, plus an additional neuter form signalling a conceptualization as an undifferentiated mass or an abstract notion (as in 'the paths of human beings'; cf. Delbrück 1893: 123–7, Brugmann 1900: 369, §429, Schwyzer and Debrunner 1950: 37, Wackernagel 1926: 88–9; Schön 1971: 39–82). Since the neuter plural is semantically motivated, these examples differ from the Italian ones in that the association with a gender carries a meaning, like a derivational affix.

The noun-forming function of gender shows up very clearly in the pluralization patterns of Somali, a typologically rather different language. Like other languages in the Cushitic branch of Afro-Asiatic, Somali has only two 'genders' in the sense of nominal agreement classes, but it freely uses them to express functional oppositions within the paradigm of one and the same noun. This results in so-called gender polarity, whereby a noun has one gender in the singular and the other in the plural. I will focus on this function of gender in Sections 3.4.1 and 3.5.2 in the next chapter; here, what is relevant is that this type of pluralization is non-deterministic. Lecarme (2002: 120) provides the following examples of one-to-many matches (note that in some cases plurality is only expressed by a change in stress pattern):

(2.17)   a. díbi 'bull' (masc.)   dibí 'bulls' (fem.)   (Somali)
                                              dibi-yó 'bulls' (fem.)
                                              dibi-yaál 'bulls' (fem.)

            b. dáas 'shop' (masc.)   daas-ás 'shops' (masc.)
                                              daas-yó 'shops' (fem.)

c.   túug 'thief' (masc.)   tuúg 'thieves' (fem.)
tuug-ág 'thieves' (masc.)
tuug-ó 'thieves' (fem.)

As we will see in the next chaper, Lecarme presents generalized multiple plurals as one of the features that characterize Somali plurals as derivational, along with the sensitivity of plural suffixes to the previous morpheme (stem or derivational affix), the possibility of adding a plural suffix to an already plural form, and the use of plural nouns in derivation and composition. It seems fair to conclude that, if coexisting alternants are generalized where pluralization consists in lexical derivation rather than grammatical inflection, episodic plural doublets signal lexicalized plurality in systems where number is otherwise inflectional.

In noun class languages, where nouns consist of a stem paired with one of several classes of prefixes, the noun-forming function of pluralization emerges more clearly, because classes have a much clearer semantic motivation than gender values (see Section 3.5.2 in the next chapter for further discussion). The classes of Bantu languages, to cite the most widely known example, typically group together nouns whose referents share a semantic common denominator like humans, animates, animals, and plants. However, classes may also have a grammatical basis, like the Swahili class 15 with prefix *ku-*, which Polomé (1967: 103) describes as 'used exclusively to form infinitives by prefixing {ku} to any verbal stem'. In many cases, the semantic opposition between 'one' and 'many' is isomorphic to an opposition between noun classes: the prefixes *wa-* and *mi-* of classes 2 and 4 in Swahili clearly stand to the *m-* and *m(w)-* of classes 1 and 3 in the relation of plural exponents to singular ones:

(2.18)   *Class 1*   *Class 2*   (Swahili; Polomé 1967: 96)
m-toto 'child'   wa-toto 'children'
Class 3   *Class 4*
mw-aka 'year'   mi-aka 'years'

Some classes are based on concepts like mass and individuation, however, and it is here that number becomes tangled with the semantic underpinnings of classes. As Contini-Morava (1999) shows, class 6 (prefix *ma-*) serves to form the plural of semantically singular class 5 nouns; for example, *ma-we* 'stone' pluralizes the class 5 *ji-we* 'stone'. However, when paired with some nouns of the 'singular' class 11 (prefix *u-*), *ma-* also expresses a collective reading that contrasts with the simple plural:

(2.19)

```
                           nyu-ta 'bows'      (Swahili; Contini-Morava 1999: 8)
                        /
         u-ta 'bow'
                        \
                           ma-ta 'bow and arrows' (a set)

                         / shanga 'beads'
      u-shanga 'bead'
                         \ ma-shanga 'string of beads'
```

Besides, *ma-* characterizes a series of mass nouns which are not paired with another class: 'water', 'millet flour gruel', 'fallen debris from trees', and also 'conversation' (Contini-Morava 1999: 6). Similarly, the 'singular' class 11 (prefix *u-*) groups nouns for individual objects, but also for abstractions like 'sleep' and for referents that, while not discrete, are more cohesive than those of class 6 (like 'cooked rice', 'honey', 'gum, glue'). Interestingly, class 11 and class 6 often stand in the relation of abstract to concrete concept: *u-lezi* 'tutorial activity' ∼ *ma-lezi* 'training, education', *u-laji* 'act of eating' ∼ *ma-laji* 'food' (Polomé 1967: 99). So, in many cases class 6 nouns are not semantically plural (or contrast with a plural, as in (2.19)); likewise, the nouns in class 11 are semantically non-plural rather than properly singular; and nouns in the two classes are not systematically paired. Similar considerations apply to the remaining classes. Contini-Morava (1999) argues that this system is based on a hierarchy of individuation, from the most individuated concepts in the classes that just pair singular and plural interpretation, to masses and abstractions. Some aspects of this analysis are open to discussion, because the vague concept of individuation leaves too much indeterminacy in the middle of the continuum (class 11 nouns with *u-* range over bounded objects, abstractions and masses, which suggests lack of discrete internal part structure rather than intermediate individuation); but the general conclusion stands, and it is that at least some Swahili noun classes encode part structure and not number. Number is still an inflectional category (in so far as it underlies agreement), but to select a number value for a nominal stem means to assign it to a class, which means to make up a noun. Unsurprisingly, attribution to a class is not a deterministic, one-to-one process. Contini-Morava (1999: 11) shows that 'not all nominal stems are uniquely associated with a single noun class prefix' (apart from the human nouns of classes 1–2). Creissels (2000: 243) summarizes as follows the resulting interaction between 'gender' (here, noun classes) and number in Niger-Congo languages generally:

gender and number interfere in a particularly intricate way: it is impossible to isolate plural markers as distinct from gender markers; nouns that belong to the same gender in the singular often belong to different genders in the plural, and conversely; alternate plural forms (with sometimes more or less subtle shades of meaning) corresponding to the same singular form are not uncommon. (Creissels 2000: 243)

In this situation, where many nouns can belong to many classes, for a noun to be plural means to fall into one of the classes with non-singular interpretation, not to be the output of an inflectional rule determining (like a mathematical function) a unique value for any input. Noun class languages have lexical plurality built in, as it were.

2.8 Lexical vs. inflectional plurals: semantic opacity

If the meaning of a plural noun cannot be completely inferred from the meaning of the base and of the plural morpheme, the plurality of that noun is lexical in a semantic sense: *books* is compositional and inflectional, but *goods* is lexical, because its sense of 'merchandise' remains inaccessible if one only knows the meaning of *good* and that of *-s*. Even though, as we saw above (Sections 2.2 and 2.3), lexical plurality cannot be reduced to semantic opacity, this lack of compositionality shows that plurality is an integral part of the semantic content of a word.

   This much is intuitively clear. Questions arise, as usual, as soon as we want to make this 'word-relatedness' more precise. Whatever one understands as lexical item, it sounds plausible to treat *goods* as a lexical item distinct from *good*, but to extend the claim to *rains* and *rain* just because the former does not admit a paraphrase like 'many a rain' seems arbitrary, given that the semantic difference between the two is slight and does not have to involve a distinction in denotation (*the rains of September* and *the rain of September* may well refer to the same occurrences). The semantic irregularity of plural is a matter of degree, and throwing every deviation from the canonical reading into the same basket of lexical exceptions would obscure the diversity of the phenomena. Besides, the semantic range of lexical plurals is wide but neither random not unbounded. A suitable empirical sample reveals semantic generalizations which show how the lexicalization of semantic plurality is a linguistic phenomenon with its own regularities and its own boundaries, far from being an amorphous collection of unpredictable meanings. In this section, I will outline the extent and the variety of semantic lexicalization of plural, leaving a proper categorization to Chapter 4. We will first consider specialized readings along the count–mass axis; then, cases where plurality

correlates with a particular part-structure conceptualization, but without a change in mass or count status; and finally, the relation between semantic and morphological oppositions, in connection with the alternation between singulative and non-singulative.

2.8.1 *Plural and mass*

Books is true of sets of things *x* such that each *x* is a book. The polysemy of the base does not matter: *books* still means 'many a book' whether this term refers to a type or to a token (*the author of three books* vs. *he burnt his books*), to a text or to a bound volume (*two books in one*). Second, the count plurals of mass nouns, as in *three coffees to table one*, or *we produce many cheeses*, are semantically transparent and so will not be further analysed here, because it is not pluralization that makes these nouns count. *Three wines* is not the plural count of a singular mass, but the semantically transparent plural of the singular count *one wine* 'one type/serving of wine'. These forms have semantic interest, but it does not lie in the use of number.

What sets apart the compositional and the non-compositional interpretation of plurals like *funds*, discussed above in Section 2.4.1, is instead the fact that, on the relevant reading, *funds* does not refer to a plurality of things each of which is a fund. This is the basic semantic property that distinguishes mass from count nouns. In Quine's (1960: 91) classic definition, only count nouns 'possess in-built modes, however arbitrary, of dividing their reference', incorporating a criterion that allows speakers to judge when one unit ends and another begins. When *funds* means 'funding', by contrast, it does not provide a criterion to discriminate what is part of one unit and what is part of another, because it does not articulate its reference into discrete units. The availability of a plural form is generally taken to be a hallmark of count nouns in English and related languages. Leaving to Chapter 4 a critical examination of the relation between mass and plurality, let us focus here on the direct evidence that a sizeable amount of nouns have a lexicalized plural mass reading. In the relevant sense of 'funding', *funds* does not admit numerals or other count modifiers: *I have some funds* can mean 'I have some money available', but the count context *I have three/a few funds* enforces reference to collections of discrete units. To take some other examples, *holidays* in the sense of 'holidaying time' (American *vacation*) is incompatible with count contexts like *three holidays* or *she enjoyed each one of her summer holidays*, which must refer to distinct years (contrast *she enjoyed all of her summer holidays*). Similarly, *I have plans for tonight* and *I have a few plans for tonight* have different meanings ('I'm busy' vs. 'there are various things I intend to do'). The most obvious examples of mass plurals in English come from *pluralia tantum* (*a few

suds, **how many furnishings*), and this has no doubt contributed to the impression that mass plurals are just unsystematic exceptions. Note, however, that only a few of the English semantically irregular plurals mentioned above in Section 2.4.1 lack a singular.

One may be tempted to equate the semantic lexicalization of plurality with a massifying function. But not all mass plurals have a count singular, as shown by *depths*, *heights*, *rains*, *snows*, *waters*. Mass plurals of mass singulars are common in languages other than English, and we have already seen in Section 2.4.2 many substance-referring nouns that have both singular and plural; further examples are Latin *harena* 'sand' or *nix* 'snow', Greek *psamathos* 'sand' and *haima* 'blood', and Lithuanian *kraūjas* 'blood' (Delbrück 1893: 146–72). In these cases, plurality does not act as a massifier, because the singular is mass already. However, it still expresses a lexicalized reading, which differs from the canonical 'many *x*' reading that is the normal interpretation of inflectional plurality in all of these languages.

Having made clear that a mass reading is not incompatible with plurality, and that it is not what lexicalized plurality 'means', we can better appreciate the way number can express lexical semantic variations along the mass–count axis. This is particularly clear with Indo-European languages of ancient attestation, where apparently either number value could express either reading. To take just one example, consider the nouns for 'meat/flesh' (Delbrück 1893: 151). The Sanskrit *masá* is said to occur also as plural, with an example where *masani* is translated as 'meat bits' (*Fleischpartien*, in the context 'the meat bits of a piece of fat'). Similarly, Latin has the singular *caro* but the plural *carnes* as 'pieces of meat' (*Fleischstücke*). Things are reversed in classical Greek, where the singular *kréas* means 'piece of meat', while the plural *kréa* is used 'wenn es im kollektivem Sinne, als Fleisch zum Essen gebraucht wird' [when it is used collectively, as meat in the sense of food] (p. 151). But the pairing of singular or plural with mass or count varies with the word and does not remain constant in a language: Greek has many mass plurals like *sárkes* 'flesh' ('on a living body': Brugmann 1900: 369), *háles* 'salt', and *ksýla* 'wood', the singulars of which occur with the meanings 'piece of flesh', 'grain of salt', and 'plank of wood' (Schwyzer and Debrunner 1950: 42–3).

2.8.2 *Neither mass nor count*

A shift along the mass–count dimension is not the only effect of the semantic lexicalization of plural. I will devote the whole of Chapter 4 to examining the full range of possible readings, but some examples are in order here to show that change in the mass or count status is a subcase of change in part-structure conceptualization.

The reading of *funds* as 'funding' qualifies as mass rather than count because, among other things, *a few funds* cannot mean the same as 'some funding'. Now, the same applies to a noun like *gates*: *a few gates* must refer to a small-numbered collection of discrete entities, and cannot mean 'a gate (of a certain extension)'. But we can also use *gates* to refer to one entrance, as in *the pearl gates*. This reading is incompatible with a grammatical count determiner like *a few* or *several*, and yet it would be very counterintuitive to assert that *gates* is mass in *pearl gates*. In fact, it would be wrong, because the expression refers to one entrance and not to an arbitrarily large collection or an arbitrarily small subpart of it. The function of the plural in *gates* as 'one entrance' is to conceptualize the referents as non-simplex. Yet its linguistic effects are very sharp, because there is nothing elusive about the intuition that *a few gates* cannot mean 'one entrance'.

Mostly, this use of plurals is mentioned in connection with *pluralia tantum*, which not infrequently refer to bounded individual objects; examples, in addition to those in Section 2.4.2 above, include Russian *nosilki* 'stretcher', *očki* 'spectacles', and Latvian *ragavas* 'sledge' (Koptjevskaja-Tamm and Wälchli 2001: 629, 632). But nouns that, while referring to single individuals in the plural, also have a singular form, like *gate* ∼ *gates*, provide even stronger evidence for the re-conceptualizing function of plural. The philological descriptions of lexicalized plural in Indo-European, mentioned above in Section 2.4.2, complemented the lists of *pluralia tantum* with several instances where both the singular and the plural can refer to just one unit of reference: cf. singular–plural doublets like, in Latin, *castrum* ∼ *castra* 'fortified camp', *currus* ∼ *currus* 'a chariot', *naris* ∼ *nares* 'a nostril', and in Greek *domos* ∼ *domoi* 'a house', *harma* ∼ *harmata* 'a chariot', *prosopa* ∼ *prosopata* 'face', *sternon* ∼ *sterna* 'chest, breast'. This sort of pairwise listing should not suggest that the choice was always freely available to speakers. But the evidence suffices to show how frequently these languages could employ a grammatical category like plurality to express some aspects of the lexical semantics of a nominal lexeme. Being lexical in the sense of affecting the core descriptive content of the noun, this use of plurality is also lexical in lacking the generality of inflectional categories. As Löfstedt (1928: 33) noted, we find the Latin plurals *currus* and *pectora* for, respectively, a single 'chariot' and the 'chest' of one man, but never the plurals *equi* or *viri* for a single 'horse' or 'man'.

2.8.3 *Morphological reflexes of lexicalized readings*

In the languages we have considered so far, certain nouns associate particular readings to the singular or the plural form, but this does not affect the morphological system in any way. Other languages, by contrast, express the

part-structure conceptualization of nouns by means of a special morphology. I am not referring to systems in which nouns have 'plural' forms that in fact only express cohesion or proximity, like those mentioned in Section 2.5 above, but to systems where an inflectional number opposition interacts with a morphologized opposition between mass and count readings. Semitic and Slavonic provide a clear illustration. These languages feature a class of count nouns traditionally called singulatives, derived from so-called 'collective' bases; a typical example is Russian *gorošina* 'a pea' derived from *gorox* 'peas as a mass' (singular). Singulatives can be singular or plural like any other count noun, and the relation between form and meaning is perfectly regular and transparent. To make these remarks more concrete, consider the following examples from Syrian Arabic and Ukrainian:[16]

(2.20) 'COLLECTIVE' DERIVED COUNT (*singulative*)
     samak 'fish'   samake 'a fish'   (Syrian Arabic; Cowell 1964: 369)
     pisok 'sand'   pisčyna 'a grain of sand'   (Ukrainian)

Recall that the derived singulatives also have a plural, here *samakaat* and *pisčyny*. The interesting thing is that some of the mass 'collectives' can be pluralized too, resulting in another mass noun with a different conceptualization, typically combining non-denumerability and large size (see Corbett 2000: 30–5 for more examples of so-called greater plural). The plural of *pisok* is *pisky*, roughly 'a lot of sand', but also 'sandy territories'. That of *samak* is ʾasmaak, which Cowell (1964: 369) renders as 'many or various fish', generally describing this type of plurals as indicating 'abundance, variety, or indefinite quantification'. Note that the plural of the Arabic 'collective' *moož* 'wave', namely ʾamwaaž, is glossed as '(many or extensive) waves', where multiplicity shades into extension (cf. further Chapter 4). Because both singulatives and 'collectives' can occur as singular or plural, this case clearly shows that the singulative–non-singulative opposition is orthogonal to the number opposition.

  Breton illustrates another way in which plurality can modify the conceptualization of a noun form that is supposed to be non-pluralizable. Here, the morphologically basic form of many nouns is not the singular, but a syntactically plural 'collective'. Some of these plurals support further pluralization, with a shift in meaning that must be precisely determined for each noun (cf. Trépos 1956: 264–5):

(2.21) tud   tud-ou   (Breton; Trépos 1956: 225)
        people.PL people.PL-PL
        'people'   'peoples, nations'

        dilhad   dilhaj-ou
        clothes.PL clothes.PL-PL
        'clothes' 'suites of clothes; garments'

This sort of non-canonical pluralization will figure prominently in the case studies of Arabic and Breton (Chapters 7 and 8). The relevant conclusions here are that the phenomenon is well attested, that it is not a morphological oddity but a process that systematically alters the conceptualization of the reference domain of a noun, and that this semantic function is not reducible to that of regular inflectional plurality. In short, morphological plurality quite often has a semantic function distinct from the transparent value associated with inflectional number. Whether this sense is associated with specific words, or is more generally an interpretive option for the plural category, or identifies a group of plurals distinct from the corresponding count versions (as in Arabic), it represents an invaluable window on the semantics of lexical plurality. Before turning to that, however, we will have to make more precise what it means for a plural form to be lexicalized, beyond the non-inflectional traits reviewed in this chapter.

2.9 Conclusion

The two parts of this chapter in different ways make the case for addressing lexical plurality as a significant linguistic issue: the first (Sections 2.2–2.4) by showing how much would missed by reducing it to listedness, non-compositionality or lack of a singular; and the second (Sections 2.5–2.8) by showing that the prototypical properties defining inflection can fail to hold in a great many cases, resulting in forms and/or readings that can all be traced to the notion of plural number being lexicalized. The facts considered so far suffice, by themselves, to show that an account of plurality limited to the regular, productive, and transparent opposition exemplified by *book* ∼ *books* would be seriously incomplete. Besides, the evidence just reviewed shows that the instances of plurality that transcend this simple opposition constitute much more than a list of exceptions, but form instead a complex network of phenomena of morphological and semantic nature. This chapter has surveyed and catalogued various facets of this empirical complex. It will be the purpose of the next two chapters to describe how these facets belong together as morphological and semantic expressions of lexicality.

3

Plurals and morphological lexicality

3.1 Introduction

Having reviewed the non-inflectional properties that plural nouns may have, we will now focus on what makes them lexical. The descriptive overview contained in this and in the next chapter will explicate in what ways a plural noun can involve knowledge of a word rather than knowledge of grammar.

A plural like *teeth* is lexical in the sense that the exponent of pluralization (here, stem revowelling) is conditional on the choice of a word instead of another. In a different sense, it is also lexical because the exponent of plurality is an alternant of the lexical stem (*teeth* versus *tooth*) rather than a separate morpheme attached to it. A suppletive plural would lay an even stronger claim to lexicality, since to know it means to know a totally independent, not just distinct, lexical stem. From a different perspective, we could justifiably call lexical those plurals embedded inside a complex stem, in such a way that the exponent of plurality bears no relation with the syntactic context. This is the case when a plural noun appears inside a derivational affix, as in the Breton verb *evn-et-a* 'to hunt for birds' (from the plural *evn-et-* 'birds'), or inside a compound like the Dutch *dak-en-zee* 'roof-s-sea' (Booij 1996: 6), from the plural *dak-en* 'roofs'. With the exception of Booij's (1994, 1996) pioneering work on what he called 'inherent inflection', the morphological literature has not generally attempted a systematic overview of these phenomena. A comprehensive categorization of lexicalized plural morphology is therefore a desirable goal in itself. But even more important is the relation with semantic lexicalization. There seems little doubt that *look* as 'act of looking' and *looks* as 'physical appearance' are distinct semantic listemes, not related by the general scheme 'one x ~ more than one xs' exemplified by *book* ~ *books*. Other cases are more elusive, like *rain* ~ *rains* or *brain* ~ *brains*, but here too the semantic contribution of pluralization is somehow irregular and word-dependent. In a crosslinguistic perspective, these cases turn out to be quite common, and very often related to morphological lexicalization. This chapter and the next will chart the whole empirical domain of lexicalized plurality, distinguishing

morphological from semantic lexicalization and bringing into focus the relation between the two. This amounts to a typology: a reasoned catalogue of the phenomena in which plurality is not reducible to a grammatical specification of a lexical base, but is part of the base itself, whether the latter is a morphological object (a stem, a morpheme, a word form minus contextual inflection) or a semantic one (a listed interpretation, a particular 'sense' of a word, a conceptualization of its reference domain). We will start in this chapter by examining the morphological side.

To begin with, Section 3.2 clarifies that number is in a sense always lexical, in so far as it encodes information about part structure that is part of every noun *qua* noun. Inflectional number is the grammaticalized expression of this information. Plural as a value of this grammatical feature is lexicalized when it becomes an integral part of a base, in form and/or meaning. Since 'base' can mean many things, lexical plurality can take many shapes, which are described in the remaining sections. Section 3.3 deals with plurals whose morphological expression makes them akin to autonomous words: free or relatively autonomous plural morphemes and suppletive stem forms. Section 3.4 examines plural morphology inside stems, understood as lexical bases minus the contextually determined inflectional material: plural affixes inside derivation, doubly pluralized nouns, and plurals inside compounds. Finally, Section 3.5 focuses on plurality as an inherent property not of single nouns but of classes, as manifested in minor numbers, noun classes and stem-inherent number.

3.2 Is number lexical on all nouns?

There is a sense in which number is inherent in all nouns, in so far as they are agreement controllers and not targets (in the terminology of Corbett 1991), determining the number of their environment without being determined by it. Unlike gender, however, the number of nouns does not specify the feature value (singular or plural) which determines agreement. If the only evidence for number being non-inflectional had to do with this value being fixed, one could identify lexicalized number with *singularia* and *pluralia tantum*. The facts reviewed in Chapter 2, however, should have made it clear that nominal number can be lexicalized in ways that have nothing to do with fixed values. How to reconcile, then, the fact that number is an inherent (not context-determined) characteristic of every number-inflected noun, with the intuition that some plurals are more lexical than others?

To answer this important question, this section distinguishes number as a grammatical category, present not only on all nouns but also on pronouns, from the information on part structure which is contained in the lexical

representation of nouns as lexemes. Number as a grammatical category is added to a lexeme; when lexicalized, it is part of a lexeme. The discussion takes as its point of departure a critical evaluation of Beard's (1995) view that number is a derivational category.

3.2.1 *Beard (1995) on number as a derivational category*

Beard (1982, 1995: 112) draws an important distinction between number as a morphosyntactic category and number as an abstract category that enters into word formation. The first is what is shared by agreeing expressions, and is obviously an inflectional concept. However, number exists independently of agreement, if only because it defines a noun's denotation in isolation, regardless of the syntactic context. Beard (1995: 112–14) points out that this property ties in with four other observations suggesting lexical rather than inflectional status. First, inflectional categories have general application: if a noun inflects for case or a verb for tense, they all have the case or tense forms. However, *singularia* and *pluralia tantum* show that nouns often do not have the full range of the number values made available by the grammar. Since lexical items appear to make choices as to the number values available to them, number seems a lexical category. Second, number morphology typically survives in languages that have otherwise lost much of their inflectional morphology (like case distinctions, for example), with a regularity that suggests that the two types of categories do not belong together. Third, number affixes can be borrowed as easily as derivational ones, while genuinely inflectional affixes like single case or tense endings typically resist borrowing. Fourth, when there is a difference between gender and case in agreement structures, number tends to pattern with gender, which is unquestionably a property of specific lexical items.

Overall, the case is strong for regarding nominal number as an intrinsic property of nouns, not reducible to the purely contextual category defined by agreement, and therefore different from a contextual inflectional category like case. Beard goes further, however, and squarely places number among the lexeme-forming categories (1995: 115–21, and especially 159–63), proposing that the lexicon generates nouns complete with features not just for number, but also for values like singular or plural. The crucial question is how to account for the fact that nouns can in principle appear in all available number values, subject to semantic restrictions (**funs*) and lexical exceptions, while lexical properties are generally fixed; for example, speakers cannot change the inflectional class of a verb at will.

Beard addresses this question without compromising on the stance that number and specifically number values are always inherent to nouns. For him,

nouns exit the lexicon with a specified choice for the number features [± sin-gular, ± plural]. In a language like English, *pen* is [+singular, –plural] as a lexical property, while *pens* is [–singular, +plural]: the two are two distinct lexical formations. The reason why the overwhelming majority of nouns (except *pluralia* and *singularia tantum*) occur in all available values is simply that a L(exical)-derivation rule can freely apply to nouns, switching the values of the number features. So, all nouns have inherent number, but most can change it. The obvious objection is that the marked nature of nouns with fixed number should be derived, not restated. Beard (1995: 179) defines mass nouns (incor-rectly equated with *singularia tantum*; cf. instead *oats*) as not possessing any values for the features [± singular, ± plural]; collectives like *committee* are [+singular, +plural], and *pluralia tantum* are [–singular, +plural]. Since mass nouns are already grouped in a class of their own, and *pluralia tantum* lie outside of this class, the latter are wrongly predicted to be count and not mass, contrary to fact (*three arrears, **the fumes are numerous*). All remaining nouns fall into the class of inherent singulars, [+singular, –plural], and all can change to [–singular, +plural] (an unexpected generality for what is supposed to be a lexical word-formation rule), while the opposite switch just happens to be impossible. So, inherent plurals are doubly exceptional: they are the only nouns that start as plural rather than singular, and they are the only nouns that cannot change their feature values. Besides, they are entirely detached from the class to which mass nouns belong, and two unrelated reasons must be given why neither class can change feature values. All these problems, in my opinion, are due to a failure to distinguish morphosyntactic from semantic plurality: mass nouns are claimed to have no set values for the two number features, so that the singular appearance of most of them is neither accounted for nor even stated at any level, but morphosyntactic appearance is all that counts for *pluralia tantum*, regardless of their interpretation. Moreover, the feature-toggling rule that turns singulars into plurals is a word-formation rule that applies in the lexicon, independently of syntax. This commits Beard to the view that nominal number is not syntactically represented, an unnecessarily strong stance at odds with much work in the syntax of noun phrases (cf. Ritter 1991, 1993; Carstens 1993; Bernstein 1993).

3.2.2 *Inherent number as part structure*

Plainly, the idiosyncrasy of *pluralia tantum* requires more than a convenient feature notation. All the evidence we have seen so far, and all that which will come, points to the mutual relatedness of *pluralia tantum*, mass plurals, non-compositional plural readings, and idiosyncratic exponence. None of these would be related if the only way to make number 'more lexical', as it were,

consisted in blocking an otherwise free feature-toggling rule. We can preserve Beard's important point that nominal number is not contextually determined, without having to claim that *pen* and *pens* are related like *pen* and *penmanship*. I will spell out my position by means of the following four propositions:

(3.1)   NUMBER AS GRAMMATICALIZED PART-STRUCTURE INFORMATION

   (i)   the interpretation of nouns and verbs includes a determination of the part structure of their denotation;

  (ii)   many languages have a morphosyntactic number category, realized through inflectional morphology;

 (iii)   inflectional number encodes information on the part structure of the denotation of its base; and

 (iv)   on certain nouns, the information expressed by inflectional number transcends its grammatically fixed content, and modifies the meaning of the base.

The first point in (3.1) states that semantic number is but a special case of those interpretive specifications qualifying a referent as a kind, an abstraction over instances, a mass, a manifold, a set, a whole, a bounded object, or whatever part-structure conceptualization accompanies a noun in any of its realizations. The mention of verbs refers to pluractional verbs, especially common in North American languages (Mithun 1988, 1999: 83–7), but it also connects the part-structure information of nouns to verbal *Aktionsart* and Aspect (an important insight extensively articulated in Rijkhoff 1991, 2002). This is the sense in which number is inherent in every noun.

The second point spells out the notion of inflectional number as a grammatical category. This is the category that underlies agreement. Notice that I am not saying that this category is semantically inert (which it is not). The point is rather to distinguish number as a semantic determinant of lexical nouns from the grammatical notion relevant for syntax. The latter obviously has semantic import, which is especially prominent in pronominal systems.

The third point states that the semantic oppositions defined by grammatical number constitute a regimentation of the part-structure information mentioned in (i) for lexical categories. This encompasses the basic one–many contrast and can be formalized in terms of features like [± augmented] or [± group] (Noyer 1997; Harley and Ritter 2002).

The fourth point makes the connection with lexical plurality. Noun forms inflected for the grammatical category of number may express information that is additional or alternative to the grammatically regimented semantic values of singular and plural (and dual). This information concerns the structure of the noun's reference domain, and revolves around some core

concepts which will be discussed in Chapter 4 (unity, identity, functional cohesion, perceptual boundedness, instantiation). Every noun, as a noun, contains some information about the structure of its reference domain, which either determines or at least makes available a certain conceptualization of what the noun is true of (cf. Wierzbicka 1988, and the 'conceptual properties' of Borer 2005: 106). A featurally regimented set of part-structural oppositions (typically 'one' ∼ 'more than one') constitutes the meaning of inflectional number as a grammatical category, also operative on the grammatical category of pronouns. Whatever part-structure information on a noun is not reducible to the meaning of inflectional number is lexical information, and constitutes knowledge of a word and not of grammar. Sometimes this word-related part-structure information has no special morphology, as in formally regular mass plurals like *fumes* or *arrears*. Some other times, the lexicalized interpretation of a noun has a noun-dependent morphological realization, often irregular. But, on the level of form, the morphology of number does not have to be necessarily exceptional or idiosyncratic to count as noun-dependent rather than grammatically determined. In particular, every morphological expression of number, no matter how regular, is lexical if it is embedded into a noun without affecting its syntactic context. This covers plurals inside compounds or inside derivational affixes, which Booij (1994, 1996) recognized as 'inherent inflection'.

3.2.3 *The morphology–semantics connection*

The two broad senses in which number can be lexical, as part-structure information that transcends the meaning of inflectional morphemes, and as morphology independent of the syntactic context, might appear to be unrelated. However, their mutual connection is first of all an empirical fact. If the lexicalization of plurality as a property of linguistic form had nothing to do with meaning, semantic lexicalization should spread evenly between morphologically regular and irregular plurals. This is plainly not the case, and in language after language the same semantic categories underlie formally irregular plural nouns, as we will see in some detail in the next chapter and in Part II. The same semantic categories also underlie the choice among alternative plural forms for a singular noun. Besides, part-structure conceptualization is precisely what optional plural markers affect in languages like Korean (cf. Section 3.3.1). In sum, semantic and morphological lexicalization of plurality overlap significantly. However, they do not coincide. For example, *brains* is morphologically regular, but its mass reading as 'brain matter', as opposed to the count reading 'brain organs', amounts to the same conceptualization as that of the Italian

irregular *cervella* 'brain matter' (see Chapter 5). Ignoring the parallel between *brains* and *cervella* just because one is morphologically regular and the other is irregular would miss an important connection.

The approach I am proposing captures this connection by viewing semantic and morphological lexicalization of plurality as closely related but distinct reflexes of one core phenomenon, which is the use of a grammatical category (number) to encode part-structural properties of a specific lexical item. This requires a notion of lexical item as an abstract entity with a specified syntactic category, underlying the range of its possible inflectional forms, which is the classic definition of a lexeme (Matthews 1972, 1974; Aronoff 1994: 9–11). A certain part-structure conceptualization is part of the semantic content of nouns as lexemes (see Acquaviva 2004 for the obvious relation of this with the 'nominal aspect' proposed by Rijkhoff 1991, 2002 and Meisterfeld 1998). Languages with inflectional number, as explained in Section 3.2.2, morphologize a canonical semantic opposition as the content of a grammatical category. We speak of lexicalized number when this grammatical category enters into the definition of lexeme-specific conceptualization, and its content shades into lexical semantics. In many cases, this lexicalization has no morphological reflex, and plurality has the same exponence and syntactic function as when it is not lexicalized. In many other cases, the lexeme-specific function of plurality is reflected in an idiosyncratic morphology. Stem-internal plurals represent a particular case in which number is detached from the syntactic context through derivation or compounding (see Section 3.4). Being thus part of an enlarged lexeme, they qualify as lexicalized both morphologically and semantically.

In conclusion, a lexeme-inherent characterization in terms of part structure is what unifies morphologically diverse types of lexical plurals. This does not mean that every morphologically irregular plural is also semantically idiosyncratic, because what is lexeme-inherent can be, but does not have to be, idiosyncratic (as elucidated by Aronoff 1994: 18–22). In addition, pluralization may have a genuine grammatical function while making use of non-predictable morphological exponence, such as the choice of one stem over another (as I will argue in Chapter 7 is the case for most Arabic broken plurals). When that happens, plurality is lexical only in the sense that it manipulates lexical stems as opposed to grammatical morphemes. As can be seen, the many senses of the word 'lexical' correspond to many varieties of lexical plurality. But these varieties are all related, and revolve around a core intuition: grammar can express properties of individual words. It is true that number is in some sense intrinsic to all nouns, as Beard (1995) recognized, but only on some nouns is it a truly grammatical category which also has a truly lexeme-forming function.

3.3 Lexicality as morphosyntactic autonomy

This section considers those plural forms that resemble words, rather than grammatical formatives, for the autonomy they display. They can be autonomous syntagmatically, as free morphemes or clitic-like particles, or paradigmatically, as forms that cannot be reduced to modified versions of the singular.

3.3.1 *Plural words*

Plurality can be lexical in the rather straightforward sense of being a word in itself, rather than a property of a word. This happens in so-called plural words, which express the plurality of a noun by means of a free morpheme, as exemplified by (3.2) and (3.3) (respectively from Dryer 1989: 875 and Corbett 2000: 134),

(3.2) a. ha   fanga   pulu   (Tongan)
           indef   PL   cow
           'some cows'

       b. àwọn   ọkùnrin   méta   yǐǐ   (Yoruba)
           PL   man   three   this
           'these three men'

(3.3) a. mga   bahay   (Tagalog)
           PL   house
           'houses'

       b. mga   tubig
           PL   water
           'cups/units of water'

The precise syntactic status of plural words varies across languages, but they are often clitics occupying a fixed position in the clause or in the noun phrase (see Corbett 2000: 135 and Dryer 1989: 867). What is significant is the frequent correlation of this syntactic autonomy with clearly non-inflectional properties. Most of the languages discussed by Dryer's (1989) typology of plural words do not have obligatory number marking, and many of them are classifier languages. Inflectional number and numeral classifiers may not be as absolutely incompatible as is generally believed (see Aikhenvald 2003: 249), but it remains significant that classifier languages which definitely lack inflectional number on nouns may still have plural words, even in co-occurrence with classifiers, as in the following example:

(3.4)   nhũng   sáu   con   tôm   (Vietnamese; Dryer 1989: 877)
        PL   six   CL   shrimp
        'six shrimps'

In fact, plural words and plural marking may even appear side by side in the same structure. Corbett (2000: 135) discusses one such case in Dogon, where *nndɛ-m* 'persons' is one of the few nouns that have an overt plural marker (unlike non-human-denoting nouns such as ɛnɛ 'goat'). This plural suffix can co-occur with the optional plural word *mbe*:

(3.5)   a.   ɛnɛ   mbe   (Dogon; Corbett 2000: 135)
             goat   PL
             'goats'

        b.   nndɛ-m   mbe
             person-PL   PL
             'people'

In most cases, then, plural words are not just the free-morpheme version of inflectional plurals. The two types of plural marking differ not only in their morphosyntax, but also in their distribution across the vocabulary, in their obligatoriness, and in their compatibility with classifiers. Plural words, it would appear, encode plurality as a property of the NP, not strictly of a noun, and are therefore akin to phrasal affixes, not to word formatives. While sensitive to the choice of the head noun, they are not determinations of its stem, but are themselves stems. This marks a clear difference with inflectional plurals, which are the forms assumed by lexical stems in a syntactically plural context. In a purely morphological sense, plural words are lexical not as opposed to grammatical (they are grammatical morphemes), but to affixal. This characterization will play an important role in the analysis of Irish counting plurals in Chapter 6.

3.3.2   *Clitic-like plural affixes*

Some of the morphosyntactic autonomy displayed by plural words also appears on certain affixal plural markers, which are much more word-like than typical inflectional affixes despite being bound morphemes. In many cases, the affixes in question belong to the class of 'optional' plural morphemes mentioned in Sections 2.5 and 2.6.2, which express plurality as concomitant with other semantic characterizations (like functional or space-temporal cohesion) in languages that lack inflectional number on nouns. The Korean morpheme *-tyr* is an example, according to the description in Unterbeck (1993) and Song (1975). This language has no inflectional

number, but -*tyr* can be suffixed to nouns to express a particular plural reading. As those authors report, -*tyr* is in fact obligatory when the noun phrase refers to a specific plurality of individuals in the context. It therefore forces a specific (plural) reading in (3.6b), disambiguating the interpretation of the noun in (3.6a):

(3.6) a. sangò-ryr   po-ass-sypnita   (Korean; Unterbeck 1993: 203–4)
           shark-ACC   see-PAST-FORMAL
           'I saw a shark/the shark/sharks/the sharks'

       b. sangò-tyr-ryr   po-ass-sypnita
           shark-PL-ACC   see-PAST-FORMAL
           'I saw the sharks'

The semantic function illustrated by this contrast clearly differs from the mere expression of plurality. Indeed, -*tyr* can co-occur with classifiers, as we have seen that plural words can do:

(3.7) tasòs   mjòng-yi   haksäng-tyr   (Korean; Unterbeck 1993: 183)
       five   CL-GEN   student-PL
       'five students'

However, like the Chinese 'collective' pluralizer -*men* (which does not co-occur with classifiers), -*tyr* also has a purely grammatical function as plural morpheme on third person personal pronouns: *ky* 'he/she' ~ *ky-tyr* 'they' (Unterbeck 1993: 190–2). The very fact of combining with pronouns beside nouns shows a certain degree of morphological autonomy, typical of clitics rather than affixes in a strict sense (see Zwicky and Pullum 1983). In fact, the autonomy of -*tyr* goes much further. As a 'floating' pluralizer for an overt or null argument (Song 1975), it can attach to a host of categories in the sentence:[1]

(3.8) a. ir   car-tyr   hajòra   (Korean; Unterbeck 1993: 209–10)
           work   well-PL   do
           'work (pl.) well!'

       b. muòs-tyr   hap-nikka
           what-PL   do-then
           'what are you (pl.) doing?'

       c. wä-tyr
           why-PL
           'why?' (addressed to several people)

---

[1] Unterbeck (1993: 220) adds that younger speakers accept sentences with *tyr* in absolute initial position, which makes of this particle a free morpheme like plural words.

In this function, *-tyr* exemplifies a truly grammatical plural suffix, enjoying a degree of morphological and syntactic autonomy comparable to that of plural words.

Similar considerations apply to plural markers in agglutinating languages where number is a grammatical category integrated in the morphosyntax of nouns, unlike in Korean. The Turkish plural marker *-lAr*, for instance (a representation which subsumes the two possible harmonic alternants *-ler* and *-lar*), attaches to nouns, to third person pronouns, and also to verbs. In addition, it can attach to proper names to express comitative readings, as in *Mehmet-ler* 'Mehmet and related people' (cf. Lewis 1967: 26, Göksel and Kerslake 2005: 169). Besides, *-lAr* also behaves like a clitic like rather than an affix in attaching only once to a conjoined base, as *-ler* distributes over both *kitap* and *sozluk* in the following example (courtesy of Leyla Zidani-Eroğlu):[2]

(3.9) kitap ve sozluk-ler var-di (Turkish)
      book and dictionary-PL arrive-PAST
      'the books and the dictionaries arrived'

For the purposes of a morphological typology of lexical plurals, what these examples show is that there is a middle ground between the realization of plurality as a syntactically autonomous word and as a phonological modification of a stem (affixal or otherwise).

3.3.3 *Suppletion*

Plural words are lexical because they are autonomous words syntagmatically, as they co-occur with a head noun in the same structure. Suppletive plurals are instead autonomous words paradigmatically. The Russian pairs *čelovek ~ ljudi* 'person ~ people' and *rebjonok ~ djeti* 'child ~ children' are typical examples (the latter somewhat blurred by the existence of a collective plural *rebjata* 'children', and an emotive singular *ditja*). Note that the English *people* does not qualify as the suppletive plural of *person*, because of the presence of *persons* and *peoples*.

---

[2] Turkish is an interesting case where the plural affix combines absolute morphosyntactic regularity with the semantic latitude of an autonomous plural word; see Ketrez (2004) for a detailed study. Notice that the comitative reading 'x and x's people' apparently cannot distribute over two conjuncts, as opposed to the non comitative reading in (3.9): *Orhan le Mehmet ler* can mean 'Orhan and the people related to Mehmet', that is [O + [M.PL]], or possibly 'the people related to both Orhan and Mehmet' (that is, [[O+M]PL]), but comitative plurality cannot distribute over both conjuncts to give 'the people related to Orhan and the people related to Mehmet' as two separate collections ([O]PL + [M]PL). I am grateful to Leyla Zidani Eroğlu for discussion on this point, which deserves closer investigation. On comitative readings in languages where plural is not inflectional, see note 22 below.

On the face of it, suppletion is an extreme form of allomorphy, whereby a cell in the inflectional paradigm of a lexical item involves not just a modified form of the base stem, but a different stem altogether (see Corbett 2000: 155, 289; Maiden 2004). There would appear to be little to add, apart perhaps from noting that the difference between two stems is really a matter of degree and that, therefore, suppletive forms shade into irregular forms. However, suppletion typically has a connection with lexical semantics. It is not just that, like other irregular forms, suppletive plurals tend to have a high frequency relative to the singular (see Corbett 2000: 284–94). They also tend to be restricted to referents high in the animacy hierarchy, typically humans and culturally salient animals. If it was only a matter of greater frequency of the plural over the singular, one would expect to find many instances of suppletive plurals for a concept like 'eyes', certainly more than for a concept like 'man'; yet this does not seem to be the case. Schuh (1989: 174) cites the following pairs from Miya:

(3.10) Singular Plural Gloss   (Miya)
        sə̀m   sə̀bə̀   'person'
        ˀám   tə̀vàm   'woman'
        yàsə̀   dàvàn   'brother ∼ siblings'
        ˀáfúw   cùw   'goat'

Although in Miya plural marking can be suspended on inanimates, all nouns can be morphologically pluralized (Schuh 1989: 173), so the number opposition itself is not restricted to a segment of the animacy hierarchy; it is just suppletive plurals that are so restricted, namely to animates. The same applies to Breton. Trépos (1956: 72–5) lists suppletive plurals for 'man', 'horse', 'cow', 'dog', 'pig', and 'son', adding some revealing comments about the meaning of some of these forms. *Tud* means 'people' (males or females), which is not the compositional plural of *den* 'man'. The plural of *marc'h* 'horse' is *kezeg*, morphologically a stem-revowelled plural of *kazeg* 'mare' (distinct forms express 'mares' in the dialects). And *bioc'h* 'cow' has in many dialects the suppletive plural *saout*, historically derived from the late Latin *soldus* 'money' (singular).

Clearly, no firm conclusions should be based on a few examples. However, the fact that suppletive plurals cross-linguistically show a definite tendency to affect certain concepts and not others, and the fact that their semantic relation with the respective singulars can be less than transparent, suggest that suppletive plural nouns have a lexicalized status, which, for instance, is lacking in unpredictable verbal forms like *is* or *am*. A human person, arguably the most individuated referent, loses some of the most salient traits, including sex and age, when conceptualized as a member of a group or a mass. A plurality of

domestic animals is not just a non-singleton set, it is an economic asset when considered as a mass. That is why suppletive plurals tend to cluster around the highest end of the animacy hierarchy: they name concepts salient enough in the speakers' experience to justify distinct (if obviously related) conceptualizations for singular and plural. Relative frequency goes hand in hand with the salience of the plural as a distinct concept, but it is its effect, not its cause.

In a language with only a few suppletive plural nouns, it might not seem necessary to distinguish between having an unpredictable form and being a separate word. But matters stand differently if most nouns in a language show no automatically predictable correlation between the shapes of singular and plural, as is the case in Nilotic languages (a family of Nilo-Saharian); see Dimmendaal (1983: 223–58) for Turkana and Corbett (2000: 155–9) for Shilluk. The generalized irregularity of such systems (which may be very complex rather than unpredictable; cf. Dimmendaal 1983: 223) clearly cannot imply that every form corresponds to a distinct semantic listeme. Being a distinct lexical item cannot be equated with having an idiosyncratic form.

To avoid confusion, we may retain the term 'suppletion' for the purely formal notion of one extreme of the irregularity continuum, but this should not obscure the fact that many suppletive plural nouns are semantic listemes partly distinct from the singular. Consider in this connection the conjecture, attributed to Marantz by Harley and Noyer (2000: 370), that truly suppletive pairs like *bad* ∼ *worse* or *go* ∼ *went* 'reflect the spell-out of universal syntactico-semantic primitives'; that is, that suppletion only involves grammatical formatives expressible as feature bundles. While I find it implausible that suppletive plurals could be reduced to a list of culture-independent basic concepts (and while 'cows' is certainly not a grammatical morpheme), Harley and Noyer's hypothesis proves sustainable, if very strong, when reinterpreted as follows: only grammatical elements can have totally unrelated base forms in the same paradigm. Word forms that seem to violate this generalization either realize in fact grammatical elements (as hypothesized for the verb GO), or they are still synchronically related (albeit idiosyncratically) by phonological readjustment rules, or do not really belong to the same paradigm, that is, they are not forms of one and the same lexeme. The last possibility amounts to saying that totally suppletive alternants like Breton *den* 'man' ∼ *tud* 'people' are in effect two distinct bases for morphological purposes, not two forms of the same paradigm. According to this hypothesis, then, suppletive plurals of lexical nouns must be lexical plurals. I will not try to verify this claim here, but it must be noted that the frequent presence of alternative regular plurals, not blocked by the suppletive ones, is immediately accounted for if the suppletive and the regular alternant are not competing realizations of one cell in the paradigm of the same lexical base.

3.4 Plural nouns within the base for inflection

The plurals we have considered have the properties of autonomous words. In other cases, a plural noun is lexical because plurality is an integral part of the base for inflection, that is, of that part of the word to which grammatical determinations are added according to the syntactic context. This covers not only markers inside the stem of a lexeme, but also inflected nouns embedded inside a compound, as in *mice catcher*. We will now consider them in turn.

3.4.1 *Plural nouns as part of a lexical stem*

Plurality is inside the base for inflection when it occurs inside derivational affixes, contradicting the generalization formulated by Greenberg's (1963) Universal 28:

(3.11)   If both the derivation and inflection follow the root, of they both precede the root, the derivation is always between the root and inflection.   (Greenberg 1963: 93)

Booij (1994, 1996) has considered this phenomenon among the instances of what he called inherent inflection, where 'inherent' is opposed to 'contextual'. In (3.12), for example, the suffixes -*dom* and -*achtig* derive a noun and an adjective from *helden* and *boeken*, morphologically identical with the plural of the corresponding nouns *held* and *boek*:

(3.12)   [held-en]-dom 'heroism'   [boek-en]-achtig 'like books'
                                                 (Dutch; Booij 1996: 6)

But are *helden* and *boeken* effectively plural? Booij (1996) provides some evidence that the plural-looking bases of Dutch examples also have a plural interpretation, although his most convincing examples, such as *stadsraad* ∼ *stedenraad* 'city council' ∼ 'cities council', involve compounds and not derived stems. Since the use of affixes as stem extensions and linking elements is common in Germanic languages, the status of examples like (3.12) is not completely clear. For Dutch, van Marle (1996: 76–8) proposed that only one of two available plural suffixes appears inside compounds, because only that affix can be reinterpreted as a (bleached) linking element. Chapman (1996: 175) argued that only some plural markers can appear within derivational affixes, namely those least semantically transparent and least paradigmatically regular (in practice, those most removed from a biunique sound–meaning map). In the light of this plausible suggestion, the Georgian adjectival suffix -*ian*- 'provided

with' is all the more remarkable, because it can attach to stems pluralized by
-*eb*-, which is the plural affix for all nouns:[3]

(3.13) a. k'omp'iut'er-eb-ian-i   otax-i   (Georgian)
            computer-PL-ADJ-NOM   room-NOM
            'room endowed with computers'

        b. c'ign-eb-ian-i   otax-i
            book-PL-ADJ-NOM   room-NOM
            'room endowed with books'

As Alice Harris points out (pers. comm.), the presence of the marker -*eb*- is
particularly significant. Georgian has other cases of plural inside derivation,
but these involve the Old Georgian oblique plural marker -*ta* (-*t*); for in-
stance, *nav-t-sadgur*, 'harbour', literally 'boat.PL.station'. The suffix -*eb*- of -*eb-
iani*, by contrast, is unambiguously the inflectional plural marker of modern
Georgian. We are dealing, apparently, with a morphologically and semantic-
ally transparent use of a plural marker inside a derived word, contrary to
Chapman's predictions. However, these brief examples warrant no immediate
theoretical conclusion, because of the typological difference between Geor-
gian, where nominal inflection is agglutinative, and fusional-inflectional
languages like Germanic.[4] More theoretically significant and better known
are the examples from Yiddish and Breton, as discussed by Bochner (1984),
Anderson (1986), Perlmutter (1988), and Stump (1989, 1990). We will now
briefly consider these languages in turn, and add other examples of base-
internal plural that have been less widely discussed.

3.4.1.1 *Yiddish*   In Yiddish, base-internal plurals involve diminutive morphemes
suffixed to an already plural form.[5] There are two relevant diminutive suffixes,

---

[3] I would like to thank Pëtr Arkadiev and Alice Harris for suggesting these examples. Note that
(3.13b) cannot be interpreted as something like 'a roomful of books': Léa Nash, whom I also wish to
thank, confirmed that *Merabma c'aik'itxa c'ignebiani otaxi* 'Merab read a room with books' is un
acceptable.

[4] Consider for instance the syntactic phrase 'more than two' embedded inside the compound noun
orzemet'marculiani 'polysyllabic'; this example is again due to the courtesy of Alice Harris:

(i) or ze met' marcul ian i   (Georgian)
     two than more syllable with NOM
     'polysyllabic'

Ze, literally 'on', is the preposition selected by the comparative 'more'. The whole complex is a single
noun, formed by the derivational affix *ian* 'provided with' and marked by the case ending *i*.

[5] Yiddish is not isolated in this respect; cf. Bobaljik (2005) on Itelmen. Within Germanic, Chapman (1996:
177) has cited parallels from East German dialects, for example *stil er chen* 'chair PL DIM' 'small chairs'.
Another case comes from Portuguese (Rainer 1996: 88), where the diminutive plural *cãezinhos* 'small dogs'
expresses plurality twice, on the noun stem and on the inflectional ending of the diminutive affix. However,

-*l* and -*ele*, which in the plural become -*lex* and -*elex* (only the plural forms will be indicated here). These attach to nominal stems marked as plural by root revowelling (*cep*), by suffix (-*er* or the loan suffix -*im*), or both (*dern-er*):

(3.14) Gloss   Singular   Plural   Dimin. plural (Yiddish; Perlmutter 1988: 80)
        'braid' cop   cep   cep-lex, cep-elex
        'child' kind   kind-er kind-er-lex
        'thorn' dorn   dern-er dern-er-lex
        'body'   guf   guf-im   guf-im-lex

Whatever the precise status of evaluative morphology, diminutives certainly do not take part in syntactic agreement, unlike plural. So, in examples like (3.14) an inflectional marker (plural) occurs inside a word-formation, non-inflectional marker (diminutive). This seems to be an unexpected reversal in the order of application of inflection and derivation, assuming that derivation creates words and inflection then adapts them to a syntactic context (Anderson 1982, 1992).

   No theoretical problems arise if the bases of diminutive suffixation in (3.14) are not inflectional plurals. This can be true in two cases: if they are not plural, or if they are plural but not inflectional. Perlmutter (1988) justified both of these conclusions. First, he showed that the form the diminutive attaches to is the ablaut stem of a noun, whether or not it happens to express plurality. For instance, some nouns have ablaut with diminutives but not with plurals: *oyg* 'eye' and *oyg-n* 'eyes' have a different stem from the diminutive *eyg-ele* 'eye.PL-DIM.PL' (Perlmutter 1988: 88). So, the ablaut stem is only a stored form and does not by itself express plurality. Since ablaut stems, not being phonologically predictable, must be 'listed in the lexicon', they are available inputs for word-formation rules. The other cases in (3.14) are plural but not inflectional. The plural forms in -*er* and the loanwords with plural in -*im* must all be listed, and therefore constitute legitimate inputs for derivational affixation 'in the lexicon' (Perlmutter 1988: 89).

3.4.1.2 *Breton*   Anderson (1986) proposed a similar explanation for parallel Breton data. As in Yiddish, diminutive suffixes (here -*ig*-) can attach to already plural nouns, and may be then pluralized in their turn by the plural suffix -*où*:

Rainer also shows that diminutive affixes likewise repeat the plural marker when they attach to agreeing adjectives:

(i)   uns   exercícios   facei zinhos/*facil zinhos   (Portuguese; Rainer 1996)
      some.PL exercises.PL   easy.PL DIM.PL/easy.SG DIM.PL

Since the plurality of adjectives is not lexical but due to syntactic agreement, its double plural marking must be a matter of inflectional realization.

(3.15) GLOSS SINGULAR PLURAL   DIMIN. PLURAL (Breton; Stump 1990: 105)
     ‘bird’ labous   labous-ed labous-ed-ig-où
     ‘cat’ kazh   kizh-ier kizh-ier-ig-où

Besides, Breton also features plurals inside a number of derivational formations (from Stump 1990: 107, 109, 113):

(3.16) GLOSS SINGULAR PLURAL DERIVED VERB
     ‘bird’ evn   evn-ed evn-et-a ‘to hunt for birds’

     GLOSS SINGULAR PLURAL DERIVED ADJECTIVE
     ‘rock’ maen   mein mein-ek ‘rocky’

     GLOSS SINGULAR PLURAL DERIVED AGENTIVE NOUN
     ‘girl’ merc’h   merc’hed merc’het-aer ‘womanizer’

The most striking phenomenon is the possibility of suffixing a plural marker to an already pluralized noun. We saw two examples in Section 2.8.3, where the bases were collectives; the same can happen when both markers are affixal (Stump 1990: 114):

(3.17) GLOSS SINGULAR PLURAL   DOUBLE PLURAL
     ‘child’ bugel   bugal-e   bugal-e-où
     ‘girl’ merc’h   merc’h-ed merc’hed-où

The double plural is not glossed in (3.17), as it is not in Stump (1990), because the interpretation of these forms is far from being predictable; and, without going into details before Chapter 8, I must point out that the pluralization of a plural is not freely available.

    The question is how much of an inflectional marker the inner plural really is. Anderson (1986) pointed out that many semantically plural nouns in Breton have monomorphemic forms with no markers of plurality. On this basis, he argued that they are morphologically non-plural and therefore tolerate further affixation and even pluralization. However, Stump (1989, 1990) was able to prove conclusively that (1) such monomorpehmic ‘collectives’ behave like any other plural in morphological and syntactic respects; and (2) very often the same plural markers that can appear inside derivation also express inflection, so that even the most regular affixes can find themselves as part of the base. The Breton facts represent a genuine conundrum, for they show unambiguously inflectional affixation put to unambiguously derivational use. This formulation foreshadows the answer I will propose in Chapter 8, namely that the morphological means are the same but the functions are different.

3.4.1.3 *Double plural in Somali and in other languages*   The Breton facts find an interesting parallel in Somali, as noted by Lecarme (2002). The systematic availability of more than one plural form for a given noun was discussed in Section 2.7.3 as a non-inflectional trait shared by the two languages. But the similarity goes further. As in Breton, Somali plurals can be input to further pluralization:

(3.18) Gloss   Singular   Plural   Double plural (Somali; Lecarme 2002: 121 2)
        'man'   nín   nim-án   nim-an-yaál
                                     nim-an-yów
        'girl'   gabádh   gabdh-ó gabdh-a-yów
        'tooth' ílig   ilk-ó   ilk-a-yáal
        'rain'   róob   roob-áb roob-ab-yów

Lecarme's glosses indirectly reinforce the parallel between the two languages. The double plurals for humans (including 'women', not reproduced here) are glossed as '(groups of) N', where N is the plural; so, *gabdh-a-yów* may mean 'groups or girls' or 'girls', with an interesting indeterminacy. Not so for the inanimate examples, which have the same gloss 'rains' and 'teeth' in the simple and in the double plural. Without reading too much into the glosses of a handful of examples, it is clear that the double plural does not have the function of turning a plurality of N into a plurality of groups of N. Such groups readings do arise, but not as the regular, automatic consequence of double pluralization.

Besides, Somali plurals appear in nominal derivations, with a striking resemblance to the Breton examples in (3.16):

(3.19) buug-a-g-shéeg 'bibliography' (buug-ág 'books') (Somali; Lecarme 2002: 123)
        geed-o-aqóon 'botany' (geed-ó 'plants')
        xagl-o-gooy-é 'diagonal' (xagal-ó 'angles')

The parallel with Breton has some theoretical interest. The derivational character of plural morphology in Somali is well known (Lecarme 2002: 119); Somali clearly resembles Breton; but Stump (1989, 1990) marshalled strong evidence *against* the derivational status of the Breton plural markers in question. In Chapter 8, I will argue that these markers are not derivational by themselves, but because of their function. Whether the same applies to Somali, I cannot say.

Leaving the theoretical question aside for the moment, let us turn to some more instances of plural inside plural. We are not concerned here with cases where a plural marker attaches to a stem that was an inflectional plural at an

earlier stage (like *child-r-en*) or, in cases of borrowings, in the source language (like Russian *dzhins-y* 'jeans.ᴘʟ'). Genuinely synchronic double pluralization appears to be rare. Corbett (2000: 37) discusses two examples from Khamtanga (Cushitic) and Warekena (Arawakan). The latter is particularly interesting because it allows the plural marker *-pe* to attach to a noun already suffixed by *-nawi*, which marks not just plural, but emphatic plurality: Corbett's example *abida-pe-nawi* 'very many pigs' (from Aikhenvald 1998) is morphologically pluralized twice, but semantically three times—or better, the compound plural emphasizes an already emphatic indication (*-nawi*) of numerical greatness. Corbett makes it clear that *-pe* is a genuine plural and not a paucal, because it would have been less surprising to interpret a structure *N-pe-nawi* as meaning schematically [[a few N] many]. Mithun (1999: 84), for example, reports just such a juxtaposition of paucal and plural in the North American language Koasati, restricted to a few kinship terms formed with the diminutive *-(o)si*:

(3.20) Gloss   Singular   Paucal   Paucal+plural   (Koasati)
        'nephew'   icofó:si   icofós-ki   ikofós-ki-ha

The paucal is used to refer to 'between two and six', and the plural morpheme is suffixed, rather than substituted, to refer to larger groups. Apparently, in a sequence STEM+PL+PL, the internal plural affix is more often than not a non-inflectional marker, or in any case a plural exponent with a specialized reading (collective, paucal). What makes Breton double plurals so interesting is that each of the two exponents can occur separately, marking just plural. In the following examples too, each marker occurs elsewhere as a simple pluralizer:

(3.21) a. noyan   noya-d   noya-d-ud
           'prince' 'princes'   'princes'
                      (classical Mongolian; Grønbech and Krueger 1976: 20)

        b. de   de-ge   de-ge-ge
           'person' 'people'   'parties of people'
                      (Sara Mbai, Nilo-Saharian; Tucker and Bryan 1966: 69)

This state of affairs, however, does not exclude the fact that the two plural morphemes could have distinct semantic functions. A correct interpretation would require a detailed study of plurality in the language in question, and one of the results of the analysis of Breton in Chapter 8 will be that distinct semantic functions can be expressed by a plural exponent with one and the same form.

This means that, even if each morpheme PL can mean simply 'many *x*', their juxtaposition gives rise to a distinct shade of plurality—a lexicalized one.[6]

Our final example is rather different. In this case, the double plural arises through the simultaneous application of two patterns of plural formation, and the resulting form selects one particular sense of the base, rather than qualifying the numerical size of the referents. The examples come from modern written Arabic (Badawi *et al.* 2004: 767):

(3.22)	GLOSS	SINGULAR	PLURAL-1	PLURAL-2	(Arabic)
	'house'	bayt	buyuut	buyuut-aat 'houses, great families'	
	'man'	rajul	rijaal	rijaal-aat 'men of distinction'	

The form here indicated as 'plural-2' arises through the simultaneous application of the stem rearrangement and -*aat* suffixation, each of which can be an exponent of regular inflectional plural. What happens, then, is that a stem alternant functioning as the plural of a given base is re-interpreted as a base in its own right, and then pluralized by a semantically transparent -*aat* suffixation. As we can see, a semantic shift accompanies the reanalysis of the base.

In sum, double pluralization is by no means a unitary phenomenon, and its significance as an indicator of lexicalized plurality varies not only across languages, but also across nouns. What is undeniable is that at least some plural markers, on at least some nouns, can appear inside the stem that serves as a base for contextual inflection. Any further conclusion must take into account the whole system of plural marking and the whole nominal morphology in the languages in question. Chapters 7 and 8 provide such case studies for Arabic and Breton.

3.4.2 *Plural nouns inside compounds*

In English, non-head nouns in compounds are normally singular, that is, morphologically unmarked: *rat catcher* is well formed, but **rats catcher* is not. Even obligatorily plural nouns respect this constraint: *trouser leg* contrasts with **trousers leg*. The plural seems to be much more acceptable, however, if it

---

[6] To avoid misunderstandings, I stress that the examples in (3.21) are not meant to illustrate regular and semantically transparent double pluralization. The precise value of these forms requires in depth studies of the respective languages. The gloss 'parties of people', for instance, reproduced from Tucker and Bryant (1966: 69), does not prove in any way that Sara Mbai can routinely apply a sort of second order pluralization ('one ∼ many ∼ groups [of many]'); first, because it is the only example of a doubly pluralized noun in that source and, second, because the same page also cites the doubly pluralized first person plural pronoun *ji ge ge*; now, if 'we' means 'speaker plus others', the double plural surely cannot mean 'groups of [speaker plus others]', but has presumably an emphatic value (possibly 'speaker plus many others'). So, Sara Mbai's double pluralization is not semantically transparent on pronouns, and there is no reason to think that it is on nouns. No theoretical conclusion can be based on a single gloss.

is listed. Pinker (1999: 180–4) reports that *mice catcher*, while not as universally acceptable as *mouse catcher*, contrasts with the unacceptable **rats catcher*. In these cases, the plurality of *mice* does not count for syntax, because it is embedded within a grammatically singular complex noun. Booij (1994, 1996) highlighted the parallel between this type of plurality-inside-a-word and *pluralia tantum*: in both cases a grammatical property that normally reflects the syntactic configuration appears as an intrinsic characterization of a lexical item, independently of the context. This is plausible, especially in the light of the observation that morphologically and even syntactically complex elements can be embedded in a compound, if they are listed enough (*end-of-the-world feel*, etc.). But it would be wrong to reduce the lexicality of plurals inside compounds to their listedness, because the lack of syntactic relevance also correlates with a particular interpretation.

As long ago as 1926, Wackernagel (1926: 84–5) pointed out that compounds like the classical Greek *hippo-damos* 'horse-tamer' or the Latin *au-ceps* 'bird-catcher' refer to someone who tames horses and catches birds, not a single horse or a single bird, even though there is no mark of plurality on *hippo-* and *au-*. In Germanic languages, where the singular form of a noun is often identical to the bare stem, the same contrast appears between what looks like a singular form and a generic plural interpretation:

(3.23) flea infested   *fleas infested   (Borer 2005: 133)
        shelter burning   *shelters burning

Borer points out that a single flea does not represent an infestation, and the denotation of *shelter burning* does not include events where only one shelter gets burned. She also notes that nouns occurring in such compounds do not even have a fixed count or mass interpretation: *stone* in *stone throwing* must refer to discrete entities, but in *stone carving* it refers to a material. This is compelling evidence that, as Borer concludes, the first element in *flea infestation* is 'not singular in any grammatically meaningful way' (p. 133); it is rather a bare stem, interpretively distinct from a noun inflected for number.

By contrast, a form like *mice* is certainly grammatically plural. However, and this is the crucial fact, its interpretation approaches that of numberless bare stems like *flea* in (3.23) above. In fact, all non-head nouns in compounds have a non-specific and non-referential interpretation. This includes plurals, which do not support pronominal anaphora nor allow for a distributive interpretation (**a real mice_i catcher knows them_i one by one*). The role of the non-head noun in compounds like *head hunter* or *mice catcher* seem very close to that of the object in constructions like *he plays piano*, where the object NP does not introduce a

discourse referent.[7] It is apparently the listedness of the form which allows some plurals to occur as constituent parts of a compound word; but when this happens, plurality is lexicalized in a semantic as well as in a morphological sense.

Languages other than English provide more illustrations of inherent plurals inside compounds. Most examples cited by Booij (1996) come from Dutch. The semantic plurality of nouns suffixed with *-en* can be dubious in some cases, but it seems clear enough in *dak-en-zee* 'sea of roofs', *huiz-en-rij* 'row of houses', or *docent-en-kamer* 'teachers' room' (Booij 1996: 6), and above all in the pair *stadsraad* ∼ *stedenraad* 'city council' ∼ 'cities council', cited above in Section 3.4.1. These forms are morphologically and semantically plural, but they do not correspond to inflectional plurals in all respects. Van Marle (1996: 77) speaks of a 'collective' rather than plural sense of the base. He points out that in the Dutch pair *priester-dom* 'clergy' (priest.SG-ABSTR) vs. *lek-en-dom* 'laity' (lay-PL-ABSTR), the suffixed *lek-en* must have the same abstract–collective nuance of the formally singular *priester*.

Two German examples show more clearly the lexical, non-compositional effects of number inside compounds. *Gasthaus* 'inn' corresponds to the English *guesthouse*; but German also has *Gästehaus*, built with the plural form of *Gast*, and this refers to a residence with apartments for non-permanent occupiers. Evidently, the interpretive difference between the two compounds cannot derive from the one–many contrast of inflectional singular and plural. The same goes for *Gastzimmer* 'room for private functions' ∼ *Gästezimmer* 'guest room'. Romance has many such compounds; I will only cite two Italian cases. One is *calzascarpe* 'shoehorn', composed of the stem of the verb *calzare* 'to fit (of shoe)' and the plural *scarpe* 'shoes'. This is an exocentric compound, which does not inherit features from *scarpe* even though this noun provides the word ending. Note that a shoehorn cannot be applied to more than one shoe at any one time. The doublet *asciugamano* ∼ *asciugamani* 'towel' has the same structure V–N, where the verbal stem means 'to dry' and the noun 'hand'—or 'hands'. In fact, usage varies between the singular *mano* and the plural *mani* in this compound, but obviously the overall meaning remains the same. Grammatical number evidently has no semantic import here, but it is required because Italian nouns of these classes have no numberless stem form comparable to the English 'singular'. Once more, the use of plural nouns inside a lexical base must be seen in the context of the morphological system of the language.

---

[7] See Krifka *et al.* (1995: 88) for an analysis of this type of non referential construction in terms of noun incorporation, and Farkas and de Swart (2003) for a detailed semantic analysis of incorporation. Ketrez (2004) analyses the interesting case of Turkish, where pluralized nouns can systematically appear in compound like syntactic constructions with the verb with a distinctive non referential interpretation, pluralizing the events or the types of referents but not the referents themselves.

3.5 Plurals as inherent class feature

We have considered cases where plurality is a property of specific nouns, and where it is the morphological structure that makes it part of a stem, and therefore lexical. But plurality can also be an inherent to whole classes of nouns. I will review three ways in which number can be lexeme-inherent as a class feature: so-called minor numbers, like duals, paucals or collectives; noun classes, which pair each noun stem with a choice of classes; and stem-inherent number, where noun stems are intrinsically specified for a number value, and morphology can change these lexically specified defaults.

3.5.1 *Minor numbers*

Not infrequently, languages with a fully developed grammatical number have an additional number value that is available only for some lexically restricted set of nouns. Such minor numbers, as Corbett (1996, 2000: 95–101) calls them, typically add to the canonical singular–plural opposition a value like dual or paucal, which however differs from the fully grammaticalized values in ways that reveal their true nature of lexical derivational classes. The Hebrew 'dual' -*ayim*, discussed in Section 2.6.4, illustrates well the semantically motivated lexical restriction, and the idiosyncratic but not random interpretations (dual with unit nouns, plural with body parts) that are the hallmarks of derivational number marking. Apart from the semantically motivated restrictions, the categorial restriction to nouns also suggests that the dual is not a category on a par with singular and plural (which apply to pronouns and verbs too), but rather a word-forming operation. Unsurprisingly, its use involves semantic and pragmatic factors, a far cry from the determinism of inflectional number marking. Much the same applies to the use of dual in Arabic (Brustad 2000: 45–52) and in Maltese (Fenech 1996; Plank 1996).

Beside the dual, Corbett (1996: 102–5) mentions paucal, mass, and collective as semantic values of minor numbers. I will have nothing to say here about paucals, except noting that the semantic motivation of their lexical restriction again suggests a derivational type of pluralization, rather than an additional number value; cf. the Koasati example in Section 3.4.1.3 above, where a paucal and a plural affix sit side by side. Collectives are exemplified by Maltese and Italian, the latter with the irregular plurals that will be the empirical focus of Chapter 5, which will confirm their derivational nature. Maltese collectives do not essentially differ from a class of semantically plural (and possibly kind-referring) underived nouns in Arabic, to be analysed in Chapter 8. An observation that is relevant now, however, is that Italian and Maltese so-called collectives are grammatically fully integrated in the inflectional number

system of the respective languages, the former as irregular plurals, the latter as singulars. Gil (1996) and Corbett (1996, 2000) treat the Maltese collective as a third number value beside singular and plural, mainly because collectives are typically employed as indefinite mass plurals of countable singulars, distinct from the countable plurals used after numerals (examples from Mifsud 1996: 37):

(3.24) GLOSS   SINGULAR   PLURAL   COLLECTIVE   (Maltese)
        'lark'   alwetta   alwettiit   alwett
        'fly'   dubbiina   dubbiniit   dubbiin

In fact, the singular and the plural used after numerals are simply the regular inflectional forms of a singulative (hence count) lexeme derived from the basic unsuffixed form, which occasionally can itself be pluralized, as in *dbiiben* (as shown by Mifsud 1996: 37, with more examples). The schema in (3.24) is misleading, and should be redrawn as in (3.25):

(3.25) GLOSS   SINGULAR   PLURAL   (Maltese)
        'lark'   alwett
        'lark'$_{count}$ alwetta   alwettiit   (singulative)
        'fly'   dubbiin   dbiiben
        'fly'$_{count}$ dubbiina   dubbiniit   (singulative)

While the singulative semantically ranges over countable units, the basic forms *alwett* and *dubbiin* are non-countable and express the lexical predicate as an abstract property or as a concrete sum of its instantiations, displaying the same concrete–abstract ambiguity often found in mass terms like *artillery, neighbourhood*, or *police*.[8] If the referent consists of salient discrete entities, the collective can play the role of a mass plural, much as the denotation of *livestock* (singular mass) can overlap with that of *cows*. This does not differ substantially from Gil's (1996: 54) interpretation of the collective as a fundamentally 'non-singular noun', which can denote a mass or a plurality depending on the referent; but its semantic non-singularity is a lexical conceptualization that opposes it to the singulative, true of atoms, and not an additional number value. Syntactically, collectives generally trigger singular agreement, and the plural agreement they can trigger if their referents are high in the Animacy Hierarchy is simply an instance of semantic-based agreement, common in Semitic (see Corbett 1996: 111–14, and Chapter 7). Since no morphological or grammatical property distinguishes collectives from singulars, the lack of any specific marker for

---

[8] Greenberg (1974: 29) points out that nouns like *police* (*faculty, personnel*) can be numerically quantified by some speakers, but the numbers must be large: *fifty police* sounds indeed better than *five police*, because it does not enforce a segmentation into individual units.

collectives is entirely expected, because they are morphologically underived singulars.[9] They therefore differ from lexical derivations like the Hebrew 'dual', which encroaches on the morphological number opposition by producing forms distinct from both plural and singular.

3.5.2 *Noun classes*

In noun class languages, the whole stock of nouns is partitioned into a number of morphologically distinct classes, each of which is associated with one or more morphemes defining a concord pattern. Nouns belong to one or to more than one class as a lexical specification, mostly semantically motivated. The relevance of this for lexical number is that, as we have seen in Section 2.7.4 of the last chapter, nouns that are singular in one class X are not always regularly mapped to their plural counterparts in another class Y, as would be the case if number cut the whole class system in two. Such a biunique relation holds only for the classes involving humans, animates, and other very salient individual referents. For the rest, as we have seen, class membership has more to do with diffuseness and cohesion than with singularity and plurality, and nouns may have the morphology of one class or another according to their part-structure conceptualization. In a way, noun class systems turn basic stems into nouns by means of a restricted set of grammaticalized nouns: a 'super-lexicon' that partitions the lexical noun stock in terms of individuation (the term and the idea are due to Contini-Morava 1999: 23). This formulation brings out the intrinsically lexeme-forming nature of number in noun class languages. Kihm (2001), whose proposal I follow here, argues more precisely that plurality (or, rather, part-structure information) is an intrinsic property of the class prefixes, and is not brought about by a grammatical operation. Nominal stems become nouns by being associated with one or another prefix, and the association that gives rise to Manjaku *u-ndali* 'cat' is not grammatically related to that which gives rise to *ngë-ndali* 'cats' (Kihm 2001: 29). In noun class languages, then, number is always lexical, not because it is always unpredictable or idiosyncratic (which it is not), but because it is an ingredient of lexeme formation.

If one should (imprecisely) equate noun classes with gender values, the lexicality of number in those languages could be expressed as follows: every gender value is intrinsically associated with a number value, in such a way that in order to change number, a noun must change gender. Interestingly, this

---

[9] Gil (1996: 77) notes that collectives are like basic singulars in having no marker, but attributes this to the relative unmarkedness of the mass interpretation relative to the count singulatives (singular and plural). He also mentions that collectives can be pluralized (note 36, p. 83).

description corresponds with some approximation to what happens in some Cushitic languages (see Corbett and Hayward 1987 for Bayso; Corbett 2000: 165–6, 181–5 for Somali, Bayso, and Qafar; Oomen 1981 for Rendille; Serzisko 1982, Saeed 1987: 114–16, Lecarme 2002, and Aïm 2003 for Somali). Without going into details, these languages have a clear grammatical opposition between two genders, evidenced in the relation between nouns and agreeing categories. These two values have a semantic basis in the distribution of animate nouns that justifies the labels of masculine and feminine. The crucial property is that some nouns (not all), change gender when they change number:

(3.26) | GLOSS | SINGULAR | PLURAL | |
|---|---|---|---|
| 'father' | áabbe (masc.) | aabbayáal (fem.) | (Somali: Lecarme 2002: 112) |
| 'mother' | hooyó (fem.) | hooyó-oyin (masc.) | (Somali: Lecarme 2002: 109) |
| 'male sheep' | waráab (masc.) | waraáb (fem.) | (Rendille: Oomen 1981: 54) |

This statement of what is traditionally known as polarity (cf. recently Baerman 2007) is a simplification, but not an oversimplification. As the agreement patterns show, plurals genuinely take the 'wrong' gender, rather than homophonous affixes (notice that in the Rendille example pluralization is signalled by a change in stress and not through an affix). Yet the morphological restrictions on this phenomenon, elucidated and discussed in the studies cited, speak against a generalized gender switch as a rather abstract exponent of plurality. Lecarme (2002) and Aïm (2003), in particular, analyse the Somali facts as a consequence of the status of the plural suffix, which they both view as intrinsically endowed with a gender feature that can override that of the noun. Oomen (1981) views the slightly different situation in Rendille as due to the part-structure information conveyed by the feminine. For her, the only real instances of polarity involve masculine nouns changing gender in the plural. The use of the feminine as a marker of plurality, she notes, is mirrored by the feminine derivations that often turn a 'collective' masculine into a singular in Cushitic. This, she argues, suggests that the feminine gender has a derivational use and contrasts with the unmarked masculine in expressing countability, singularizing unmarked collectives and pluralizing unmarked singulars.

Over and above the differences, these analyses converge in attributing polarity to a morpheme that displays the same properties of noun class markers: fused exponence of number and of inherent 'class' (here, gender). The relation of lexicalized plural with gender is a theme that will re-emerge repeatedly in the following chapters.

3.5.3 *Stem-inherent number*

The last type of inherent plurals we will examine inheres to the noun stem itself, rather than to a derivational affix. This state of affairs is common in Nilotic languages (Nilo-Saharian), here exemplified by Turkana. Noun morphology is extremely complex, but Dimmendaal's (1983) description clearly isolates the main features: any given noun (with one interesting exception, mentioned in Section 4.5.3 in the next chapter) has a prefix that fuses gender, number, and case, and some nouns, but not all, have a suffix that agrees in number with the prefix. Since number opposes singular and plural, one would normally expect singular and plural forms for both prefixes and suffixes. Some nouns behave like that, following a pattern whose productivity is demonstrated by loanwords like *e-kùk-ut* (from Swahili):[10]

(3.27) SINGULAR PLURAL GLOSS (Turkana; Dimmendaal 1983: 239–40)
   e-kùk-ut ŋi-kuku-*ì* 'chicken'
   a-kwàŋ-at ŋa-kwaŋa-ɪ 'lip plug'

The prefixes *e-* and *a-* realize masculine and feminine singular, which become *ŋi-* and *ŋa-* in the plural. The suffixes *-ut/-at* and *-ì`/ɪ*, here used from among several others, spell out the same feature combinations. As against this pattern, the nouns we are concerned with do *not* uniformly have suffixes. Instead of two suffixes for each noun, one for singular and one for plural, there is just one, so that the formal contrast involves not two suffixes but a suffixless and a suffixed form. Which of them expresses singular and which plural depends on the noun, and this is where lexical number comes into play. In one of the two groups, the singular is basic, and the plural is suffixed:

(3.28) SINGULAR PLURAL GLOSS (Turkana; Dimmendaal 1983: 233–4)
   e-putìr ŋi-pùtìr-o 'warthog'
   a-mosiŋ̀ ŋa-mòsìŋ-o 'rhinoceros'

The pattern is reversed in the second group, where it is the plural that has no suffix:

(3.29) SINGULAR PLURAL GLOSS (Turkana; Dimmendaal 1983: 228)
   e-turkàna-ɪt ŋi-tùrkanà 'Turkana person'
   a-lagà-ɪt ŋa-lagà 'copper necklace'

---

[10] I will ignore the phonological readjustments that often differentiate singular and plural stems, like *kwàŋ/kwaŋa*. Grave accents mark low tone, and a transcription like *ì`* indicates a word final floating low tone, in this case after *i* (cf. Dimmendaal 1983: 39).

Schematically, then, the expression of number on nouns in Turkana can be summarized as follows, where px stands for the prefix:

(3.30)   Singular   Plural
       Class I   px-stem-sg   px-stem-pl   (=(3.27))
       Class IIa px-stem   px-stem-pl   (=(3.28))
       Class IIb px-stem-sg   px-stem   (=(3.29))

Each prefix has one and only one form for any gender–number combination. Crucially, this form is the same across all classes. Therefore, what decides whether a noun falls in class I, IIa, or IIb is not the prefix, which is class-invariant, but the stem itself. It is the choice of one noun instead of another that determines whether a suffix is needed to express singularity, plurality, or both. In sum, the morphology of Turkana distinguishes nouns that are intrinsically neither singular nor plural (class I), inherently singular (class IIa), and inherently plural (class IIb). Unsurprisingly, this partition has a semantic motivation. Inherently singular terms refer to abstract concepts and to some mass notions (Dimmendaal's examples suggest a preponderance of masses with no natural articulation into subparts: liquids, but also 'sand' and 'flour'), and of course to referents that can be experienced as clear atomic units—what for us are count nouns. On the other hand, 'nouns denoting entities that normally occur in unspecified numbers, such as, for example, "hair", "grass", "word", and entities that normally occur in pairs, like "ear", "breast", most often have a plural which occurs as the basic form' (Dimmendaal 1983: 227). Notice the dangers of English translation: what occur in pluralities are single hairs, ears, or words, but not single 'grasses'. Despite the semantic motivation, these noun classes remain grammatical entities, as is shown by some glaring examples: Dimmendaal (1983: 224) notes the pair 'fly' (basic plural) ∼ 'bee' (basic singular), and we have 'buttocks' among basic singulars (p. 235) and 'anus' among basic plurals (p. 247).

  Interestingly, the very same suffix can serve to mark either singularity or plurality, specifically -*ì* (Dimmendaal 1983: 240–9; cf. also Corbett 2000: 162 for other Nilotic languages). This should not come as too big a surprise, since for a large part of the nominal vocabulary what counts is not the identity of a suffix, but whether a suffix is there or not. This way of using one and the same affix to mark the opposite of the inherent value of a noun is generally called inverse marking, and finds a particularly wide application in North American languages of the Tanoan family (cf. Corbett 2000: 161; Noyer 1997: 167–95; Harbour 2003, 2007). While invariant suffixes occur alongside single-value suffixes in Turkana, the Tanoan languages Jemez and Kiowa make use of a single suffix. This, in conjunction with a system of distinct noun classes, suffices to express the values singular, plural, and dual. Kiowa, to which I will limit my exemplification,

distinguishes four noun classes, with a semantic basis but grammatically defined by the interpretations they assign to noun forms with or without the inverse marker -*gɔ́*. Class I comprises animates (and mobile artifacts like 'wheel' or 'car': cf. Noyer 1997: 180–1, from which all examples are taken), and a bare form is either singular or dual; the suffix -*gɔ́* inverts this value and brings about a plural reading: *cê:* 'one horse, two horses', but *cê: -gɔ́* 'horses'. The same mechanism is at work in class II, consisting of inanimates. Here the bare form is, however, non-singular, which means it can refer to two or more referents but not one; the inverse marker produces the singular: *tʰɔ́:sè* 'two or more bones' ∼ *tʰɔ́:sè-gɔ̀* 'one bone' (where -*gɔ́* becomes -*gɔ̀* by regular tone lowering: Noyer 1997: 181). Class III likewise comprises inanimates, but this time the bare form is exclusively dual: affixation with -*gɔ́* therefore returns both non-dual forms, namely singular and plural, as in *kʼɔ̂n* 'two tomatoes' ∼ *kʼɔ̂n-dɔ́* 'one or more than two tomatoes' (-*dɔ́* being the allomorph of -*gɔ́* after/n/). Finally, class IV groups together nouns to which the inverse suffix never applies, like *tó:* 'teepee' and *có:* 'rock'. These nouns fall into three subclasses distinguished by the different agreement prefixes they determine on verbs, depending on their interpretation as singulars, duals, or plurals (one class always triggers the plural prefix). None of the nouns grouped together into class IV has a stem-intrinsic number value which may be switched by the inverse marker. For morphology, if not for syntax, they are effectively numberless, as distinct from the inherent-number nouns of classes I–III.[11]

   As can be seen, stem-inherent number, which is the prerequisite for inverse marking, allows a very economical representation of a complex set of oppositions. Since all number values are expressed in this way on all nouns, the Tanoan languages represent an extreme example of how inherent number can be integrated in a grammatical system. And because the inherent number values of nouns are not mediated by any marker outside the root, this system must also rank as the one where number is most inextricably associated with a lexeme. From the point of view of morphological form, number does not get more lexical than this.

3.6 Conclusion

The variety of phenomena examined in this chapter fall into a few basic types of morphologically defined lexicality. First, there is a notion of word as a syntactic unit defined by its distribution; this encompasses plural words and, under a

---

[11] The qualification 'if not for syntax' refers to purely syntactic *pluralia tantum* like *tó:* 'teepee', which trigger the plural agreement marker *gyà* whether their reference is singular, dual, or plural. For a more general view of the number system of this language, and for a featural analysis of inherent number values, see Noyer (1997: 167 95) and Harbour (2003, 2007).

broad enough construal, clitic-like plural affixes. Second, plural morphology may be lexical if it affects the lexical stem of a word, rather than its grammatical appendages. This connects suppletive plurals and plural markers that appear word-internally, inside derivational affixes or compounding. In some languages, the lexical stems of nouns involve class diacritics, either in the form of concrete affixes or as stem-inherent specifications that determine grammatical oppositions. When plurality is expressed through these diacritics, it contributes to make up a lexical stem too, but for a whole morphological class of nouns rather than for individual bases. The plural types illustrated in the course of the chapter differ substantially from one another, but all of them can be, and indeed have been, viewed as 'lexical', although no one study has considered them all as a whole category. The notions of syntactic word and lexical stem enable us to see the underlying unity of this vast and diverse morphological domain.

4

The meaning of lexical plurality

4.1 Introduction

The morphological overview in the previous chapter brought out the importance of lexemes as bases for inflection, and we have seen that every noun, as a lexeme, encapsulates a conceptualization of its reference domain in terms of its constituent parts. Inflectional number, in languages that have it, turns this lexeme-inherent information into a grammatically fixed set of functionally opposed features, primarily singular vs. plural. The content of this grammatical opposition appears in its pure form on pronouns, which have no lexemic content but consist entirely of grammatical information. With lexical nouns, on the other hand, the precise interpretation determined by number is sensitive to the conceptualization inherent in the lexeme: for example, *eggs* and *wines* are both plural, but their different countability preference implies that *wines* can mean 'wine portions' or 'wine varieties' much more readily than *eggs* can mean 'egg portions' or 'egg varieties'. The interaction between grammatical and lexeme-inherent information is especially close when plurality brings about a distinct shade of meaning, or when a particular form of plural selects nouns with a specific lexical content. In these cases, plurality remains a grammatical category but its content shades into lexical semantics, determining or being determined by the information inherent in the lexeme. This chapter will chart the empirical content of this interaction between grammatical and lexical meaning.

In order to achieve a revealing categorization, the survey will structure the whole range of lexicalized plural readings according to a small number of semantic common denominators. These revolve around the conceptualization of a reference domain in terms of its constituent parts; in particular, in terms of what counts as 'one'. In a nutshell, plural does not mean 'many' but 'not-one', and its precise semantic value on a noun depends on how that noun defines 'one'. As we will see, all instances where grammatical plurality affects, or is affected by, the lexical semantics of a noun, can ultimately be traced back to the following notions: unity, identity, boundary, cohesion, and instantiation. Unity and identity are the most fundamental concepts, which

name respectively the qualities of what is a complete whole in itself, and of what is re-identifiable as one and the same. Boundary is what characterizes something as 'one' by reference to its boundedness, as opposed to a diffused continuum or to scattered fragments. Cohesion refers to the mutual bond between elements viewed as naturally co-occurring and conceptualized as parts of a larger whole. Instantiation, finally, is the relation between a property and an entity with that property; as we will see, plurality may bring about a concrete reading in terms of multiple instantiations.

By structuring the empirical domain in terms of these categories, this chapter aims to show that when a noun has a special meaning because it is plural, or when it has a special plural because of its meaning, the relevant semantic properties systematically have to do with criteria for oneness and granularity, which define what primitive entities the noun is true of. In other words, lexicalized plurality, being part of the descriptive content of a noun, co-determines the ontology in which the noun takes its reference. For this reason, I will not deal with pluralization as a formal operation defined on basic elements whose value is taken for granted, but focus instead on the basic elements themselves.

Spelling out these ideas will lead to a number of specific claims, which it is useful to summarize here:

- the interpretation of plural is not necessarily a function of that of the singular;
- nouns vary in the way they conceptualize the primitives of their reference domain, not just in binary terms (as atoms or non-atoms), but along the dimensions of unity and identity;
- these dimensions are continua, so that conceptualization is a matter of degree;
- the conceptualization of what is 'one' and 'not-one' is sensitive to the perceptual salience of boundaries, and to non-linguistic knowledge about functional cohesion;
- plurality is related to the contrast between concrete and abstract, because a conceptualization as 'not-one' may convey a concrete reading as manifold instances rather than as an abstract undivided property; and
- some plural nouns conceptualize their reference only as instances of a contingent property, without any information about the entities themselves; this corresponds to the philosophical concept of *tropes*.

The use of some technical notions from philosophy, formal ontology, and semantics (unity and identity, parts and pieces, kinds, instantiation, tropes) serves to clarify the interpretive nuances by reference to precise theoretical

concepts. I will not, however, attempt to formalize the interpretations of lexically plural nouns, chiefly because that would entail a formalization of the descriptive content of lexemes, that is, of the most elusive kind of information encapsulated in a word. For the same reason I will not use features to express the semantic contrasts that we will come across. One of the main contentions of this chapter is that, when plurality is lexicalized, its semantic value transcends the clear-cut systematic contrasts which we represent by means of features, and which are the hallmark of knowledge of language as a grammatical system. The value of plurality in cases like *waters* or *depths* is lexical, not grammatical, because it fuses with the meaning of the lexeme and so depends on its descriptive content. While features may certainly prove useful in lexical semantics, an attempt to regiment the meaning of lexicalized plurals into a system of semantic features would not do justice to the observed gradience and variability. This does not mean that lexical knowledge escapes scientific theorizing, of course; only that the conceptualization of lexical plurals is more revealingly explicated by relating it to extralinguistic ontological categories, rather than by restating the attested contrasts in the form of grammatical features.

The chapter is structured as follows. Section 4.2 argues, against a common attitude, that semantic plurality is a linguistic property in its own right, not reducible to singularity or atomicity. This perspective allows us to see the many generalizations and cross-linguistic parallels across lexicalized plural readings, instead of viewing them as unrelated exceptions. Section 4.3 traces the diversity of lexicalized plural readings back to the two fundamental aspects of the concept of 'one', namely unity and identity, elucidated by reference to the notions of part and piece. Section 4.4 examines readings where the elements of the denotation have low individuality because of the lack of perceptual boundaries or because of some cohesive relation between them. Finally, Section 4.5 will show how grammatical plurality can mark concrete extension in space or time, or refer to a kind as a plurality of instantiations, sometimes with no information about the single entities themselves.

4.2 Plurality without singularity

Many of the plural nouns we have seen so far, and many more that will come, feature an interpretation that, while undeniably plural, does not reduce to the familiar concept of a collection of individuals. While such non-canonical readings appear throughout the book, this section focuses specifically on justifying the following claim: plural does not reduce to the canonical count interpretation 'many singular individuals'. This is what makes it possible to speak of semantic plurality beyond the canonical grammatical opposition in number—that is, of lexical plurality.

4.2.1 *Pluralities and masses: the received view*

The singular *book* takes its reference over a domain of individual entities, each of which is a book (in any of the available senses), while the reference of *books* ranges over pluralities of books. But the concept of plurality is far from self-explanatory. For a start, it can be taken to mean a collection with an arbitrary number of elements, or a collection with a number of elements equal to or higher than two. When speaking of a collection, we may have in mind a single entity made up of its members, like a single wall is made up of its bricks, or we may employ the term as a shorthand for the list of its members; some predicates refer to collections as unitary entities, as in *these books fill the whole shelf* (where no one book fills the shelf, but the whole collection does), while other contexts force a distributive interpretation, as in *these books are bound in leather* (where each book has a separate binding). Moreover, predicates that apply to a collection as a group may or may not apply to all of its subcollections; for example, *these books are alike* may be true of subcollections of any size (from two up), while *these books are numerous* can only be true of subcollections of an appropriately large size. All of these issues, and related ones like quantification, numerical modification, reciprocity, and the semantics–morphosyntax map, form the subject matter of a wide research domain, probably one of the main subfields of natural language semantics (see Link 1983, 1998; Landman 1989, 1991; Barker 1992; Eschenbach 1993; Carlson and Pelletier 1995; Schwarzschild 1996; Winter 2002, to name just a few directly relevant references). What we are interested in here, however, is not one or another plural construction, but the semantic value of plurality in itself. This question comes into focus when comparing plurals with mass terms.

Suppose there are two books, *a* and *b*. The singular term *book* is true of *a* and also of *b*, but not of *a* and *b* together: a book plus a book does not make another book. Things stand otherwise for plurals and mass terms: if *books* is true of two collections (not individuals) *a* and *b*, it will also be true of *a+b*, because a quantity of books and another quantity of books, taken together, are still describable as *books*. Likewise for mass terms: if *a* is water and *b* is water, *a+b* is still water. This property of mass nouns and plurals, which sets them apart from singular count nouns, is traditionally known as cumulative reference, and its introduction into modern semantic studies goes back to the classic discussion in Quine (1960: 90–100; cf. also Goodman 1977 [1951]: 39). Mass and plural terms also share, to a degree, the inverse property, namely divisibility of reference: if *a* is water, a part of it is still water, and likewise if *a* is a collection of books, a subcollection *b* of *a* will still be *books*; not so for count singular terms, because if *a* is a book, a part of it is not a book. Divisibility stops, or becomes unclear, when

we reach a certain level: *books* is true of a subcollection of a collection *a* of books, but not (or not obviously) of a single book, which is likewise a part of *a*. Likewise, smaller and smaller amounts of water are still water; but if we accept that water is defined as a molecule, it does not seem appropriate to call water a quantity so tiny that it corresponds to a single chemical atom. As can be seen, the so-called minimal parts problem affects both mass and plural terms (though not in the same fashion). What, then, is the semantic difference between plurals and masses which correlates with the morphosyntactic distinction in number? The answer that most obviously suggests itself is that pluralities are made up of singular elements, while masses are not. In different ways, this is indeed the received view, which I will argue to be wrong. But even this view cannot simplistically hold that plurals and mass terms denote different entities. As is well known, in many cases a plural and a mass term can be true of one and the same entity in the world: *jewels* and *jewellery*, *rice grains* and *rice*, *cows* and *livestock* are just a few examples. The difference, intuitively, has to do with the way we view the referent, and in particular with the way it is structured (see Wierzbicka 1988: 542–8 for a perceptive discussion). The distinct conceptualizations underlying nominal reference along the mass–count opposition parallel the aspectual distinctions in the verbal domain, which correlate with the boundary properties of events. In an influential synthesis of much previous work in this line, Jackendoff (1991) categorized part-structural distinctions for nouns and verbs by means of a system of features. In the nominal domain, divisibility correlates with the feature [± bounded]: *apple* as a singular count noun is [+ bounded], because an apple does not remain an apple if it is cut in half, while *apple* as a singular mass noun is [–bounded], as it refers to the same foodstuff whether or not a part of it is removed. As noted, the plural *apples* is [–bounded] too. However, Jackendoff distinguishes singular mass and plural count terms by the feature [± internal structure], which has a positive value for count plurals, interpreted as aggregates of discrete elements, and a negative one for masses. As Jackendoff (1991: 20) clearly states, 'the value [–internal structure] does not mean lack of internal structure, but rather lack of necessary entailment about internal structure'; in other words, being articulated into discrete elements is a matter of conceptualization, so one and the same entity can be viewed as a mass or as an aggregate.

The same basic distinction between masses and pluralities characterizes work in formal semantics. Link (1983, 1998: 16) distinguished at the source the denotation of plurals and of mass terms, by having their reference ranging on two distinct domains, respectively that of individuals in the model and that of 'the set of all individual portions of matter in the model'. In this way, pluralities and masses are made up of formally distinct primitive entities (individuals vs.

divisible portions) as a matter of definition. The main reason for this radical separation is the choice of mereology rather than set theory as the basic formal tool. In this influential approach, plurals do not refer to sets (in the mathematical sense) but to formal objects defined as plural individuals, like the sum $a+b$ (distinct from the material fusion of a and b). Formally, these plural individuals are related to each other by the 'part-of' relation and by sum formation, in such a way as to define the algebraic structure of a partial order. Since the algebraic properties are the same for the reference domain of masses and plurals, the two cannot be distinguished on the basis of the formal properties; so, Link takes them to define distinct ontological domains. In a similar vein, Ojeda (1993: 119–21) identified the difference between pluralities and masses with that between atomic and atomless mereologies, that is to say, algebraic structures defined by ordering relations and ultimately consisting either of discrete atoms (for plurals) or of infinitely divisible portions (for masses). From a rather different perspective, but within the same algebraic framework, Moltmann (1997) characterized the difference between pluralities and masses purely in terms of part-structural properties (relative to an interpretation in a context). For her, plurals but not mass terms encode what she calls the whole-properties of individuals in their reference, making them accessible to the semantic requirements of some predicates (as we will see in Section 4.2.2).

Other approaches model the reference of plurals by means of sets rather than mereological sums (see especially Landman 1989, 1991, and Barker 1992 for collective readings). Accounts differ on the way to represent the reference of mass nouns, ranging from Bunt (1985), who argued that they are always conceptualized as made up of infinitely divisible parts, to Chierchia (1998*a*), who claimed instead that they are formally represented as consisting of individual minimal parts, and that the impression of infinite divisibility is a matter of semantic vagueness (for instance, the noun *rice* does not define the level at which we may legitimately stop calling an entity 'rice'). What matters here is that if plurals denote sets of individual entities, they are ultimately defined on the basis of singulars, as in Chierchia (1998*a*: 59–60): '[the plural operator] must map a set of atoms into the set of pluralities constituted by those atoms'. This amounts to positing a biunique relation between grammatical plurality and aggregates of atomic individuals, enumerable one by one and therefore count, as opposed to the continuous denotation of mass terms. Even some recent and otherwise innovative analyses share this traditional assumption. Heycock and Zamparelli (2005) posit a 'pluralizing' semantic operator both for plurals and for mass terms, but in such a way that morphosyntactic plurality is equated with reference to sets of individuals, while mass terms take their reference over the parts of singular individuals:

(4.1)   If P is the denotation of a predicate over singulars (a property of type
        $<<e,t>,t>$), the notation *P ('star P') stands for the set product of the
        elements in P, minus the empty set (the closure under union)

        If P is a singleton property of a singular individual (e.g. {{a}}), the
        notation $\div$ P ('div P') stands for the set of parts of that singular individual
                                        (Heycock and Zamparelli 2005: 220–1)

Over and above differences among semantic approaches to plurality, therefore,
the consensus view equates morphosyntactic plurality with a reference to discrete
aggregates, so that plurals must be count and non-count terms cannot be plurals.
The evidence in the rest of this section will show this view to be inadequate.

   Claiming that plurals are always count is not the same thing as claiming that
their meaning is based on that of singulars. I will argue against both of these
views, but there are approaches that uphold the first while rejecting the second.
Ojeda (1993), for instance, sees plurality as the default way to organize a reference
domain into subcollections of arbitrary size, while singulars require a special
stipulation restricting their reference to individual atoms. Plural reference,
which ranges over all subcollections, is encoded on the noun stem itself, not
on plural morphology (interpreted as identity function), and in particular it is
not computed from the interpretation of the corresponding singular.

   A similar stance has resurfaced independently in Borer's (2005) analysis of
the syntax–semantics map in the nominal domain. In this view, noun stems
provide a reference domain which is then partitioned according to the
information encoded on a functional head that expresses 'division' (Dɪᴠ); a
higher head 'quantity' (#), quantifies over the parts so defined, and the
association with a discourse referent takes place at the determiner level:

(4.2)
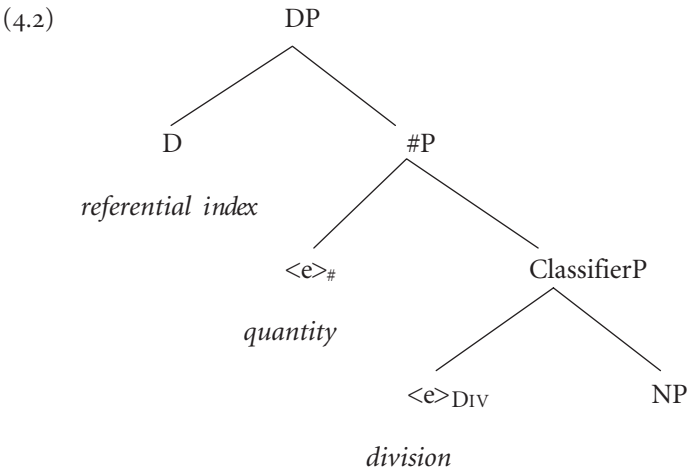
For Borer, the division of the noun's reference may provide a stable and uniform criterion defining what counts as a unit, or alternatively a range of all possible divisions, amounting to the possible portions of a mass. In the first function, the division acts as a classifier; singulars require this unit-defining stable partition. As for plural reference, Borer (2005: 115, 120, 128–9) states quite clearly that it comes about through the division of the reference of the bare stem, not by an operation on the reference of singulars. In other words, the meaning of a plural is not a function of the singular. The evidence from lexicalized plurality points to exactly this conclusion. Still, even in Borer's innovative approach (2005: 109), 'stems which are marked as plural now become count by definition'. But the facts do not bear out this unquestioned assumption.

4.2.2 *The semantic reality of lexical plurals*

The discussion of *pluralia tantum* in Section 2.4 and of semantic opacity in Section 2.8 has already made it clear that plurals can be mass, which is not a novel discovery (see McCawley 1975; Wierzbicka 1988: 499–560; Corbett 2000: 173; and Tamm 2004). However, since the strongest evidence comes from languages where the use of number differs from English, and the semantic literature is mostly English-centred, mass plurals are usually sidelined in works on formal semantics, and viewed as an exception, or as a genuine *explanandum* but only in the context of typologically rather different languages—not as a part of a semantic account of plurality. That is why it is important to recognize the properties of mass plurals as a class, not as unsystematic exceptions.

At first sight, the idea that plural nouns denote aggregates of discrete elements seems based on sound evidence. Although *furniture* and *pieces of furniture* may refer to the same entities, only the plural NP can act as antecedent of a reciprocal:

(4.3) a. those pieces of furniture are leaning
           against each other   (Chierchia 1998*a*: 86)
       b. *that furniture is leaning against each other

For Chierchia, this contrast shows that 'reciprocal predicates are sensitive to being plural vs. being singular'. Similarly, Moltmann argues that the semantic requirement of predicates like *distinguish* is 'selection of plural as opposed to mass NPs':

(4.4) a. John cannot distinguish the rice grains   (Moltmann 1997: 87)
       b. *John cannot distinguish the rice

It is easy to show, however, that mass plurals align themselves with singular mass nouns, and against count plurals, with respect to these properties (cf. also Wierzbicka 1988: 499–562 for more examples from several languages):

(4.5) a. these notes are different from each other (plural, count)
      b. *this money is different from each other (singular, mass)
      c. *these monies are different from each other (plural, mass)

(4.6) a. I enumerated the rainy days (plural, count)
      b. *I enumerated the rain (singular, mass)
      c. *I enumerated the rains (plural, mass)

(4.7) a. John cannot distinguish the rice grains (plural, count)
      b. *John cannot distinguish the rice (singular, mass)
      c. *John cannot distinguish the dregs (plural, mass)

As is well known (see Allan 1980 and Pelletier and Schubert 1989), the mass–count distinction is not always linguistically clear-cut, partly because the same noun may admit some determiners but not others (*three clothes* but *a few clothes*), partly because different nouns have different semantic requirements (*he counted the furnishings* but *he counted the cattle*), partly because usage varies. Thus, although none of the nouns in the following examples tolerate numeral modification (*three head of cattle*, but not *three cattle*), their acceptability varies in contexts that force a count reading (% indicates variable judgements):

(4.8) a. a few   cattle/clothes
                   %belongings/%embers/%oats
                   *arrears/*furnishings/*suds/*fumes

      b. I enumerated cattle/clothes
         the   %belongings/%embers/%oats
                    *arrears/*furnishings/*suds/*fumes

      c. These   cattle/%clothes   resemble each
                  %belongings/%embers/%oats   other
                  *arrears/*furnishings/*suds/*fumes

That even in English mass nouns are not all singular emerged clearly from Allan's (1980) detailed study, and the observation was anticipated, in the philosophical literature, by Cartwright (1970: 25), who commented: '"Groceries are..." is right; at any rate, "groceries is..." is wrong. But "two groceries" is wrong too, and I at least am quite unsure about "many groceries".' This simple observation suffices to prove that plurals with a mass

reading are not syntactically singular nouns with an irregular and semantically void plural affix, as claimed by Heycock and Zamparelli (2005: 228): 'plural mass nouns like *brains* or *funds*, which we can assume to have a – Plur value despite the plural suffix (hence *he hasn't got much brains/funds*)'. These nouns may well agree in the singular, but that is no precondition for mass reading: *these funds/resources/monies are insufficient* can mean 'this amount of money is insufficient', while triggering plural agreement on *these* and *are*. That reading clearly contrasts with the count reading of *this collection of funds/resources/*monies* (the latter term being ungrammatical because the *plurale tantum* admits no count interpretation).

A much weaker objection consists in claiming that nouns like *brains* are indeed plural morphosyntactically, but not semantically. This is the position of Ojeda (1993: 120) and Borer (2005: 105), who identify mass plurals with *pluralia tantum*. But the force of this objection is inversely proportional to its contentfulness. If all it means is that mass plurals do not have the semantics of aggregative plurals, then it is obviously true—they are not plural, if being plural means to be count by definition. What else could one mean by claiming that *brains* is semantically not plural? Certainly not that it is singular like *book*, which is true of atomic individuals. Perhaps that it is singular the way *water* is singular, that is, as a mass noun. But then the claim that mass plurals are semantically 'not really plural' reduces to the restatement that they are mass. However, why should mass terms be semantically singular rather than plural? If anything, the shared semantic properties of cumulativity and divisibility make all mass nouns (regardless of morphosyntactic number) akin to plurals. This is indeed what Chierchia (1998*a*) proposes, and it is fully confirmed by the relative acceptability of some of the examples in (4.8), where the context enforces reference to more than one entity. If terms like *cattle*, *belongings*, and *fumes* were simply not plural in a semantic sense, they should all pattern alike, and nothing like the shaky and variable judgements in (4.8) would be predicted. In my view, the very variability and fuzziness of some of the data show instead that grammatical plurality has a semantic content, but one that fuses with the lexical meaning and therefore changes with the choice of the word.

One may be excused, at this point, for suspecting that this is just a matter of terminology. It is, but terminology is not neutral. One may well decide to call semantically plural only count nouns; my point is that this is a bad terminological choice, because it implies that all mass terms are non-plural, not only *water* and *brains*, but also *cattle* or *waters*, in a way that lumps together very different types of referents, overlooks the similarities between mass and plural reference, and effectively views the choice of morphosyntactic number for mass terms as arbitrary.

This last point is crucial. Even when it does not have a canonical aggregative reading, grammatical plurality is not semantically irrelevant, nor is its distribution arbitrary. The exemplification we have seen so far is enough to show not only that mass plurals are relatively common (especially across languages), but also that they correlate with certain concepts and not others, expressing certain readings and not others. Consider the English pair *water* ∼ *waters*. Both are mass, so the two number values cannot express a countability opposition. If plurality were just a matter of morphosyntax, the two forms should be semantically interchangeable. And so they are, in some contexts—but not all, as (4.9) shows:

(4.9)   a. the river discharges its water/waters into the lake
       b. the formula of water/*waters is H_2O

Apparently, there is still a semantic difference between the singular and the plural version, which rules out the plural in (4.9b). The point generalizes to all semantically lexicalized plurals, whether or not the singular is count. To stay with English, recall the nouns of the type *fund* ∼ *funds* 'many a fund' ∼ *funds* 'funding' (discussed in Sections 2.4 and 2.8). There are many such cases where the plural has one reading which does not reduce to 'many a *x*': more examples are *ashes*, *brains*, *crops*, *debts*, *depths*, *dimensions*, *directions*, *foundations*, *gates*, *heavens*, *heights*, *holidays*, *intricacies*, *loyalties*, *manners*, *mists*, *plains*, *plans*, *preparations*, *proofs*, *rains*, *resources*, *results*, *skies*, *snows*, *suspicions*, *thoughts*, *times*, *views*, *waters*, *winds*, *woods*. The claim that plurality 'isn't really there' semantically seems a very poor characterization of the relevant readings. In many cases, these readings incorporate a notion of manifoldness or complexity; in others, a sense of concreteness; or the plural may simply serve to eliminate the boundedness of the singular if this is count (cf. for instance *woods* meaning 'wooded landscape form'). But plurality does affect the interpretation, no matter how elusive its function may be.

Of course, having shown that morphosyntactic plurality has a semantic value even when lexicalized is not enough. Unless we can say what this value is, lexical plurality may mean just about anything, which is as much as saying that it means nothing. The rest of this chapter will flesh out a substantive semantic characterization based on a few basic concepts.

4.3 Ontological categories for a semantic typology

The previous section has led us to reject the identification of plurality with the reading 'consisting of many individuals'. A first step towards a better understanding must consist in cataloguing the alternative readings. However, such a typology remains ultimately arbitrary without an explanation of why certain

semantic properties are involved and not others. To address this question, this section traces back the semantic range outlined so far to its ontological underpinnings. The label 'ontological', as opposed to 'semantic', reflects the claim that lexical plurality expresses a conceptualization of part structure; this is not a specifically linguistic notion, but one that anchors linguistic meaning into the categorization of thought and reality. Of course, describing linguistic conceptualization is not the same as describing reality (or whatever name one wants to call the object of metaphysics); but both rest on a basic conceptual vocabulary. The fundamental categories needed to understand how lexical plurals conceptualize part structure are unity and identity, as two related but distinct ingredients of the notion 'one'.

4.3.1 *Unity*

Unity is what defines something as a discrete whole in itself, rather than as a fragmentary part of a continuum or of a larger whole. Usually, language seems to wear on its sleeve what counts as a unit, via grammatical means like number and the mass–count distinction: *shoe*, unlike *footwear*, encapsulates a criterion for what counts as one in the extension of the predicate. If plurals really referred to collections of units as a matter of definition, the linguistic conceptualization would be clear, no matter how inadequate or arbitrary it may prove under scientific or philosophical scrutiny. But, as we have seen, plurality, countability, and conceptual discreteness do not always coincide, so one notion cannot be used as a diagnostic for the other.

Ontology is useful in this situation because it gives us an independent way to talk about the conceptualization encapsulated in words. To see how, consider the following example. Intuitively, a leg and a splinter are both parts of a table, but not in the same sense. The splinter is one among an undefined number of wood fragments, we have little or no expectation as to its appearance, and we may seriously question whether it exists as a splinter before being detached from the rest of the table. By contrast, we can identify a leg among a small number of similar parts, associate it with a function and with some expected traits of its appearance, and readily admit its existence whether detached or undetached. One might sum up these differences by saying that the splinter is part of the table, while the leg is *a* part of the table. However, this is just a way to recast the perceived distinction in linguistic terms (through countability), without stating what it lies in. The properties I listed, by contrast, make more precise the intuition that a leg is more 'one' than a splinter. Krecz (1986) proposed to distinguish the two concepts by means of the terms 'part' and 'piece'. Both are subdivisions of a greater whole,

but for pieces the division is arbitrary (a splinter of a table, a shard of glass, a slice of pie), while parts correspond to specific ways to subdivide the whole (like a leg of a table, the stem of a glass, the crust in a pie): 'The cut of a part in a whole is *locus-specific* and therefore nonarbitrary. Cut in the wrong place and you have failed to distinguish the part. Instead you will end up with a piece' (Krecz 1986: 383).[1]

To see the relevance of this for plurality, consider a multitude of parts and pieces, in this technical sense. The former consists of a collection of well-defined elements, each of which counts as one because of its intrinsic properties. By contrast, pieces like splinters count as one only in so far as they are detached in a particular situation; they have no context-independent criterion for unity. Besides, since any detached and suitably small bit of wood qualifies as a splinter, there is no way to judge whether one such bit is a whole splinter or only a piece of it. We can tell one splinter from two, but not from a half. In a way, then, a subpart of a splinter is still a splinter, like a subpart of water is water. Yet *splinter* is count, and it does not refer to a continuous substance but to discrete elements. What it shares with mass terms is the lack of a stable criterion of granularity (a term borrowed from Chierchia 1998*a*). So, we conceptualize the reference of both *legs* and *splinters* as a plurality, and both are count, but the former is more count than the latter. There are differences in part-structure conceptualization, then, that the linguistic mass–count opposition does not express.

This example shows how two count nouns may differ in what Moltmann (1997) calls 'whole properties', which characterize the reference of a noun as made up of whole, discrete, single elements. If being a whole was a yes–no property, it would amount to a simple restatement of the mass–count contrast. Instead, (Moltmann 1997: 20–3) argues that it can be a matter of degree, beside being relative to a particular interpretation and to a specific dimension; her example is Kurdistan, which has whole properties culturally but not politically. The relative fuzziness inherent in such an information-based account may be criticized for its indeterminacy (Pianesi 2002), but it is justified on two grounds: first, it reflects the blurred and variable nature of much of the data themselves, and second, unity as an ontological notion must be interpretation-dependent, and so ultimately relative rather than absolute.[2] This emerges quite clearly from the fact that attempts to formalize unity must

---

[1] Lowe (1998: 73) uses the two terms in exactly the opposite fashion.

[2] The indeterminacy of these semantic intuitions transpires very clearly when considering the transitivity of the part of relation. The concept of part ensures that, if *x* is a part of *y* and *y* is a part of *z*, then *x* is a part of *z*. Now, Jim's nose is a part of Jim, and Jim is a part of Jim's family, so it should follow that Jim's nose is part of Jim's family but this inference, while formally unexceptionable, conflicts with our intuitions. Moltmann (1997: 27 9) proposes that transitivity is only blocked when the middle term is a 'functional part' of the last term. For example, *page>book>written work* goes through (with > for 'is a part of'), but *page>book>library* does not, because libraries, but not *written work*, have books as

base it on some unifying relation (Simons 1987: 324–38; Moltmann 1997: 24–6; Guarino and Welty 2000; Gangemi *et al.* 2002), so that an entity can only be a whole relative to some relation, which mutually connects all its parts and nothing else. As Simons (1987: 331) makes especially clear, recognizing that being whole is a matter of degree does not mean that there are degrees of individuality, but that 'there are degrees of warrant for accepting the existence of an individual composed of certain parts', depending on the relations chosen and on decisions about the relative importance for determining individual integrity. Simons goes on to argue that in some cases there is no fact of the matter as to whether an entity is one whole or a manifold complex: his examples are multicellular organisms such as sponges.[3]

It is generally assumed that language imposes its own categorization regardless of these ontological fine points: a sponge, like a splinter, is viewed as 'one' as the referent of a singular count noun, the objection would run. But linguistic conceptualization is not always so transparent, and in many cases— more than is usually assumed—it requires elucidation; this brings us back to lexical plurals. What is a singular count noun in English may be a plural (with or without a singular) in other languages, and forms like *pliers* or *resources* should remind us that even English can express through plural nouns concepts that straddle the line between single wholes and manifold complexes. Moreover, alternative number values may convey the complexity or the unity of one and the same entity (think of *the crowd/crowds dispersed* referring to the same collection of people). So, it is precisely these ontological fine points that underlie the uses of plural which go beyond what is grammatically regimented; and the whole properties encoded in number must be information-based and relative to a specific interpretation, if they can be back- or foregrounded by a choice of number or of a particular type of plural.

Nouns may conceptualize their referents as lacking unity. For *splinter*, this has no grammatical effect. Other nouns show some limited effects of this conceptualization: as we will see in Chapter 4, some Italian plurals are count but resist direct numerical modification. Many more are the cases where lack of unity corresponds to lack of grammatical countability. Obviously, the prime examples are nouns denoting homogeneous substances or abstractions (cf. the examples

---

functional parts. However, an arm is a functional part of a body, yet *forearm*>*arm*>*body* does go through, as does *platoon*>*regiment*>*army*. The crucial factor is the uniqueness of the part of relation: libraries are made up of books and not of pages, while the constituent parts of bodies or armies are not so uniquely specified. But then, semantic judgements depend on aspects of lexical semantics which can be quite elusive: *nucleus*>*cell*>*organism* (Moltmann 1997: 15) is often cited as an incorrect inference, but the judgement is just as clear as one's understanding of the concept *organism* (cf. Pianesi 2002: 88).

[3] As Simons (1987: 326) points out, the fact that being a whole is a relative notion was recognized by Aristotle (*Metaphysics* △, 6, 1016a10), who observed that a straight line is 'more one' than a bent line, because we are prepared to say that a line which forms a corner is one and not one at the same time.

in Sections 2.4.2, 2.7.4, and 2.8.1). The linguistic import of conceptualization shows through more clearly in those mass plurals whose reference does actually make up a multiplicity of entities, discrete enough to be 'many', but not so 'one' to qualify as individual wholes. English examples are *belongings*, *entrails*, *foundations*, *furnishings*, or *intricacies*; we will see more examples in the chapters to come. In most of these cases, the lack of unity criteria for the parts that make up the noun's reference correlates with a certain semantic vagueness, because the lexical predicate describes the function or the salient trait shared by the multiple entities involved, without however describing what these elements are; I will consider this aspect more closely in Section 4.4.2 below. It must be noted, however, that it is not vagueness by itself that causes grammatical non-countability. *Furniture*, for instance, is not too vague although it ranges over distinct types of objects; and *cattle* non-equivocally identifies a single type of animal, yet both are non-count (the Arabic 'collectives' discussed in Chapter 7 are particularly instructive in this respect). Conversely, nothing could be vaguer than the description associated with *thing* or *entity*, yet these are count nouns and so, no matter how little we may know about something designated by them, we know that it is one. If anything, their descriptive content reduces to unity.

In sum, the characteristic interpretation of certain plurals is best described in terms of the degree of unity of the parts making up their reference. Without this notion there would be no middle ground between totally homogeneous masses and totally discrete aggregates. In addition, unity also provides the semantic generalization behind some morphological phenomena. Part II will feature several illustrations, so one example will suffice here. Kihm describes as follows the choice of plural noun class prefix in Manjaku (West Atlantic):

Take 'fingers', for instance: if the plural refers to a discrete number of fingers that does not usually exceed ten, i.e. the normal number a human is endowed with, then it is expressed in noun class 10 (e.g. *kë konj kë wants* 'three fingers'); if it refers to an unknown and/or indefinite number, generic interpretation (fingers in general) included, it is expressed in noun class 8 (*i konj* 'fingers'). (Kihm 2001: 9 10)

(See also Corbett 2000: 35.) There is no need to argue that the conceptualization of such cases is the same as that of homogeneous substances, or that plurality 'doesn't count'. An indefinite multitude is a concept distinct both from a continuous mass and from a denumerable collection; we can describe it as a plurality of entities that have a low degree of unity.

4.3.2 Identity

If unity is what makes an entity one, identity is what makes it one and the same. While the former describes what characterizes a discrete and self-connected

whole as opposed to an arbitrary cut in a continuum, the latter is what allows for (re)identification of an entity as one and the same. The two are very closely related, but should not be confused. To fully appreciate this point, let us go back to one of the most cited passages in Quine (1960: 91): 'To learn "apple" it is not sufficient to learn how much of what goes on counts as apple; we must learn how much counts as *an* apple, and how much as another'. So, a grasp of the term requires, first, knowing what it applies to; second, since this is a count term, knowing what counts as one. This in turn requires not only an ability to distinguish a unit from what is more or less than a unit, but also to distinguish units from one another, by identifying those properties which allow us to establish that two descriptions refer to one and the same referent. These identity properties underlie what we understand to be an object falling under the relevant concept. For example, we are prepared to accept that what we call *river* can remain the same object although its material parts change with time, while other nouns, like *apple*, require their referents to preserve their material composition through time. The two readings of *book* as physical object and as abstract publication determine distinct ways to define what counts as the same item: several copies of a publication are distinct objects under the former definition, but one and the same under the latter. Similarly, *passenger* and *person travelling in a vehicle* are distinct ways to define what counts as one, because one and the same person may stand for several passengers in a sentence like *National Airlines served at least two million passengers in 1975* (Gupta 1980: 23), even though passengers are certainly persons. But the two terms, while ranging over the same referents, differ in conceptualization—not in what counts as a whole, because the granularity of the domain is the same for both concepts (single human beings), but in what counts as one and the same, and so in identity properties.[4]

In the strongest form, identity properties provide a criterion for discriminating objects by means of necessary and sufficient conditions. The clearest example is that of a set in the mathematical sense, for which the axiom of extensionality states that sets A and B are one and the same set if and only if they have the same members. Identity criteria, then, spell out what it means to be a particular entity falling under a general concept, and that places them at the very heart of the philosophy of being.[5] But apart its central role in

---

[4] Gupta does not distinguish between unity and identity, but between principles of application (stating what entities a noun is true of in a world) and of identity (stating what counts as the same entity in and across worlds).

[5] The logical status of identity, the proper formulation of identity criteria and their role in metaphysics and ontology are the object of vigorous debate; see Strawson (1997) [1976], Gabbay and Moravcsik (1979) [1973], Griffin (1977), Wiggins (1980), Hirsch (1982), Brennan (1988), Lowe (1998), Guarino and Welty (2000), and Gangemi *et al.* (2002). On the linguistic side, Krifka (1990) addressed the problem of

metaphysics, identity also has an important role to play in clarifying linguistic conceptualization. Unlike formal constructs like sets, the concepts of natural language do not usually allow for explicit *criteria* of identity, but we can list properties enabling us to track the same entity across repeated experiences. Building on the logical analysis of common nouns of Gupta (1980), Larson and Segal (1995: 128–30) and Baker (2003: 101–9) have made the very important point that only nouns can apparently encapsulate this type of information. Baker (2003), in particular, has taken the encoding of identity conditions as a defining property of nouns as a lexical category: by carrying a criterion of identity, nouns and only nouns convey a notion of objecthood. However, Baker inherits from Gupta (1980) the lack of a principled distinction between unity and identity (cf. instead Hirsch 1982: 236–263, and especially Lowe 1998: 58–83 and Guarino and Welty 2000). This leads him, in agreement with Griffin (1977: 51), to directly link identity and countability:

An essential precondition for counting a group of things is the ability to distinguish which of those things are the same. In order to count a group of dogs, I must not count the same dog twice. Therefore, I must know if X (the one I am focusing on now) is the same dog as Y (the one I just counted). In other words, I must use *dog*'s criterion of identity. (Baker 2003: 106)

But identity properties are not immediately tied to countability. As Lowe (1998: 33) observes, '*mass* terms such as *gold* and *water* appear to convey criteria of identity but not necessarily countability—one can meaningfully ask whether the gold in this room . . . is *the same* as the gold that formerly composed a certain ornament'. Besides, many nouns like *object*, *thing*, or *part* are grammatically count and denote discrete entities, but simply do not provide enough information to allow for an enumeration of their referents. As is well known (Lowe 1998: 31, Chierchia 1998a: 71), *how many books are here?* admits a definite answer, but *how many things are here?* does not, because the term *thing*, while grammatically countable, is too vague to convey what must be counted—not only is every object a thing, but also every collection or part of an object, and every part of an infinitely divisible mass.[6] So, some nouns with identity properties are not count, and some count nouns do not carry identity properties.

determining what counts as an individual referent in contexts like *4,000 ships passed through the lock* (true of ships or of ship passing events, with different truth conditions). Barker (1999) qualified this analysis, arguing that the identity conditions associated with a noun should be relativized to the context of interpretation.

[6] Griffin (1977: 60) calls these terms *dummy sortals* or *sortalizing auxiliary nouns* and Pelletier (1979: 12 [1975]) *secondary sortals*. Sortals are grammatically count terms that describe *sorts*, or types of individual particular entities, as opposed to amorphous masses or abstract kinds; cf. Wiggins (1980: 63 4), Hirsch (1982: 34 71). Philosophers who uphold the thesis of 'relative identity' claim that every identity statement of the form *a is the same as b* is relativized to a sortal, as in *a is the same S as b* (Geach

In order to encompass the whole spectrum of conceptualizations express-ible by nouns, and not just those that encapsulate a criterion for enumeration, we need to distinguish unity from identity. For Guarino and Welty (2000), a property is said to carry identity or unity if there is some condition for deciding what should count as 'the same' and as a whole, for all instances of that property. Apples can be identified, since it is possible to tell whether a twice experienced apple is or is not the same object on the basis of having the same material composition over time, and 'biological unity' (p. 4) provides a basis for judging what counts as a whole apple. As a count noun, then, *apple* names a property whose instances can be identified and can be recognized as wholes. In the mass reading, *apple* can still be identified, in so far as two masses can always be identified if they have the same parts ('mereological extension-ality'), but it provides no basis for judging when an instance of apple-mass is a whole, such that each of its parts are connected by some relation to all other parts and to nothing else. Mass concepts, then, carry identity but not unity. In sum: 'When something is an instance of a property *carrying* identity, it *can be identified*. If something is an instance of a property that carries unity, it is a *whole*. If something can be identified *and* is a whole, then we say it is an *individual*' (italics in the original).[7]

This partition makes room for count nouns that express properties carry-ing only unity but not identity. A property in this category would provide a criterion for deciding when its instances are whole, but not for deciding that one is the same instance as another. Lowe (1998: 62) exemplifies this case with electrons: as they are conceived of in quantum physics, the electrons in a helium atom are definitely two, but we cannot keep track of which is which, not just because we cannot tell, but because there is no fact of the matter about the question. We may label them *a* and *b*, but neither has individual properties guaranteeing that an observed electron is *a* rather than *b*.

This is, I believe, a less than compelling example, but the usefulness of identity as distinct from unity emerges much more clearly when elucidating the conceptualization of certain linguistic expressions, in particular of plurals, to which we finally return. Certain expressions conceptualize their referents as

---

1962; Griffin 1977; Wiggins 1980). Baker (2003) effectively reduces nouns to sortals, leaving out mass terms and dummy sortals. This follows Gupta (1980), who however claimed that his analysis of common nouns extends to mass terms because, for him, mass terms are sortal (p. 25, note 17)   a position at odds with general usage, as Moltmann (1997: 38) notes. The use of a mass term like *gold* in *this is the same gold as that* masks an implicit reference to some (count) sortal: the statement has different truth conditions depending on whether it means 'the same type of gold' or 'the same piece of gold', with *type* and *piece* relativizing identity to a sort.

[7] Adjectival predicates like *red* are said to carry neither unity nor identity; what can be identified as the same are not two reds, but two red objects, or two colours, in any case entities expressed by nouns and not adjectives.

token-units indistinguishable from each other: this applies to all nouns used as standards of quantification, not only those exclusively used as units of measurement like *meter*, *kilo*, or *yard*, but also otherwise referential nouns in classifier-like constructions like *three packets of cigarettes* or *ten sacks of coal*, which have the morphosyntax of the corresponding referential nouns even though in this construction they do not presuppose the existence of the containers they name. Nouns used in this function often have distinctive morphosyntactic properties, like the exceptional lack of plural morphology displayed by *head* in the construction *three head of cattle*, or by the German *Sack* 'sack, bag' in *drei Sack Mehl* 'three bag[s] of flour'; or they may have special forms for units of counting, like the Irish plurals discussed in Chapter 6. Lack of identity is the common semantic trait shared by the various morphological irregularities of unit expressions (nouns and classifiers) across different languages. What is especially useful about using identity as the relevant concept, rather than for instance individuation, is that identity clearly contrasts with unity. Nothing can be more unitary than a unit of measure, which provides the very standard for oneness along a given dimension. When we say that *head* in *three head of cattle* lacks individuality, therefore, what we mean is not that it lacks a clear determination of what counts as one, but rather than it consists *only* of this determination, lacking instead the ability to refer to one particular item.

Abstract standards of measurement have no identity properties as a matter of logical necessity. More numerous, and more interesting, are those terms referring to entities that are distinct in principle but conceptualized as practically indistinguishable. We will see in Chapters 5 and 6 that the morphologically irregular plurals of Italian and Irish comprise not only units of quantification, but also other terms like divisions of time (like 'week' or 'year'), body parts, and entities like eggs or shouts. Divisions of time come closest to abstract measures, and have the same tendency to morphological or syntactic irregularity cross-linguistically. As we saw in Section 2.6.4, they emerge as a semantically coherent group among the Hebrew plurals that irregularly take the dual ending; and even in English, the following contrast, pointed out by Frank Anshen, suggests that the determiner *a pair of* requires nouns whose referents can be distinguished from one another:

(4.10) a. a couple/pair of students
        b. a couple/*pair of hours

For referents other than divisions of time, the lack of identity properties is more a matter of convention than of objective semantic characteristics. Every day we experience entities as discrete members of aggregates without being

able to tell the difference between one element and the other in isolation: small objects or animals, modular elements forming a collective whole, atoms of granular masses, or simply entities that we are not used to differentiating. English *pluralia tantum* like *oats* fit this description. Other languages have various ways to express this kind of conceptualization. When there is a morphological class typical of mass nouns, it often includes the plurals of such non-individual entities, as class 6 in Swahili (see Section 3.7.4). Other languages distinguish basically singular nouns for intrinsically individuated entities from basically non-singular ones for masses and aggregates of undifferentiated entities (cf. the inherently plural or singular noun stems in Turkana or Kiowa discussed in Section 3.5.3). Welsh, for example, has a large class of monomorphemic 'plurals' indistinguishable from singular mass nouns, including *adar* 'birds', *coed* 'trees', *dillad* 'clothes', *graean* 'gravel', *plant* 'children', *pys* 'peas', and *pysgod* 'fish' (King 1993: 67–9). Chapter 8 will discuss similar facts in Breton.

For such concrete entities, the role of identity as opposed to unity in bringing about a non-count interpretation may appear dubious. After all, conceptualizing a plurality of individuals as undifferentiated does not seem too far from conceptualizing them as a mass. However, there are mass plurals for which objective linguistic phenomena clearly rule out a conceptualization as a substance-like continuum. For example, *cattle* and *clothes* support at least some count determiners even though they resist direct numerical modification (see (4.8) above). If a determiner like *a few* is acceptable, then, as in *a few clothes/cattle*, the denotation is structured more like an aggregate than like a homogenous mass. Much stronger evidence comes from so-called collective plurals in Arabic, to be discussed in Section 7.5.2 in Chapter 7. These are non-count plural nouns for aggregates of undifferentiated tokens like animals, small plants, or other perceptually indistinguishable objects. While definitely non-countable, these plurals can occur as arguments to predicates that unequivocally distribute over individual elements (thanks to Jamal Ouhalla, whose judgements correspond to those reported by Zabbal 2002):

(4.11)   a. hut   'fish'   (Moroccan Arabic)
         b. *xamsa hut   'five fish'

(4.12)   hseb   l-hut   wahda b wahda
         counted   the-fish   one by one
         'he counted the fish one by one'

Since (4.12) is grammatical, *hut* 'fish (pl.)' cannot denote fish as a continuous mass that blurs the boundaries between individual animals. On the contrary,

this form makes the individual fish accessible to semantic interpretation. But then, the reason for the uncountability of (4.11b) cannot be that *hut* refers to elements without unity. The only possibility is that this form conceptualizes fish as undifferentiated, and so unable to support enumeration, which requires fixing the identity of the counted elements (except for abstract units; cf. Chapters 6 and 7). In order to understand these facts, we need both identity and unity as distinct dimensions of conceptualization.

4.4 Conceptual/perceptual categories

Distinguishing unity from identity has shown how oversimplistic it would be to interpret lexicalized plural readings in terms of a simple binary choice between discrete aggregate and continuous mass. Sharpening the analysis, we will now see how some finer-grained interpretive categories situate the interpretation of lexical plurals along the dimensions of unity and identity.

4.4.1 Kontur/*boundary/shape*

What Langacker (1987) and Jackendoff (1991) call *boundary*, Meisterfeld (1998) *Kontur*, and Rijkhoff (2002: 50–2) *shape*, identifies the unity of count nouns by relating them to the perceptual discontinuity delimiting what is perceived as a unitary whole;[8] see Simons (1987: 354–60) and Hirsch (1982: 236–63) for references to *Gestalt* psychology, and Meisterfeld (1998: 37) for antecedents in the philosophical tradition. Moltmann (1997: 39) rightly points out that this intuitive notion does little to elucidate the semantic distinction over and above formal properties like cumulativity. Actually, her own account is based on what is ultimately an intuitive notion of whole properties (Pianesi 2002: 101). But, as I have already observed in Section 4.3.1, unity is relative to unifying relations, and these are ultimately based on the lexical meaning of natural language predicates, with all the fuzziness this implies.

While perceptual boundedness may be less than enlightening as a basis for the mass–count distinction, there are some cases where its role is decisive. Consider English count nouns like *cliff*, *mountain*, or *hill*. Their plurals seem to have a transparent interpretation: for whatever is a cliff, *cliffs* refers to a corresponding plurality. But a moment's reflection shows that this is not totally accurate: *cliffs*, *mountains*, or *hills* often mean something like 'cliff-area', 'mountainous/hilly landscape', 'wave-filled body of water'. A house in

---

[8] This is a simplification. Rijkhoff (2002: 50 5), in particular, bases his system of oppositions on the syntactic behaviour of nouns with respect to classifiers and determiners, and achieves a much more complex classification.

the hills does not have to lie between distinct hills, but generically on high ground (and asking 'how many hills?' sounds silly); falling from some cliffs is not appreciably different from falling from a cliff; and a trip to the mountains is not necessarily made up of individual single-mountain stages. At first sight, plurality has here a massifying function. But notice that the single entities which make up the reference of *cliffs* or *mountains* are still cliffs or mountains, under any interpretation, while the things that constitute *brains*, *funds*, or *resources* (in the mass reading) are not single brains, funds, or resources. The contrast hinges on boundedness as a purely perceptual notion. Terms such as *cliffs*, *clouds*, *hills*, *mountains*, or *waves* name concrete entities that have a criterion of unity provided by a salient focal point (an edge, a top, or a crest), but they do not define clear boundaries. We have no problems telling one entity from two, but there is no non-arbitrary line at which one ends and another begins. In the plural, such entities can of course be viewed as atomic wholes, in so far as their focal points are perceptually distinct (as in *these two waves/hills are very different from each other*). But this individualizing conceptualization requires a small number of referents, each with an informationally foregrounded focal point. When referents are many, or an indefinite quantity, shifting the focus away from the focal points of waves or hills amounts to depriving them of their whole properties; cars, by contrast, still have boundaries that define each element as a unit. In sum, the plural of these terms has an in-built vagueness which results in a weakly individual interpretation, not so much for the lack of identity or unity properties, but specifically for the lack of perceptual boundaries.

Recognizing the role of perceptual boundaries in conceptualization thus enables us to see where otherwise regular count nouns like *cliff* overlap the semantics of mass nouns, without for that shifting the meaning of the lexical predicate, as happens instead in *brain* ∼ *brains*. Besides explaining why this kind of reading affects the plural of certain nouns and not others, it also brings to light an interpretive trait that has morphological reflexes in languages other than English (for instance, the noun for 'wave' can have three different plurals in Arabic; cf. Section 7.5). Finally, the linguistic relevance of perceptual boundaries is firmly rooted in the conceptualization of entities as wholes, as Meirav (2003: 260–2) emphasizes. The notion of 'battlement', to cite his example, 'would often seem not to determine whether a certain stone which is a part of the castle is or is not also a part of the battlement. One is likely to be unsure where the battlement ends and the rest of the castle begins'. If the identity of a concrete object were determined by its material parts, this vagueness about constituent parts would automatically be vagueness about the whole. And yet, we can clearly tell whether some stones are a whole

battlement or just a part of it, and identify a battlement as one and the same, despite being unsure about its boundaries and its material constitution. This is because what we understand as an entity falling under a sortal concept is not defined by the extensional sum of its parts, but by a notion of unity compatible with distinct material constitutions (see Meirav 2003 for further discussion of what he calls unities, as opposed to sums). The mass-like interpretation of plurals like *cliffs* stems from the lack of perceptual boundaries, not from a conceptualization as a continuous mass.

4.4.2 *Cohesion and collectiveness*

As we have seen in Section 4.3.1, the notion of unity is based on relations that connect all the parts of a whole with each other and with nothing else. A whole may of course be composed by other wholes, in which case unifying relations hold between the individual members of collective concepts, like *orchestra* or *committee* ([+internal structure] and [+bounded] in the system of Jackendoff 1991). To quote Gil (1996: 63): 'even though *the boys* may be coextensive with *the team*, the latter NP says more, namely that the boys are organized in a particular fashion. In other words, the whole is more than the sum of its constituent parts: it is non-additive'. I will refer to these unifying relations as cohesion.

As I have just described it, cohesion is the unifying principle of a single entity, whether atomic or collective, referred to by a grammatically singular term. But pluralities may be conceptualized as cohesive even in the absence of a corresponding singular term. Typical examples are body parts that naturally occur in pairs or sets, like arms or teeth. There is no need for a singular noun meaning 'arm-pair' or 'denture' for each element to be viewed more as a part of an organ than as an organ in itself. In this way, cohesion directly affects the whole properties of a plural's reference: the greater the unity of the complex (a pair of limbs, a set of teeth, the legs of a chair), the lesser the unity of its constituent parts. For this reason, the degree of unity of a tooth is less than that of a head, although both are body parts. We therefore expect (correctly) to find more frequent reflexes of cohesion in the morphology and semantics of nouns for 'tooth' than for 'head'. Languages often provide a morphological expression for the cohesive reading of plurals, either by special lexeme-forming affixes like the German collective prefix *Ge-* as in *Geschwister* 'siblings' (from *Schwester* 'sister'),[9] or by means of lexically listed plural forms like the final -*a* of Italian *ossa* 'bones' and the ablaut-plural of Breton *dent* 'teeth', or by

---

[9] *Geschwister* is syntactically plural, although most German collectives in *Ge* are singular, like *das Gestirn* 'the star collection'.

'minor numbers' like dual and paucal; see Corbett (2000: 23–6, 97–101), and the discussion of modern Hebrew in Section 2.6.4. There is a whole literature on the cohesive overtone of the dual in Indo-European and especially in classical Greek (see Delbrück 1893: 133–46; Brugmann 1900: 371, and the reassessment of Meisterfeld 1998: 109–13), which Wackernagel (1926: 83) concisely summarized as follows: 'wenn es sich darum handelt, Zusammen-gehöriges zusammensufassen' [when it is a matter of putting together what belongs together]. The cohesion associated with the dual can sever the connection with pairwise reference, turning it into a paucal; this has happened in Russian, for example, where nouns governed by 2–3–4 take a special form historically derived from a dual (Corbett 1993), as well as in modern Hebrew and in some Arabic dialects (cf. again Section 2.6.4 and Brustad 2000: 45–52), and in Akkadian dual forms like *šinnaan* for 'teeth' (von Soden 1969: 76); see Section 8.3.3.1 in Chapter 8 for similar facts in Breton.

The grammatical importance of cohesion comes to the fore in North American languages, which often feature distinct forms for nouns and verbs according to whether their referents are somehow related or not (Mithun 1988, 1999: 79–94; Corbett 2000: 111–20). The cohesive forms may denote referents that are simply close to each other in space or time, or related in some other way. Ojeda (1998) has discussed in some detail the semantics of distributives and collectives (non-distributives) in Papago. In this language, nouns (and verbs) are categorized on the basis of the part structure both of their referents and of the locus they occupy. There are four classes of nouns: individual types 1 and 2, mass, and aggregate. Each class distinguishes at least two forms, a singular, a non-singular and/or a 'distributive'. In the words of Mathiot (1983: 205), cited by Ojeda:

The nominal singular indicates a single entity at a single locus. The nominal non distributive indicates a single locus without specifying the number of entities. Both the nominal nonsingular and the nominal distributive indicate several entities at several loci, without specifying the number of entities per locus. . . . Entities referred to by mass nouns or individual type 2 nouns are viewed as being coterminous with their respective loci. The locus of entities referred to by aggregate nouns is viewed as being the groups to which these entities belong, for example, a herd. The locus of entities referred to by individual type 1 nouns, which typically refer to domesticated animals and tools, is viewed as being their owners or makers. (Ojeda 1998: 248)

When locus means spatial proximity, the number opposition indicates whether the referents occupy one or more than one locus, and the noun classes define different ways to conceptualize occupancy (by one or an unspecified number or items, or by a mass). Locus, however, refers to

different types of cohesion as well, not only proximity but also membership in a collection (for aggregates) and relation to salient humans (for domesticated animals and tools). As a result, locus defines the granularity that is relevant for the number opposition:[10]

(4.13)   a. háiwañ   (collective)   'one or more head of cattle belonging
                                              to the same herd'

         b. háhaiwañ   (distributive)   'cattle belonging to more than one herd'
                                              (Papago; Ojeda 1998: 248)

Ojeda provides an explicit mereological model for the various interpretations of Papago nouns, but he takes as given the property of sharing the same locus ('Let us furthermore assume that two head of cattle are equivalent if and only if they belong to the same herd' (p. 258); 'But the notion of equivalence involved in the semantics of (non)distributives is identified not by grammar but by other aspects of Papago culture' (p. 260)). This major assumption allows a precise definition of the denotation of plurals: this is made up of every sum of x and y such that $x \neq y$, where the latter relation means either that x and y are not the same individual or that they do not belong to the same locus (distributive reading). But taking for granted the mutual equivalence of referents at the same locus means taking for granted what makes this system so peculiar. Only the most individuated referents (individual type 1 nouns) support a singular–plural distinction independently of locus, and even then they have a third form for referents scattered across loci. For all other classes, what counts as 'one' or 'more than one' for grammatical purposes is not the referent itself but its locus, which is a dimension of cohesion that changes according to the class of the noun (spatial proximity, membership in the same stable collection, shared relation to a salient human), and so a lexically encoded notion. In sum, cohesion as a lexical semantic property in Papago is so intimately integrated with the number system that, in certain noun classes, distinct entities that share a locus are not just conceptually indistinguishable, but also grammatically singular.

   Apart from its reflexes in the grammatical oppositions themselves, cohesion also has semantic effects in the readings of some lexicalized plurals. Despite the ubiquitous use of the term 'collective' (justly rejected by Gil 1996) to refer to variously cohesive plural readings, the property of being related together should not be confused with that of referring 'collectively' to

---

[10] See also Mithun (1999: 88 9), who explains that distributives differ from plurals because they indicate that the reference is scattered over different places or different types.

one whole.[11] Within the literature on formal ontology, the point has been made by Gangemi *et al.* (2002), who have distinguished the property of being unified by a relation from the property of being *maximally* unified by it; only the latter defines a whole. But the conclusion that the members of a plurality may be mutually related without necessarily making up a bounded whole is also required by a careful analysis of some linguistic facts. To anticipate an important result of Chapter 5, consider Italian *-a* plurals that refer to natural sets of body parts, such as *dita* 'fingers' or *braccia* 'arms'. As we will see, it is true that these plurals are irregular because they name entities occurring in cohesive aggregates; but it is *not* true that they must denote these collections. *Dita* may refer to any number of fingers above one, not just to a natural set of five or ten. The same applies to *ossa* 'bones', which denotes bones viewed as parts of a skeleton, but not necessarily arranged into skeletons.[12] As opposed to the regular form *ossi*, which conceptualizes single bones as wholes, *ossa* names the same entities conceptualized as related parts, and therefore with a lower degree of individual whole properties. It is true that the regular *ossi* typically describes disconnected bones; but *ossa* can refer to disconnected bones too, provided they are viewed as partial fragments; it can be true of one and a half skeletons, for example. To claim that *ossa* denotes bones arranged in skeletons would be like claiming that the expression *mechanical parts* denotes complete appliances.

Although Italian 'collective' plurals do not denote entire natural collections, they are cohesive because their referents are parts of such larger wholes. In other cases, cohesion holds between the elements of something which is not a bounded whole. In fact, once detached from collectiveness, the notion of cohesion emerges as an important factor in a great many instances of lexical number, especially (but not only) *pluralia tantum*. Plurals like *belongings*, *contents*, *foundations*, or *depths* name the cohesive principle that unifies heterogeneous entities or events. As Wierzbicka (1988: 539–45) noticed, forms like *groceries* or *leftovers* describe entities that belong together not on the basis of their intrinsic properties, but merely on the basis of contingent groupings in space and time. These nouns lack a descriptive content that might identify a sortal, and label instead the contingent property on which the grouping is

[11] Not all plurals denoting natural collections qualify as collective. *Parents* or *heirs*, for instance, necessarily refer to a closed group bound together by a clear cohesive relation; but, unlike *orchestra*, they do not allow ascribing a property to the group as a single entity, as in **my parents have two members* or **the heirs shrank in size*. Yet, like all plurals, they can allow for non distributive predication as in *my parents are two in number* or *the heirs met*, where the predicate is true of the group but not of its members. This type of predication is often called 'collective', but it does not make *parents* any more collective than *things*.

[12] Cf. instead Corbett (1996: 104): '*Le ossa* is used of bones which belong together, a particular person's bones a skeleton'. The first characterization is correct, the other two are not.

based. Consider English gerundial formations like *beginnings*, *belongings*, *findings*, and *furnishings*. Their interpretation may be dubbed 'the whatever reading': *beginnings* are whatever events mark a beginning, *belongings* are whatever objects make up someone's property, *findings* whatever results have been found, *furnishings* whatever objects furnish a closed environment.

Some cross-linguistically recurring semantic classes of *pluralia tantum* likewise name what fuses distinct events into a ritualized occasion, without describing the single events themselves. Plural names of ceremonies like the Latin *nuptiae*, the Lithuanian *vesuves*, or the Finnish *häät* (all 'nuptials'), names of games like the Russian *prjatki* 'hide-and-seek', or names of eventualities like the Russian *rod'i* 'childbirth' or *zamorozki* 'light frosts', can all be construed as referring to whatever happens during the corresponding occasions (cf. Section 2.4.2).

English gerundial nominalizations like *belongings* and event-referring *pluralia tantum* thus illustrate the same type of reading, one which denotes a plural reference domain leaving its constituent elements unspecified. Like other lexicalized plurals such as *foundations*, these nouns are interpretively plural, but they do not express what they are pluralities of. What makes this reading possible is the mutual cohesion holding between the single elements, inherently or contingently. Cohesion turns these elements into subcomponents of a larger concept, which alone is described by the lexical predicate. Being conceptualized as parts of a larger domain, and so radically underspecified as to be unidentifiable, the single elements denoted by cohesive plurals like *belongings*, *nuptials*, or *foundations* lack both unity and identity.

4.4.3 *Split levels of oneness*

The granularity defined by a number system assigns the status of singular or plural to entities that we may perceive as one and not-one at the same time, according to different criteria. Morphological systems that have special markers for cohesion and/or collectiveness alongside a number opposition can express the two conceptualizations at the same time, as in the following examples from the North American language isolate Zuni:

(4.14) a.

Noun		Noun-Coll-Sg	(Zuni; Mithun 1999: 92)
lu	'ash'	lu-ł-'e	'ashes in an ashtray'
sa	'bone'	sa- łpo-n	'skeleton'
he	'metal'	ha- łpo-nne	'bridle bit'
c'ina	'paper, letter'	c'ina-p-'e	'papers in a drawer'

b.

Noun		Noun-Coll-Pl	
mo	'spherical object'	mo-pa-we'	'truckloads of melons'
sa	'bone'	sa-pała-we'	'very skinny people'

The collective suffixes, attached to the root inside the number suffixes, shift the criterion for unity from single entities to cohesive collections, which then feed the number opposition. I quote Mithun in full:

A single set of objects is treated grammatically as singular, as can be seen from the singular suffixes *ʼe*, *n*, and *nne* on 'ashes in an ashtray', 'skeleton' ('bundle of bones'), 'bridle bit' ('metal in an arrangement'), and 'papers in a drawer'. Collective nouns with plural suffixes designate multiple sets, as in 'truckloads of melons' ('spherical objects in deep containers') and 'very skinny people' ('bones wrapped in bundles'). (Mithun 1999: 92)

The simultaneous presence of distinct levels of granularity can also be expressed through number alone, in languages without a dedicated morphological expression of cohesion. Plurals interpreted as multiple parts making up a single bounded whole represent the most obvious example: English nouns like *pliers*, *pyjamas*, *scissors*, *trousers*, Russian ones like *sani* 'a sledge' or *sutki* 'one day+night cycle', and those plurals which can refer to a single entity as well as to a multiplicity, like the classical Greek *harmata* 'a chariot', *domata* 'a house', or Latin *cunae* 'a cradle' (Brugmann 1900: 170; Meisterfeld 1998: 115). In some cases, the ambivalent unit conceptualization has syntactic effects, as in the tendency to accept singular agreement in constructions like *a garden shears* (see Section 2.4.1), or, more dramatically, in the possibility of numeral modification with the determiner 'one' showing plural agreement, as in the Russian *odni sani* 'one.PL sledge.PL', *odni nosilki* 'one.PL stretcher.PL' (Wade 1992: 199).[13]

Breton offers a particularly clear example of different criteria of unity for one and the same form of a noun. This involves some nouns whose plural denotes a cohesive collection. The plural of *bot-ez* 'shoe', for instance, does not mean 'shoes' but 'a pair of shoes' (and can indeed agree with a singular determiner). Unlike in *scissors*, here the constituent parts of the cohesive wholes are wholes themselves, representing the denotation of the singular; an English counterpart could be *wheel* ∼ *wheels* 'a car'. Now, the unity imposed on this cohesive plural is not just a matter of world knowledge, but has the grammatical consequence of allowing for a further pluralization, as if *wheels* 'a car' had the plural **wheelses* 'cars' (something that has happened in *summons* ∼ *summonses*):

---

[13] The cohesion inherent in plurals like *scissors* also establishes strong unity criteria. *Scissors*, without a disambiguating context, can refer either to one pair or to several; but a heap of loose scissor blades would not qualify as *scissors*, because the reference domain only includes blades connected to form a functional whole.

(4.15) *SG (an atom)* *PL (a natural set)*   *PL+PL (sets, natural or otherwise)*
      troad 'foot' treid 'a pair of feet'   treit-ou 'feet (in pairs or not)'
      bot-ez 'shoe' bot-ou 'a pair of shoes' bot-ey-er 'shoes (in pairs or not)'
                                                      (Breton; Trépos 1956)

Notice that the meaning of the doubly pluralized form in the third column does not arise compositionally from the cohesive reading as pair in the second column. *Treit-ou* does not have to mean 'many pairs', but it can also mean 'many feet' without cohesive reading. So, this is not a simple reanalysis of *treid* as a singular with an irregular plural form. Instead, *treid* maintains both senses 'one as a pair of feet' and 'two as single feet'; further pluralization has access to the second as well as to the first sense, and can therefore refer to pluralities of single feet or of pairs. Leaving to Chapter 8 the task of placing this phenomenon in its proper context, we can conclude by once more noting the connections that emerge across languages with different morphological systems, once the categories that underlie the semantics of plural are made clear in terms of unity, identity, and cohesion.

4.5 Plural and instantiation

The discussion on cohesion has brought into focus a type of semantic irregularity that plays a major role in many lexicalized plurals, namely the fact that the lexical predicate does not describe the units of the denotation. In some cases, these units are simply unspecified; in others, the elements of the denotation are named but what counts as one is determined by other factors, like cohesion. This contrasts with the canonical reading of regular plurals, where the granularity of the domain is given by the lexical predicate. In the cases to which we now turn, what defines a noun's denotation as plural, that is, what divides up its reference in such a way that the noun is true of collections and not of atoms, is instantiation, or the relation between a property and what has that property. In the case of plurally conceptualized mass nouns, to which I turn first, an intrinsically continuous referent is viewed as 'not-one' because its instantiation is structured into spatial or temporal parts. Next I will interpret certain lexicalized plural readings, in particular the 'whatever reading' (cf. Section 4.4.2), as denoting instances of a property, rather than entities defined by the lexical base.

4.5.1 *Situated extension*

A significant share of lexicalized plural readings involves mass nouns where plurality contributes one or more of the following characterizations: concreteness, abundance, dispersal, and reference to the spatial region occupied

by a substance, rather than to the substance itself. What all these characterizations have in common is a division of the reference not on the basis of its atomicity, but of the way it is instantiated in space and time.

Abstract nouns provide a convenient point of departure, because the semantic effect of pluralization is mostly quite clear. English is probably one of the worst languages in which to investigate the pluralization of abstracts, but even here it is clear that *feelings*, *loyalties*, and *sorrows* do not (or do not necessarily) denote collections of abstract objects, or subtypes of the corresponding concepts. Nor does the plural refer to the abstract property itself: a sentence like *I respect her loyalty* means 'I respect her quality of being loyal', but *I respect her loyalties* means 'I respect her choice of what to be loyal to'.[14] The difference stands out even more clearly in a context like *loyalty is/*loyalties are hard to come by* (used by Moltmann 2004 to identify reference to properties as abstract objects). An even better example is *depths*, which sharply differs from *depth* in denoting not an abstract property, but some area or areas having that property.[15] Pluralization here causes a concrete interpretation, where the property denoted by the noun is situated in space—not as *one* deep place, but as places, with the vagueness caused by the plural's characterization as 'not-one'. Other examples involve a concrete reading situated in time as well as space. *Pains*, for instance, describes the sensations caused by pain, not as denumerable events but as making up a state. *Times* acts as a generic name for the eventualities (events and states) that identify stages in the chronological flow: if time is a river, *times* describes its *waters*. Possibly for historical reasons (the influence of the Latin *tempora*), this particular example of lexicalized pluralization is generalized across European languages (perhaps elsewhere); see Plungjan's (1997) detailed study of its Russian version, which opens by referring to its Latin antecedent. Lexical pluralization is not always so uniform cross-linguistically, however. The French *amours* 'loves', for instance, refers to episodes involving love ('love affairs'); this reading is not prominent with the English *loves*.[16] In other languages, the pluralization of abstracts is less sporadic, but causes the same kind of situated interpretation. Landgraf (1906), for instance, observes that Latin abstract nouns tend to be

---

[14] Thanks to Jennifer Petrie for discussion on this point.

[15] The following contrast, for which I am indebted to Josef Bayer, highlights this semantic difference (the star indicates that the singular is not ungrammatical, but odd and leading to a clearly unintended reading):

(i)   I admire the depth /*depths of his knowledge.

[16] Lewis (1960) entitled a book *The four loves*, that is the four types of love; I find this reading of the plural totally unacceptable in the Italian *amori*. Although this is tangential to the discussion, note that pluralization does not turn a mass term *X* into a count term 'types of *X*' with the ease some authors imply. Chierchia (1998a: 57) presents as grammatical *in this lab we store three bloods*, but I find the Italian counterpart totally impossible, with 'blood' as well as 'milk'   which is quite surprising, given that these substances have culturally recognized subtypes.

pluralized when the expression refers to multiple manifestations, as in *irae* 'bursts of rage' or *morae* 'displays of hesitation'.

Concrete mass nouns provide a much richer exemplification. English terms like *fogs*, *rains*, *sands*, *snows*, or *waters* in contexts like *the October fogs* or *lost in the sands* illustrate a use that finds parallels in many other languages, not only within Indo-European. These pluralizations filter out the reading as abstract kind, which conceptualizes a mass as one abstract individual, and bring out the concrete sense as mass with a spatial extension (see Pelletier and Schubert 1989 for this differentiation, which goes back to Quine 1960):

(4.16) a. the river discharges its water/waters into the lake
        b. the formula of water/*waters in H_2O

Other observations are in order. An entity described as *waters* must be a large body of water:

(4.17) a. the waters in the sea
        b. *the waters in a bottle/pool

Besides, any type of quantification seems to be excluded, including the existential interpretation that goes with null determiner constructions:[17]

(4.18) a. *the lake lost some waters through evaporation
        b. *there were waters everywhere

Definite descriptions are by far the most usual syntactic context, possibly the only one for this reading. This strongly suggests that the mass plurals in question conceptualize their reference as a kind; more precisely, not as an abstract individual but as a plural kind, in the sense of Chierchia (1998*b*), corresponding to the maximal sum of water-instances (in a contextually relevant situation, not necessarily in a whole possible world).

In sum, these mass plurals refer to a kind viewed as a concrete sum, extended in space and/or time, and large enough to be a feature of the environment. The necessary extension typically implies abundance. The denotation is divided, but the division lacks a standard unit, being instead derived from the extension into multiple spatial and/or temporal portions.

Descriptions of single languages have documented this reading of plural mass nouns in a variety of languages. A typical example is the comment by Waltke and O'Connor (1990: 120) on Biblical Hebrew: 'Plurals of extension

---

[17] Edgar Allan Poe spoke of 'the voice of a thousand waters' in the close of *The Fall of the House of Usher*; rather than being a genuine counterexample to the claim that *waters* cannot be quantified, this example of poetic style merely shows that large numbers can act as non quantificational indications of size. Baudelaire translated it as *la voix de mille cataractes*, or 'waterfalls'.

indicate that the referent of the noun is inherently large or complex; the plural quality is the result not of countable multiplicity, but of a multiplicity that is nonetheless perceived as real'. The details vary, of course, and some languages bring out more than others different aspects of the situated sum interpretation. Cowell's (1964) monograph on Syrian Arabic, for instance, pinpoints with the aid of the following example the contrast between the generic/abstract reading of the singular *ramʔl* in (4.19a) and the concrete reading of the plural *ramlaat* in (4.19b) (I reproduce Cowell's glosses):

(4.19) a. har-ramʔl maa byəswa ləl baaṭoon   (Syrian Arabic; Cowell's 1964: 370)
             'this sand (i.e. this kind of sand) is no good for concrete'

        b. xood har-ramlaat mən hoon
             'get this sand (i.e. this batch of sand) out of here'

I will return in Section 7.5.5 to this use of pluralization in Arabic, which evidently differs from that of English.

The non-canonical values of plurality have also been carefully investigated in comparative Indo-European philology (cf. Delbrück 1893: 146–72 and Wackernagel 1926: 84–96, and more recently Meisterfeld 1998: 102–27). Numerous studies have focused specifically on Latin (Maas 1902; Landgraf 1906; Löfstedt 1928: 11–68; Hofmann and Szantyr 1965: 13–21; Kühner and Stegmann 1971: 67–86), and they have established that, factoring out the conventions of poetic diction, pluralization often targeted mass or abstract nouns, to express manifold instantiation and extension in space and time. Regarding this last aspect, Landgraf (1906: 66) made it quite clear that plurals like *harenae* 'sands' or *nives* 'snows' refer not just to the substance, but to occurrences of the substance extended in space: 'Wenn wir von Regionen des ewigen Schnees und ewigen Eises sprechen, so drückt das der Lateiner durch den Plural aus' [while we speak of places of eternal snow and eternal ice, the Latin speaker expresses this by the plural]; Italian, incidentally, works the same. The expression of repeated instance could apparently license the plural even for a noun like 'sun', as in the following example from Lucretius (*De Rerum Natura* 5.252):

(4.20)   pars terrai nonnulla, perusta solibus adsiduis (Latin; Landgraf 1906: 66)
            'a large part of the earth, scorched by continuous suns'

The literal 'suns' seems to refer to sun-stages, as ever-returning sun-events; Bailey's (1966) commentary notes: 'The plural denotes the effect of the sun on many occasions'. The granularity induced by division into stages parallels that illustrated by *times*, or *tempora*.

The equivalence between extended substance and the region it occupies is well attested elsewhere. In the following Ukrainian example, the plural *piski*

'sands' (which also in Russian has the value 'a great amount of sand') explicitly refers to a landscape form, and according to the informants I could consult, it contrasts with the singular *pisok*, which would be inappropriate:

(4.21) sered   landšaftnix   form   Avstralii
        among   landscape.GEN.PL   form.GEN.PL   Australia.GEN
        golovnim činom   prevaljujut'   piski   (Ukrainian)
        main.INSTR way.INSTR prevail.3.PL   sand.PL
        'among the landscape forms of Australia, deserts are prevalent'

Finally, there are languages where mass nouns systematically alternate between singular and plural depending on the conceptualization required by the context. According to the experimental study conducted by Sharifan and Lotfi (2003) on modern Persian, substance nouns like 'water', 'oil', and 'rice' favour the plural with predicates that foreground spread and spatio-temporal extension, like in the following examples:

(4.22) a. berenj-â-ro   invar-o   unvar   na-pâš
           rice-PL-ACC   here-and   there   not-throw
           'don't throw the rice here and there!'

        b. roɣan-â   dâre   mirize
           oil-PL   is   dripping
           'the oil is dripping'   (Persian; Sharifan and Lotfi 2003: 231, 236)

Dispersal is tightly linked to instantiation as a non-cohesive substance, which takes up the part structure of the environment it appears in; plurality is not simply a result of there being multiple discrete parts. Consider the following contrast between the plural *âb-â* 'waters' and the singular *âb* 'water':

(4.23) a. âb-â-ro   az   kaf-e   âšpazxune   jam   kon
           water-PL-ACC   from   floor-PRT   kitchen   gathering do
           'wipe away the water from the kitchen floor'

        b. âb-e   daryâ   bâlâ   umad
           water-PRT   sea   high   came
           'the sea level rose'   (Persian; Sharifan and Lotfi 2003: 235, 236)

The denotation of both terms is concrete, and the substance has the same physical properties. But water scattered on a floor, unlike that of the sea, is not one body of water, regardless of whether it forms one patch or several. The authors' comments on (4.23a) are revealing: 'One might propose that the plural form here refers to patches of water on the floor. However, our intuitions suggest that *âb-â* 'water-PL' could be used to refer to one big patch of water'

(p. 235). So what favours the plural is specifically the lack of cohesive integrity, not multiplicity. Persian thus highlights the overlap of concreteness, scatter, and abundance that makes a reference domain perceptually plural—or 'not-one'.

4.5.2 *Plurality as manifold kind-instantiation*

The previous section has focused on plurals whose reference is conceptualized as manifold because it is extended in space and/or time. In other cases, it is not a certain type of instantiation that motivates plurality, but instantiation itself. This emerges when considering the use of number in reference to kinds, that is, to the general concepts of which entities are instances.

The central insight at the basis of the notion of kind (as a linguistic term) is that we can speak about the regularities we perceive among referents of linguistic expressions, as if the regularities themselves were objects of reference (cf. Section 2.5). Certain contexts clearly distinguish what Krifka *et al.* (1995: 86) call 'object-oriented' and 'kind-oriented mode of speaking'. As is well known, predicates like *to be extinct* can only be true of kinds, not of the individuals they are instantiated by; something similar holds of *to be widespread* or *to be rare* (but cf. the discussion in Krifka *et al.* 1995: 95–8). Likewise, the object of verbs like *to invent* must refer to kinds: if we say that Edison invented the electric bulb, we do not mean that he invented this or that bulb, nor that he invented a set or even the whole set of them; what constitutes the object of an invention is a type, distinct from the token entities that instantiate it. Or again, taxonomic statements like *tigers are a subspecies of wild cats* (Link 1998: 221) refer to an abstract kind, because tigers are not subspecies. These predicates highlight the specificity of kind readings, but a noun can be kind-referring when combined with other predicates too. *To reach Europe after the Crusades*, for instance, is a property that can be true of an individual. But in (4.24a–b) it is predicated of a kind, in (4.24d) of individuals, and in (4.24c) it can apply to a kind or to a set of individuals:[18]

(4.24)   a. the black rat reached Europe after the Crusades
         b. *Rattus Rattus* reached Europe after the Crusades
         c. black rats reached Europe after the Crusades
         d. some black rats reached Europe after the Crusades

Insofar as it consists of its instantiations, a kind is at the same time one and many. The sentences *the dodo is extinct* and *dodos are extinct* both say something about a multiplicity of animals, and the difference in number

---

[18] The relation between the kind interpretation and its linguistic manifestations is the subject of ongoing debate (see, among a rich literature, Carlson 1977; Carlson and Pelletier 1995; Chierchia 1998*b*; Moltmann 2003; Zamparelli 2003).

value on the subject certainly does not correlate with a collective or distributive reading of the predicate, given that *extinct* cannot apply to individuals. Ojeda (1993: 24, 33) proposes reducing kinds to mereological sums, that is, to all logically possible aggregates of basic elements in the denotation of a term. But to talk about a kind is not to talk about the sum (or a sum) of its instances: the plural individual formed by the sum of all lions has properties that cannot be predicated of the kind *the lion*, and vice versa (see especially Pelletier and Schubert 1989):

(4.25)   a. the sum of all lions   has many individual members
         a'. the sum of all lions   weighs thousands of tons
         b. *the sum of all lions   has a bushy tail
         b'. *the sum of all lions   is a mammal

(4.26)   a. *the lion   has many individual members
         a'. *the lion   weighs thousands of tons
         b. the lion   has a bushy tail
         b'. the lion   is a mammal

In this sense, kinds refer neither to single individuals nor to their collections. Consider now the following pairs:

(4.27)   a. *the bear is numerous here
         b. bears are numerous here

(4.28)   a. *the bear used to be killed in its hundreds here
         b. bears used to be killed in their hundreds here

(4.29)   a. *the bear outnumbers man in this valley
         b. bears outnumber humans in this valley

(4.30)   a. *the expedition will count and tag the bear
         b. the expedition will count and tag bears

These are all instances of kind-predication; kinds are beyond the singular–plural opposition; and yet, the choice of one grammatical number value over the other decides the acceptability of these sentences. This is not a question of morphosyntactic agreement, as it would be if *the bear* was in construction with the syntactically plural *a few* or *many*. What matters is instead the semantic relation between the subject's number and the predicate.

These contrasts confirm the hypothesis, proposed by Dayal (1992) and developed by Chierchia (1998*b*), that kinds can be semantically singular or plural. The difference is a matter of part structure: although both *the bear* and *bears* refer to a collective entity, the plural alone conceptualizes its constituent

elements as discrete wholes, which predicates like *to count* can refer to.[19] Viewed as plural, a kind has the internal structure given it by the multiplicity of its instances. When plural in this sense, a kind is still a kind, as opposed to a simple plurality of individuals, and it admits predicates that can only be true of kinds and not of their instantiations (*bears are extinct*, *bears are rare*). But it is conceptualized as having a part structure made up of distinct wholes accessible to semantic interpretation (following the approach of Moltmann 1997), which is not the case when we refer to a kind as internally undifferentiated (effectively a mass, as Chierchia 1998*b* showed). Importantly, the articulation into discrete wholes depends on the lexical semantics on the nouns involved and is not an automatic consequence of grammatical plurality: *ashes* or *looks* are as plural as *bears* is in (4.27), but **ashes are numerous* or **his good looks are numerous* are clearly deviant. We must conclude that kind readings are available as a matter of grammar, in different guises across languages (cf. Krifka 1995; Chierchia 1998*b*), while the part-structure conceptualization depends on the lexical item, as well as on the particular interpretation.[20] The use of plural for kinds thus relates grammatical plurality, lexical semantics, and part-structure conceptualization.

4.5.3 *Kind interpretation and number morphology*

Besides affecting the use of grammatical number, the way kinds are conceptualized may affect its morphology as well. Consider an example from Turkana (cf. Section 3.5.3). In this language, all nouns have a gender prefix, which also marks number and case, with two exceptions: nouns occurring as bound stems in compounds, like *mosiŋ* in (4.31c), and nouns used as animal names in folk-tale contexts, like *ŋatùŋ* 'lion' and *kalèes* 'ostrich' in (4.32c):[21]

(4.31)   a. e-wur̀   'smell, scent'   (Turkana; Dimmendaal 1983: 221)
        b. a-mosiŋ   'rhinoceros'
        c. e-wur mosiŋ̀   'kind of tree'

(4.32)   a. ɛ-ŋatùŋ   'lion'
        b. ɛ-kalèes   'ostrich'

---

[19] Chierchia (1998*b*: 381 2) makes much the same point when he says that singular kinds as *the bear* are mass. In his account, however, this is due to a mass operator composed with (and distinct from) the definite determiner. Since nothing prevents the definite determiner from composing with nouns that are already mass, without the need of an extra mass operator, his account predicts, I believe, that a mass term like *fire* can be turned into a singular kind, yielding the incorrect **the fire*.

[20] I must therefore disagree with the characterization of the unacceptable **the bear is numerous* as a 'morphological' matter and not 'conceptual semantic', advanced by Krifka *et al.* (1995: 91).

[21] I have simplified Dimmendaal's gloss to (4.32c), merging in *ùne* the imperfective aspectual marker *e* and the 'ventive' *ùn*, which is a deictic marking direction towards the speaker.

c. ɛ-à-yò-ùne-tè   kɔ̀lɔ̀ŋ̀   ŋi-rùkutosì ŋɪ-àrèy, ŋatùŋ
    3-PAST-live-ASP-PL   long.ago PL-friends   PL-two,   lion
    kà   kalèes
    and   ostrich
    'long ago there were two friends, lion and ostrich'

In the compound 'rhinoceros-scented', the noun for 'rhinoceros' lacks its feminine prefix *a-*, and 'lion' and 'ostrich' used as individual-referring names lack the masculine *ɛ-*. The two exceptions clearly correspond to contexts where the nominal stem refers to the kind as a whole, not to particular instances. In addition to these exceptions, the kind reading can emerge with its own morphology. Dimmendaal (1983: 240) reports that a few nouns add a third form to the customary singular–plural contrast, which may be an additional singular or an additional plural depending on the noun.

(4.33) a. a-ɪtɛ̀   'cow'   (Turkana)
            ŋa-àtuk`   'cows'
            ŋa-àtùk-o   'flocks, cows in general'

        b. ɪ-tʋanì   'person'
            ŋɪ-tʋŋa`   'persons'
            ŋɪ-tʋŋà-si-ney   'mankind'

(4.34) a. a-kwàp-ɪt   'a piece of land'
            a-kwap`   'land, country'
            ŋa-kwàp-ɪn`   'countries'

        b. a-kʋàr-ɪt   'one night'
            a-kʋar-ì   'night'
            ŋa-kʋàrɪ-sìà   'nights'

The special (suffixed) plurals in (4.33) have 'a special collective meaning', and the special (likewise suffixed) singulars in (4.34) have 'a special individualizing meaning' (Dimmendaal 1983: 240). The glosses, reproduced from Dimmendaal, strongly suggest that the additional plural is a generalizing form, referring not to sets of individuals of arbitrary size (above one), but to the whole kind. The suffixed singular, by contrast, is the true counterpart of the plural form, thus pushing the unsuffixed singular into the role of a generic description of the kind.

In other cases, the opposition between kind- and object-level interpretation makes use of derivational morphology. Akkadian provides a convenient illustration. In this language, masculine nouns form the plural with the ending *-uu*, which replaces the singular *-um*, as in *il-um* 'god' ∼ *il-uu* 'gods'. The plural of some nouns has however the form *-aanu/-aanuu*, which is traditionally viewed

as expressing a concrete and countable reading, while the form in -*uu* has a generic interpretation (Goetze 1946; von Soden 1969: 77). As Buccellati (1976) pointed out, the plural ending -*aanu(u)* is none other than the regular -*uu* preceded by -*aan*-, independently recognized as a nominalizing affix (von Soden 1969: 70). This derivational affix has an individualizing function and brings about a specific reading: *naadin-aan-um* is not just a 'vendor', but a vendor in a given situation, or a vendor of the cited article; *šarraaq-aan-um* is 'the individual characterized by a particular act of stealing' (Buccellati 1976: 12). The interpretation of denumerable plurality simply follows from the particularizing function of this affix. In this way, Akkadian nicely illustrates how the morphological expression of lexicalized plural readings can arise through the interplay of inflection and derivation.

4.5.4 *Property instances* vs. *entities*

Finally, viewing plural conceptualizations from the angle of how a property is instantiated opens up a revealing perspective on the central semantic problem posed by lexical plurals: how a noun can possibly denote a plurality if it does not specify what it is a plurality of. To bring the problem into focus, consider the 'whatever reading'. It may seem that describing the meaning of *belongings* as 'whatever makes up someone's property', or *foundations* as 'whatever acts as basis for a construction' is simply too vague. But, for a start, the vagueness is a fact: exactly what entities make up the denotation of *belongings*, *furnishings*, *foundations*, and other nouns like those discussed in Section 4.4.2, must be gleaned from the context and from world knowledge—and semantic descriptions must reflect this fact. What is essential, however, is not the lack in descriptive content. *Thing* also lacks descriptive content (what is not a thing?), yet *furnishings* lacks the whole properties associated with the count plural *things*, and cannot therefore support predicates that refer to the individuals in its denotation, as in **these furnishings are numerous* or **these furnishings are different from each other*. In exactly the same fashion, the unrestricted *everything* differs from *every thing*, which restricts the admissible values for the variable to those falling under the sortal *thing*. Over and above the lack in descriptive content, therefore, these lexical plurals do not even define the elements of their reference as having unity or identity. How can they be different from a continuous unstructured mass?

The notion of instantiation comes into play here, because it allows us to talk about instances of a property without further defining what entities instantiate that property. The relevant concept is that of trope, as the philosophical

literature calls a concrete instance of a property, as opposed both to the property and to the entity that instantiates it: the sphericity of a particular ball, for instance, as opposed to sphericity, which is a property, and to the ball itself, which is a spatiotemporally situated object characterized by the property of being spherical, but distinct from it (see Bacon 1995; Lowe 1998: 78–83; McDonald 1998: 329–50). The particular sphericity instantiated by a certain ball, for instance, is what ceases to exist when the ball loses its spherical shape; neither the object not the universal property cease to exist for that, but only that particular instantiation of that property (the example comes from Lowe 1998: 78). Moltmann (2004) has made use of this concept to elucidate the semantics of abstract nominalizations like *wisdom*. As she notes, such terms display the interpretation of names of kinds in contexts like *wisdom is rare to come by* or *I have experienced wisdom*, where what is infrequent or what I have experienced is not the abstract individual concept of wisdom, but instances of the property of being wise. Likewise, *I have often met generosity* means that I have met frequent instances of generous people, and so covertly quantifies over the bearers of the named property. Moltmann therefore proposes that such terms name in fact kinds, but kinds with a special ontological status: '*wisdom*, I will argue, stands for what I will call a kind of trope, a universal whose instances are concrete property manifestations, but which does not have the status of an object' (p. 1).

The key point is that the instances themselves are undefined, yet they can be conceptualized as many. Consider Moltmann's comments on the following example:

(4.35)   a. John admires Mary's beauty
         b. ??John admires Mary's being beautiful

John can admire Mary's beauty, but not Mary's being beautiful, because Mary's beauty refers to the various features of Mary that make her beautiful, whereas Mary's being beautiful refers to the simple state characterized by the property of being beautiful holding of Mary. (Moltmann 2004: 11).

The 'various features' that make someone beautiful are just whatever constitutes a particular instance of beauty. The plural *features* in this informal characterization does not so much express a denumerable collection of features (feature *a*, feature *b*, . . .) as bring out the fact that the property results from a complex of overlapping instantiated properties—in this case, from all those that make up Mary's appearance. The property instances cannot be counted, as one cannot exhaustively enumerate the properties holding of an object: to count, we need a standard of unity, and a description as vague as 'whatever property is instantiated' provides no

such measure.[22] Tropes, in sum, can be many, but they cannot be identified nor counted. This makes them excellent candidates for the role of basic elements of the denotation in the 'whatever reading', where the noun only expresses a property holding of the overall denotation and not of its single elements. To take another example, *depths* denotes 'instances of the property of being deep', where the nature of the property implies that its instances have spatial extension; as Greene (1974: 192–3) noted in a totally different context, 'there is no great difference between *deeps* and *depths*'. More broadly, tropes fit the intuitive idea of non-identifiable entities which other lexical plurals also seem to refer to, like some of the 'collective' Arabic plurals to be considered in Chapter 7 below.

These somewhat speculative remarks do not amount to a semantic analysis; but if they are on the right track, a semantic analysis for non-aggregative plurals cannot merely distinguish atomic individuals from infinitely divisible portions of mass. Readings like 'instances of the property *P*' cannot be modelled on the basis of individuals (grouped into sets or mereological sums), because individuals are entities and can be identified, while instances are not entities and have no distinctive identity. No distinct sets can be defined on the basis of such nameless instances, because a set is entirely determined by the identity of its members, and that is precisely what tropes fail to express. The same applies to mereological sums. And yet the intuition is clear that morphosyntactic plurals like *depths* or *waters* or those with the 'whatever reading' are indeed semantically plural in some sense, as distinct from singulars and continuous masses. Tropes give us a way to model this intuition, accounting for its vagueness on principled grounds. Plurals with this sort of reading therefore bolster the case for tropes, as distinct from entities, as part of the denotation of linguistic expressions.

4.6 Conclusion

This chapter has offered a semantic typology of lexical plurals by identifying the basic concepts underpinning the main interpretations. What it has not attempted to do was to account for (some of) the attested readings compositionally, relating the meaning of morphemes with each other through functional application. The problem is not how to arrive at a formula like, for instance, $\lambda x[\mathbf{belong}\,(x)\;\&\;\text{not-Atom}\,(x)]$ ('being an x such that x belongs and x is not an atom') from the grammatical structure underlying *belongings*, but how to qualify the domain over which x takes its reference. Two extreme

---

[22] Cf. also Lowe's (1998: 79 82) arguments to the effect that tropes (which he calls 'modes') have no identity criterion.

positions are possible, both unsatisfactory: to force all non-canonical plurals into the procrustean bed of continuous homogeneous masses, or to pretend that they denote aggregates of atomic entities, like every other plural. The first alternative is unsatisfactory because it blinds us to the special manifold conceptualization encoded by plurality on a great many plural mass terms, and because it wrongly predicts that terms like *water* and *waters* are interchangeable (both being mass). The second alternative is closer to the mark, but it gives us no way to move forward beyond the statement that these terms denote a plurality. Why, then, can we not specify the atoms of their denotation? Do all such plurals structure their denotation in the same way? What does the lexical predicate characterize, if not the atoms? In addition, playing down the peculiarities of lexical plurals may lead to the wrong empirical prediction that plural mass nouns do not exist. Taking for granted what *x* stands for, in short, amounts to giving up on an analysis of lexical plurals—which probably explains why they have never been described as a coherent semantic-morphological class.

Since the nub of the problem is determining how these plurals characterize what they are true of, I have focused on conceptualization, and specifically on the way denotation is structured into parts. The foundational concepts are to be found outside linguistic semantics; metaphysics and formal ontology provide the notions of unity and identity, which in tandem (as distinct concepts) clarify a great deal of lexicalized plural readings. In some cases, the denotation is conceptualized as a plurality of parts lacking unity, but this characterization holds to variable degrees and is tightly linked to lexical semantics and encyclopaedic knowledge, typically having to do with cohesion among the parts, and sometimes with the salience of perceptual boundaries. Parts conceptualized as fragments rather than wholes also lack individual identity, but there are cases where the denotation is clearly structured into discrete atoms, and yet these atoms cannot be distinguished from each other: units of measurement, which have no individual identity on logical grounds, but also entities like eggs or small insects. Reference to the relation of instantiation links together several apparently distinct types of lexical plurals; some denote concrete manifestations of abstract notions or masses, while others just name a property and denote 'whatever' instantiates that property at a given situation, without giving a sortal characterization of the elements in the denotation. The vagueness and lack of countability of such unspecified plurals is expected if they are true of tropes, or property instances, as opposed to entities. The lack of identity would then descend from the ontological category of the denotation.

While not providing a formal analysis, this chapter has brought out the linguistic facts which distinguish lexical plurals from regular aggregates and

masses, and it has provided the main conceptual tools to categorize the various readings cross-linguistically attested. This completes the typology of lexical plurals begun in Chapter 2. Having delimited and clarified the ways in which plurality can be part of a noun, we now turn to analysing how the general concepts discussed in this typology emerge in specific instances.

Part II
Four Case Studies

This page intentionally left blank

5

Italian irregular plurals in -*a*

5.1 Introduction

Our first case study involves a small class of Italian irregular nouns featuring a mix of morphological and semantic idiosyncrasies, the most obvious being that they agree as masculines in the singular but as feminines in the plural. These plurals display several of the morphological and semantic properties we have seen in Part I as typical of lexicalized plurals.

The origin of these irregular forms ultimately goes back to the generalized reassignment of Latin neuters to either of the two genders of Italoromance.[1] In synchronic terms they make up a homogeneous morphological class, defined by an idiosyncratic ending and by a gender value that is idiosyncratic in so far as it is opposed to that of their respective singulars. This class, though small, comprises some highly frequent nouns, several of which have a distinctive interpetation. Perhaps unsurprisingly, this class is usually taken to represent a marked, exceptional pattern in the inflectional morphology of nouns—but still a pattern, although one that does not fit in the rest of the inflectional system.[2] I think the most candid statement about plurals in -*a* is also the most accurate: 'While their existence is undeniable, attempting to explain them may be more trouble than it's worth' (Riente 2000: 32, referring both to -*a* plurals and to the gender-changing *eco* 'echo').

I will argue instead that plurals in -*a* do not belong to the inflectional system at all, and not just for the embarrassment they cause there. My proposal is that they are lexical plurals: distinct, inherently plural nouns, related to the base noun by a word-formation process. The feminine gender, the special ending and the idiosyncratic semantics are all aspects of this

---

[1] The fate of the Latin neuter in Romance is a classic theme in Romance historical linguistics; see Löfstedt (1928: 46 51), Spitzer (1941), Lausberg (1963: 26 33), Väänänen (1963: 107 11); for Italoromance, Rohlfs (1968: 59 63), Tekavčić (1972: 92 100), and Maiden (1995: 108 9).

[2] See Merlo (1952), Hall (1956), Rohlfs (1968: 35 6, §368), Bonfante (1973), Tekavčić (1972: 92 9), Regula and Jernej (1975: 104 6), Brunet (1978: 30 90), Santangelo (1981), Lepschy and Lepschy (1988: 110), Schwarze (1988: 9), Marcantonio and Pretto (1988: 326), Bach and Schmitt Jensen (1990: 117 21), Ojeda (1995), Maiden (1995: 103 4), Dressler and Thornton (1996: 16 18), Maiden and Robustelli (2000: 25 8), Riente (2000: 32 4), Acquaviva (2002), and Thornton (2003*b*: 212, 219).

derivation. After laying out the relevant facts in Section 5.2, I will present the reasons for a non-inflectional analysis in Sections 5.3 and 5.4, which focus on the morphosyntactic and semantic dimension respectively. The argument built in these two sections involves a number of phenomena characteristically associated with -*a* plurals, some never previously analysed, and includes an analysis of the differences with an apparently similar class of ambigenerics in Romanian. The concluding Section 5.5 spells out the proposal that -*a* plurals are the output of a (non-productive) derivational process which creates intrinsically plural lexemes.

5.2 Description

This section will present the main facts to be accounted for. It will introduce the relevant plurals, place them in the context of the Italian nominal inflection system, and outline their morphological and semantic peculiarities.

5.2.1 *The morphological irregularity of plurals in* -a

Italian nouns and adjectives are organized in a series of inflectional classes that can be pre-theoretically described as follows:

(5.1)   Canonical inflectional classes of Italian nouns:

Singular	Plural	Example
-a	-e	zia, zie 'aunt, aunts'
-a	-i	poeta, poeti 'poet, poets'
-o	-i	zio, zii 'uncle, uncles'
-e	-i	fiore, fiori 'flower, flowers'
invariable		città 'city, cities'

Invariable nouns may end in a vowel or, much more rarely, in a consonant (as in the borrowings *sport* or *bar*). All other nouns end in a vowel, according to the pattern set out in (5.1). Since nominals do not inflect for case, the category of a noun is completely determined by the endings for singular and plural, or by the single ending in the case of invariables.

   Gender correlates with the classes thus defined in a complex but systematic way (for details, see Dressler and Thornton 1996; Riente 2000; Thornton 2001, 2003*a*). Class I, -*a/-e*, contains only feminines; class II, -*a/-i*, only masculines (except *ala* ∼ *ali* 'wing ∼ wings' and *arma* ∼ *armi* 'weapon ∼ weapons', which are feminine); and class III, -*o/-i*, only masculines (except the feminine *mano* ∼ *mani* 'hand ∼ hands'). The remaining two classes include nouns of either gender.

The nouns we will consider in this chapter deviate from this paradigm. One example is the noun for 'finger': its masculine singular form is *dito*, but the plural is *dita*, and it is feminine:

(5.2)   il   tuo   dito
        the.MASC.SG your.MASC.SG   finger.SG

(5.3)   le   tue   dita
        the.FEM.PL   your.FEM.PL   finger.PL

The plural version of the masculine prenominal determiners *il tuo* is *i tuoi*, as for example in *i tuoi zii* 'your uncles' (literally 'the your uncles'). *Le tue* is instead the feminine plural version, as exemplified in *le tue zie* 'your aunts'. For all grammatical purposes, therefore, the plural *dita* is feminine, while the singular *dito* is masculine. That a noun should change its gender depending on grammatical number is odd enough; but the form of this plural noun is also irregular by itself, since *-a* is nowhere else an exponent of plural (many plural nouns end in *-a*, like *cinema*, but these are invariables and therefore *-a* is not an exponent of plurality).

If the irregular plural of *dito* were an isolated exception, it would not have a great significance. But an isolated exception it is not: at least twenty more nouns in common use, as well as a few more limited to regional, archaic, or idiomatic use, share the pattern with a masculine singular ending in *-o* and a feminine plural in *-a*. So perhaps we should recognize *-o/-a* as a marked variant of the *-o/-i* class, as in (5.4):

(5.4)   SINGULAR   PLURAL   EXAMPLE
        -o   -i   zio, zii 'uncle, uncles'
        (-o, masc. -a, fem.) dito, dita 'finger, fingers'

This is indeed the received view. With some variations, all references listed in note 2 (except Ojeda 1995 and Acquaviva 2002) state or imply that the class of *dito ~ dita* is an irregular inflectional class, a lexically restricted set of nouns with a marked plural. This chapter will argue against this view. To see why *-a* plurals are more than irregular inflectional forms, we must first of all consider in detail the members of this class.

5.2.2 *Overview of the class*

Among the nouns with a singular masculine in *-o* and an irregular feminine plural in *-a*, many have a concurrent regular plural, masculine and in *-i*. Unsurprisingly, the two plural forms tend to have distinct meanings; but just how distinct they are varies with the choice of the noun, as well as with the

idiolect. The lists that follow give an idea of the size of this group and the proportion of -*a* plurals with a regular alternant in -*i*; the details of course vary from one speaker to another. The plural doublets that can be reasonably ascribed to the active or passive vocabulary of all speakers of Italian are at least the following:

(5.5) braccio 'arm'   braccia 'arms'   bracci 'arms (of objects)'
       budello 'gut', 'alley' budella 'intestines'   budelli 'alleys'
       corno 'horn'   corna 'horns'   corni 'horns'
       fondamento 'basis'   fondamenta 'foundations'   fondamenti 'bases'
       membro 'member'   membra 'limbs'   membri 'members'
       muro 'wall'   mura 'walls (perimeter)'   muri 'walls'
       osso 'bone'   ossa 'bones'   ossi 'bones'

A second group comprises nouns that have a plural in -*a* but for which the regular alternant in -*i* is not uniformly available for all speakers and in all dialects and registers:[3]

(5.6) cervello 'brain'   cervella 'brains (mass)'   cervelli 'brains (organs)'
       ciglio 'eyelash/edge'   ciglia 'eyelashes'   cigli 'edges'
       ginocchio 'knee'   ginocchia 'knees'   ginocchi 'knees'
       labbro 'lip'   labbra 'lips'   labbri 'edges (of wound)'
       lenzuolo 'sheet'   lenzuola 'sheets, bed linen'   lenzuoli 'sheets'
       sopracciglio 'eyebrow' sopracciglia 'eyebrows'   sopraccigli 'eyebrows'
       urlo 'shout'   urla 'shouts'   urli 'shouts'

Another group lacks a regular -*i* plural altogether. The details depend on individual usage, but it seems fair to separate *dito* and *grido*, which have the regular plurals *diti* and *gridi* in some varieties (usually as stigmatized forms), from the nouns that pluralize exclusively in -*a*:

(5.7) dito 'finger'   dita 'fingers'   %diti
       grido 'shout'   grida 'shouts'   %gridi

(5.8) centinaio 'hundred'   centinaia 'hundreds'   *centinai
       midollo 'marrow'   midolla 'marrows (human)'   *midolli

---

[3] The doublet *frutti* ∼ *frutta* 'fruits', plural of *frutto*, also belongs here; however the feminine plural *le frutta* is exceedingly rare in contemporary Italian, having being replaced by the singular mass *la frutta* 'fruit, fruitstuff', of the regular feminine class a/ e. For this reason, I will leave it out in what follows.

migliaio 'thousand'	migliaia 'thousands'	*migliai
miglio 'mile'	miglia 'miles'	*migli
paio 'pair'	paia 'pairs'	*paii
riso 'laughter'	risa 'peals of laughter'	*risi
strido 'shout'	strida 'shouts'	*stridi
uovo 'egg'	uova 'eggs'	*uovi

In addition, three plurals in -*a* lack both a singular and a regular plural:

(5.9) gesta 'deeds' interiora 'entrails' vestigia 'relics'

Actually, the singulars *gesto*, *interiore*, and *vestigio* do figure in the vocabulary of Italian, each with a regular plural in -*i*; but their meaning and that of the plural forms in (5.9) are too distant to be considered grammatical variants of the same lexical entry. The singular *gesto* unambiguously means 'gesture', never 'deed'. *Interiore* is in fact an adjective, meaning 'inner', and it has no lexicalized nominalization in the singular.[4] As for the singular *vestigio* 'trace, footstep', it is obsolete in modern Italian.

Finally, there are four nouns with a perfectly regular singular–plural paradigm, supplemented by -*a* plurals that have drifted semantically and have become fossilized. The -*a* plurals of *cuoio* 'hide', *fuso* 'spindle', and the literary *calcagno* 'heel' are only used in idioms:

(5.10) alle calcagna (di) 'on the heels (of)'
        tirare le cuoia   'to die' (lit. to pull the hides)
        fare le fusa   'to purr' (lit. 'to make the spindles')

The additional plural of *filo* ∼ *fili* 'thread ∼ threads', namely *fila*, has instead a less fixed distribution, but it means 'ranks, main threads' (in a metaphorical sense) and tends to be used in phrases like *le fila della congiura* 'the threads of the plot', or *serrare le fila* 'to close ranks'.

The pattern involved many more nouns in the past (Santangelo 1981). However, earlier stages of Italian, or rather Tuscan, did not have the same nominal morphology as contemporary Italian, and for this reason it is important clearly to restrict the investigation to the synchronic status of -*a* plurals, without drawing synchronic conclusions on the basis of data from earlier stages.

---

[4] A synchronic relation with *interiora* is made even less likely by the fact that the singular form ends in *e*, not *o*, which would make this an unparalleled exception to the morphological pattern under consideration.

5.2.3 *Variation and acceptability*

The need for a careful appreciation of what counts as a synchronic pattern is nowhere more apparent than in judging acceptability. Traditional grammars like that of Regula and Jernej (1975: 104–6) erred on the side of too little information on the status of each plural variant, in that they merely presented a list of plurals in -*a*, often with no indication of the historical and stylistic diversity; but they did not attempt to account for speakers' judgements of acceptability. More ambitious accounts, like that of Ojeda (1995), aim at a precise characterization of the semantics of the irregular -*a* ending; in so doing, it is easy to sneak in an unwarranted deterministic relation between a certain meaning and a certain form (admittedly, a relation often suggested by descriptive grammars). For instance, Ojeda (1995: 222) illustrates the meaning of the regular *lenzuoli* ('sheets') '…when a "multitude" of sheets is taken "one by one"' with the aid of the following example (from Brunet 1978: 62):

(5.11)   Mafalda stendeva i lenzuoli sul terrazzino
          'Mafalda was hanging out the sheets on the little balcony'

It is true that *lenzuoli* fits in this singulative context; but it is not true that speakers of Italian who have the plural doublet in their vocabulary deterministically prefer the regular *lenzuoli* to the irregular *lenzuola*. Some do, some plainly do not, and some accept either. In my idiolect, *lenzuola* is the preferred plural overall, but the following context (suggested by Giulio Lepschy) makes *lenzuoli* strongly preferable:

(5.12)   un uomo coperto di lenzuoli/??lenzuola
          'a man covered with sheets'   (said of a Ku Klux Klan member)

This shows that even people who regularly use only one member of the plural doublet *lenzuoli* ∼ *lenzuola* may have intuitions about both forms. Specifically, my judgement (and that of others) contradicts the deterministic relation between form and meaning implied by Ojeda: a singulative context like (5.11) is not enough to rule out *lenzuola*; it takes something like (5.12), and even there, it is a matter of preference, not of downright unacceptability. This is the kind of datum a theory should explain. Consider also the following:

(5.13)   a. Maria-Nunziata…si stringeva la sottana tra i ginocchi
             'Maria-Nunziata…was clutching her petticoat between her knees'

         b. anche lei, a collo inclinato, con le mani tra i ginocchi
             'she too, her neck bent, with her hands between her knees'

Ojeda (1995: 216) comments: 'As Rocchetti (1968) pointed out in relation to [5.13a], these sentences would simply make no sense with *ginocchia*.' This is just not true. Even for speakers who feel a definite contrast between collective *ginocchia* and singulative *ginocchi*, these sentences do not sound degraded with *ginocchia*, much less nonsensical. Clear contrasts in acceptability do indeed arise in other contexts, but not here, so that positing too clear-cut a semantic opposition between collective and not-collective sense may well lead to factually incorrect descriptions. The observed semantic tendencies should not be linked to the irregular morphology in too deterministic a fashion; final *-a* does not mean 'collective'in this sense.

5.2.4 *Recapitulation*

Summing up, the following features define *-a* plurals as an irregular class:

   (i) they change gender in the plural;
  (ii) their ending is nowhere else in the language an exponent for plurality;
 (iii) some of them coexist with a regular plural;
  (iv) they are restricted: categorially, to nouns alone, and lexically, to a small group of around twenty nouns;
   (v) four of them (*calcagna*, *cuoia*, *fila*, *fusa*) occur only in listed phrases with non-compositional interpretation;
  (vi) they are all inanimates (coming from Latin neuters)
 (vii) they display semantically homogeneous subgroupings: body parts, measurement units, and mass concepts.

A satisfactory analysis should also account for a fact that descriptions tend to play down:

 (viii) the choice between regular and irregular plural in the same context varies significantly, and even in the same idiolect choosing one often does not entail that the other is unacceptable.

Properties (i)–(iv) are morphological in nature, (v) concerns the relation between morphology and meaning, (vi)–(vii) have to do with meaning and (viii) calls into play the status of these marked plurals in a speaker's competence. When analysed closely, these properties call into question the traditional view that plurals in *-a* are marked exponents of an inflectional class. Let us start by considering the morphosyntactic evidence.

5.3 The morphological evidence

We will now consider in detail some morphological and morphosyntactic evidence suggesting that pairs like *dito* ∼ *dita* do not have the properties of nouns that in other Romance languages constitute a coherent paradigm. In particular, it will be instructive to compare Italian and Romanian, which has a class of ambigeneric nouns displaying the same gender–number correlation as *dito* ∼ *dita*.

5.3.1 *The exceptionality of gender inversion*

In the inflectional system of Italian, like more generally in Indo-European, the categories of gender and number classify nouns in two very different fashions. Number defines an opposition that applies to most nouns and has a predictable semantic value, as evidenced by the possibility to construct and interpret the plural of invented nouns (as in the *wug* test: *this is a wug, these are*   ). By contrast, cases where the same nominal stem occurs in both genders are a minority, and the morphological opposition in this case reflects one of several possible semantic oppositions, ranging from sex-based pairs like *zio* 'uncle' ∼ *zia* 'aunt' to encyclopaedic oppositions like *mela* 'apple' ∼ *melo* 'apple-tree', all the way to pairs that only differ in details of the conceptualization, like *buco* 'hole (bi-dimensional)' ∼ *buca* 'hole (tri-dimensional)', or purely stylistic variants like *orecchio* ∼ *orecchia* 'ear'. The meaning of the lexical base co-varies with the choice of gender, so that we can say that a noun keeps its gender value fixed across the inflectional paradigm defined by number. In this situation (to be contrasted for instance with cases of gender polarity; cf. Sections 3.7.3 and 4.3.3.2 in the previous chapters), forms like *zio* and *zia* must be viewed as distinct nouns, not as the same noun with alternative gender values (see the classic discussion in Matthews 1974: 45–7).[5]

Given this relation between gender and number in the language, the received view according to which the masculine *dito* and the feminine *dita* are respectively the singular and the plural of the same noun ascribes to them a very exceptional status. Subordinating the choice of gender to that of number would effectively run counter to the architecture of Italian nominal morphology and syntax. To fully appreciate this point, consider the characteristics of those very few nouns other than those with *-a* plurals which 'change gender in the plural'.

---

[5] This brief reminder of the different status of gender and number may appear superfluous; it is necessary for those who deny that inflection and derivation are distinct even in an abstract sense. Marantz (2003), for example, claims that the existence of many nouns without singular or plural shows that an inflectional category like number allows for gaps just like derivational categories. But the lack of one number form is restricted enough to warrant discussion, while having only one gender value is the norm for nouns.

With the usual provisos about idiolectal variability, the nouns that are exceptional in this way are limited to the following, including archaic ones:

(5.14)   Italian nouns with different gender values in singular
         and plural

Singular	Plural	Gloss
carcere (masc.)	carceri (fem.)	'prison'
eco (fem.)	echi (masc.)	'echo'
gregge (masc.)	greggi (fem.)	'sheep herd'
rene (masc.)	reni (masc./fem.)	'kidney'

In each case, the relation between singular and plural is not only morphologically irregular, but also semantically idiosyncratic. *Carcere* 'prison' has an abstract reading that admits no pluralization, and in the concrete reading ('building with the function of prison') both the singular and possibly the plural may refer to a single jail building. The meaning of 'echo' makes the plural *echi* more easily conceived as a manifold 'echoing' than as a discrete collection of denumerable sounds, just like a plurality of shouts (*grida*, *strida*) or of peals of laughter (*risa*). Likewise, the plural *greggi* is typically used to refer to masses of sheep, as opposed to a denumerable plurality of discrete herds (a bit like *crowds*; cf. Section 4.5.2). *Reni* refers to a two-membered body part on a par with several -*a* plurals, and the semantic differentiation between the regular masculine plural and the (by now archaic) irregular feminine centres on cohesion, just like -*a* plurals of body-part nouns: the feminine *le reni* is restricted to humans and conceptualizes kidneys as a whole manifold (twin) organ, as opposed to a collection of individual objects. Moreover, at least in my judgement, none of these plurals can be modified by a numeral, and all occur more comfortably with the definite article than with any other determiner, just like mass plurals such as *the rains*. In one case the relation between singular and plural is also formally idiosyncratic. The final -*o* of *eco* is not a regular ending for feminine nouns, so that it is plausible to assume that the singular form is listed rather than being the output of a regular rule. If it were also feminine in the plural, *eco* would be the only other example apart from *mano* 'hand' of a feminine noun with singular in -*o* and plural in -*i*.[6] The fact that the plural *echi* has masculine gender avoids this major irregularity, at the price of making the relation between singular and plural much looser than it normally is: the singular must be listed for its form, and the plural must be listed in the sense

---

[6] What is exceptional is not for feminines to end in -*o* in the singular, but for such forms to have a plural in -*i* which is likewise feminine. Clipped feminine forms like *auto* 'car/cars' (from *automobile*) are invariable; so is the learned loanword *parodo* 'parodos' (a sung intermezzo in Greek tragedies), which retains the feminine gender and the -*o* vocalism of its classical Greek antecedent.

that it is not predictable from the singular; that is to say, it does not fall into any of the inflectional classes schematized in (5.3), exactly as for -*a* plurals.

All in all, these cases reinforce the impression that a plural with a different gender value with respect to the singular must be somewhat different from a regular plural. If anything, they emphasize the peculiarity of -*a* plurals in the inflectional system.

5.3.2 *The lack of a syntactic 'neuter' in Italian*

The fact that the class of -*a* plurals comprises nouns but no adjectives or pronouns, by itself, is not an argument for the non-inflectional status of this type of pluralization. It is amply known that, in Corbett's terminology (1991, 2000), agreement controllers like nouns may instantiate more inflectional classes than targets. Still, this characteristic distinguishes Italian -*a* plurals from other phenomena which, within Romance, are often regarded as parallel because of the common diachronic origin.

First of all, we should distinguish Italian -*a* plurals from the so-called 'neuter' of Ibero- and Rhaeto-Romance. In Spanish, a definite article governing nominalized adjectives takes the form *lo*, distinct from the masculine *el* and from the feminine *la*:

(5.15) lo   curioso   de esa situación (Spanish; Corbett 1991: 215)
         the.NEUT curious.MASC of this situation
         'the curious thing about this situation'

Contrasting at the same time with the masculine and with the feminine form of the definite article, *lo* is indeed outside the gender opposition; the same applies to the demonstrative *eso* 'this'. The existence of special inanimate pronouns and determiners, by itself, is nothing unusual (cf. the Italian *ciò* 'this thing', another relic of Latin neuter); in any case, this phenomenon is restricted to determiners in contexts lacking a head noun, as opposed to lexical nouns like the Italian -*a* plurals.

In other Romance varieties, a mass or kind reading has a morphological reflex not limited to determiners. In some Asturian dialects of Ibero-Romance, nouns with mass interpretation trigger masculine agreement on a following adjective, even if the noun itself has a feminine ending and otherwise triggers feminine agreement on a pre-nominal modifier (see Windisch 1973: 76–83 and the literature there cited; cf. also Ojeda 1992*a* and Corbett 2000: 124–5). This leads to mismatches like *pera maduro* for 'ripe pear-stuff' in (5.16b), contrasting with *pera madura* for the object count reading in (5.16a):

(5.16) a. nun   comí   más   qu'una   (Asturian; ALA 2000: 115)
           neg   eat.1.SG.PAST   more   than a.FEM
           pear.FEM ripe.FEM
           pera   madura
           'I only ate a ripe pear'

       b. llegó   un   camión cargáu de
           arrive.3.SG.PAST a. MASC truck   laden   of
           pera   maduro
           pear.FEM ripe.MASC
           'a truck arrived laden with ripe pears'

This amounts to a third agreement class (gender value) alongside masculine and feminine. Even more significant is the presence of a third ending beside masculine and feminine for a set of pronominals and for adjectives that distinguish masculine from feminine:

(5.17) bunu   buna   buno   (Asturian; Windisch 1973: 78)
        good.MASC good.FEM good.NEUT

Similar facts appear in a number of dialects in the centre and south of the Italian peninsula (Rohlfs 1966: 185, §145, 1968: 108–10, § 409; cf. also Maiden 1991: 177–9; Haase 2000; Bigalke 1996: 33–4). In other cases, the mass interpretation determines special endings, occasionally leading to minimal pairs like *lo peššo* 'the fish (substance)' ∼ *lu peššu* 'the fish (object)', (dialect of Servigliano, Rohlfs 1968: 109, § 419).[7] It must be noted, however, that such formal alternants do not make up a separate morphological paradigm for mass nouns (cf. Maiden 1991: 178). Rhaeto-Romance features a slightly different type of neuter. Haiman (1988: 364–6, 382–3) reports that the dialects of Switzerland have a special class of non-referential expletive subjects, illustrated by *iλ* in (18):

(5.18) iλ   ej   veniw   priw   numerusas mesiras
        it.NEUT is   come   taken.NEUT numerous   measures
                                          (Surselvan; Haiman 1988: 383)
        'there were many measures taken'

The Surselvan Romansh *iλ* (and its positional allomorphs *ej* and *i*) is formally distinct from the masculine *el* and from the feminine *ela*. Moreover, this expletive subject (along a few other 'neuter pronouns'; see Haiman 1988: 385)

---

[7] This quick overview of Romance neuter survivals simplifies a complex domain. Giammarco (1979: 129) reports that some Italoromance Abruzzese dialects actually distinguish *two* plural neuter endings, one derived from Latin *a* with a 'collective' reading, and one derived from *ora* for countable aggregates. This parallels the Breton facts in Section 8.4 in Chapter 8.

triggers a distinctive zero ending on agreeing predicative adjectives, thus defining a match between a closed class of words and a distinctive agreement pattern.

Italian *-a* plurals differ from all these survivals of the neuter in two important respects. Morphosyntactically, the ending *-a* is strictly limited to lexical nouns, that is to a small group of agreement controllers, while the main evidence for neuter typically rests on a special form assumed by pronouns, determiners, and adjectives—that is, by agreement targets rather than controllers. But even where a morphologically 'neuter' ending characterizes lexical nouns, as in *pešǒ* 'fish', what makes these constructions neither fully masculine nor fully feminine is their agreement pattern, revealed primarily by agreeing modifiers. By contrast, a form like *dita* 'fingers' is syntactically fully equivalent to any other feminine plural. The only reasons to see it a member of a third gender (as argued by Bonfante 1973) have to do with its paradigmatic irregularities, namely the *-a* ending in the plural and the gender mismatch between singular and plural. This is very different from the syntagmatic irregularities of 'neuter' structures like (5.16).

Second, *-a* plurals differ sharply from the interpretation of these neuter survivals. They have nothing to do with the nominalization of adjectives, as in Spanish *lo hermoso* 'the beautiful'; and they are certainly not restricted to the expression of mass. Although some *-a* plurals cannot be counted, like *cervella* 'brains', others are countable, like *braccia* 'arms', and the lack of countability is neither an exclusive nor even a predominant trait for this type of plurals (contrary to Marcantonio and Pretto 1988: 326). What is more, the neuter as a distinct morphological category is not generically mass, but refers more specifically to a (natural) kind; recall *lo pešǒ* '[the] fish (as a natural kind)' ∼ *lu pešǔ* '[the] fish (as an entity)'. This is not the meaning of *-a* plurals, as is shown by contrasts like the following:

(5.19) a. un pettine di corno
            'a comb of horn' (i.e. made of horn)

        b. *un pettine di corna/corni
            'a comb of horns'

In sum, *-a* plurals are always restricted to lexical nouns and are irregular paradigmatically, not syntagmatically. Other Romance survivals of neuter typically appear on agreement targets, and involve peculiar syntagmatic patterns. The obvious historical relationship must not obscure the synchronic difference between these two sets of phenomena.

By contrast, the change of gender according to number represents a genuine connection between Italian *-a* plurals and a different type of 'neuters', to which we now turn.

5.3.3 *Differences between Italian and Romanian ambigeneric nouns*

The same gender–number correlation of Italian nouns with -*a* plurals also characterizes a much better known class of Romanian nouns.[8] As (5.20) shows, a noun like *pahar* 'glass' requires modifiers of the same form as those agreeing with the masculine *pom* 'tree' in the singular, and of the same form as those modifying the feminine *casă* 'house' in the plural:

(5.20)   *Romanian noun genders*

		MASCULINE	FEMININE
SG.		*un/*o pom* a.MASC/*FEM tree *un/*o pahar* a.MASC/*FEM glass	**un/o casă* a.*MASC/FEM house
PL.			**doi/două pahare* two *MASC / FEM glasses
		doi/*două pomi two.MASC/*FEM trees	**doi/două case* two.*MASC/FEM houses

In fact, we will now see that it is not the same phenomenon, and that the differences are instructive.

5.3.3.1 *Differences within the morphological system*   The first macroscopic difference lies in the extension of this phenomenon. The plurals in -*a* having at least some currency in contemporary Italian cannot be many more than those listed in (5.5)–(5.10) above, even for the most sophisticated speaker. In Romanian, on the other hand, the agreement pattern defining what is traditionally called either the neuter or the ambigeneric gender concerns hundreds of nouns. But sheer size is less important, in this regard, than productivity: while the Italian list has been steadily shrinking and cannot conceivably be enriched, Romanian neuter nouns include recent loanwords like *televizor* 'television set', pl. *televizoare*;

---

[8] On this topic, see Mallinson (1986: 244 6), Corbett (1991: 288), and especially the book length study of Windisch (1973). Farkas (1990) and Lumsden (1992) analyse the feature system of Romanian noun morphology in terms of feature underspecification; see most recently Kihm (2005) (whose analysis of plural as uniformly derivational I will not follow). Windisch (1973), Bonfante (1973), and Kihm (2005) also consider similar facts in Albanian (the former two mention a more distant parallel in Tocharian).

radio 'radio', pl. *radiouri*; Hall (1983: 195) cites *bombon* 'piece of candy, *bonbon*', pl. *bomboane*, and Mallinson (1986: 247) even mentions *inputurile şi outputurile* 'the inputs and the outputs'. We are clearly dealing with an open class, vital enough to provide a morphological template for the adaptation of loanwords.

There are also differences in the relation between agreement pattern and exponence. In Italian almost all nouns with different gender values in singular and plural have a plural in -*a* (the exceptions are discussed above in Section 5.3.1). In Romanian, by contrast, there is no one ending for all neuters (beside the canonical -*uri* and -*e*, some nouns have -*ii* like *seminar* ∼ *seminarii* 'seminar ∼ seminars', or zero like *nume* ∼ *nume* 'noun ∼ nouns');[9] and no ending is entirely and exclusively neuter: -*e* is both neuter and feminine, and even -*uri*, which is specifically associated with neuters, occasionally appears on feminine nouns (i.e. syntatically feminine in the singular as well as in the plural), like *carne* 'meat, flesh', pl. *cărni* or *cărnuri*; *vreme* 'time', pl. *vremuri*, or *iarbă* 'grass', pl. *ierburi*.[10] The indirect connection with exponents we observe in Romanian neuters, in sum, is the behaviour to be expected from a gender value understood as a systematic relation between exponents, grammatical functions and agreement; an inflectional category firmly established in the morphological system of the language, in stark contrast with the totally irregular -*a* plurals of Italian.

5.3.3.2 *Syntactic differences: Agreement*   The strongest indication that Romanian ambigenerics and Italian nouns with -*a* plurals are not the same type of class comes from agreement. In both languages, feminine plurals with a masculine singular have the same syntax as any feminine plural. But when two singular nouns are conjoined, they trigger *plural* on an agreeing pronoun or participle. What gender appears then on the plural agreeing element?

As is well known, gender resolution rules vary across languages (see Corbett 1991: 261–306), so that we must compare Italian and Romanian in this respect before we can test the behaviour of gender-switching nouns. The Italian pattern is simple enough: masculine is the default, and feminine only appears if all conjuncts are feminine.

---

[9] Different neuter endings may convey distinct interpretations, like *vise* 'dreams' and *visuri* 'ambitions' from the singular *vis*. This shows that neuter agreement by itself does not correlate with a particular reading (thanks to Irina Tarabac for pointing this out).

[10] Windisch (1973: 20 3) points out that when the *uri* ending pluralizes a noun whose singular is feminine, the resulting plural systematically has a mass or collective interpretation; his claim that these are lexicalized plurals without a proper singular coincides with my view of Italian *a* plurals. The Romanian class that really parallels Italian *a* plurals, then, is not constituted by the plurals of ambigeneric nouns, but by the much fewer plurals like *cărnuri*. The point is that the ending *uri* also extends to the large productive class of neuters, and therefore does not coincide with lexical plurals as *a* does in Italian.

(5.21) Gender resolution rules in Italian coordinated NP
    a. [masculine + masculine]   . . .   masculine
    b. [masculine + feminine]   . . .   masculine
    c. [feminine + masculine]   . . .   masculine
    d. [feminine + feminine]   . . .   feminine

(5.22)   a. [il divano e il tavolo] che sono stati acquistati/*state acquistate
        'the sofa.masc.sg and the table.masc.sg that were
        bought.masc.pl/*fem.pl

    b. [il divano e la poltrona] che sono stati acquistati/*state acquistate
        'the sofa.masc.sg and the armchair.fem.sg that were bought.
        masc.pl/*fem.pl

    c. [la poltrona e il divano] che sono stati acquistati/*state acquistate
        'the armchair.fem.sg and the sofa.masc.sg that were bought.
        masc.pl/*fem.pl

    d. [la poltrona e la sedia] che sono *stati acquistati/state acquistate
        'the armchair.fem.sg and the chair.f.sg that were bought.
        *masc.pl/fem.pl

The picture is considerably more complicated in Romanian. Corbett (1991: 288–90) shows that gender resolution is identical with Italian when the nouns refer to animate concepts; with inanimates, however (which include neuter/ambigeneric nouns), feminine is the default choice, overridden by masculine only if all conjuncts are masculine. I reproduce Corbett's examples involving inanimates:

(5.23) Gender resolution rules in Romanian coordinated NP
    a. [feminine + masculine]   . . .   feminine
    b. [masculine + neuter]   . . .   feminine
    c. [neuter + feminine]   . . .   feminine
    d. [masculine + masculine]   . . .   masculine
    e. [neuter + neuter]   . . .   feminine
    f. [feminine + feminine]   . . .   feminine

(5.24) a. uşa   şi   peretele . . .   ele
        door.fem.def   and wall.masc.def . . .   they.fem.pl

    b. peretele   şi   scaunul . . .   ele
        wall.masc.def   and chair.neut.def . . . they.fem.pl

    c. scaunul   şi   masa . . .   ele
        chair.neut.def   and table. fem.def . . .   they.fem.pl

    d. nucul   şi   prunul . . .   ei
        walnut-tree.masc.def and plum-tree.masc.def . . . they.masc.pl

e.	frigiderul	şi	televizorul ...	ele
	refrigerator.NEUT.DEF	and	television.NEUT.DEF ...	they. FEM.PL
f.	uşa	şi	masa ...	ele
	door. FEM.DEF	and	table.FEM.DEF ...	they. FEM.PL

As we can see, in Romanian a coordinated noun phrase containing one or more ambigeneric singular nouns requires the feminine on a long-distance agreeing element. This is plausibly related to the default nature of feminine in Romanian, evidenced by the pattern schematized in (5.23) as well as by uses of the feminine on pronouns whose antecedent has no gender, like *asta* in (5.25) from Farkas (1990: 541):

(5.25) Petru   e   acasă,   dar asta   nu   o   (Romanian)
       Petru   is   at.home, but that.FEM.SG not it.FEM.SG
       ştie   decît Maria
       knows   but   Maria
       'Petru is at home, but only Maria knows that'

Whatever the reasons for this behaviour, Romanian, ambigeneric nouns show unambiguously that the correlation 'feminine in the plural' holds even when no one noun is marked [plural]. The crucial example is (5.24e), where the larger coordinated NP triggers feminine plural agreement, proving that the whole complex NP has the agreeing properties of a neuter, like the singular nouns out of which it is made. All descriptions agree on this (cf. Windisch 1973: 41 and Mallinson 1986: 84), and the native speakers I could consult did not show the least trace of hesitation. The correlation 'feminine in the plural' is grammatically encoded as a property that can be passed on to a complex noun phrase.

Italian nouns with plural in -*a* are different. When a gender-changing noun like *dito* is conjoined with a regular masculine like *piede* 'foot', plural agreement is in the masculine, not the feminine (contrast (5.23b) and (5.24b) above):

(5.26)   [il dito e il piede] che sono stati amputati/*state amputate
          'the finger and the foot that have been amputated.MASC.PL/*FEM.PL'

When it is conjoined with a feminine like *gamba* 'leg', we might expect feminine plural agreement, because each of the two conjuncts, when pluralized, would require feminine agreement by itself. Instead, what we have is again masculine:

(5.27)   [il dito e la gamba] che sono stati amputati/*state amputate
          'the finger and the leg that have been amputated.MASC.PL/*FEM.PL'

These cases falls under the generalization (5.21b): if the conjuncts have different genders, the complex noun phrase triggers masculine agreement. The gender-switch specification of *dito*, so to speak, does not count. What is more, even when both conjuncts are gender-switching, the complex noun phrase triggers masculine plural, not feminine:

(5.28)   [il dito e il braccio] che sono stati amputati/*state amputate
         'the finger and the arm that have been amputated.MASC.PL/*FEM.PL'

There is no conflict to resolve here; each conjunct has the property of changing gender in the plural. Yet, the complex noun phrase does not inherit this property.

The different agreement patterns of Italian and Romanian ambigeneric nouns show that what makes them special has a different status in the grammar of the two languages. In Romanian, it has the common traits of an inflectional property: it affects a sizeable number of nouns, it has a range of distinct exponents, it is obligatory in that it does not compete with an alternative pattern of plural formation, it has syntactic relevance in the sense that it can mark syntactic constituents that are not themselves words, and it is automatic in the sense that any constituent endowed with this property automatically switches to [feminine] in the presence of the feature [plural]; the complex noun phrase *[frigiderul și televizorul]* in (5.24e) is one such constituent. And of course there is nothing strange in a grammatical property being passed on from two conjunct elements to the complex phrase: 'When two 'neuter' singulars are conjoined, agreement is 'neuter' plural, *as might be expected*' (Mallinson 1988: 410; emphasis added). Italian contrasts with Romanian in all these respects.

It is important to frame the agreement facts in the context of the morphological properties of ambigeneric nouns in the two languages.[11] In Romanian, ambigenerics differ from masculines only in the agreement pattern they induce when coordinated. A complex [MASCULINE + MASCULINE] triggers masculine agreement (cf. (5.24b)), while [MASCULINE + AMBIGENERIC] triggers feminine, as shown in (5.24d) and (5.24e). The paradigmatic property of having a feminine form in the plural is thus reflected in the syntagmatic property of triggering an agreement pattern distinct from that of normal masculines. There is, then, full justification for seeing these masculines that have a feminine plural as constituting a distinct agreement class from masculines, whether one calls them neuters or ambigenerics. With this third gender label, the gender resolution rules schematized in (5.23) can be captured

---

[11]   I would like to thank Frank Anshen for helpful discussion on this point.

straightforwardly: in Romanian, a complex NP triggers masculine only if the conjunct NPs are all masculine, otherwise feminine. Crucially, masculine singulars that become feminine when plural must count as non-masculine.

Not so for Italian. The pattern schematized in (5.21), which does not consider nouns like *dito*, amounts to the following generalization: a complex NP triggers feminine only if all of its conjuncts are feminine, otherwise masculine. Suppose we add ambigeneric nouns like *dito* ∼ *dita*, labelling them as 'neuters' to allow for a direct comparison with Romanian. The patterns exemplified in (5.26–5.28) result in the following schemes:

(5.29)   a. [NEUTER + MASCULINE]   . . .   MASCULINE   (cf. (5.26))
         b. [NEUTER + FEMININE]   . . .   MASCULINE   (cf. (5.27))
         c. [NEUTER + NEUTER]   . . .   MASCULINE   (cf. (5.28))

Unlike in Romanian, such hypothetical 'neuters' have exactly the same properties as masculines. The simple reason is that they are masculines—as opposed to members of an ambigeneric gender. The singulars *braccio*, *dito*, *membro*, etc. are masculine, and so are their regular plurals, when available; and the irregular plurals *braccia, dita, membra*, etc., are feminine, not neuter/ambigeneric. This is the most economical interpretation of the agreement facts, but one that can be sustained only by accepting that *dito* and *dita* are two distinct nouns each with its gender value, not two grammatical forms of the same noun.

5.3.4 *Feminine and plural in distributive constructions*

The data from coordination have shown one way in which the correlation between feminine and plural fails to hold for nouns with a plural in -*a*. A different construction illustrates the opposite type of mismatch, where a pronoun agrees in gender but not in number:[12]

(5.30) a. le uova costano sessanta centesimi l'una/*l'uno
            'the eggs.FEM.PL cost 60 cents each.FEM.SG/*each.MASC.SG'

        b. le braccia di Ugo sono una più lunga dell'altra/*uno più lungo dell'altro
            'Ugo's arms are one.FEM.SG longer.FEM.SG than the other.FEM.SG/
            *one.MASC.SG longer.MASC.SG than the other.MASC.SG'

---

[12] Not many accounts of *a* plurals have discussed this important point: see Brunet (1978: 95 6), Lepschy and Lepschy (1988: 110), Bach and Schmitt Jensen (1990: 121) and Maiden and Robustelli (2000: 28).

As we will see in greater detail in Section 5.4.3, only some of the countable plurals in -*a* admit a fully distributive interpretation. Where distributive pronominals like *each, one . . . the other* can apply to these plurals, however, their gender is feminine. Taken separately, the singular number and the agreeing gender are just what one would expect of distributive expressions anaphorical on feminine antecedents; taken together, they lead to another violation of the correlation between feminine and plural. Once again it is irrelevant what gender would be there if the noun were singular instead of plural. If agreement in Italian (unlike Romanian) is insensitive to the feminine–plural and masculine–singular correlation, then this correlation (as opposed to the features themselves) is not an inflectional, syntactically relevant property.

5.3.5 *Blocking (and the lack of it)*

We know that a plural in -*a* is related not only to its singular, but often to another plural as well: for instance, the singular *membro* 'member' has *membra* 'limbs' beside *membri* 'members' (see (5.5)–(5.6) above). Put crudely, this happens too often in this class. The overview in Section 5.2.2 lists fourteen plural doublets, contrasting with just eight cases where the form in -*a* is the only admissible plural (without counting fixed idioms and the three *pluralia tantum* ending in -*a*). Evidently, -*a* plurals are particularly liberal when it comes to admitting alternative forms of the plural—certainly more so than Romanian neuters, for example. One might object that a small class, constituting a minor subregularity in the system of nominal morphology, is likely to allow for competition from the unmarked patterns, like the English *louses* 'contemptible individuals', *mouses* 'pointers for computer', and *oxes* 'large uncouth men'. But the concomitant semantic differentiation only goes to show that the choice of plural realization is also a choice between distinct senses, or lexical bases—between more than just ways to spell out one and the same grammatical content.

 In fact, the usual analysis of -*a* plurals as inflectional plurals with a special meaning is strictly speaking contradictory. One cannot at the same time claim that these plurals are inflectional and that they escape blocking because of semantic differentiation. In the most restrictive sense, 'inflectional' identifies the property of being constructed according to the same principles as regularly, productively inflected word forms: in the case at hand, the claim that *membra* is an inflectional plural of *membro* entails that it is derived from the same base as that of the singular, by adding a plural morpheme with the same content and combined in the same way as any other plural morpheme. To be inflectional in

this strict sense, then, *membra* should be interchangeable with *membri*. The very difference that is invoked to justify the presence of two plural alternants, therefore, shows that at least one of them cannot be inflectional.

This conclusion is quite straightforward really, but the ambiguity of the term 'blocking' usually obscures it. As originally made explicit by Aronoff (1976), this notion applies to word formation, not inflection, and the synonymy required for blocking refers to lexical, not grammatical, synonymy. A doublet like *membra* ∼ *membri* is usually explained on the ground that the semantic difference between the two words ('limbs' ∼ 'members') prevents blocking. The fallacy in this argument is that it purports to account for the lack of *inflectional* blocking (the presence of two inflectionally equivalent forms for one base) by appeal to the properties of *lexical*, or non-inflectional, blocking. This latter notion refers to the paradigmatic relation that rules out morphologically well-formed words like *gloriosity* or *cheater* (as 'someone who cheats') because of the presence of *glory* and *cheat*. Whatever the real linguistic import of such phenomena,[13] the fact remains that the synonymy avoidance exemplified by *cheater* ∼ *cheat* is rather different from that of *childs* ∼ *children*. Morphosyntactic features certainly encode meaning (cf. above all Carstairs-McCarthy 1994), but the grammatically relevant meaning is regimented in categories that are mutually exclusive, often semantically arbitrary, and defined by opposition. The realization of such grammatical categories is strongly disjunctive, in such a way as to leave no room for uncertainty or ambiguity (unless it consists in stem-forming operations, as we will see in Chapter 7 for Arabic). This is the notion of blocking that is relevant in the deterministic choice of exactly one exponent for a grammatically specified word, thanks to which *children* blocks *childs*, *childen* or *childrens*. As we have seen in Chapter 2 (see especially Section 2.7), the lack of a uniquely determined realization typically characterizes plurals that are lexical in one sense or another. 'Lexical' is here a cover term for whatever cannot be reduced to the grammatical system, that is to the knowledge of how to put linguistic formatives together, as opposed to the knowledge of formatives themselves. In a given doublet like *membra* ∼ *membri*, therefore, the semantic differentiation that is usually taken to justify two inflectional plurals must lead to the conclusion that one of them is non-inflectional (namely *membra*).

Clearly, *membra* is also inflectional, but in the much weaker sense that it expresses inflectional features. In a similar vein, *cattle* is also inflectionally

---

[13] Marantz (2003) argues that this lexeme to lexeme blocking is an illusion caused by the un acceptability of certain derived words. *Stealer* as 'someone who habitually steals' appears to be blocked by *thief*; but, as he points out, *breaker* is as unacceptable as *stealer* in this habitual reading, yet there is no synonymous noun to block it.

plural, and can function as the plural of *cow*, but it remains a distinct noun, no matter what functional relations may connect it with the singular *cow*.

The lack of blocking therefore suffices to reject the view that a plural like *membra* is an inflectional form on a par with *membri* and distinguished by a particular interpretation. The plurals in *-a* that admit a concurrent regular plural in *-i*, then, are all non-inflectional, even those that are semantically reasonably transparent, like *corna* 'horns (cohesive)'. But if they are non-inflectional as a class, the possibility is open to regard as non-inflectional even those plurals which, like *uova* 'eggs' and the others listed in (5.7–5.8) above, have no concomitant masculine plural in *-i*. We will see in the following sections that other facets of this plural class, morphological and semantic, strongly suggest just this conclusion.

5.3.6 *Plurals in -a as lexical plurals: the morphological case*

The irregularities we have discussed are not a random collection of morphological properties, but naturally complement each other as aspects of the same phenomenon: the derivation of a lexically restricted class of plural-only nouns. This section will recapitulate the morphosyntactic evidence for this conclusion, before focusing on the morphosemantic side.

5.3.6.1 *Listedness*   Listedness alone, of course, does not denote non-inflectional status. But the idiosyncrasy of both the membership and the exponence of the class of *-a* plurals (properties (iv) and (ii) in Section 5.2.6) acquires a particular prominence when considered in tandem with the special status they have with respect to idioms. As shown in (5.9–5.10) in Section 5.2.2 above, some *-a* plurals only appear in phrasal idioms; moreover, when an idiom involves a plural that has both an *-a* and an *-i* form, it takes systematically the former (property (v)). There is no reason why it is the irregular *calcagna*, *fila*, and *fusa* that appear in these idioms, rather than their regular counterparts *calcagni*, *fili*, and *fusi*. This correspondence between listed form and listed structure becomes more understandable if we see the listed forms as autonomous lexical entries, related to but detached from the inflectional singular–plural pair. As distinct plural-only lexical items, such forms in *-a* have an independent existence and can become fixed in a fossilized expression, possibly losing the semantic connection with the inflectional singular–plural pair. In fact, this conclusion is independently necessary for the *-a* plurals that have no singular at all (cf. (5.10) above).

5.3.6.2 *Categorial restriction*   The restriction to nouns alone, as opposed to adjectives (which otherwise inflect like nouns), is likewise easier to reconcile with non-inflectional than with inflectional status. If *-a* pluralization is a derivational

operation that creates plurals with certain properties, it is entirely expected that the output of this operation should be categorially restricted to nouns, that is to a category whose number value is not determined by agreement. By contrast, it would not conform to the prototypical generality of inflectional classes, irregular or otherwise. In particular, the lack of anything comparable to *-a* pluralization on determiners and other functional (i.e. non-lexical) modifiers underscores the difference with respect to those Romance survivals of grammatical neuter we considered in Section 5.3.2.

5.3.6.3 *Idiosyncratic ending*   The truly exceptional status of *-a* as a plural ending (property (ii)) points to the same conclusion. On its own, it proves nothing; but a comparison with the minor inflectional class of archaic Romance varieties (Section 5.3.2) shows that the formal idiosyncrasy of *-a* plurals differs because it is a property of certain nouns, not a minor agreement pattern. In fact, we can be more precise. It is not just that *-a* as a plural ending is restricted lexically and categorially (two unusual properties for an inflectional ending); what is more, it is only associated with certain atomic lexical stems, and does not seem to occur on any derivational affix. All *-a* plurals are synchronically non-suffixal.[14] This clearly shows that the appearance of *-a* in Italian nominal morphology is inseparable from the choice of certain specific lexemes. An analysis of *-a* plurals as lexically derived lexemes captures this observation and naturally connects it with all others: a plural like *braccia* 'arms' constitutes a distinct way to derive a noun from the same base as that underlying the regular *braccio* ∼ *bracci* 'arm ∼ arms':

(5.31)   ROOT: *bracc-*
      NOUN 1:   *braccio* ∼ *bracci* (masc., class *-o/-i*)
      NOUN 2:   *braccia*   (fem., pl.)

The final *-a* in *braccia* is so exceptional as an exponent of the plural because it is not an exponent of plural; in a very concrete sense, it is part of a derived stem. What is plural, as well as feminine, is the whole noun derived from the same base as that of *braccio*.

It may be objected that nothing in my account makes the *-a* ending a necessary counterpart of this type of lexical plurals. But this seems right: we know from Part I that lexical plurality is not biuniquely associated with certain patterns of exponence, and Section 5.3.1 featured cases of gender-switching nouns without the *-a* ending. I do claim, however, that plurals in *-a* are the output of a word-formation operation which defines them as a class, not as isolated doublets like German *Worte* ∼ *Wörter* (see Section 3.3.5). The historical development from what used to be an inflectional class (Latin *-a* neuter

---

[14] Unless we decide to consider the final *aia* of *centinaia* and *migliaia* as a suffix, synchronically as well as diachronically (< Lat. ARIA), which seems very dubious.

plurals) is obviously responsible for this state of affairs. But there is no reason for predicting that all lexical plurals in Italian should share the same ending, and that this ending should be *-a*.[15]

5.3.6.4 *Lack of inflectional disjunctivity*   The evidence from blocking (Section 5.3.5) shows most directly that plurals in *-a* are not forms automatically and regularly assumed by a certain class of nouns in [plural] contexts. The simple fact that a number of plurals in *-a* do not block their regular alternants in *-i* is enough to prove the point, if we take seriously inflectional disjunctivity. However, this conclusion hinges on the assumption that the morphology relevant for syntax (inflection, or whatever one calls it) deterministically generates one form per inflectional cell and no more. The common view that *braccia* and *bracci* are an inflectional doublet, made possible by semantic differentiation, effectively posits an inflectional pattern whereby certain nouns become feminine in the plural (yielding *il braccio* ∼ *le braccia*), alongside the regular inflectional pattern whereby change in number does not affect gender (*il braccio* ∼ *i bracci*). The singular *braccio* would be ambiguous between these two hypothetical inflectional classes. Belonging to two distinct inflectional classes, the plurals *braccia* and *bracci* would no longer be in a disjunctive relation, while being both inflectional.

If this hypothetical analysis could be sustained, the lack of blocking between *braccia* and *bracci* would no longer be an argument that one of them is lexical. That is why it is important to add that this analysis is wrong, because we have seen in some detail in Section 5.3 that the syntax is blind to the observational correlation 'masculine in the singular, feminine in the plural'. In Italian (unlike Romanian), nouns are never marked [changes gender in the plural], not even optionally (this would be a way to bypass the theoretical consequences of the lack of blocking). Otherwise, they would trigger feminine agreement when coordinated under a formally plural complex noun phrase, at least optionally (cf. Section 5.3.3), or masculine agreement when in construction with a formally singular distributive pronoun, again at least optionally (cf. Section 5.3.4). What we see instead is that the syntax is sensitive to the features actually present on a given form, without any trace of an inflectional deterministic rule linking plural with feminine. There is no inflectional ambiguity on the singular *braccio*, and the only regular plural according to the workings of inflection is *bracci*; the feminine *braccia* can only exist because it does not arise through these operations. Saying that *braccia* is a noun derived from *braccio* ∼ *bracci* amounts to

---

[15] Not all Italian lexical plurals end in *a*. Gender changing nouns like *echi* are one example that falls outside that morphological class; but also other formally irregular plurals, like *buoi* 'oxen' and *armi* 'arms', are suspiciously similar to collectives, as Maiden (1995: 105) points out.

claiming that the correlation between masculine singular and feminine plural has no status in the nominal inflection of Italian.

5.3.6.5 *Gender*   The most telling piece of evidence is obviously the feminine gender value that systematically accompanies -*a* plurals (property (i)). In itself, a change of agreement class from singular to plural is compatible with inflectional status, as shown by Romanian. But I have taken some time to show that Italian -*a* plurals are very different from Romanian ambigenerics: they form a small unproductive class, they have some kind of semantic motivation, they often coexist with regular plurals, they have a privileged relation with listed idioms, they all have the same exceptional plural ending, and, above all, they display none of the agreement effects which motivate a third gender value for Romanian. All this finds a simple explanation by viewing forms like *braccia* as derived inherently plural lexemes, which like all nouns (in a system like Italian) are marked for gender. All we need to say is that the output of this derivation is morphologically feminine. If any grammatical feature characteristically distinguishes distinct nouns, this is gender. The difference between the feminine *braccia* and the masculine *braccio* (plural *bracci*) is as unremarkable, from a morphological viewpoint, as that between *mela* 'apple' (plural *mele*) and *melo* (plural *meli*) 'apple tree'. Of course many questions arise as soon as we broaden the perspective to include the semantic counterpart of this opposition: fruit–tree is an unremarkable lexical opposition, but singular–plural is not. However, before attempting to shed some light on the relation between morphosyntactic and semantic properties, we should take a closer look at the latter.

5.4 The semantic evidence

The morphological reasons for viewing -*a* plurals as intrinsically plural nouns find confirmation in a number of semantic observations. The first piece of evidence to be considered (Section 5.4.1) concerns an unexpected agreement pattern and is thus closely linked to the previous section. After that, I will review the nuances of semantic plurality associated with -*a* plurals (Section 5.4.2). The semantic variety brought to light by a careful inspection undermines the traditional view that -*a* plurals are in some sense collective; at most, something similar may hold for a subset of them, but cannot be the characteristic reading of -*a* plurals. In fact, no shade of meaning can be isolated as *the* semantic contribution of the -*a* ending to plurality. However, we can discern a common denominator: the property of denoting weakly differentiated entities (Section 5.4.3). Crucially, this is a property of the lexical

meaning, not a special type of plurality. The only common denominator shared by *-a* plurals, therefore, emerges when we consider the whole word.

5.4.1 *The role of plurality in gender agreement*

Section 5.3.4 featured some examples of number mismatch between plurals in *-a* and coreferent pronouns in the singular. Consider now the following structure:

(5.32)   volevo due uova, e me ne hanno date tre
       'I wanted two eggs.FEM.PL, and they have given.FEM.PL me three'

The clause 'they have given me three' contains an empty category referring back to 'eggs', and preceded by the numeral 'three'. The pronominal features of this empty category, which obligatorily surface as agreement on the past participle (*date*) via the partitive clitic *ne* 'of-them', depend partly on the antecedent, partly on its determiner. In (5.32), both the antecedent and the object have a plural determiner ('two' and 'three'), and 'two eggs' agrees with 'three [of-them]' in feminine plural. What will happen if we replace 'three' with 'one', thus causing a mismatch between the plural of the antecedent and the singular imposed by 'one'? Surprisingly, neither form is as good as the examples in Section 5.3.4 above:

(5.33)   a.   %volevo due uova, e me ne hanno data una
          'I wanted two eggs.FEM.PL, and they have given.FEM.SG me
          one.FEM.SG'

      b.   %volevo due uova, e me ne hanno dato uno
          'I wanted two eggs.FEM.PL, and they have given.MASC.SG me
          one.MASC.SG'

All speakers I consulted agree with me in finding neither version fully acceptable, although opinions vary about which is better and how deviant it is with respect to the impeccable (5.30a), here repeated:

(5.30)   le uova costano sessanta centesimi l'una/*l'uno
       'the eggs.FEM.PL cost 60 cents each.FEM.SG/*each.MASC.SG'

This state of affairs must reflect some difference between the two sets, and the only one is that the singular pronouns were all distributive in (5.30), while none is in (5.33). The singular of distributive structures is semantically plural: the truth value of (5.30) must be evaluated by assigning to *una* each one of the anaphoric indices of the objects referred to by the plural antecedent 'the eggs'. Grammatical number, by contrast, matches semantic number in (5.33): *uno* in

(5.33b) means one (egg), not one at a time. Inevitably, this leads to contradictory requirements when the number of the antecedent is the opposite of that of the pronoun.[16]

This is a significant empirical result. A feminine plural in -*a* can be resumpted by a feminine singular pronoun without any problems, but only if the pronoun has a distributive interpretation and is therefore assigned distinct anaphoric indices, as in (5.30); otherwise, where the pronoun is not distributive, as in (5.33), a structure where either its gender or its number value fails to match that of the antecedent has shaky or at least variable acceptability (noticed in Maiden and Robustelli 2000: 26). This means that the feminine gender of -*a* plurals has an intrinsic association with plurality. If the feminine of certain nouns were just the automatically triggered consequence of their being plural, the distributive status of a morphologically singular pronoun should be irrelevant, and all mismatches in number between antecedent and pronouns should be equally acceptable or unacceptable. The observed crucial role of distributivity follows instead from the assumption that both the gender and the number value in *uova* are lexeme-inherent specifications. When the reading is distributive, a singular pronoun can of course refer to a plural antecedent, with agreement in gender routinely marking the anaphoric relation; but in a non-distributive structure like (5.33), the plural 'eggs' and the singular 'one [egg]' are arguments of different predicates, and the pronoun only depends on the noun for the identification of the lexical predicate (eggs), without creating an anaphoric dependence. The gender of the pronoun must therefore be chosen on the basis of what the antecedent would be if it were singular—and the resulting mismatch with the actual antecedent gives rise to the uncertain and variable judgements.

The behaviour of gender in such pronominal dependencies, then, provides an independent argument for thinking of -*a* plurals as intrinsically plural nouns, distinguished from their bases by a different gender value. This coincides with the results of the comparison with Romanian, but extends the conclusion to all -*a* plurals, not just those competing with a regular plural: the examples above involve *uova*, which has no **uovi*. We must therefore conclude that a seemingly straightforward paradigmatic relation like *uovo* ∼

---

[16] Notice that Romanian is different. Here, the correlation 'masculine if singular, feminine if plural' holds good regardless of distributivity, and there is no uncertainty in the judgements (for which I thank Carmen Dobrovie Sorin and Donka Farkas):

(i) vroiam   două   ouă   dar nu mi a   dat   decît
     wanted.1.SG   two.FEM   egg.FEM.PL   but not to.me has given but
     unul/   *una
     one.MASC.SG/ one.FEM.SG
     'I wanted two eggs but (s)he only gave me one'

uova 'egg' ∼ 'eggs' involves two minimally distinct nouns: a singular-only *uovo* and a plural-only *uova*, functionally 'the' plural of *uovo* but morphologically a distinct noun. This view rests on the idea that the plurality of *uova* is as lexically inherent as its gender. There is indeed ample evidence that the plurality of -*a* forms is part of their lexical meaning.

5.4.2 *The semantic categories of -a plurals*

I will now examine the 'special' meanings of -*a* plurals, with a view to showing the inadequacy of the traditional notion of collectiveness, and indeed of any single semantic correlate for the -*a* ending. The obvious interpretive regularities and subregularities, it will be argued, can only be captured at the level of lexical meaning. Specifically, -*a* plurals constitute a class of semantically related lexemes, referring to units of measurement, objects like eggs, elements of cohesive aggregates, and parts of masses. Collectiveness, cohesion, mass are all relevant for some members of this class, but none is 'the' meaning of -*a*.

5.4.2.1 *Cohesion, collectiveness, and semantic diversity*   The traditional idea that plurals in -*a* uniformly have a collective sense is plainly inadequate, as Italianists know at least since Brunet's (1978) detailed study.[17]

Irregular plurals with the meaning of units of measurement are neither cohesive nor collective, in any of the senses this term may have. At most, it is the singular of some of these nouns that may be viewed as collective, naming an indefinite amount as a plurality of units. This applies to *centinaia* 'hundreds', *migliaia* 'thousands' and *miglia* 'miles'; *braccia* 'arms' and *dita* 'fingers' belong here too in their use as measures. Nouns of containers used as measures, like the obsolete *sacca* 'bagfuls' or *carra* 'cartloads', also fell into this category. Finally, we should include *paia* 'pairs', which is not strictly a measurement but has no descriptive content beside expressing a numerical size. One might call its singular a collective, but not its plural. Beyond measurements, *uova* 'eggs' is not collective either. Nor does its reference have particular cohesion: it may be true that eggs are more often referred to in the plural than in the singular, but that alone does not turn 'eggs' into a collective noun, unless we bleach the notion of collectiveness so much that it ceases to have any use.

Grida 'shouts', *strida* 'cries', and *urla* 'cries, shouts' are cohesive rather than collective. They refer to pluralities of shouts connected to each other by belonging to a single complex shouting event (whether or not the single shouts overlap). The cohesion between the single shouts may be provided by a single agent, or just perceptual contiguity (in time and/or space), or by

---

[17] Since nouns limited to fixed idioms have an idiosyncratic and largely non compositional interpretation, the survey in this section will not include them.

a combination of these factors when the agent of the event is plural—as in screams from a large audience. The meaning of these forms ranges from 'individual related shouts' to a mass-like 'shouting' (which would be incompatible with the form *gridi*). In any case, their denotation is not limited to well-defined bounded sets of shouting events.

The notion of cohesion is particularly relevant for a group of -*a* plurals denoting naturally related referents, mostly body parts: *braccia* 'arms', *ciglia* 'eyelashes', *corna* 'horns', *dita* 'fingers', *ginocchia* 'knees', *labbra* 'lips', *lenzuola* 'sheets', *membra* 'limbs', *ossa* 'bones', *sopracciglia* 'eyebrows'. This set provides the basis for the traditional definition of -*a* plurals as collectives. This is why it is important to clarify that cohesion and collectiveness are distinct (cf. Section 4.4.2). Plurals in -*a* belonging to this group may well imply that the noun refers to parts of an organic whole, but this does not mean that every such plural denotes natural collective wholes. *Dita* 'fingers', for example, does not exclusively refer to a standard ten- or five-membered set, but simply means 'fingers', regardless of how many hands they belong to. *Braccia* 'arms' likewise can (in fact, must) be used when talking about many people's arms, as in the advertising slogan *la forza di venti braccia* 'the strength of 20 arms' (said of a dough-mixer). The arms in question are human, but twenty human arms do not form a naturally occurring collection: two do. In conclusion, it is plainly wrong to attribute to body-part plurals like *dita* or *ginocchia* reference over naturally occurring sets of fingers or knees (as expressly proposed by Ojeda 1995). Natural aggregates do provide the basis for conceptualizing -*a* plurals as constituents of larger wholes; but the -*a* plurals do not denote these larger wholes.

By contrast, *mura* is a real collective, which refers to a walled perimeter, typically of a walled city or of a room. A perimeter is one individual entity, not just a cohesive aggregate; crucially, it is not a mass noun either, in that it refers to a bounded whole.

All remaining nouns are mass. None of *budella* 'guts, intestines', *cervella* 'brains', *fondamenta* 'building foundations', *gesta* 'deeds', *interiora* 'entrails', *midolla* 'marrow, core', *risa* 'peals of laughter', and *vestigia* 'relics' admits a numeral in front of it. As the glosses should make clear, the body-part nouns in this group refer to organs perceived as complex structures but without whole parts (the convolutions of *budella*, for example, cannot be subdivided into individual members in a non-arbitrary way). They are no more collective than mass plurals like *suds* or *fumes*.

In sum, -*a* plurals fall into the following quite distinct semantic categories: non-collectives (measures, 'eggs'); cohesive aggregates (discrete body parts, contiguous events); the single collective *mura* 'walled perimeter'; and mass terms (non-discrete body parts, artefacts, manifold events, 'deeds').

5.4.2.2 *Are only some* -a plurals semantically idiosyncratic?   The four cate-
gories just identified define a scale of decreasing individuality: discrete elements
viewed as non-identifiable, logically (measurements) or conceptually (eggs);
elements forming a larger whole; 'walls' as a true collective; and masses
without discrete entities at all. Of these four classes, only the first contains
plurals whose referents are not conceptualized as parts of some cohesive larger
entity. Now, the core elements of this category (*centinaia* 'hundreds', *miglia*
'miles', *migliaia* 'thousands', *paia* 'pairs', *uova* 'eggs') also have no competing
plural in -*i* (cf. (5.8) above). It might seem plausible to relate the two facts and
conclude that nouns whose only plural is in -*a* have no semantic idiosyncrasy,
while the cohesive reading arises in plural doublets. If so, -*a* would have a special
meaning only as a second plural in a doublet, not when it is the only plural:

(5.34)   SINGULAR   PLURAL
                        Semantically   Semantically
                        regular   lexicalized (cohesive)
          -o   -a   //
          -o   -i   -a

This seemingly plausible approach, it must be noted, does not explain why it is
systematically the -*a* member of each doublet that has a special reading. Since,
by hypothesis, the -*a* ending is a transparent pluralizer in the case of measures
and *uova* 'eggs', nothing would explain why it is invariably the semantically
marked member of doublets: contrast the irregular English *oxen*, for instance,
which is semantically unmarked with respect to the regular *oxes*.

   But even apart from this, the suggested approach would overlook two
important facts. First, some -*a* plurals without a counterpart in -*i* have a
mass reading, like *midolla* 'marrow', *risa* 'peals of laughter' and *strida* 'shouts'.
According to (5.34), these should be semantically regular; but their mass
reading does not arise from a transparent pluralization of the singular.
Second, a form like *braccia* 'arms' has a secondary measure reading 'arm-
lengths' which, like 'eggs', 'pairs', or 'hundreds', has nothing to do with
cohesion; but, unlike these plurals, *braccia* has a regular alternant *bracci*
(typically for 'arms' of objects, or to pluralize *braccio di mare* 'sea channel').
The same applies to *dita* 'fingers', which has the alternant *diti* in some registers
but never with the measurement reading ('the thickness of a finger'). And the
ancient language had several plurals like *sacca* 'bagfuls' and *carra* 'cartloads',
where -*a* expressed the measure reading in contrast to the regular plurals *sacchi*
'bags' and *carri* 'carts'. What (5.34) implies is that -*a* plurals for cohesive body
parts can make up doublets with regular -*i* alternants, while units of measure
(like *miglia*) do not; but some plurals for body parts double up as expressions

of measure. In conclusion, a close examination does not really justify separating a lexicalized group (masses, cohesive aggregates and *mura*) from a non-lexicalized group (measures and *uova*); nor does it justify treating all and only *-a* plurals in doublets as lexicalized, as (5.34) would suggest.

One way to skirt all these issues consists in simply claiming that some plurals in *-a* have a peculiar interpretation, which has to do with cohesion, and that several among them have a semantically and formally regular alternant. This would not greatly differ from the received view, and resembles the approach of Ojeda (1995). He restricted his analysis to plurals with a competing *-i* alternant, and argued that the *-a* forms differ from them not in the value of the plural morpheme, but in the semantics of the stem. Ojeda regards the plural as the unmarked number, so that plural endings are just identity functions, which do not affect the reference of the stem. It is rather the singular endings that limit it to the domain of atomic individuals (a point developed in Ojeda 1993: 70–9). The difference between *ginocchi* and *ginocchia* (both 'knees'), then, resides in their stems: there are in fact two homophonous stems for each plural doublet, one denoting the set of all sets of single referents, including singletons, and the other denoting 'the set of natural groups of knees of the universe of discourse—the set of pairs of knees of each individual and the group formed by these pairs' (Ojeda 1995: 218). This mereology takes as basic elements, then, neither single knees nor pairs of knees (which would make it a lexical dual),[18] but natural pairs of knees, that is two-membered body parts.

The shortcomings of this approach are evident: systematic homophony between stems seems the least revealing way to account for the different readings a noun can support (count, mass, kind, collective), and taking 'natural set' as given might work for 'knees', but not for 'building foundations' (what are they a set of, natural or otherwise?).[19] Indeed, Ojeda acknowledges (1995: 235, note 29) that his analysis does not work for plurals in *-a* if they are mass, a limitation that, coupled with the exclusion of *-a* plurals which happen to have no competing *-i* form, means that the proposal concerns only a few among all *-a* plurals, missing the correlation between cohesive aggregates and manifold masses.

---

[18] Plurals in *a* resemble duals in their peculiar restrictions to 'natural' aggregates (as Ojeda 1995 has it) or more generally to cohesive pluralities. Hall (1956) related this to the irregular development of Latin *duo* 'two' into *dua* in early Tuscan. For the special cohesive interpretation carried by dual morphology, see Chapter 3 above (Section 3.6.4).

[19] It might work for knees or similar twin organs, but I don't think it does. If *ginocchia* refers to the set of natural knee pairs, it cannot refer to an odd numbered set of human knees. But (i) below sounds perfect in a context where a photograph shows three knees without revealing the faces of their owners; at least, no speaker would say it is ill formed because the *a* plural was used instead of *ginocchi*:

(i) di chi sono queste tre ginocchia?
     'who do these three knees belongs to?'

For all that, Ojeda's (1995) analysis has the merit of making it clear that plurals in -*a* are lexical in the sense that what plurality means on them is a function of the lexical entry.[20] It is not that -*a* adds something to the semantics of certain nouns, be it cohesion or collectiveness; rather, this ending marks a particular interpretation of the whole word, which emerges as cohesion with certain nouns but not with others.

5.4.3 *A common semantic denominator*

What measurements and eggs have in common with cohesive aggregates and masses is that the parts making up the denotation are conceptualized as undifferentiated, in different ways according to the lexical semantics of the noun.

Consider the first of the semantic categories into which we have grouped -*a* plurals: numerical measures ('hundreds', 'thousands', and 'pairs'), measurements of linear distance ('miles', 'arm's lengths', and 'finger's lengths'), and the homogeneous class of -*a* plurals derived from names of containers, which express quantities according to the scheme 'X → amount contained in a X'. As we will see in greater detail in Chapter 6, *three miles* does not refer to sets of miles of cardinality 3, as *three books* does to sets of books. Miles express a conventional segmentation of the dimension in which amounts are quantified (here, linear distance) and, preceded by a quantifying expression that determines a value on a scale, yield an amount in this dimension.[21] Being abstract standards for measure functions, such units have the grammar of atomic entities but are not conceptualized as possessing identity properties that could enable a speaker to refer deictically to them, contrast one with another, or determine the identity of one in isolation.[22]

What of *uova* 'eggs', which certainly is no unit of measurement? Surely we can point to an egg, distinguish it from another one, and attribute properties

[20] In fact, Ojeda (1995: 231) does not treat *a* plurals as lexical, and terms the *a* ending a 'plural inflection' on a par with *i.*

[21] In a sentence like *these three miles have been easy going*, where we seem to point to three particular miles, reference is in fact to a three mile long stretch, corresponding not to a measure but to a stretch of land of that length.

[22] Revealingly, the few 'collective' feminine nouns in old Provençal display a morphosyntactic split that becomes understandable in the light of our semantic characterization of measure terms: 'Graf ström [1968] distinguishes between such terms as *paira*, *charra*, and *semoia* ['(ox)pairs', 'cartloads', and 'half bushels'] which may be used as plurals: *doa paira* ['two ox pairs'], *tria semoia* ['three half bushels'], *.vii. charra* ['seven cartloads'], and the other collective nouns which, formally, remain singulars, such as *la brasa* ['the.sg pair of arms, arm length'], *la ossa* ['the.sg set of bones'], although clearly denoting plurality' (Jensen 1976: 32). Apparently in old Provençal, *a* forms remained plural only when they denoted non cohesive units of quantification, while those denoting collections of parts of a larger whole became singular. Notice that, as a measurement, *brasa* corresponded 'to the span of the arms' (Jensen 1976:35), not to *one* arm's length as in Italian, and therefore fell into the second class.

to it. Still, eggs are interchangeable in a way that deprives them of distinctive individuality. An n-numbered quantity of eggs can consist of any eggs, because our encyclopaedic knowledge of the world makes it exceedingly unlikely that the truth value of a sentence could depend on the identity of the single eggs in question. Eggs, of course, do have individual properties that may distinguish one from the other even independently of the space occupied, for instance, colour and dimensions; my point is that the inherently plural *uova* disregards these differences and conceptualizes eggs as interchangeable tokens. It stands to reason that this kind of conceptualization should apply to referents with low perceptual salience as individuals: inanimate objects, with uniform appearance, and seldom experienced outside of aggregates.

The other types of -*a* plurals correspond to different ways to conceptualize certain referents as weakly individual, this time not so much because they are interchangeable (lacking identity) as because they are not whole (lacking unity). The isolated collective *mura* does not require too much justification to be considered a plurality with non-individual parts. In a walled perimeter conceptualized as a unit, the distinctive properties of its constituent parts are backgrounded. The parts that make up the denotation of the mass nouns do not have the interpretive properties of individual parts either, as we have seen in some detail in the preceding chapter. Portions of masses can be pointed to, but they are defined by their size and by their spatio-temporal coordinates, not by any property that might distinguish each portion from all others. A plural like *interiora* 'entrails' refers to organs that are nameable and distinguishable, but it neither names the sorts into which they fall as units (i.e. the types of organs: heart, liver, and so on), nor provides a criterion for deciding when one member of its denotation ends and the other begins (like *offal* or *offals*, and unlike *orchestra*, it does not refer to a collective whole).

Consider now cohesive aggregates. When sheets or human limbs are in question, the lack of individuality associated with the -*a* ending means that the individual parts are seen as if they were mutually indistinguishable. This explains the following contrast, from an insightful observation of Lepschy and Lepschy (1988: 110) and Brunet (1978: 40, 43):

(5.35) a. due dita lunghe/affusolate
        b. ?*due diti lunghi/affusolati
            'two long/slim fingers'

(5.36) a. *due dita mignole
        b. due diti mignoli
            'two little fingers' (as opposed to thumbs or pointing fingers)

In contemporary usage, *diti* is substandard at best and cannot be said to compete with *dita*. Just how deviant it is varies with the speaker, but *dita* is unambiguously the preferred option in a context like (5.35). Having said that, *dita* in the minimally different (5.36) is sharply ungrammatical, and *diti* is the obligatory choice with no stylistic connotations. But (5.36) differs from (5.35) in a crucial respect: *mignole/mignoli* refers to one specific finger-type among those included in a hand, and thus provides a criterion for articulating the concept 'finger' into subtypes. This particular modifier, unlike 'long' or 'slim', prevents the conceptualization of fingers as undifferentiated between each other, and thus enforces *diti*. The closely related *ossa* describes bones bearing the relation of being parts of an organic whole, but crucially this cohesive relation may hold without the bones being arranged into such a whole: compare the perfect acceptability of *le ossa dell'avambraccio* 'the bones of the forearm', referring to a small subpart of a skeleton. *Ossa* would be perfectly appropriate to describe a set of bones consisting of one complete skeleton plus an incomplete one, or even a heap of scattered bones, provided they are seen as parts of natural wholes (likewise, bricks may be viewed as building material without being part of a building). In short, *ossi* means 'bones as wholes', and *ossa* 'bones as parts'.

As for *grida*, *strida*, and *urla* 'shouts, screams', we can now see that what really counts is that shouts in a complex shouting event are mutually indistinguishable. More precisely, to conceptualize factually distinguishable shouts as parts of a shouting event means to disregard the properties that could distinguish one shout from the other.[23]

An important trait of -*a* plurals now falls into place, namely the dimness of some grammatical intuitions, repeatedly noted in the descriptive literature and unjustly played down in some attempts to categorize this class. First, this indeterminacy does not apply to all -*a* plurals. There is nothing dim or fuzzy about the use and distribution of -*a* plurals for units of measurement, because their semantics is clear: they define unequivocal criteria for unity (corresponding to the respective singulars), and the abstract units they denote cannot have individual distinctive properties, otherwise they would not be abstract units.

---

[23] Ojeda (1995: 234, note 15) came very close to this conclusion in an aside: 'According to Goidánich (1967: 142), *corna* is used for horns 'of animals that have two', whereas *corni* is used for horns 'of animals that have one'. When Goidánich attempts to illustrate the latter he speaks, however, of the horns of a rhinoceros. But a rhinoceros has two horns rather than one! Perhaps what this grammarian meant was that *corna* is used for horns of animals that have two horns *of the same shape*' (emphasis in the original). This particularly good example teases apart the notions of cohesion and lack of distinctiveness, because there is no sense in which the horns of a bull should form a set any more cohesive than those of a rhinoceros.

For all other -*a* plurals, however, the lack of individual distinctive properties is a matter of how the lexical predicates are conceptualized, and this often leads to variation among speakers and uncertain intuitions for one and the same speaker. This indeterminateness does not only manifest itself in the choice between -*a* and -*i* form, which does not apply to all -*a* plurals, but much more clearly in countability, and specifically in numerical modification. Names of specific body parts can all be counted; but the acceptability of numeral modification for other -*a* plurals of this group is much less homogeneous. I myself find a numeral determiner odd with *grida*, *urla*, and *ossa*,[24] and downright unacceptable for *membra*, but perfect with *lenzuola*. Usage varies, however, and depends significantly on the availability of an alternant in -*i* for a speaker's active vocabulary. Besides, the syntactic placement of the numeral also plays a role, as shown by the following slight but noticeable contrast:

(5.37)   a. le ossa dell'avambraccio sono due: ulna e radio[25]
              'the bones of the forearm are two: ulna and radius'

         b. ??ci sono due ossa nell'avambraccio: ulna e radio
              'there are two bones in the forearm: ulna and radius'

Membra 'limbs', expresses something like 'body parts', where 'body' is the criterion for cohesion. Limbs are discrete entities, but the noun does not provide a criterion for uniquely segmenting a body. Since the description is so vague, contexts supporting a numerical quantification of limbs, without a specification of which limb, are hard to come by.

Similar considerations apply to other -*a* plurals. *Ossa* 'bones' surely provides a well-defined criterion for countability. Yet, for speakers who share the judgements described in (5.37), *ossa* is dispreferred (not ungrammatical) after a numeral because the competing *ossi* better fits an enumeration, in which no item bears a particular relation to any other.[26]

Likewise, for many speakers, the form *lenzuola* 'sheets' must refer to sheets functionally arranged to form bed linen (not necessarily of a single bed), and countability is restricted to the non-cohesive *lenzuoli*. For others, including myself, the form in -*a* does not require this degree of cohesion, and therefore countability is not a problem. But the sheets covering a Ku Klux Klan member (cf. (5.12) above) stray so much from the usual conceptualization of sheets as

---

[24] *Quattro ossa* 'four bones' is an idiom meaning something like 'poor bones', and does not impinge on the countability of *ossa*; replacing *quattro* with *tre*, for instance, sounds quite odd.

[25] Unlike *braccio* 'arm', *avambraccio* 'forearm' only has the regular *avambracci*.

[26] Cf. the following minimal pair: *le ossa del gatto* 'the cat's bones (remains)' vs. *gli ossi del gatto* 'the cat's bones (as her food)'. The English translation *remains* is itself a *plurale tantum*, as I claim *ossa* is.

linen that even for me that reading favours *lenzuoli*, which expresses plurality without any additional nuance. The same applies to *urla* and *grida* 'shouts', cohesively interpreted as parts of a complex shouting event.[27] Whether they can be counted depends on the context and on the speaker's vocabulary. If *urli* and *gridi* are available, these forms are better in a singulative context. But the availability of a form for active use and the conceptualization of an event as part of a larger event are a matter of degree and vary contextually, dialectally, and idiolectally. These are hallmarks of properties of the conceptualization, as opposed to grammatically regimented categories, and conceptualization applies to the semantics of lexical items.[28]

In sum, Italian plurals in *-a* are restricted to concepts whose reference contains elements perceived as equivalent to one another, which fits both the logical equivalence of units of measurement and the perceptual equivalence of weakly differentiated objects, like eggs. A noun like *books*, by contrast, has a denotation based on individual objects for which there is a criterion of atomicity, conceptualized in a way that makes it possible to tell one book from the other. *Eggs* too has a denotation based on individual objects, but they are conceptualized as equivalent to each other, like bricks in a wall. Hence the class of *-a* plurals in Italian includes the counterpart of *eggs* but not of *books*. Historical accident played a large role in ensuring that today the word for eggs but not for, say, spaghetti, has a plural in *-a* (it so happened that *ovum* was neuter in Latin and that spaghetti were yet to come), but the weak individuality of the single referents is a necessary, not sufficient condition.

5.5 Conclusion: plurals in *-a* as derived lexemes

Braccia 'arms' is not the plural of *braccio* 'arm'; it is an inherently plural lexeme, derived from the same root as *braccio/bracci* and provided with a gender value like any other noun:

---

[27] *Risa* 'peals of laughter' is different because, like its English counterpart, it does not provide a clear criterion for segmenting its units (where does a peal start and finish?). The mass reading, which is available for *urla* and *grida* as well, is here the only one possible. Notice that the singular of *grida* and *urla* refers to a single shout, while the singular *riso* means 'laugh, laughing, laughter' as an activity predicate, not as a single event.

[28] Countability alone is not a reliable indicator of part structure conceptualization, because it is inhibited by cohesion. Other tests help discriminate between *a* plurals that conceptualize their denotation as aggregates of individuals and those that refer to manifolds whose parts cannot be semantically accessed. *Membra* and *ossa*, for instance, cannot usually be preceded by a numeral, but

(5.38)   lexeme 1:   ARM$_1$
form: *braccio* [sg]/*bracci* [pl]
meaning: **arm**
morphosyntactic information: noun, masculine, class *-o/-i*

lexeme 2:   ARM$_2$
form: *braccia*
meaning: **arm** (pl) as complex with functionally non-distinct parts
morphosyntactic information: noun, feminine, plural

The shared root accounts for the identity of the core meaning; and the status of derived noun accounts for the non-inflectional traits of these plurals, including not only the possible semantic distinctness from the singular and/or the regular plural, but also their behaviour with respect to agreement in coordinate and distributive constructions (Sections 5.3.3, 5.3.4, 5.4.1). In addition, at least some *-a* plurals are input to further derivation. The meaning of *membra* 'limbs' differs enough from that of *membri* 'members' to make it plain that the verb *smembrare* 'to dismember' must be based on the *-a* form: it means 'to take the *membra* apart', not the *membri*. Similarly, *corna* 'horns' has an idiomatic reading unavailable to *corni*, namely the horns of a cuckold. The verb *cornificare* means 'to make a cuckold', and thus must be based on the *-a* form. These plurals can therefore act as lexical bases for derivation, as would be expected of autonomous lexical entries rather than inflected word-forms.

The paradigmatic relation between *braccio* and *braccia* is thus that between *cow* and *cattle*, with the crucial twist that the two nouns are based on the same root; the Slavonic plurals mentioned in Section 2.7.3 are a much closer parallel. Being the result of a lexeme-forming operation, *-a* plurals do not block inflectional plurals, nor must they have a different denotation to justify their existence. Since the plurality they express is a property of the derived lexeme, it often has a special sense, and the variation according to noun, context, speaker, and register is much stronger and ranges across a much greater span than the possible range of inflectional, compositional plural. At the same time, since these plurals lie outside inflection, they do not have to be semantically special at all: *uova* or *miglia* simply refer to pluralities of eggs and miles. The lexicality of *-a* plurals, in this precise sense of being distinct lexemes in their own right, thus accounts in a unified fashion for their diversity, their grammatical and semantic properties, and also for the indeterminacy that often blurs acceptability contrasts.

they can be antecedents of distributive expressions, unlike mass terms like *budella*:

(i)   le membra / ossa / *budella di un corpo umano hanno ognuna una funzione diversa
'the limbs / bones / *intestines of a human body have each a different function'

Invoking lexicality does not mean that anything goes. The plurals in -*a* are a morphological class, and they remain a single class despite the semantic diversity we have uncovered. To capture this distinctive trait of -*a* plurals, I propose to view the plurality of weakly differentiated parts, the morpho-syntactic properties of feminine gender and plural number, and the ending -*a*, as the semantic, morphological, and phonological sides of a single lexeme-deriving process that creates plural lexemes with a specified morphology.

The input of this derivation must of course be lexically restricted, as the class is closed and unproductive. But the lexical restriction is not arbitrary; the derivation creates nouns interpreted as pluralities (including masses) with weakly differentiated elements, and the choice of inputs reflects this. The notion of weak individuality I have suggested correlates in a natural way with the category of inanimate concepts, and this in turn with the historical source of these -*a* forms, the Latin neuter. The common semantic denominator of Italian inherently plural -*a* lexemes derives historically from the conceptual-ization of inanimate pluralities that was grammatically encoded in Latin (and in Indo-European: cf. especially Schön 1971). Evidence that the peculiar 'collective' brand of plurality remained a synchronically active semantic category after the demise of the Latin gender system is provided by the well-known fact that the -*a* plural was extended to nouns which were not neuter in Latin, as in the case of *mura* or *dita*. On the formal side, the old neuter plural ending -*a* remained the formal correlate of this conceptualiza-tion, now related to a derivational process and no longer to an inflectional category. The fact the all the singulars corresponding to -*a* plurals end in -*o* is also due to historical reasons (neuter singulars generally went into this category); but while -*a* is synchronically the exponent of this derivation, there is no need to state that the input nouns are in the inflectional class -*o*/-*i*, partly because it would be unusual for a lexical derivation to target a specific inflectional class, partly because at least one plural in -*a* (*interiora* 'entrails') is related to the adjective *interiore* (Latin *interior*), which is not in the -*o*/-*i* class.[29]

The output of the derivation consists of plural nouns interpreted as plural-ities of non-distinct elements. According to the lexical meaning of the base, this non-distinctness takes one of several forms: units of measurement are by definition abstract equivalence classes, while eggs are conceptualized as non-distinct on the basis of their lack of perceptual salience, natural aggregates on the basis of spatial, temporal, or functional cohesion, and mass terms (a kind of super-cohesive aggregates) on the basis of the lack of clearly segmentable parts.

---

[29] Thanks are due to Frank Anshen for discussions on this issue.

But all -*a* plurals are reducible to this common denominator: what sets measures and 'eggs' apart from the rest is that this type of plurality for them represents the only one, without contrasting with a 'singulative' in -*i*. In fact, lexicalized plurality does not require semantic idiosyncrasy, nor for that matter does it require formal idiosyncrasy (see Chapter 2).

Finally, the -*a* ending correlates with a choice of gender and number values. Normally, a noun is specified for a gender only, but -*a* plurals are a class of *pluralia tantum*; recall that, for example, *uovo* 'egg' is strictly speaking a distinct noun from *uova* 'eggs', although they are functionally equivalent to inflectional pairs like *book* ∼ *books*. This makes of -*a* plurals a morphological class by themselves, defining a distinctive match between exponents, grammatical information and an agreement patttern: I see no objection to calling them an inquorate gender, following Corbett (1996: 105); but they are a gender because they are plural, not because they are neuter (as argued instead by Bonfante 1973). Since plurality is inherent to them as a gender value is, -*a* plurals are, if anything, a plural gender, more or less like the inherently plural noun stems of Kiowa and Turkana discussed in Chapter 3 (Section 3.5.3).

The last question concerns the status of -*a*, which is the formal correlate of the nexus of morphological and semantic properties defining these lexical plurals as a grammatically relevant class. All the evidence reviewed in this chapter suggests that -*a* plurals are a class for the traits they share as lexemes, not because -*a* as a plural ending 'means' cohesion, or lack of distinctness. I think the evidence warrants a stronger conclusion, namely that -*a* by itself does not even 'mean' feminine or plural. The form *braccia* 'arms' can be suffixed with the diminutive -*ino*, an evaluative affix of class -*o/-i* in the masculine and -*a/-e* in the feminine. Since the affix takes on the gender and number values of the base, the result is *braccine* 'small arms'. But this feminine plural form has no final -*a*: as usual in Italian, the vocalic ending does not appear in the base for suffixation (Scalise 1994: 154–60). On the assumption that the vowel is genuinely absent rather than phonologically deleted (an assumption supported by Peperkamp 1995), the necessary conclusion is that *bracc-* is feminine and plural even without the final -*a*. This follows if, as I argued, the distinctive properties of -*a* plurals are properties of the lexical base. The derivation of *smembrare* and *cornificare* from *membra* and *corna* points to the same conclusion.

The -*a* of *braccia*, then, does not carry inflectional information at all, and its status is that of word marker (in the sense of Harris 1991), a vowel marking the right edge of a word for morphophonological well-formedness. This conclusion accords well with my previous observation (Section 5.3.2) that -*a*

differs from Romance 'neuter' endings in being exclusive to nouns and not affecting the morphology of agreeing elements.

These last cursory remarks do not purport to replace an analysis of the *-a* of *braccia* within the context of Italian noun morphology, but they do suggest a general conclusion. Of the features that identify *-a* plurals as lexical, those related to semantics and to agreement are abstract properties. If *-a* is no more than a word marker, as opposed to an affix, there remains no element to directly spell out the status of *braccia* as a feminine plural lexeme. The hypothesized derivation has several (indeed, many) observable reflexes in semantics and morphology, but it has no dedicated exponence comparable, for instance, to the collective morphemes of Amerindian languages (cf. the Zuni examples in Section 4.4.3). This underscores, I think, the abstractness of the notion of lexicality needed to tease apart inflectional and lexicalized plurals. *Braccia* as a whole spells out an autonomous lexeme, but—apparently—none of its parts does.

6

Irish counting plurals

6.1 Introduction

Our second case study features a special type of plural nouns in Irish. In this language, nouns governed by a numeral between three and ten must be, or at least can be, singular. The admissibility of the plural in this context varies lexically and dialectally, but no dialect follows the usual Indo-European pattern which mandates plural in this context. The (modern) Irish equivalent of 'three books' is thus literally 'three book', as a possibility if not always an obligation.

Not all nouns conform to this scheme, however, and some take the plural instead of the singular. They make up a class whose membership varies dialectally but centres around measures, units of time, and objects conceptualized as non-distinctive, like 'eggs' or 'times' (as in *four times*). In the dialects where this class is relatively small and closed, moreover, the form of the plural used after 3–10 is often distinct from that used in all other contexts.

These special plurals used after 3–10 are what we will focus on. Like Italian -*a* plurals, they refer to weakly individual concepts and often contrast with regular alternants; however, their limitation to a specific context makes them closer to positional allomorphs than to alternative lexemes. Besides, their irregularity must be seen in the context of the Irish nominal morphology, where their endings are not idiosyncratic by themselves. For these reasons, treating Irish counting plurals simply as another lexically restricted group of irregular plurals would be misguided. On the basis of their semantics, of their morphological structure, and of the syntax of number in Irish numeral constructions, I will argue that these counting plurals are lexical nouns acting as grammatical classifiers, and segmenting the reference domain in a way that is alternative to that of inflectional number. For this reason, counting plurals never have unambiguously plural suffixes, but always involve non-affixal modifications of the basic stem: they are inherently plural stems, not stems turned into plural nouns by inflectional number. But they fulfil a grammatical function, and only appear in one construction. Apart from their

intrinsic interest, therefore, Irish counting plurals are theoretically instructive because they show how a stem may be plural 'lexically', as a result of morphological stem formation rather than syntactic inflection, while at the same time acting as a grammatically conditioned allomorph of a lexeme.

The structure of this chapter is as follows. Sections 6.2 and 6.3 lay down the empirical evidence, the former describing Irish counting plurals and the latter placing them in a comparative context, highlighting the special status of unit nouns with respect to number. Section 6.4 focuses on the semantic specificity of unit nouns and other weakly differentiated concepts, proposing that they structure their domain of reference as a non-free, rather than free, join semilattice. Having thus made precise the semantic bases for viewing counting plurals as a natural class and not just a collection of surviving old plural allomorphs, I turn in Section 6.5 to the morphological side of the question. That section will argue that counting plurals systematically lack a morpheme analysable as an exponent of plurality, and are instead unsuffixed stem alternants. Section 6.6 finally merges the semantic and the morphological evidence into a morphosyntactic analysis of counting plurals as inherently plural stems in a classifier structure.

6.2 Numeral constructions in Irish

The syntax of numeral constructions in Irish is a rather complex matter, with different structures for different types of numeral expressions (see Duffield 1995: 323–41). However, plural nouns in the special counting forms only occur in the most basic context NUMERAL (3–10) + NOUN. After a quick overview of the other numeral constructions, this section will present the data that represent the core empirical concern of this chapter.

6.2.1 *The background*

As Corbett (2000: 211) remarked, numerical quantification makes grammatical number functionally redundant. It is no surprise, therefore, that many languages deviate from the correlation 'singular–one, plural–more than one' in that context. The peculiarity of Irish (and with it of other Celtic languages) is not that nouns can be singular when quantified by numerals above one, but that the use of grammatical number depends on the numeral and the precise pattern varies considerably on a dialectal basis.

The number 'one' of course governs the singular. However, Irish has no proper numeral for 'one'. The etymologically corresponding form would be *aon*, which still means 'one' in Scottish Gaelic; in modern Irish, however, *aon*

has this meaning only when the noun is followed by the adverb *amháin* 'only'; in fact, *amháin* can replace the numeral altogether.[1] Besides, 'one cat' can appear simply as 'cat':

(6.1)   a.   (aon) chat amháin
                (one) cat only

         b.   cat
               'one cat'

This use of the singular would not call for further comment, if the examples did not introduce the issue of word-initial consonant mutations. Notice that *aon* triggers the initial mutation known as lenition on *cat* in (6.1a), in this case turning the velar stop [k] into the homorganic continuant [x], orthographically *ch*. All numerals from one to ten trigger one of the two initial mutations that in Irish serve as morphological markers for a great variety of syntactic and morphological functions (for an overview, see Ó Siadhail 1989: 111–34). The official standard, following the most common dialectal pattern, prescribes lenition from one to six and nasalization from seven to ten. However, dialects differ from each other not only with respect to which kind of mutation is triggered by which numeral, but also with respect to which nouns exceptionally deviate from the respective general pattern. The essential observation for our purposes is that the patterns of mutation after numerals do not exactly coincide with those of grammatical number (see especially Ó Siadhail 1982). Since the considerable dialectal and lexical (and possibly idiolectal) variation with respect to mutation does not appear to match patterns of morphological irregularities, I will not rely on mutation patterns in the analysis of counting plurals, limiting myself to point out that dialect-specific accounts should throw some light on this complex interdependence.[2]

    The next thing to note is that 'two' governs the singular. This derives from a former stage of the language, in which nouns had a special form after the numeral 'two' (Thurneysen 1980 [1946]: 154–5). In fact, the dual still has a morphological residue in the form of special allomorphs of some nouns,

---

[1]   In fact, *amháin* has a much more subtle interpretation, approximating 'alone' rather than 'only'. When not interpreted as a numeral, *aon* is a polarity determiner like the English *any*.

[2]   The synchronically capricious nature of mutation patterns with numerals ultimately go back to language change (lenition typically represents the preservation and partial generalization of the pattern of older neuter plurals), and the dialectal variation reflects different patterns of morphologiza tion. Notice that a correlate of the lack of lenition is the insertion of /h/ before a vowel initial noun such as *uibhe*, as in *trí huibhe* 'three eggs'.

exemplified by the palatalized form *cois* in (6.2b); see Hughes (1994: 633), de Bhaldraithe (1977 [1953]: 29), Ó Sé (2000: 223) for data from the three main dialects of Ulster (North), Connacht (West), and Munster (South), respectively.

(6.2) a. dhá chat   (cat sɢ, cait ᴘʟ)
         'two cat.sɢ'
       b. (mo) dhá chois   (cos sɢ, cosa ᴘʟ)
          '(my) two feet'

If 'one' introduced mutations, 'two' introduces another noteworthy feature of Irish numeral constructions: the possible mismatch between the number of the noun governed by a numeral and the number of the DP headed by it. Note that the noun phrase headed by the 'dual' is—or at least can be— singular, as illustrated by the singular article *an* in the following example, taken from Ó Dónaill's (1981 [1978]) standard Irish-English dictionary (*s.v. dhá*):

(6.3) tá an   dá   Bhrian ann
       is the.sɢ two   Brian there
       'there are two sides to Brian's character'

This mismatch will play a role in our analysis (see Section 6.6 below).

    Before turning to the units three to ten, a word is in order about the rest of the numeral system. Following a common typological trend, large round numbers govern the singular: *fiche/céad/míle/milliún cat* '20/100/1,000/1,000,000 cat.sɢ'. As for unit numerals in between round numbers, like 23 or 1,224 for example, the noun they govern is a direct complement of the unit, and so any conclusion about, say, 'three N' carries over to 'four hundred and twenty three N'. Numbers between eleven and nineteen deserve a special mention because of their syntax, characterized by the linear order 'numeral—N—teen' (see Duffield 1995: 323–32 for the complex syntactic issues raised by this structure):

(6.4) trí   chat   déag
       three cat.sɢ teen
       'thirteen cats'

Besides cardinal and ordinal numerals, Irish possesses a series of so-called 'collective' numerals, which fuse unit numerals with a suffix originally derived from the noun *fear* 'man' and are used as pronouns or determiners. In the latter function, they modify a noun in the genitive plural, as if (6.5b) below read literally 'a three-strong human collection of women':

(6.5)   a. triúr
             'three people'

         b. triúr   ban
             three.COLL   women.GEN.PL
             'three women'

A collective like *triúr* effectively combines a numeral and a classifier. To anticipate, this is the analysis I will extend to all instances of unit numerals, to the effect that 3–10 must be followed by a syntactically realized classifier. Counting plurals, I will argue, are nouns whose meaning allows them to function a such a classifier.

6.2.2 *Unit numerals three to ten: regular and irregular patterns in the standard*

Since this chapter focuses on a set of exceptions to the regular pattern, we must have an idea of the regular pattern. But modern Irish as a native language is fragmented in geographically discontinuous dialects, coexisting with English and with the standardized Irish of education and modern communication in a complex sociolinguistic situation that often makes for an unstable and variegated 'norm'. In this context, some generalizations are problematic, and the generalization as to the number value a noun must have after 3–10 falls into this category.

The official standard prescribes the singular (*An caighdean oifigiúl* 1975 [1958]: 39), and dialect descriptions confirm that this value is dominant or even near-exclusive in some dialects and at least widely acceptable in others.[3] Ó Siadhail's (1982, 1989) claim that the singular is also dialectally the rule after 3–10 seems to be too heavily influenced by the Connacht dialects, though, and for this reason I will discuss the dialectal facts in some detail in Section 6.2.3. The fact remains, however, that all dialects admit the singular after 3–10 as an option for many or most nouns, and no dialect generalizes the plural, while some generalize the singular. It seems then justified to take it as a genuine generalization that nouns after 3–10 are singular in Irish, as here exemplified (recall that the series 3–6 triggers lenition while 7–10 triggers so-called nasalization):[4]

---

[3] Pedagogical grammars differ on points of grammar such as the use of number after numerals; GGBC prescribes the singular as the default, while Mac Congáil (2004: 190) allows for either singular or plural.

[4] Corbett (2000: 212) states that in Irish 'the plural is used with numerals "three" and above, and the normal singular with "one" and "two", according to Dochartaigh (1992: 62, 77)'. In fact, Ó Dochartaigh asserts that 'in enumeration…plurals begin with "three"' (p. 62) and 'with both *aon* and *dhá* the singular is used' (p. 77). Ó Dochartaigh is obviously right, because the plural, *when it is used at*

(6.6) trí/ceithre/cúig/sé   chat   seacht/ocht/naoi/deich   gcat
      'three/four/five/six   cat.sɢ   seven/eight/nine/ten   cat.sɢ'

Some nouns, as stated above, fail to respect the rule that 3–10 govern the singular. The description of the standard in the Christian Brothers grammar edited by Ó hAnluain (1999:70, henceforth cited as GGBC), distinguishes two lists of such exceptions: one with nouns that appear as plurals after 3–10, and one with nouns that appear in that context with a special plural form:

(6.7) a. Nouns taking the plural after 3–10:   (GGBC 1999: 70)

Singular	Plural
ceann 'head (as a unit), one'	cinn
cloigeann 'head (counting persons)'	cloigne
troigh 'foot (measure)'	troighthe/ troigheannaí
slat 'rod (measure), yard'	slata

   b. Nouns taking a special plural form after 3–10:

Singular	Plural	Plural after 3–10
bliain 'year'	blianta	bliana
fiche 'twenty'	fichidí	fichid
pingin 'penny'	pinginí	pingine
seachtain 'week'	seachtainí	seachtaine
scilling 'shilling'	scillingí	scillinge
uair 'time, occasion'	uaireanta	uaire

These two lists obviously group together nouns with a similar meaning: measures, units of the counting system and units of time, as well as *uair* 'time, occasion'. The role of the unit interpretation is highlighted by the observation (GGBC 1999: 70) that *ceann* and *cloigeann* cease to be an exception when they literally mean 'head':

(6.8) a. bhí   trí   cheann   ar   an   bhfathach
         be.ᴘᴀsт three head.sɢ on   the giant
         'the giant had three heads'

   b. ceithre   shlat   sailí
      four   rod.sɢ willow.ɢᴇɴ
      'four rods of willow'

---

all, appears only from 'three' up. But he did not make explicit that, in most cases, the plural is not used. Probably in order to save space and gloss over the complications, he chose to state two exceptionless generalizations without adding that the singular is a default after 'three' and above.

6.2.3 *Plurals after 3–10 in the dialects*

An examination of the living dialects in the three main areas reveals a more complex picture. The Connacht dialect of Cois Fhairrge has the following special plurals, according to the thorough morphological description of de Bhaldraithe (1977 [1953]: 7); non-standard forms are underlined for convenience:[5]

(6.9) | SINGULAR | PLURAL | PLURAL AFTER 3–10 |
|---|---|---|
| bliain 'year' | blianta | bliana[6] |
| ceann 'head (as a unit), one' | <u>ceanna</u> | cinn |
| pingin 'penny' | <u>pinginneachaí</u> | <u>pinginne</u> |
| scór 'score (measure)' | <u>scórtha</u> | <u>scóir</u> |
| uair 'time, occasion' | <u>uaireantaí</u> | uaire |
| ubh 'egg' | <u>uibheachaí</u> | <u>uibhe</u> |
| scilling 'shilling' | scilleacha | scillinge |
| slat 'rod (measure), yard' | (<u>slatrachaí</u>) | slata |
| troigh 'foot (measure)' | (<u>troigheannaí</u>) | troighthe |

Ó Siadhail (1982: 102–4, 1989: 167–8) lists more nouns that can take the plural after 3–10 in the usage of Connacht (not necessarily all from the Cois Fhairrge dialect):

(6.10) | SINGULAR | PLURAL |
|---|---|
| bord 'load' | boird |
| ceathrú 'quarter' | ceathrúnaí |
| cloch 'stone (of weight)' | clocha |
| dual 'strand' | duail |
| galún 'gallon' | galúin |
| lámh 'hand' | lamha |
| mála 'bag(ful)' | málaí |
| punt 'pound' | puint |
| stór 'storey' | stóir |
| ualach 'load' | ualaí |

Although the lists are longer than in the standard, these data from Connacht confirm the choice of plural after 3–10 as a lexically restricted phenomenon.

---

[5] De Bhaldraithe (1977 [1953]: 7) explains that the plurals *slateachaí* and *troigheannaí* alternate in use with their shorter counterparts *slata* and *troighthe*; however, only the shorter forms are admissible after numerals. Ó Siadhail (1989: 166) gives the plurals for 'penny' as *pingne* and *pingneachaí*.

[6] I adopt the modern spelling *bliana* for de Bhaldraithe's *bliadhna* (/blʲíənə/).

The presence of the non-measures *uair* 'time' and especially *ubh* 'egg', however, means that this class does not consist entirely of standards for measurement.

With the Ulster and especially Munster dialects, by contrast, the acceptability of plural after 3–10 becomes too widespread to be called exceptional. For Ulster, Ó Baoill (1999: 104) states that 'singular or plural may be used although the usual number is singular'; he adds that some nouns are mostly plural, giving ten examples with morphologically regular plural and eleven with a special counting plural (almost all matching those in use in the standard or in the Connacht dialect). More exceptions, some with special plurals, are specific to Munster dialects (Ó Siadhail 1982: 102–4, 1989: 166–8):

(6.11) | Singular | Plural | Plural after 3–10 |
|---|---|---|
| bád 'boat(ful)' | báid | |
| fear 'man' | fir | feara |
| nóimint 'minute' | nóimintí | |
| punt 'pound' | púint | |
| réal 'sixpence' | réalacha | réal(t)a |
| seachtain 'week' | seachtainí | seachtaine |
| turas 'time, occasion' | turais | |

Many more are uncovered by a thorough investigation of a single Munster dialect. In his detailed description of the Corca Dhuibhne dialect, Ó Sé (2000: 225–8) mentions many nouns as appearing in the plural after 3–10; the following are among those that do not figure in any of the previous lists:

(6.12) Nouns taking the plural after 3–10 (*Corca Dhuibhne*)
Feminine:
bó 'cow', cnámh 'bone', maidean 'morning', méar 'finger', naomhóg 'small boat'

Masculine:
capall 'horse', cos 'foot', dosaen 'dozen', focal 'word', fód 'strip of land', piunt 'pint', saoitheach 'vessel, jar', tigh 'house', bua 'life, gift', suíochán 'seat', gort 'field'

It seems clear that plural is broadly available after numerals in Munster dialects, and that the choice of number in that context varies across dialects.

6.2.4 *Generalizations over the irregular pattern*

Despite the dialectal variability, some broad generalizations are discernible. First of all, as Ó Siadhail (1982: 104) recognized, the nouns that appear as plurals after 3–10 cluster around a core made up of terms of measure: 'It

therefore seems reasonable to assume that, apart from nouns forming an integral part of the counting system (i.e. *ceann, fiche, fear, cloigeann*), there was a small group of "measure units" which took a plural form after cardinal numbers. This group was analogically expanded to various other units of measurement'. This conclusion must be somewhat qualified, however. Even in the restricted lists of the standard and of the Connacht dialect, units of measure are accompanied by concepts like 'years', 'eggs', and 'weeks', which are referential terms and not abstract standards of countability like for instance 'yard' or 'shilling'. And if the category of measures and unit nouns is a little too narrow for these varieties, it clearly cannot explain the vast and possibly open class of nouns that can occur in the plural after 3–10 in Munster. But some notion of unit nouns must be relevant even for these dialects, because the long lists of examples of plurals after numerals systematically include the core measures that are also plural in the other dialects; besides, a look at the examples in (6.12) reveals a preponderance of inanimate objects, some of them likely to be conceptualized as undifferentiated units, like 'cows', 'fingers', 'horses', or 'strips of land'.

In addition, special counting forms also display some uniformity in the exponence of plurality. In Irish, noun pluralization takes a number of forms, as we will see in more detail in Section 6.5. Two of these involve adding to the stem a short vowel (spelled -*a* when neutral and -*e* when in the context of a palatal consonant), or palatalizing the stem's final consonant. The interesting thing about counting plurals is that, as Ó Siadhail noted, they only display these two exponents, instead of consonantal affixes and of the -*í* suffix. Combining the doublets listed above in (6.7) and (6.9), we can see that most counting plurals end in a short vowel, contrasting with the corresponding normal plural which ends in long -*í* as in (6.13a) or in a suffix as in (6.13b) (typically extended by -*í* in Connacht forms like *uibh-eachaí*):[7]

(6.13)	Singular	Plural	Plural after 3–10
a.	pingin 'penny'	pingin-í	pingin-e
	scilling 'shilling'	scilling-í	scilling-e
	seachtain 'week'	seachtain-í	seachtain-e
b.	bliain 'year'	blian-ta	blian-a
	pingin 'penny'	pinginn-eachaí	pinginn-e
	slat 'rod (measure), yard'	slat-r-achaí	slat-a

---

[7] In addition to the morphological process involving the ending, some forms involve stem adjustments, either in the singular or in the plural. For example, the stem of both plurals *blianta* and *bliana* 'years' ends in non palatal *n*, contrasting with the palatalized form of the singular *bliain*.

uair 'time, occasion'   uair-eanta   uair-e
ubh 'egg'   uibh-eachaí   uibh-e

Other counting plurals involve a palatal stem-final consonant, possibly involving stem allomorphy. This is what happens in *fiche* ∼ *fichid(-í)*, and also with *ceann* ∼ *cinn*, which looks like an ablaut form but in fact results from the automatic adjustment of the short stem vowel to the [−back] quality of the following consonant (cf. Ó Siadhail 1989: 36–7, 146, 149; Ní Chiosáin 1991: 137–44):

(6.14) | SINGULAR | PLURAL | PLURAL AFTER 3–10 |
|---|---|---|
| ceann 'head (as a unit), one' | ceann-a | cinn |
| fiche 'twenty' | fichid-í | fichid |
| scór 'score (measure)' | scór-tha | scóir |

Palatalization and the addition of a short neutral vowel are also common exponents of plurality on non-counting plurals; but on counting plurals they are the only ones. In connection with the semantic coherence of this class in particular, this further uniformity calls for an explanation.[8]

The dialectal survey undertaken in this section also raises questions about the proper formulation of the relevant generalizations. Special counting plurals undeniably centre around measure and unit nouns, but are not strictly limited to these non-referential concepts (cf. 'eggs', as in Italian); and dialects vary in the use of the plural after 3–10. More generally, we would like to know why precisely these semantic and morphological properties come together they way they do in counting plurals.[9]

6.3 Unit nouns and number in comparative perspective

Irish special plurals are alternants of nouns, which are a lexical class, but their contextual restriction makes them akin to classifiers, which are grammatical

---

[8] As is well known (Greene 1974, 1992), the consonantal affixal plurals of modern Irish noun morphology are a late development, caused by the gradual loss of the ancient plural exponents due to phonetic erosion. The singular after 3 10 arose because the new endings were not extended to numeral contexts, where plurality is semantically implied. Stem final palatalization and neutral vowel endings likewise derive from the older plural endings. This, however, does not explain why some nouns retained a form synchronically distinct from both singular and plural, why this form was restricted to counting contexts, and why this three way opposition concerns predominantly but not exclusively units of measurement.

[9] The morphological generalizations may be extended to a predominance of feminine among special counting plurals (Ó Siadhail 1982: 103; see also Acquaviva 2006). This undeniable tendency is, however, even less absolute than that concerning number, and only really observable in the Connacht dialect. In order to focus on number, therefore, I will leave it out of the picture here.

elements. No direct parallel is possible with canonical classifier languages like Chinese, which lack inflectional number; but classifier-like constructions also appear in languages with nominal number. The cross-linguistic evidence reviewed in this section shows that Irish counting plurals are an instance of a more general pattern.

6.3.1 *Unit nouns as classifiers*

That nouns referring to standards of quantification can be morphologically and syntactically irregular should not be surprising. We have considered in some detail the irregular patterns of Italian *-a* plurals, which include measurement terms as a homogeneous subclass, and we saw (Section 2.6.4) that the lexical dual *-mayim* in Hebrew applies to units of time and of the counting system. More examples are easy to find: in Russian, for instance, units of measurement are a prominent semantic class among masculine nouns with zero-ending in the genitive plural; these also include *raz* 'time', exactly parallel to the Irish *uair*. What is specific about Irish counting plurals is that they only emerge as a morphosemantic category in one grammatical context: complement to unit numerals (3–10). In this, they resemble classifiers, which encode countability by means of a free morpheme interpreted as a standard of atomicity. So-called classifier languages require such expressions whenever a noun has a count interpretation. Mandarin Chinese, for example, has a system of classifiers that encode different criteria for individuality, distinguishing for instance reference to individual instances and to species:

(6.15)   a. san   zhī   xíong   (Mandarin Chinese; Krifka 1995: 399)
             three   CL   bear
             'three bears' (objects)

         b. san   zhǒng xíong
             three   CL   bear
             'three bears' (species)

The line dividing classifiers as grammatical morphemes from lexical nouns used as standards of individuality is not always clear-cut. Languages vary in this respect, but at least in systems like Vietnamese, the same lexical item can function either as a noun made countable by a classifier, or as a classifier for another noun (cf. also Bisang 1999):

(6.16)   a. hai   cái   bao   (Vietnamese; Löbel 2000)
             two   CL.thing bag
             'two bags'

b. hai   bao   cam
   two   CL.bag   orange
   'two bag(fuls) of oranges'

Generally, languages in which classifiers are a necessary concomitant of the count reading lack inflectional number marking on nouns. This well-known correlation suggests that grammatical number and a grammaticalized classifier system are alternative morphosyntactic means to turn the basic concept associated with a lexical noun into a denotation with a discrete part structure, as required by the syntactic context (among the rich literature on this complex topic, see Greenberg 1974; Allan 1977; Link 1998: 213–21; Cheng and Sybesma 1999; Aikhenvald 2003: 243–52; Borer 2005: 87–135; and note 5 in Chapter 2).

6.3.2 *Unit nouns and grammatical number*

Occasionally, languages with inflectional number also have classifiers. In (6.17), for instance, a numeral may govern a plural noun or a classifier, in which latter case the noun appears in the bare form identical to the singular:

(6.17) a. θalaaθ   baqraat   (Omani Arabic; Greenberg 1974: 26)
         three   COW.FEM.PL
         'three cows'

     b. θalaaθit   rwaas   baqar
         three   CL   cattle
         'three cows'

See also Borer (2005: 94–5) for a similar example from Armenian. In English, too, lexical nouns can have the function of packaging a domain of reference, while at the same time exhibiting morphosyntactic irregularities which suggest a classifier-like status. The best known example is *head* in the construction *three head of cattle.* This differs from a purely semantic classifier construction like *three bottles of wine* because of the lack of plural marking on *head*. But if the lack of plural marking shows classifier status after numerals, then other cases should be included which do not (or do not necessarily) govern a noun expressing the domain being discretized. English, especially in its non-American dialects, has other instances of 'illogical' singulars in a numerative context:

(6.18) a. three dozen/score/hundred/thousand/million   (British English)
     b. three bob/quid/pound/grand/cent/euro/stone/fathom

The units of the counting system in (6.18a) can be followed by a complement noun (*three dozen horses*, *a few hundred miles*) and have an alternative use as lexical nouns outside of quantifying contexts, with regular plural morphology (*dozens of horses*, *hundreds of miles*). The measures of currency, weight, and depth in (6.18b), on the other hand, cannot govern a complement noun, and only *pound*, *cent*, and *euro* admit a plural form with the same meaning (*three euro/euros*, but *three bob/*bobs*). But these differences are much less important than the similarities between (6.18a) and (6.18b). Both lists comprise unit nouns that must or at least can appear without plural marking after semantically plural quantifiers, like numerals above one or *a few*. In a nutshell, canonical classifier structures resemble *three head of cattle*; *head* in this function resembles the unit nouns in (6.18); and among these, the units in (6.18b) parallel the Irish special plurals in that they match a unit noun semantics with an irregular number morphology.[10]

This match between unit noun interpretation and 'illogical' singular in a numerative context is nothing strange. German has a wide range of such constructions:

(6.19) a. drei Mark/Pfund/Pfennig/Kilo/Gramm/Fuss/Faden
           'three mark.SG/pound.SG/cent.SG/kilo.SG/gram.SG/foot.SG/fathom.SG'

b. drei Sack Kohle   drei Glas Wein   drei Korb Kartoffeln
    'three sack.SG coal'   'three glass.SG wine' 'three basket.SG potatoes'

c. drei Mann   drei Stück
    three man.SG   three piece.SG

The units of currency, weight and length in (6.19a) are parallel to those in (6.18a), except that in German there is no plural alternant (except for *kilos*).[11] Example (6.19b) illustrates instead lexical nouns in their use as abstract measures: *drei Sack Kohle*, for instance, refers to the equivalence class of all portions of coal with the required size, regardless of whether they occur in sacks, and the same applies to the parallel examples. The lack of plural morphology on these nouns, their non-referential reading as standards of quantification, and the lack of any preposition to introduce the complement noun, make these particularly close to proper classifier structures.

Finally, (6.19c) illustrates the classifying function of lexical nouns in a more dramatic way. *Mann* expresses both a standard of atomicity (a single person)

---

[10] Note that the singular of constructions like *three foot tall* or *six day week* is a different phenom enon, because any N can be singular in that attributive position if it has a suitable unit interpretation.

[11] The lack of plural (in fact, of number morphology) on units of measure after numerals is attested elsewhere in Germanic. Swedish, for example, features a list of irregularly 'singular' measures very similar to (6.19a) (Holmes and Hinchliffe 2003: 33).

and the domain of quantification (people), just like for instance *bob* or *quid* express a unit and what they are units of (namely currency). Unlike the English examples in (6.18b), however, *Mann* and *Stück* are also full-fledged lexical nouns, which give rise to minimal pairs like the following, where the regularly plural *Männer* and *Stücke* refer specifically to 'men' and 'pieces' while their counterparts used as classifiers refer more generically to humans and objects respectively:

(6.20)   drei Mann 'three people'   ∼   drei Männer 'three men'
         drei Stück 'three (relevant objects)' ∼   drei Stücke 'three pieces'

Since both alternants can occur in the same context, they are not automatically selected contextual allomorphs; the grammaticalized version of *Mann* and *Stück* in their classifier function are not only semantically, but also grammatically distinct from the corresponding lexical nouns. This recalls the opposition of the Irish *ceann* 'head' and *slat* 'rod' in their literal sense and as abstract units for counting (cf. (6.8) above).

  Viewed against this backdrop, the number irregularities of Irish unit nouns in numeral constructions seem much less exotic. True, they are irregular because they are plural (and irregularly so, for special forms), while the Germanic unit nouns we have seen so far are irregular because they are singular. But in both cases the irregularity concerns number morphology, which is also affected in a host of other phenomena involving standards for quantification. I mentioned in Section 2.7.2 the existence of Slavonic special 'quantifying plurals', as well as the Dutch plural doublet *stukken* ∼ *stuks*, which expresses through different plural exponents the opposition between 'pieces' and 'items' which German expresses through the use of number (cf. (6.20) above). English too, beside the exceptional singulars in (6.18), has the special plural form *pence*, selecting the reading of *penny* as an abstract unit of value rather than that as concrete coin:[12]

(6.21)   a. I inserted three pennies/??pence, but they were old and scratched
         b. these two pennies/*pence are not exactly alike

Pence can occur outside of numerical contexts (including a fixed phrase like *St Peter's pence*/*pennies*), and so more closely resembles a semantically specialized lexical plural than a grammatical classifier. But my hypothesis is

---

[12] I would like to thank Jennifer Petrie for discussion on the use of *pence*. In American English, the opposition equivalent to *pence* ∼ *pennies* would seem to be *cents* ∼ *pennies*, where the abstract value reading is encoded by a different lexical root altogether.

precisely that Irish counting plurals are semantically specialized plurals with the grammatical function of classifiers.

   One last clarification is in order. A suitable meaning is of course a necessary but by no means sufficient condition for the status of grammatical classifier. *Bob* never admits a plural form, *three euro* coexists with *three euros*, and *three dollars* cannot be replaced by *three dollar*, as far as I know. Usage varies, and although some factors may influence it (e.g. *bob* has only the abstract sense, while *euros*, like *dollars*, can refer to physical objects as well), the bases for this variation lie outside the realm of grammar. The same applies to Irish, and explains the strong variation in the use of counting plurals as well as the uncertain status of some forms. The non-deterministic aspects in the use of Irish special plurals, then, provide another parallel with unit nouns in other languages. I will now focus on this unit-noun reading.

6.4 The semantics of unit nouns

The morphological similarities between Irish counting plurals and unit nouns used as classifiers in other languages certainly aid an understanding of why measures and other standards for quantification are grammatically irregular. But 'eggs' or 'times' are neither measures nor standards for quantification; and 'years' and 'weeks' are units for segmenting the flow of time, but they do not partition a mass-like domain like *head* does, say, in *three head of cattle*. There must be a better semantic notion to capture what is shared by all Irish special plurals.

   This is the goal of this section. I will make precise the sense of 'unit nouns' as expressions whose denotation is structured in parts that carry criteria for atomicity but not for identity. A plural unit noun thus denotes something conceptualized as made up of units indistinguishable from each other; its reference domain distinguishes pluralities on the basis of their different sizes, not of the identity of their elements, forming a structure modelled by a non-free join semilattice.

6.4.1 *Identity*

To have a plurality, we need things that are not identical with each other, since something described just as *x*, *y*, and *z* cannot be termed plural if $x = y = z$. This elementary consideration does not apply to units of measure, as we have seen in Section 4.3.2. For example, while the reference of *two books* ranges over two-membered collections of books, *two miles* does not refer to a collection, but expresses a length quantified according to a certain standard. As Bunt

(1985: 74–81) explains, *two litres* is but one possible designation for an equivalence class of (liquid) matter, equivalent to *two thousand millilitres* or *3.5196 pints*. Units of measure thus name fixed standards for converting something into a numerical value; functions, that is, from extensions into numbers. They may have the grammar of count nouns, but not their reference (see Quine 1960: 244–5, and Greenberg 1974: 33), and in this light their deviations from the grammatical behaviour of count nouns is not surprising.

The problem is that we cannot reduce the denotation of all Irish counting plurals to measure functions. Apart from 'eggs', which certainly carries an ontological commitment to the existence of objects, *uaire* 'times' is as referential as English nouns like *occurrence* or *instance*, which in any case are not measure functions; and *seachtaine* 'weeks' and *bliana* 'years' can name time-segments which can be located and to which properties can be attributed, so that we can say *these weeks have been better than the previous ones*, while a pure measure function like *litre* requires a complement to express an entity: *this litre *(of beer) was better than that.*

A closer look at measure nouns, however, can reveal what they have in common with terms like 'eggs': the lack of identity properties. As we saw in Section 4.3.2, unity and identity are distinct characterizations in the formation of individual concepts. In particular, an entity conceptualized as an instance of a sort can be distinguished from other instances if it has identity properties that allow it to be re-identified as one and the same. Unity, on the other hand, refers to the properties defining wholes as opposed to masses or arbitrarily delimited cuts in a larger domain. The independently justified distinction between unity and identity implies that entities may be conceptualized as having only one of the two; in particular, as having unity but not identity. Terms of measurement, which define a standard for unity, clearly fit this description, because they define a criterion for what is one and what is not-one (whether more or less) along a dimension of measurement, but they do not refer to individual entities which may be re-identified in isolation. Nouns used as unit counters therefore represent a case where, in the words of Guarino and Welty (2000), 'we may be interested in counting tokens without caring about the possibility of distinguishing one token from another'. But the description can also fit the conceptualization of terms for actual objects, if they are viewed as discrete but undifferentiated. Of course, units of measure are just abstract standards, not objects, while eggs are physical objects and years are time intervals identifiable by the events they contain. However, unity and identity refer to *conceptual* properties that hold both for objects and for abstract units; what the two have in common is that they both define ways to be 'one' (unity). Both, therefore, can also lack identity: abstract units

necessarily, for their very meaning; similar-looking objects like eggs possibly, if they are conceptualized as interchangeable tokens. Counting is the context that most naturally favours a conceptualization of the counted objects as tokens, which can be told apart and thus counted because of the unity conditions encapsulated in the noun; but these entities are not conceptualized in a way that might allow re-identification of any of them in isolation. Counting a domain involves treating its members as equivalence classes, keeping the elements distinct from each other but disregarding which is which. Irish counting plurals are specialized for this interpretation.

6.4.2   *The reference domain of unit nouns*

Guarino and Welty's principled distinction between discreteness (based on unity) and identifiability as the same individual (based on identity) gives us the conceptual tools to characterize Irish counting plurals as unit nouns. This is precisely the function of classifiers. Now, a typical classifier construction involves a complement expressing the domain of partition, like 'pen' in 'three-unit-pen'. At first sight, this seems different from 'three-eggs-as-indistinguishable-tokens'; but the difference is only apparent. It has been noted (Greenberg 1974: 31–2, Allan 1977: 288, Link 1998: 217; cf. Aikhenvald 2003: 98–124 for a typological overview) that classifiers are grammatically more strongly linked to numerals (or other quantifiers) than to the complement noun expressing the domain of partition. There is nothing typologically or semantically unusual, then, in a construction like 'numeral + criterion for atomicity', where the domain of partition is expressed by the unit noun itself, or by the morpheme corresponding to it in cases of 'collective' numerals like the Irish *triúr, ceathrar, cúigar*...'three/four/five...people'. Our definition of unit nouns as lacking identity criteria, in sum, enables us to treat them as classifiers even when they are not strictly speaking measure functions, as 'eggs' would be if it could measure out a mass of non-eggs, as in a hypothetical 'three eggs of water'. This also clarifies the sense in which a construction like the German *drei Mann* (Section 6.3.2 above) differs from *drei Männer*. The denotation of both expressions ranges over three-membered groups, but the former description discounts the individual properties that unequivocally distinguish each person from all others (the identity criteria associated with the lexeme MAN, realized as *Mann*, plural *Männer*). This results in a conceptualization of objectively distinguishable entities as indistinguishable (but discrete) tokens, just like *pence*. Small wonder, then, that *Mann* used as a mere criterion for granularity refers to humans, not specifically to men; human beings are the prototypical individuals. Indeed, the Irish 'collective'

numerals *triúr, ceathrar, cúigar*… 'three/four/five…people' (not 'men') seem
to have arisen from *trí*, etc. + *fear* 'man' (Greene 1992: 513–16), like the
German *Mann* in *drei Mann.*

The intuition that Irish counting plurals lack individuality can thus be
made precise in terms of lack of identity criteria. Moreover, it can be correl-
ated to a specific formal property. Consider first a domain consisting of the
entities in the set {*a,b,c*}. If the atoms are individual entities and their
combinations are sets, then the set {*a,b,c*} has the subsets {*a,b*}, {*a,c*}, and
{*b,c*} (disregarding the empty set). Considering the atoms as (singleton) sets
themselves would allow one to see also {*a*}, {*b*}, and {*c*} as related to the other
sets by the inclusion relation (see Landman 1989). Alternatively, the combin-
ations of atoms may be treated as sums, that is not sets but individual
elements in their own right, related to each other by the part-of relation
(this is the mereological approach inaugurated by Link 1983). The atom *a* is
then a part, not a member, of *a+b*, which is a part of *a+b+c*. Intuitively, this
defines a hierarchy where the basic atoms are parts of larger two-membered
sums which in turn are parts of the sum of all three atoms:[13]

(6.22)
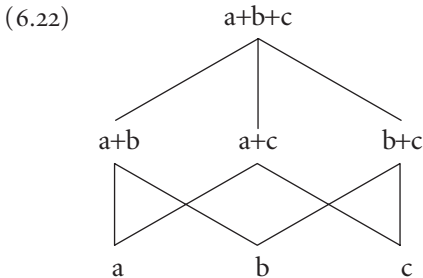
This hierarchy (the atoms and their sums) has specific formal properties.
Stretching the intuitive notion of 'part' to 'part-of or equal to' makes the
relation reflexive (every element is part-of or equal to itself, namely equal-to),
transitive, and antisymmetric (if two elements are parts-of or equal-to each
other, they are the same element). A relation with these properties, defined
over the set comprising atoms and their sums, makes it a partial order. Every
element of this set stands in this relation with some element from the same set,
possibly itself. The intuitive stricter notion of part, which relates *a* to *a+b* but

---

[13] Since the purpose of this section is merely that of showing how unit nouns differ from referential
nouns, I very briefly introduce here the notions necessary for distinguishing free from non free
semilattices; see Simons (1987), Ojeda (1993), and Link (1998) for the analysis of reference domains in
terms of the algebraic properties of their part structure. Zabbal's (2002) analysis of Arabic plurals also
contains a very clear exposition of the relevant notions.

not with itself, defines the corresponding strict order (irreflexive and transitive). Representing as $\leq$ the 'part-of or equal-to' relation, we can say that $a+b$ is the smallest element x belonging to this set such that $a \leq x$ and $b \leq x$. Formally, this is the *join* of a and b, or the smallest sum containing them. Summing in this way any two members of the set, their join is still in the set (for instance, the join of a and $a+b+c$ is $a+b+c$ itself). This means that the partial order corresponding to the set of atoms and their sums is closed under join: it is then a *join semilattice* (technically, this is complemented by a *meet semilattice* based on the relation opposite to $\leq$ to form a lattice).

Consider now the following way to group $\{a,b,c\}$ into a unique sum:

(6.23)

```
              a+b+c
             /  |  \
            /   |   \
           /    |    \
          a     b     c
```
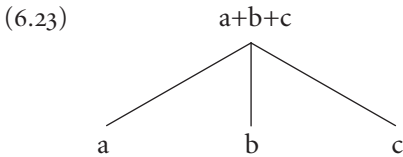
Formally, this is too a join semilattice: the $\leq$ relation orders all elements, and every subset of the set (which now comprises a, b, c and the only available sum, namely $a+b+c$) has a minimal sum, which happens to be always $a+b+c$. This does not correspond to what we would intuitively see as the reference domain of a plurality; it must be possible to refer to intermediate sums which are not the total $a+b+c$. As Landman says (1991: 262), in order to formalize a plurality with a lattice structure, 'what we want to insure is that if individuals a, b, and c are distinct, then the sums $a \lor b$, $a \lor c$, and $b \lor c$ should all be distinct' (where $\lor$ represents the join operation). This is lacking in (6.23); therefore, that is not a proper representation for a domain of reference that ranges over individuals. Intuitively, what is missing are the intermediate sums, a fact reflected in the formal property that the different subsets of $\{a,b,c, a+b+c\}$ do not have a distinct sum for each distinct subset; $\{a,b\}$ and $\{b,c\}$ for instance are distinct subsets, but they share $a+b+c$ as the *smallest* sum comprising them. A lattice where every distinct subset has a distinct sum is *free*; (6.23) is non-free ('freedom means that whenever two pairs of elements are distinct, their unions are distinct'; Szabolcsi and Zwarts 1993: 264). But precisely because the need for separate sums for distinct subsets is related to the individuality of the basic atoms, a non-free join semilattice can model the reference domain of non-individual terms. This is the way Szabolcsi and Zwarts (1993) have analysed the reference of amount expressions. Consider an amount corresponding to three times the amount of a unit:

(6.24)

$$[x+y+z]\ (=3)$$

$$[x+y]\ (=2)$$

$$[x]\ (=1) \qquad [y] \qquad\qquad [z]$$
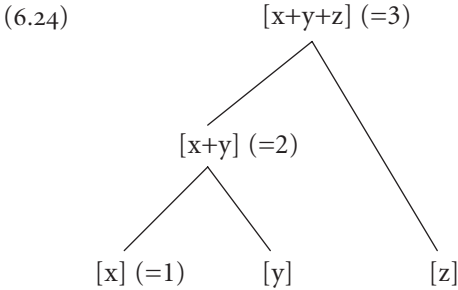
The sum of *x* and *y* defines an amount corresponding to two units, and the sum of that plus another unit *z* makes up an amount of three units. But there is no sense in which *y+z* or *x+z* would form a two-sized amount different from *x+y*. These intermediate sums are therefore absent from the graph, making the structure a non-free join semilattice. 'Here [x], [y] and [z] are all unit-sized, though they are not unit-sized bits of concrete stuff, but arbitrary (and therefore abstract) unit-sized bits' (Szabolcsi and Zwarts 1993: 267). This is the status of units of measurement; but it is also the status of factually distinct objects conceptualized as undifferentiated tokens. Applied to humans, (6.24) describes the reference domain of the German *Mann* in *drei Mann*, where the noun's descriptive content restricts what the units must be true of (the domain of partition). The regular plural *Männer*, by contrast, refers as in (6.22), because it ranges over sets of men made distinct by the identity of each member. Like *Mann* in *drei Mann*, Irish counting plurals conceptualize their referents as interchangeable; to make this intuitive characterization precise, we can now say that their reference domain has the structure of a non-free join semilattice.

We have thus reached a precise characterization of what it means for a noun to be semantically a standard for measurement, or a criterion for atomicity. What unifies measure nouns, classifiers and collections of undifferentiated units is the lack of distinctive individual properties associated with the lexical predicate. In Irish as well as in English and German, and in many other languages, measure terms represent the core case because they are the most prototypical equivalence classes.

6.5 Counting plurals as unsuffixed stems

Having reached a better understanding of the function of special counting plurals, we will now concentrate on how this function is grammatically represented. As anticipated in Section 6.1, the morphology of counting

plurals, analysed within the context of Irish noun morphology, suggests that they are inherently plural stems rather than bimorphemic constructs of stem and a number affix. This section will present the evidence for this claim.

6.5.1 *The exponence of regular and counting plural nouns*

Our preliminary description of counting plurals in Section 6.2.4 has already brought to light their characteristic exponence with respect to other patterns of nominal plural formation. As we have seen, the forms in question express plurality in one of the following three ways, the third of which is a combination of the first two:

(6.25)   a. palatalization (*scór* ∼ *scóir* 'score ∼ scores')
          b. vowel extension (*bliain* ∼ *blian-a* 'year ∼ years')
          c. vowel extension + palatalization (*ubh* ∼ *uibh-e* 'egg ∼ eggs')

Recall that cases like *ceann* ∼ *cinn* are the result of palatalization of the stem-final consonant. As shown by Ní Chiosáin (1991: 140–51), short vowels in Irish are not specified for the feature [± back] but receive this specification from a following consonant; the backwards spreading of [–back] automatically raises a preceding low short vowel, but not a long one like that of *scór*, plural *scóir* /skoːrʲ/ (not **scír*). Other such cases include *fear* 'man', plural *fir, mac* 'son', plural *mic, leanbh* 'child', plural *linbh*.[14]

Many other plural nouns share these three patterns of exponence with counting plurals. But, and this is crucial, many nouns also display distinct endings which are systematically absent among counting plurals. Consider a few examples, taken from Ó Siadhail (1989: 149–64):

(6.26)   a. palatalization
             bád ∼ *báid* 'boat ∼ boats', *bord* ∼ *boird* 'table ∼ tables',
             fear ∼ *fir* 'man ∼ men', *punt* ∼ *puint* 'pound ∼ pounds'

          b. vowel extension
             ceann ∼ *ceanna* 'head ∼ heads', *lámh* ∼ *lámha* 'hand ∼ hands',
             muc ∼ *muca* 'pig ∼ pigs'

          c. palatalization + vowel extension
             éadach ∼ *éadaighe* /ˈediː / 'garment ∼ clothes',
             Sasanach ∼ *Sasanaighe* /ˈsasəɴiː/ 'Briton ∼ Britons'

---

[14] The assignment of [back] to a short vowel apparently raises an underlying low vowel, but not a mid one: cf. *scoil* 'school', where a fronted mid vowel precedes the final palatal consonant (Ní Chiosáin 1991: 140).

d. consonant extension (+syllabic readjustment)
   ainm ~ *ainm-eacha* 'name ~ names', *am* ~ *am-anna* 'time ~ times',
   bliain ~ *blian-ta* 'year ~ years', *leabhar* ~ *leabhar-tha* 'book ~ books',
   ní ~ *ni-the* 'thing ~ things', *slí* ~ *sli-te* 'way ~ ways'

e. *-í* suffix
   cóta ~ *cóta-í* 'coat ~ coats', *rása* ~ *rása-í* 'race ~ races',
   seachtain ~ *seachtain-í* 'week ~ weeks'

f. combined consonant extension
   leaba ~ *leab-tha-cha* 'bed ~ beds', *uair* ~ *uair-ean-ta* 'time ~ times'

g. consonant extensions + *-í* suffix
   carraig ~ *carraig-re-acha-í* 'rock ~ rocks', *uibh-each-aí* 'eggs'

Over and above the numerous complications arising from stem allomorphy and from dialectal variation, and factoring out the possible co-occurence of plural exponents (a distinctive trait), the system of Irish nominal plural formation can be schematized as follows:

(6.27) a. stem-final palatalization

b. extension through a short neutral vowel with no inherent [–back] value

c. extension through invariant affix, underlyingly a consonant or the long palatal *-í*

Counting plurals systematically express plurality by the first two strategies alone. It so happens that the two classes thus defined, (6.27a–b) and (6.27c), have distinct status in the morphology of Irish.

6.5.2 *Two types of plural exponents*

Neither palatalization nor vowel extension are restricted to the expression of plurality. The former is also used for genitive singular, plural, dative singular, vocative, and comparative in nominal morphology, as well as third person masculine in prepositional inflection (Ó Siadhail 1989: 135–9). Moreover, Ní Chiosáin (1991: 162–5) notes that since suffixes beginning with a palatal vowel do not usually palatalize the previous consonant (e.g. *bonn* ~ *bonnaíocha* 'coin ~ coins', with coronal /ɴ/ throughout), those that do (e.g. the comparative in *bán* ~ *báine* 'white ~ whiter', with palatal /ɲ/ in the latter) must involve vowel extension and palatalization as two distinct processes. This in turn shows the relevance of both processes outside of plural formation. As for vowel extension, it is also used for verbal subjunctive, plural on adjectives,

genitive singular on nouns, vocative plural on petrified expressions such as *feara* 'men!' (Ó Siadhail 1989: 139–41), as well as for the derivation of abstract nouns (Doyle 1992: 28–31). The other endings, namely *-í* and those involving consonantal extensions, are instead almost exclusively exponents of plurality, and indeed on nouns only (the exceptions being *-ta* and *-tha*, which enter in the formation of the verbal adjective: *scríobh* 'write', *scríobhtha* 'written').

That (6.27) is in fact a two-way distinction is confirmed by another observation, which again separates (6.27a–b) from (6.27c). Irish nouns distinguish two cases, a default (traditionally the nominative) and a genitive, with traces of other cases (vocative and dative) restricted lexically or surviving in frozen expressions. On examination of the plural case forms, two patterns emerge. One category, traditionally called strong plurals, morphologically contrasts two case forms in the singular but neutralizes the case distinctions in the plural. This is exemplified by the nouns *bliain* 'year' and *carr* 'car':

(6.28) Strong plurals:

	SINGULAR	PLURAL	SINGULAR	PLURAL
NOMINATIVE	*bliain*	*blian-ta*	*carr*	*carr-anna*
GENITIVE	*blian-a*	*blian-ta*	*cairr*	*carr-anna*

So-called weak plurals, here illustrated by *fear* 'man' and *muc* 'pig', preserve instead the opposition between nominative and genitive plural:

(6.29) Weak plurals:

	SINGULAR	PLURAL	SINGULAR	PLURAL
NOMINATIVE	*fear*	*fir*	*muc*	*muc-a*
GENITIVE	*fir*	*fear*	*muic-e*	*muc*

Strong plurals systematically correlate with those with exponents other than palatalization or vowel extension, that is with (6.27c). The correlation is not perfect, because even the genitive plural itself is increasingly rare as a distinct morphological form, and so there are cases where a noun has an innovating plural with a consonant extension and an older genitive plural limited to fixed phrases; for example, *tobar* 'well' has a consonantal plural *toibreacha* but maintains *tobar* in petrified phrases (Ó Siadhail 1989: 150, 165). Even considering these exceptions, however, the correlation between strong plural pattern and exponents of class (6.27c) is synchronically extremely tight.

6.5.3 *Affixation vs. stem modification*

To sum up, then, plural exponents fall into two natural classes: one, comprising vowel extension and palatalization, can express other grammatical functions and when it appears on a plural noun may fail to do so for all its grammatical

forms (i.e. they feature in weak plurals). The other class comprises all the rest and has the following properties: its exponents are fully specified for all phonological features (including [±back]), they only express nominal plurality, and when they do so they appear on both case forms (i.e. they form strong plurals).We can make sense of these regularities by a simple hypothesis: only the exponents in (6.27c) are nominal plural suffixes, the others are not. This means that only the consonantal extensions and -*i* 'mean' plurality if and only if they appear on a noun; they alone are category-specific mappings of sound and one meaning, namely plural. The other exponents, vowel extension and palatalization, are morphological tools that can express one of several grammatical meanings, as vehicles for functional oppositions. Unlike the affixes in the previous category, these do not 'mean' plurality; there is no special palatalization nor vowel extension dedicated to plurality. There are, instead, stems with a palatalized and non-palatalized alternant, vocally extended and not extended. But the final -*a* of *blian-a* is not a plural suffix, even though it looks similar to the final -*i* of *cóta-i*. If it was, the properties just illustrated that align it with palatalization and oppose it to the other plural endings would be a coincidence. Claiming that they are not accidental amounts to claiming that palatalization and vowel extension belong together, as ways to modify a stem for grammatical purposes. Consonantal extensions and -*i* are instead plural suffixes.

What of counting plurals? Recall that not all weak plurals are special counting forms; but all counting forms are weak plurals. That is, among the two main strategies for expressing plurality on a noun in Irish, through suffixes or through stem alternants, counting plurals systematically choose the second. This, again, can hardly be a coincidence, and is a further reason for separating the two types of exponents. All counting forms are stem alternants rather than suffixed stems. The next section will propose an explanation for this convergence of semantic and morphological properties.

6.5.4 *The syntactic representation of affixal and non-affixal plurals*

Of course, an added vowel can always be analysed as a suffix; and palatalization, too, can in principle be regarded as a suffix, a formant with phonological information attached to the right edge of the stem. But just because it is possible to treat them as suffixes does not mean it is correct. In particular, an analysis of palatalization and vowel extension as suffixes should explain why it is precisely the putative plural 'suffixes' whose segmental make-up is not entirely determined until they combine with a stem, that serve as exponents for a variety of other grammatical values beside plural, and *correspondingly* fail to appear in all plural case forms, as real suffixes regularly do instead.

The hypothesis that palatalization and vowel extension are properties of monomorphemic stem alternants is better equipped to account for the much looser connection between these formal traits and plurality. No subpart of *muca* 'pigs', for instance, specifically 'means' plurality, unlike the plural suffix -*ta* in *blianta* 'years'. *Muca* fills the inflectional cell of non-genitive plural by virtue of its being distinct from *muc*. To pursue this approach, however, I must clarify two things: the nature of what I have been calling stems, and their precise relation with number as a morphosyntactic property.

As regards the first point, which will be pursued in greater detail in the next chapter, I analyse stem forms derived by palatalization and/or vowel extension on a par with other stem forms. They are all instantiations of a lexeme in a grammatical structure. If the word form involves a segmentable affix spelling out an inflectional category, then the stem is the part of the word without the affix; for example, *blianta* 'years' has the stem *blian-*, to which -*ta* attaches to express (in this case) plurality. The same noun features a distinct stem form in the singular: *bliain*, with final palatalization. This shows, again, that palatalization is a standard way to affect stems in Irish, not a morpheme with its own meaning. The counting form *bliana* is a distinct alternant again, based on the non-palatalized stem that appears in *blian-ta*. Schematically, then, the lexeme YEAR is realized by one of the stem alternants *blian*, *bliana*, or *bliain*, the first of which supports a suffix. In other cases it is the palatalized alternant that has an affix: EGG has a non-modified stem form *ubh* /ʊv/, which spells out the singular (nominative), and an alternant *uibhe* /ivə/, with palatalization and vowel extension; this expresses on its own the genitive singular and the counting plural, and is suffixed to give the regular plural *uibhe-acha*; compare also (6.28) and (6.29) above. These few cases suffice to show the utter lack of biuniqueness between a type of stem form (basic, palatalized, extended) and a grammatical function (see Ó Siadhail 1989: 150–68). These alternants, then, correspond to Aronoff's (1994: 44) definition of a stem as 'the sound form on which a particular form is built', with the important clarification that each stem here has exactly one form (like the Latin *b*-stem, used for all imperfect verbs and for the future of -*a* and -*e* verbs, but unlike the more abstract present, perfect, and 'third' stems; cf. Aronoff 1994: 58–9).

As for the second point, clarifying how these stems interact with number means clarifying what role they play in the morphosyntax of Irish nouns. As a first approximation, I will propose the following structures for suffixal and non-suffixal plurals respectively:

(6.30) a.

```
              Number P                      b.            Number P
          ╱────────────╲                            ╱────────────╲
      Number          NP                        Number          NP
      ╱───╲            │                        ╱───╲            │
   Nᵢ-Number           │                     Nᵢ-Number           │
   │      │            │                     │      │            │
   │      │            │                     │      │            │
  blian-ta            tᵢ                    muca-Ø              tᵢ
```
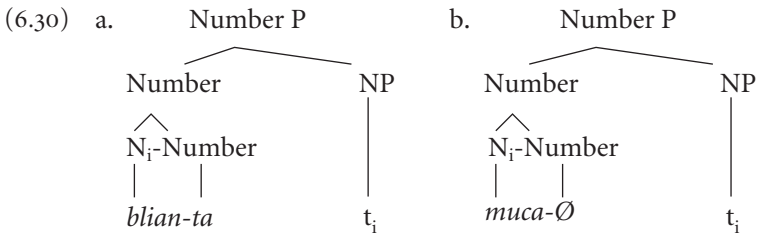
The grammatical information that defines inflectional plurality is syntactic-ally encoded in a head immediately above NP. For ease of exposition I use here the traditional label Number, but this syntactic locus of plurality corresponds more precisely to Borer's (2005) 'Classifier' (hosting a 'Division' head), and to Heycock and Zamparelli's (2005) [Plural] (cf. Sections 4.2.1 and 6.6.2 and Chapter 9). A noun acquires grammatical plural features as a result of raising and adjoining to this head. This is the case for both choices of noun, because both are associated with number features as abstract grammatical informa-tion. But for *blian-* this requires the spell out of Number as a separate head, realized as *-ta*; *muca*, on the other hand, is the form of PIG specified to spell out N° when it combines with a [plural] Number head. The bimor-phemic *blian-ta*, then, involves spelling out two heads, N and Number, while *muca* only of N; both reflect the plurality encoded on Number, *blian-ta* by realizing it directly and *muca* by selecting a stem alternant conditional on its presence. I will refine this syntactic analysis in Chapter 9; for now, what matters is not the precise phrase structure, which is more complex, but the fact that weak plurals like *muca* spell out just the noun and not the noun plus Number.

Such unsuffixed stems, then, express plurality intrinsically; they represent the form taken by the corresponding lexeme in a grammatically plural context. In a way, then, all such plurals are lexical, in the sense that plurality is a property of the stem itself. This is very different from the sense in which *brethren* or the Italian *braccia* 'arms' are lexical; *muca*, along with all weak plurals, is perfectly integrated in the inflectional system of Irish (exactly as strong verbs, which likewise do not express tense by an affix, are perfectly integrated in the inflectional systems of Germanic languages). The grammat-ical 'meaning', the plural feature in this case, is exactly the same for suffixal and non-suffixal plurals, encoded on the syntactic head Number. But non-suffixal plurals do not spell out Number; and that is the reason why all special counting plurals fall into this class. This is so, because the syntactic structure they appear in, I will now argue, has no head corresponding to Number.

6.6 Irish counting plurals as inherently plural classifiers

Irish counting plurals are semantically unit counters, and morphologically stems without a plural affix (a trait they share with all weak plurals). The two properties converge to define a class of lexical stems used in a grammatical capacity, as classifiers governed by unit numerals. Irish counting plurals, then, are lexical plurals that only arise in a particular syntactic structure. This final section will spell out this idea in syntactic terms.

6.6.1 *Irish numerals and classifiers*

To implement this intuition, two assumptions are needed: that 3–10 in Irish always govern a syntactically realized classifier, and that this classifier is incompatible with the syntactic projection of number. Jointly, this determines the pattern of special counting plurals.

The second assumption, that classifiers and syntactically projected number are mutually exclusive, does not need much justification. Although some classifier languages have a grammatical category of plurality, (cf. references in Section 6.3.1 and note 5 in Chapter 2), it is a well-established fact that classifiers and inflectional plurality are alternative ways to express countability on nouns. In a previous analysis of Irish numerals (Acquaviva 2006), I posited two mutually exclusive syntactic projections for classifiers and for number features. In the light of Borer's (2005) work, however, it seems more revealing to view the two as alternative contents for one and the same syntactic projection, which is the grammatical locus for segmenting a noun's reference. Borer (2005) calls this projection Classifier Phrase, headed by an operator which, as we saw in Section 4.2.1, encapsulates information about how the noun's reference is divided. A reading as continuous mass corresponds to a division into all possible parts of the denotation; a count reading instead corresponds to a uniform division into elements, made possible by a fixed criterion of unity (as noted, Heycock and Zamparelli 2005 make a similar proposal). A classifier is a word that performs this regular division, typically by naming one of several ways to be a discrete entity in the language; a super-category like, say, 'human being', or 'elongated object'. What I called Number in Section 6.5.4, following traditional usage, accomplishes this division by grammatical features. In a count reading, this corresponds to the information which defines the noun's reference as ranging over all atoms described by the noun (with singular value) or over all sets of such atoms (with plural value). It seems natural to view the two mechanisms for achieving a stable division of reference, by a word

(classifier), or by grammatical features (inflectional number), as mutually exclusive.

That Irish numerals should syntactically select a head with the properties of a classifier requires more justification. Since counting requires a division into units, the semantic aspect of this requirement is not in question, nor is under discussion the typological generalization that, in the sequence NUMERAL–CLASSIFIER–NOUN, the first two elements form a unit, regardless of linear order (see Sections 6.3.1 and 6.4.2). Indeed, classifiers are often morphologically merged with numerals, giving rise to special numerals for counting people, or even to several distinct series of numerals (Anderson 1999 describes reduplicated numerals for people and animals in Salish, and several series of numerals in the Paleo-Siberian language Nivkh). As pointed out in Section 6.2.1, Irish too has a series of such 'collective' numerals, which govern nouns in the genitive plural (as in *triúr ban* 'a three-people of-women'), or take no complement at all. Common numerals, by contrast, cannot occur without a complement. In Irish, a question like 'how many cats do you have?' cannot be answered literally by 'three'; the numeral must be followed by something, as in *trí cinn* (a counting plural), literally 'three heads', or 'three ones'. A suitable context may make it perfectly clear what is being counted, but this would not ameliorate the ungrammaticality of a unit numeral without a complement.[15] The only context where Irish numerals occur on their own is when they do not quantify a domain, but simply spell out the progression *one*, *two*, *three*, and so on (they must be preceded by *a* in that use, and some forms are different, for example *a ceathair* 'four' but *ceithre chat* 'four cat[s]'). I take this to mean that the Irish unit numerals 3–10 are not strictly speaking equivalent to their counterparts in other languages. The English numeral *three*, for instance, can be used for counting a domain, meaning 'three times a unit'. The Irish *trí* only gives the first part, the element in the arithmetical progression; in order to count anything, it must be combined, as a grammatical requirement, with something meaning 'times a unit'. And, as Greenberg (1974: 21) remarks, 'all classifiers are from the referential point of view merely so many ways of saying "one" or, more accurately, "times one"'. What is missing from Irish unit numerals, then, is precisely this classifier component, which requires a separate syntactic expression in a counting context.

---

[15] In Scottish Gaelic, by contrast, 'three' can be simply *a trì* (Greene 1992: 532). Note that Scottish Gaelic 3 10 govern plural nouns, unlike in Irish. However, in neither language(group) can numerals act as predicates, as in *the apostles were twelve* (thanks to David Adger for information about Scottish Gaelic).

6.6.2 *Number in Irish numeral constructions*

If Irish unit numerals require a syntactically expressed classifier, but a syntactically expressed classifier is incompatible with inflectional number, then Irish unit numerals are incompatible with inflectional number. This means that both the apparently singular *cat* in *trí chat* 'three cat[s]' and the so-called counting plural *uibhe* in *trí huibhe* 'three egg[s]' are neither singular nor plural, but numberless.

This conclusion agrees with the fact that Irish nouns have no ending to express singular number. The same happens in English, where *cats* has a plural morpheme but *cat* is a bare stem and derives its singular reading purely by opposition. Formally numberless noun forms are morphologically well formed in Irish and English, unlike for instance in Latin, where every noun must carry an indication of number, beside gender and case.[16] It is the syntactic context that makes such bare stems 'singular', when they are governed by a grammatical head specified for that value of the grammatical number feature. In (6.30) above, that head is called Number; as we have seen, it corresponds to the dividing operator which heads the Classifier Phrase in Borer's (2005) system. The evidence just reviewed suggests that unit numerals 3–10 require this division to be encoded by a classifier, not by a value for the inflectional number feature. I will therefore assume the following structure for the relevant subtree, where unit numerals appear in a position below the top Determiner Phrase but just above Classifier Phrase. This is Borer's '#P' phrase, which encodes quantity:

(6.31)   [$_{DP}$ Det ... [$_{\#P}$ numeral [$_{ClassifierP}$ Classifier [$_{NP}$ N]]]]

The base structure of *trí chat* (before any DP-internal raising by the noun) will then be the following, with the numeral governing a phonetically unrealized but syntactically projected classifier, which prevents *cat* from being assigned a value for morphosyntactic number:

(6.32)   [$_{DP}$ Ø ... [$_{\#P}$ *trí* [$_{ClassifierP}$ Ø [$_{NP}$ *chat*]]]]

The classifier position is empty, because the atomicity of the domain being counted is identified through the criterion of unity (in the sense of Guarino and Welty 2000) associated with the lexical noun. This would not be possible in canonical classifier languages, where the mass–count distinction as well as all information about the atomicity of nouns must be conveyed through

---

[16] As I argued in Acquaviva (2005), that is why languages like Irish, English, or German, but not Latin or Russian, have pseudo singular noun forms with the function of counters (*three cent*, *drei Mann*): such forms are numberless nominal stems, which are only morphologically well formed in some languages.

classifiers. But Irish is not a canonical classifier language; nouns do have clear countability preferences, and the syntactically realized classifier position reflects a requirement by the numeral, not the lack of unity criteria on nouns. The position can then be realized as zero in presence of a following count noun, and only in that context. In sum, the structure required by unit numerals (except, by hypothesis, in dialects that generalize plurals after 3–10) amounts to a classifier construction without a classifier. Numerals must govern a standard of granularity as a grammatical requirement, and count nouns license this as a value for the dividing operator.[17]

The hypothesis that apparent singulars after 3–10 are numberless explains an otherwise puzzling mismatch between head nouns, which look singular, and noun modifiers, which have plural morphology:

(6.33) trí   chat   dhubha
        three   cat.SG   black.PL
        'three black cats'

It would be very strange if the noun and its modifier had two contrasting values. The hypothesis that *cat* is numberless allows for a more plausible interpretation, namely that the DP as a whole is grammatically plural (hence triggering plural on the adjective), but N is not marked for number, as Duffield (1995: 330) recognized. Interestingly, recent work on the syntax and semantics of DPs suggests that the number of the DP as a whole has a syntactic encoding that is distinct from that of the head noun. Sauerland (2004) summarizes his proposal as follows:

I propose that there is another syntactic head above the determiner, which I call the ϕ head, and that the only semantically contentful number features are contained in this head. Furthermore, I will argue that [Plural] on nouns is not interpreted, but is a reflex of syntactic agreement with a ϕ head, just like [Plural] on adjectives, verbs, and determiners. (Sauerland 2004: 1)

In a similar vein, Heycock and Zamparelli (2005) spread the syntactic realization of number in DP over two heads, of which only the most embedded one (their [Pl], above NP) affects the noun's morphology. These recent theoretical developments, based on evidence totally unrelated to Irish, converge in treating number as a property DP and of N as two separate syntactic

---

[17] Positing instead a contentful but phonetically null classifier would not only be arbitrary and unrevealing; it would also incorrectly rule in structures like *trí Ø uisce* 'three CL water', with a mass noun made countable by the null classifier. It seems probable that the dialects allowing plural after 3 10 allow count nouns to supply the counting criterion to numerals through number, and not (or not necessarily) via a classifier position.

entities. This explains the apparent mismatch in (6.33), if the head of Classifier Phrase above N does not encode inflectional number; N is thereby number-less, while adjectival modifiers agree with the plural number value encoded higher up in the DP (I take this higher head to correspond to the DP-internal Agreement of Duffield 1995: 308–21). A numberless specification for N, then, solves the puzzle presented by (6.33), and this in turn bolsters the hypothesis that numerals require a structure that forces nouns to be numberless.

We can finally return to counting plurals. As should be clear, these are not inflectionally plural, because numeral constructions with 3–10 are incompatible with inflectional number (again, not equally across all dialects). What distin-guishes special counting plurals from other nouns is that they act themselves as classifiers. Only nouns with a suitable lexical semantics can be bent into this grammatical function; and this function only arises in numeral constructions, where the Classifier Phrase cannot encode number features. *Trí bliana* 'three years' has then the following structure, with the extended noun stem *bliana* raised from its base position and occupying the head of Classifier Phrase:

(6.34) $[_{DP} \, \emptyset \ldots [_{\#P} \, trí \, [_{ClassifierP} \, bliana \, [_{NP} \, t]]]]$

Since there is no inflectional number information, no suffix can be spelled out, and so noun stems in this position will never be 'strong' plurals with a number suffix (like *blian-ta*). Weak plurals, which do not spell out number features, may instead occur there in principle, provided their meaning is compatible with the classifier function; unsurprisingly, their choice varies considerably across dialects.

The last question, at this point, is whether counting plurals are plural at all.[18] It is not just that they lack a plural suffix; more radically, they lack number features on the head that usually encodes them (the head of Classifier Phrase, in the terminology I have followed). That is the difference between counting plurals and any other with 'weak' plurals like *muca* 'pigs' (see (6.30) above) which are likewise suffixless stems but in the context of grammatical features. All the evidence we have seen shows that counting plurals do not realize inflectional number. However, they are still plural, not because they spell out grammatical features, but because they express by the very form of their stem a non-singular interpretation—after all, they express transnumeral standards of countability, but the denotation of the DP ranges over sets, not just singletons. This is confirmed by modifying adjectives. Recall that, in a numerical construction involving 3–10, the head noun is not inflected for number but the whole DP is still plural, as shown by the agreeing adjective in

---

[18] Thanks to Christoph Schwarze for discussion on this point.

(6.33) above. Therefore, counting plurals are classifiers but are also intrinsically plural stems, and their formal identity with inflectional weak plurals is not accidental. Thanks to the distinction between number as a property of the noun and of the DP, we can refine the traditional idea that number and classifiers are alternative ways to syntactically express countability: when a language with grammatical number exceptionally replaces number features with a classifier, this affects the inner, noun-related number projection, not the outer, DP-related one. This is natural, since what gets partitioned in a classifier construction is the reference domain of a head noun, not of a whole DP. But only an intrinsically plural form will do, which can express plurality independently of the features on a grammatical head. And this is just what Irish counting plurals are: inherently plural stems.

6.7 Conclusion: Irish counting plurals and lexical plurality

This chapter has examined a lexically restricted class of plurals fulfilling three conditions: their interpretation makes them suitable as unit nouns, they appear in syntactic construction involving a classifier (after 3–10), and they are morphologically non-suffixal. The interplay of these three features gives rise to the complex empirical pattern we have examined.

As lexical plurals go, Irish counting plurals are rather atypical—and that is why they are so instructive in a general survey. They are listed, as a small class whose members must be listed; but their form conforms with one of the two main types of plural morphology for Irish nouns, and their meaning is, if anything, grammaticalized rather than lexicalized; not inextricably fused with the interpretation of specific nouns, but turned into a grammatical operator for partitioning reference domains as non-free join semilattices. This makes them rather different from exceptions like *brethren* or *looks* 'beauty', and even from a morphologically and semantically uniform class like Italian -*a* plurals. No shallow conception of lexicality as idiosyncrasy would capture the peculiarity of this class, which arises as a set of lexicalized forms at the intersection of semantics, morphology, and syntax.

In the terminology of Aronoff (1994), Irish counting plurals are stems, or more precisely invariant stem forms, in any case, realizations of lexemes—not lexemes in their own right, except perhaps where there is a clear semantic difference between counting and non-counting plural: *cinn*, for instance, is formally the plural of *ceann* 'head', but it is used as a generic classifier with numerals, for animates and inanimates alike. But in other cases, like *uibhe* 'eggs' or *uaire* 'times', we are dealing with a peculiar class of allomorphs that spell out a grammaticalized reading as unit nouns, in a classifier position

specific to numerically quantified DPs. Despite the lack of inflectional number features, they bring about a plural reading by filling the head of Classifier Phrase with a specially marked form of the stem. Their peculiarities, then, lie primarily in their syntactic function and in the meaning that derives from it. They express plurality precisely when inflectional number is demonstrably missing, that is in numerically quantified DPs. This is possible thanks to the linguistic properties speakers associate with these words and not others. Still, the classifier-like properties of numeral constructions in Irish are part of grammatical knowledge, not lexical. As lexical nouns used in a grammatical capacity, then, Irish counting plurals are lexical *qua* non-inflectional, but non-lexical *qua* grammatical. They represent the grammaticalized version of lexical plurals.

7

Arabic broken plurals

7.1 Introduction

A long and illustrious grammatical tradition has labelled 'broken plural' (*jamʕ mukassar*, *pluralis fractus*) an Arabic pluralization pattern which consists in reshaping the stem of the singular: for example, the singular *kalb* 'dog', is altered ('broken') to form the plural *kilaab* 'dogs'.[1] Most nouns and many adjectives form the plural in this way; the alternative strategy is called sound plural (*jamʕ saalim*, *pluralis sanus*), and consists of adding a suffix to the singular stem, which normally remains unchanged or 'sound'; for example, *taʕriif* 'definition' ∼ *taʕriif-aat* 'definitions'. Glossing over many complications, the opposition of broken vs. sound plural generally corresponds to that between internal vs. external, ablaut vs. suffixal forms.

The rearrangements brought about by broken plural formation follow several schematic patterns; however, the choice among these patterns cannot be reduced to automatic rules, so that broken plurals must be learned as listed forms—with notable regularities and some entirely predictable patterns, but on the whole still listed. Moreover, some patterns used for the plural of some nouns appear in the singular of other nouns: for instance, the plural *kilaab* 'dogs' has the same prosodic structure and the same vowel melody as the singular *kitaab* 'book' (plural *kutub*). In addition, a given singular can have only one sound plural but possibly more than one broken plurals, often differentiated in meaning and/or stylistic register. Finally, some suffixless stem forms represent non-countable, 'collective' plural alternants opposed to 'individuating' suffixal plurals. For all these reasons, broken plurals are traditionally regarded as lexical plurals, derived by the same stem-forming processes that relate roots (understood as the typical Semitic consonantal patterns) to word-stems. This much is amply known, thanks to a millenary

---

[1] These are the citation form of nouns, without the endings that normally accompany non pausal forms in the standard language. Other works cite nouns with the ending *un* (*u* for some nouns), which marks the default combination of nominative, singular, and indefinite.

descriptive tradition and to a significant number of modern theoretical analyses.²

Under these circumstances, there is no need to go to great lengths to argue that broken plurals (henceforth BPs in this chapter) are the output of a stem-forming process. The goal of this chapter is rather to disentangle the nexus of morphological, syntactic, and semantic properties that accompany these derived stems, with a view to bringing out what is lexical in this expression of plurality—and what is not. I will argue that BPs form a coherent class only in a morphological sense, as stem forms standing in a certain formal and semantic paradigmatic relation with singular noun forms. As a morphological class, they conflate aspects of lexicality that belong to different planes. By teasing them apart, this chapter aims to shed light on the complex interplay between listedness, stem-creating morphology, number as an abstract category, and part-structure conceptualization.

Section 7.2 places BPs in the context of Arabic nominal morphology and clarifies the status of multiple BPs in the competence of speakers. Section 7.3 focuses on the seemingly contradictory morphological status of BPs, summing up the evidence for and against viewing them as lexical plurals. The paradox is solved in Section 7.4, where their lexical properties are traced back to their nature of derived stems, and their inflectional ones to the function of these derived stems in the inflectional paradigm. Having thus clarified the morphological status of BPs, I will turn to their interpretive value, and in particular to the question of lexicalized readings. Section 7.5 describes how the basic singular–plural semantic opposition interacts in Arabic with an opposition in terms of individuation (conceptualization as discrete distinguishable parts), in the light of the notions introduced in Chapter 4 and of some recent formal semantic analyses of BPs. Similarities and divergences with other types of lexical plurals are highlighted in the concluding Section 7.6.

7.2 BPs in Arabic and its dialects

This section will sketch out the morphology of nominal number in Arabic, closing in on BPs. Recent theoretical investigations have derived the apparent

---

² Broken plurals feature prominently in several strands of research on Arabic and Semitic linguistics. Among traditional descriptions of classical Arabic, they are discussed in Wright (1967 [1896]: 199 234), Fleisch (1990: 470 500) and Fischer (1972: 51 8); there is also a wealth of specialist contributions in the philological literature, among which I have relied heavily on Fischer (1980). This tradition has been fully taken into account by recent theoretically oriented investigations: cf. Abd Rabbo (1990), McCarthy and Prince (1990), Ratcliffe (1990, 1998), Ojeda (1992), Belnap and Shabaneh (1992), Belnap (1993, 1999), Zabbal (2002), and Kihm (2003). As for transcription, a dot underneath represents velarization in 'emphatic' consonants is the glottal stop, ħ and ʕ are the voiceless and voiced pharyngeal fricatives, and *j* stands for IPA [dʒ].

unpredictability of BPs from the workings of grammatical rules; however, this cannot suffice to view them as products of automatic and deterministic inflectional processes. Most key issues have to do with the relations between BPs and singulars, or BPs and other BPs; these paradigmatic relations have had little import for purely morphological investigations, but are crucial for the issue of lexicality. For this reason, the final part of this section will bring into focus the status of BP alternations in Arabic as a synchronic grammatical system, as opposed to a historical norm.

7.2.1 *Number within the categories of Arabic noun morphology*

Arabic distinguishes two values for gender (masculine and feminine) and three for number (singular and plural, partly also dual). These are the fundamental categories that underlie the morphology and syntax of nouns and adjectives in all stages of the language. The classical language and the modern written standard also inflect for case, which determines a three-value declension, alternating with a two-value or diptotic one for some grammatical contexts and some lexical choices. Since this has no effect on the encoding of plurality as a morpheme or as a word, I will leave case out of the picture in what follows.

7.2.1.1 *Dual*   Within the number system, dual is peripheral, both for the range of forms it applies to and for its morphology. The modern dialects have reduced the applicability of the dual to nouns, and some varieties restrict it to human or kin terms, or to units of time (Fischer and Jastrow 1980: 89; Brustad 2000: 45–52; cf. Section 3.6.4). Moreover, the dual ending usually has the function of plural for paired body parts, sometimes with a distinctive form. A further sign of lexicalization is provided by the possibility to pluralize a dual, reported by Fischer and Jastrow (1980: 89) for the dialect of the Muslims in Baghdad: *sitt rijl-een-aat* 'six pairs of legs' (where the plural -*aat* attaches to the dual -*een*-). The dual is instead still alive in the contemporary standard I will call modern written Arabic, henceforth MWA (after Badawi *et al.* 2004); but even there it is marked off from the other number values by the regularity of its suffix allomorphs: -*aani* and -*ayni* for nominative and non-nominative case, reduced to -*aan* and -*ayn* in pre-pausal position.[3] For all nouns and adjectives (pronouns and verbs are distinct), dual entails this exponent, and this exponent entails dual.

---

[3] These endings are reduced to *aa* and *ay* in the so called construct state, where two nouns in a genitival relation merge into a syntactic unit, triggering morphological reductions on the first. As this (very important) aspect of Semitic morphosyntax plays no role in plurality, I will not consider it in what follows.

7.2.1.2 *Singular*   No such biuniqueness of form and meaning obtains for singular and plural. In classical Arabic, the singular is marked by a vowel after the stem, whose quality expresses case (*-u-* for nominative, *-i-* for genitive/oblique and *-a-* for accusative); this vowel is then followed by a suffix *-n* for indefiniteness (nunation, *tanwiin*).[4] While masculines have no specific marking, most feminines add a suffix, mostly *-at-* but also *-aa-* or *-a'u*. Other nouns are inherently feminine without these suffixes, mostly but not exclusively semantically motivated nouns like 'mother'. Ignoring case, this results in the following picture:

(7.1)   DEFINITE   INDEFINITE (Thackston 1994: 10)
  MASCULINE   al-malik-u   l-kabiir-u   malik-un   kabiir-un
                    the-king-DEF the-great-DEF king-INDEF   great-INDEF
  FEMININE   al-madiinat-u l-kabiirat-u   madiinat-un   kabiirat-un
                    the-city-DEF   the-great-DEF city-INDEF   great-INDEF

There is evidence that the inflectional endings and the final *-t* of *-at* had been phonetically eroded already by Koranic times, although it is possible that they survived for some time when followed by another word (see Fischer and Jastrow 1980: 15–19 and Holes 2004: 9–18; but cf. also Fleisch 1990: 281–2). In MWA, the inflectional endings are deleted when in pre-pausal position, that is in phrase- or sentence-final position and in the citation form: 'the king' is *al-malik-u*, and its indefinite form is *malik-un*, but this appears as *malik* before a pause. For such a masculine noun, the pausal form is identical with the stem; the corresponding forms for a feminine are *al-madiinat-u*, *madiinat-un*, and *madiinah*, where the pausal form shows the phonological reduction of the stem-extending *-at* suffix.[5]

The modern dialects have lost the inflectional endings altogether, and the old *-at* has become *-a* or a derivative:

(7.2) a.   al-bayt-u   l-kabiir-u   (Classical Arabic)
              the-house-DEF the-big-DEF
       b.   il-beet ik-kibiir (Egyptian, Cairo dialect) (Fischer and Jastrow 1980: 88)
       b'. əl-beet l-əkbiir   (Syrian, Damascus dialect)
       b". al-beet al-kabiir   (Sudan, Omdurman dialect)

---

[4] The distribution of the *n* ending is partly grammaticalized. Diptotic nouns (those with a two rather than three way case distinction) systematically lack it; as some proper names fall into this class, *n* often has nothing to do with indefiniteness.

[5] The addition of endings to a stem can cause phonological readjustments. Stems ending in *iy*, for instance, combine with *u* to give *ii*, not *iyu*; this accounts for forms like *muḥaamii* 'lawyer' (Badawi *et al.* 2004: 52).

(7.3)   a. bayḍ-un ∼ bayḍ-at-un   (Classical Arabic)
            'eggs ∼ an egg'

        b. biḍ ∼ biḍa   (Moroccan dialect; Harrell 1962)

        c. beeḍ ∼ beeḍa   (Syrian, Damascus dialect; Cowell 1964: 300)

With the retreat of inflectional endings, it is the stem that bears the functional burden of expressing number.

7.2.1.3 *Sound plural*   The sound plural (henceforth SP) is so called because it leaves the singular stem intact or 'sound'—although this is not always true, as we will see. The plural suffixes are *-uuna* for masculine and *-aatun* for feminine. This makes for the following contrasts in classical Arabic, again ignoring case and leaving aside pausal forms as not precisely determinable:

(7.4)   MASCULINE SINGULAR   MASCULINE PLURAL

Definite	*Indefinite*	*Definite*	*Indefinite*
as-saariq-u	saariq-un	as-saariq-uuna	saariq-uuna
'the thief'	'a thief'	'the thieves'	'thieves'

(7.5)   FEMININE SINGULAR   FEMININE PLURAL

Definite	*Indefinite*	*Definite*	*Indefinite*
as-saariqat-u	saariqat-un	as-saariqaat-u	saariqaat-un
'the woman thief'	'a woman thief'	'the women thieves'	'women thieves'

As can be seen, SPs are not suffixal in the same way as the English *book-s*. For masculine, the ending *-uuna* is suffixed not to the singular form but to its stem, that is, the form minus its inflectional ending. In the feminine, *-aat* instead replaces the feminine marker *-at-*, to be followed by the same endings. It is only with reference to the pausal forms of MWA that McCarthy and Prince (1990: 211–12) can say that 'the sound plural is formed by suffixation of masculine *-uun* or feminine *-aat* to a usually unchanged stem'. Their examples feature bare stems for masculines, as in (7.6a), and stems extended by *-at* for feminines, as in (7.6b):

(7.6)   SINGULAR   PLURAL

	Singular	Plural	
a.	šuwayʕir	šuwayʕir-uun	'poet (diminutive)'
	kaatib	kaatib-uun	'writing (participle)'
b.	taʕriif	taʕriif-aat	'definition'
	kaatib	kaatib-aat	'writing (fem. participle)'

In a footnote, McCarthy and Prince add that 'In feminines CVCC-*at* the sound plural/CVCC-aat/typically shows an epenthetic vowel in the CC cluster which is either *a* or a copy of the stem vowel', adding the example *kisr-at*

'fragment', pl. *kisar-aat/kisir-aat* (and some more specific conditions on this epenthesis).[6] In fact, the details can be quite complex, especially when all inflectional forms are taken into account. For colloquial Egyptian, for instance, Aboul-Fetouh (1969: 57, 60–1) shows that the suffixes -*iin* and -*aat* attach to the singular stem form in some cases (7.7a) and to a modified stem form in others (7.7b); notice that -*aat* attaches to a base that may or may not be suffixed by -*a* in the singular:

(7.7)   SINGULAR   PLURAL
    a. mudarris   mudarris-iin   'teacher (masc.)'
       mudarris-a   mudarris-aat   'teacher (fem.)'

    b. san-a   sin-iin   'year'
       razil   rizl-iin   'unpleasant person'
       sit   sitt-aat   'lady'
       tiin-a   tin-aat   'fig'

SPs, then, do not strictly speaking add a suffix to the singular, but more precisely to a stem form systematically related (typically identical) to that of the singular.

7.2.1.4 *Broken plural*   In contrast with SPs, BPs change the shape of the stem: *kalb-* 'dog', *kilaab-* 'dogs', *ximaar-* 'veil', *xumur-* 'veils'. The plural stem forms fall into a series of patterns; for instance, *xumur-* and *kutub-* 'books' illustrate the pattern *CuCuC*. The possible patterns are many, although their precise number depends on the criteria for grouping into a single pattern superficially distinct forms. Wright (1967 [1896]: 199–234) recognizes thirty-two patterns (twenty-nine plus three minor ones); Fischer (1972: 51–8) and Fleisch (1990: 472–84) group them slightly differently. Since detailed lists are available in descriptive grammars and in the abundant linguistic literature, I will only give a few examples from classical Arabic:

(7.8)   SINGULAR   BROKEN PLURAL   GLOSS
     kitaab-un   kutub-un   'book'
     rajul-un   rijaal-un   'man'
     ʕayn-un   ʕuyuun-un   'eye'
     himaar-un   hamiir-un   'donkey'
     rukb-at-un   rukab-un   'knee'
     jaar-un   jiiraan-un   'neighbour'

---

[6] For Ratcliffe (1998: 89), instead, 'there are compelling reasons for regarding vowel insertion as the primary marker of plural and the sound plural ending as secondary and redundant.'

kursiiy-un	karaasiiy-u	'throne'[7]
faylasuuf-un	falaasif-at-un	'philosopher'
qadam-un	'aqdaam-un	'foot'
baab-un	'abwaab-un	'gate'

Adjectives too can have BPs: *ḍaxm-un* 'large, fat', *ḍixaam-un*; *kabiir-un* 'big', *kibaar-un*.

Several traits of BPs make their relation with plurality less than straightforward. First, the plural patterns do not 'mean' plural by themselves, because they often appear on singulars; at most, there can be preferences for one pattern or another, (for instance, Kihm 2003: 148 notes that *CuCuC* is restricted to plurals), but overall the form does not carry plural meaning any more than it does in Germanic ablaut plurals like *geese* (cf. *cheese*). Second, BPs have the same inflection as singulars, as evidenced by the *-un* ending in Arabic varieties that have it. There is no special paradigm associated with the BPs in (7.8). Third, patterns can be intercalated with consonantal extensions, or 'affixes' (but remember we are talking about stems), like the *-aan* in *jiiraan-*. The feminine *-at-* has a special role to play, as it appears on some singulars but not on their plurals (cf. *rukb-at-*), or conversely on plurals but not on their singular (cf. *falaasif-at-*). Fourth, the relation between a singular and a BP is not one-to-one. For the standardized varieties, all accounts confirm the frequent presence of more than one plural for a given singular; Fischer (1972: 49) reports for *farx-un* 'young of bird' the plurals *firxaan-un*, *firaax-un*, *furuux-un*, *'afraax-un*, and *'afrux-un*, and Wright's description is especially rich in such multiple formations.

7.2.2 *The grammatical status of BPs*

The non-uniqueness of BP formation shows it is not a deterministic process. All the evidence I have briefly reviewed suggests that BPs provide a plural form for a given singular noun by effectively creating another word based on the noun's root, or more precisely by creating another stem. In spite of local subregularities, this type of pluralization therefore lacks the predictability of grammatical (inflectional) processes, and the presence of multiple outputs is one consequence. This is the unanimous position of traditional accounts, bolstered by the observation that this type of indeterminacy can be found elsewhere in the Arabic grammar: the verbal noun known as *maṣdar* (e.g. *qatl-un* 'the killing, to kill' from *qatala* 'he killed, to kill'), has a plethora of possible forms (44 in Wright 1967 [1896]: 110–12), and it is not predictable which of them a given verb takes; some verbs have multiple forms.

---

[7] The *u* ending is a sign of diptotic (two way) case inflection and characterizes, among other grammatically defined classes, plurals with the structure *CaCaaCiiC*.

However, this position has been challenged in recent years by analyses that have brought out the hidden systematicity of BP formation. In different ways, McCarthy and Prince (1990), Ratcliffe (1990, 1998) and Kihm (2003) have uncovered significant regularities in the pairing of singulars to BPs, which greatly reduce the perceived arbitrariness of the process. In particular, they have emphasized the systematic connection BPs have with the prosodic structure of their respective singular. In effect, the crucial changes in stem shape tend to affect the position between the second and the third consonant, other changes being at least partially predictable in a great many cases. Morphological rules formulated as operations that target prosodically defined domains like the first CVX of the stem (McCarthy and Prince 1990; Ratcliffe 1990, 1998), or the position CVCV   CV in an abstract phonological representation (Kihm 2003), can therefore capture the underlying regularities of BP formation. In tandem with independently justifiable claims about well-formedness constraints on the output form, these theories can bring BP much closer to the determinism of grammatical rules. What matters most is not eliminating the unpredictablity of BP formation, which would be a misconceived goal given the lack of a one-to-one relation between singulars and BPs. The main point is rather that BP formation applies to a base systematically related to the singular, just like SP formation. Ratcliffe (1990, 1998: 36–7) notes this correlation explicitly, and argues that both processes manipulate the singular stem: BP by adding a heavy rhyme (VX) to the initial CVC, SP by adding a mora to its right edge:

(7.9)   a. BP: k a l b   k   l   b   'dog / dogs'
                   <C V X >C   > <C V C V X >C

                                        i   a a   (Ratcliffe 1998: 34)

        b. SP: m u s l i m u n   m u s l i m u u n a 'Muslim / Muslims'
                   <C V C C V C>V C > <C V C C V C>V V C V

                                                          (Ratcliffe 1990: 110)

It would seem, then, that BPs and SPs are much closer to each other than the traditional literature has recognized. Their main difference may well lie in their domain of application, SPs targeting a stem and BPs a phonologically defined sequence (Ratcliffe 1998: 54, 115). This would respect the traditional view that BPs involve the formation of a stem (Level I) while SPs modify a formed stem (Level II), but it would reduce the two to the same kind of morphological operations—the same that can be shown to be at work in Arabic and Semitic morphology at large.

These recent advances concern the formal side of BP formation. However, there is more to this type of plurals than their form alone. BPs, but not SPs, can be several for a given input; why is disjunctivity suspended with them, if they have the same morphosyntactic function as SPs? Why do they differ in agreement, with BPs tolerating both strict and 'deflected' agreement while SPs deterministically demand plural agreeing modifiers? Granted that BPs appear to have a 'singular' inflection because they do not mark plurality by lengthening the stem as SPs do, why is it that the inflection in the two classes also differs in the exponence of indefiniteness, where only SPs neutralize the definiteness opposition in masculines and add a word-final -*a* (cf. (7.4)–(7.5) above)? And why can BPs be occasionally pluralized, but not SPs? Only Kihm (2003) addresses these questions in his analysis of BPs, reaching the conclusion that their observed derivational traits descend not from their formation, as per the traditional approach, but from the intrinsically derivational nature of pluralization (Kihm 2003: 115, 131).[8] I will present in the next sections my reasons for not following this interpretation, but before that a clarification is in order on the place of BPs in the vocabulary of spoken varieties, as opposed to standardized norms.

7.2.3 *The question of the data*

The non-uniqueness of BP formation seems to prove beyond dispute that they do not fall into the same category as deterministically formed SPs. This, however, presupposes that the BP alternants listed in grammars reflect the competence of each speaker. In fact, there are grounds for doubting this assumption. If the non-deterministic character of BPs was just an illusion caused by superficial descriptions, and uncritical use of them in theoretical work, the consequences would be far-reaching. This raises non-trivial questions about the relation between codified standard and internalized speakers' competence. I will now spell out the reasons for taking BP formation as a *potentially* non-deterministic process, and for rejecting classical Arabic as evidence for generalized polymorphism.[9]

---

[8] In fact, Kihm (2003: 136) adds that plurality is encoded as one of the intrinsic values of the abstract element that turns a root into a noun, other values being those for gender or noun class. BPs for him comprise such an inherently plural nominalizer; SPs instead comprise, in their abstract representation, a 'morpheme' (for him, a lexical item) encoding plurality but distinct from that encoding nominality. I accept most of this interpretation, which has shaped the ideas presented in this book, but I see it in terms of the distinction between stem formation and stem external inflection.

[9] I use *polymorphism* rather than *allomorphy* here, because I refer to alternants whose distribution is not determined by grammatical principles.

7.2.3.1 *The survival of BPs in modern vernaculars*   If one is interested in the morphology of BPs, the difference between classical Arabic, MWA and the contemporary vernaculars is not an issue: the dialects may have privileged some pattern or abandoned others, but the formal principles underlying the construction of BP stems remain the same. The same applies to the relation between BPs and SPs: the fundamental principle underlying all varieties of Arabic, past and present, is that noun pluralization primarily involves stem changes, possibly combined with affixes.[10] The SP is used instead, both in classical Arabic and in the modern dialects, for forms already manipulated by stem-creating morphology (Level I in Ratcliffe 1998): feminine derivations (singulative nominalizations, or feminine adjectives, or feminine derivatives of occupational nouns), diminutives, participles, many adjectives, proper names (including those of letters and of numbers), and recent loanwords that do not fit into a BP pattern. There is also a tendency to use -*aat* to pluralize human and kin terms (Fischer and Jastrow 1980: 90). What matters is that SPs are linked to specific morphological categories, not that they are few—they can be a great many, given the productivity of some derivations. But when plurality appears as a SP, there is a morphological reason. BPs are not thus restricted, and show their vitality with recent loanwords like *film* ∼ '*aflaam*. In short, BPs are definitely not relics.

7.2.3.2 *What BPs are available in a synchronic grammar?*   Dialectal differences, then, are not an issue for the formal patterns of BPs. They have a certain importance, however, for deciding whether or not BP formation is deterministic like an inflectional operation, perhaps with some leeway for phonological readjustments. It would be entirely possible, for instance, that the synchronic competence underlying any one dialect should define a unique BP for a given singular base, while speakers may have only a passive knowledge of obsolete or specialized alternants. The polymorphism of BPs would then be a sociolinguistic but not a grammatical fact, and any conclusions based on it about the grammatical competence of Arabic speakers would be unwarranted.

Something like that seems to be true up to a certain extent. For cultural reasons, the Koranic norm as fixed in the eighth and ninth centuries has always been the standard of prestige for Muslim speakers, as well as for non-Muslims living in the same cultural context. The strong linguistic conservatism deriving from this stability of the prestige written norm has led to a situation where speakers' native competence cannot be too sharply divided

---

[10] The dialects know prefixal plurals: Cesaro (1939: 125) mentions some very few plurals with the prefix *m* in the dialect of Tripoli, and Tsiapera (1969: 51) cites some formed with *ti* in the Cypriot dialect.

from their historical competence. For BPs, this means that the abundance of alternative forms described for classical Arabic does not faithfully reflect the linguistic reality of contemporary (and perhaps past) native speakers— regardless of which forms educated speakers may be familiar with. The collective ʕarab 'Arabs, Bedouin Arabs', for instance, is unlikely to have all the plurals ʕuruub, ʾaʕrub, ʕurbaan and ʾaʕraab (listed in Wehr's dictionary of MWA) in the active competence of any one speaker. It would be wrong to take these multiple forms as evidence for the non-disjunctive nature of BP formation as a synchronic rule.

It would also be wrong to associate certain readings with certain shapes of BP. The pattern ʾaCCaaC is traditionally associated with countable pluralities of a few elements (jamʕ al-qilla, pluralis paucitatis; cf. Wright (1967 [1896]: 234). Fischer (1980: 74–7) has argued that the specialization of certain BP patterns for count readings (typically entailing reference to a few elements) must have characterized a stage of the language immediately preceding the early standardization, but no longer valid as a synchronic rule. However, what used to be a fact of grammar has remained as a tendency, as shown by the fact that forms of *pluralis paucitatis* are not deterministically associated with certain singular patterns but are in statistical free variation with other BP patterns (Ratcliffe 1998: 79–80). It may be that certain varieties even today favour the pattern ʾaCCaaC after unit numerals (Fischer 1980: 75), but these sub-regularities are due to historical rather than grammatical competence.

In sum, then, the polymorphism of Arabic BPs typically emerges from standardized varieties, which are not reliable evidence about this particular aspect of the speakers' competence. To ascertain this point, we must turn to descriptions of spoken dialects. Indeed there is evidence, albeit on a much smaller scale, that even in the dialects BPs are not formed deterministically. Although dialect descriptions usually give exactly one BP for each singular, this biunique pairing is not absolute. Harrell (1962: 101) states that in Moroccan 'it is not uncommon for a word to have two or more plurals, sometimes equivalent in meaning, sometimes different in meaning'. Cowell (1964) says that 'many nouns have different plurals corresponding to different meanings ... sometimes different plurals are stylistically significant ... in still other cases, different plural forms may be virtually equivalent, or a matter of personal or regional variation' (Cowell 1964: 371; examples omitted). Aboul-Fetouh (1969: 65), for Egyptian, reports the variants *xaddamiin, xuddaam, xadam,* and *xadama* for the plural of *xaddam* 'servant', and *šamʕaat, šamʕ,* and *šumuuʕ* for the plural of *šamʕa* 'candle'. I will therefore assume that BP formation is non-deterministic in principle, in the sense that the system is so structured as to tolerate doublets, even if their number in actual usage may

be small—certainly much smaller than in the historically stratified standard. This potential for polymorphism can be, but must not be, functionally exploited to distinguish different readings.

7.3 The lexicality of BPs

Intuitively, BPs are lexical because they seem to derive anew the noun they pluralize, rather than add a morpheme to it. Yet, as we have just seen, they can be derived from the singular by specifiable rules—not entirely predictable, but still rules. The picture is not clear because the evidence is contradictory; to clarify the picture, I will now review in detail the empirical properties that speak for (Section 7.3.1) and against (Section 7.3.2) lexicality.

7.3.1 *Why BPs are not inflectional*

What sets BPs apart from inflectional plurals is first and foremost their non-deterministic relation with singulars. In addition, two more properties are distinctly non-inflectional: the semantic specialization of some alternants, and the possibility to be input for derivation and indeed pluralization.

7.3.1.1 *Lack of inflectional blocking*   If the competence of Arabic speakers does not fully determine a unique BP for a given singular but merely individuates a few possible output patterns, from which historical accident picks out one or more, then pluralization comes about through the interaction of both grammatical and non-grammatical knowledge. Because the possibility of one-to-many pairings between singulars and BPs has such a fundamental importance, I have taken some care in the last section to establish its existence as a genuine property of Arabic as a grammatical system, even if a given spoken variety may have only a few doublets.

The relation between BPs is in principle distinct from that between BPs and SPs, but here too disjunctivity does not apply with the absoluteness expected from a grammatical principle. For example, *saariq* 'thief' in MWA may have either the SP *saariq-uun* or one of the BPs *saraqa* and *surraaq*; and the plural of *sana* 'year' is the SP *sin-uun* or the form *sanawaat*, which combines stem change with the -*aat* ending (Holes 2004: 166 reports that the latter has become the norm today). According to Badawi *et al.* (2004: 767), this holds of 'some words' (illustrated by four examples other than *saariq* and *sana*), but only 'occasionally' does the choice of a plural serve to distinguish different meanings; semantic differentiation is thus not a precondition for multiple plurals. This applies to adjectives as well as nouns: Ratcliffe (1998: 114)

mentions two forms for the feminine elative *suhla(y)a* 'easier', namely *suhal-* and the SP *suhlay-aat-*.

7.3.1.2 *Semantic lexicalization*   If choosing between plural forms entails choosing between distinct meanings, and the choice is not driven by principles of the grammar, then these forms stand with each other in the paradigmatic relation of words with other words. Although analyses primarily interested in the formal array of BPs do not usually distinguish the senses of the various forms, dialectal descriptions make clear both that semantic differentiation is possible, and that it is not necessary. Cowell (1964: 371) and Harrell (1962: 101), cited in Section 7.2.3.3 above, provide the following examples from Syrian and Moroccan respectively: *lsaan* 'tongue', pl. *lsaanaat* (anatomical) and *'əlson/'əlsine* (languages); *xeddam* 'worker, servant', pl. *xeddama* 'workers' and *xdadem* 'units of agricultural land measurement'. A particular case of semantic differentiation is that between the kind- and the instance-reading of mass nouns, discussed in Section 4.5.1. We will examine more closely the relation between plurality and part structure in Section 7.5.

7.3.1.3 *Number-independent morphological status*   Among the arguments for the lexicality of BPs there is also the relative independence of BP from inflectional grammatical information, including plurality itself. Kihm (2003: 112–13) observes that the form of BPs is not determined by the gender of the noun or adjective; this contrasts with the gender-differentiated exponents of SP, which preserve the inflectional opposition of the singular. Besides, the often noted fact that the patterns of BPs are not restricted to plurals means that the patterns themselves are not exponents of plurality. While these facts do suggest lexicality, however, they are still compatible with a view of BPs as inflectional plurals brought about by re-shaping the singular stem by regular changes. An analysis like that of Ratcliffe (1998) could achieve this, although Ratcliffe (1998: 71) explicitly rejects the hypothesis that 'all variation in the form of the noun plural in Arabic is conditioned by phonological or morphological features in the singular' as 'too strong for a synchronic account of the Arabic data'.

The observation that a BP stem can serve as base for category-changing derivation provides at first sight a much stronger argument. Discussing the great productivity of adjectival formation through the relational (or *nisba*) suffix *-ii* in MWA (classical *-iyy-un*), Holes (2004: 161) cites examples like *barnaamaj waθaa'iqii* 'documentary programme', literally 'programme documents.REL', where the second word is the BP of *waθiiqa*. Another example is *jamʕiiya nisaa'-iiya* 'women's society', lit. 'society women-REL', with the suppletive plural of *nisaa'-un* from *imr'at-un* 'woman'. Badawi *et al.* (2004:

746) also illustrate several doublets formed by affixing the *nisba* ending to the singular or the BP, as for instance *jumhuurii* 'republican' (from the singular *jumhuur* 'crowd, mass') versus *jamaahiiri* 'of the masses' (from the BP *jamaahiir* 'crowds, masses'). However, Holes also shows that SPs as well as BPs can feed *nisba* suffixation, as in *'imaar-aat-ii* 'Emirate-PL-REL', from the singular *'imaar-at* 'emirate'. It is not the status of BPs, then, that is crucial for this adjectival derivation from a plural noun. While the fact that a plural can feed adjectival suffixation remains interesting, this argument for the special status of BPs as opposed to SPs is only apparent.[11]

7.3.1.4 *Double pluralization*   A stronger piece of evidence that BPs are plurals because of their paradigmatic relation with singulars, and not because they 'mean' plurality, is that they can themselves feed pluralization. In a few cases, this is straightforward because the formal pattern of a BP becomes detached from its grammatical value. For example, the noun *balad*, plural *bilaad*, translated as 'country; town, city; place, community, village' in Wehr's dictionary, is related to a distinct singular–plural pair *bilaad* ∼ *buldaan*, glossed as 'country' (exemplified with names of states). The same form *bilaad* can thus serve as a grammatical plural (for *balad*) and as a grammatical singular (for *buldaan*), in the latter case feeding BP formation. Of course this is to some extent exceptional, here made possible by the ambiguous conceptualization of 'towns' as a designation for a region. But the significant thing is that the BP form of *bilaad* does not prevent this admittedly exceptional reanalysis. As a morphological word, this form could be detached from plurality, because it does not contain a sign uniquely interpreted as a pluralizer. This sets BPs apart from SPs, whose suffix unambiguously identifies plural, and on a par with other morphological formations that likewise arise through stem rearrangement alone, like the participle *kaatib-un* 'writing' being reanalysed as 'writer' and feeding BP formation as *kuttaab-un* (Ratcliffe 1998: 55–6).

   However, *balad* ∼ *bilaad* ∼ *buldaan* is an isolated example in MWA. More often, BPs feed further pluralization in a formal sense only, without a plural being reanalysed as a singular. We saw in Section 3.4.1.3 that the senses associated with the singular *bayt* 'house as building/house as dynasty' and *rajul* 'man/personage' are disambiguated by twin plural forms. This is another instance of disambiguation through formally distinct plural doublets; the twist is that the two plurals (*buyuut* ∼ *buyuut-aat* and *rijaal* ∼ *rijaal-aat*

---

[11] A different instance of category changing derivation might come from Ratcliffe's (1998: 50) interpretation of deadjectival verbs in the so called ninth stem form (structure *yaCCaCCu* in the imperfective). He argues that a form like *yaḥmarru* 'become red' is constructed from the adjective 'red' on the basis of its BP *ḥumr* rather than its singular *aḥmar*.

respectively) are formally related by the SP suffixation rule *X > X-aat*, which makes this a case of double pluralization, although the meaning of the SP does not derive from the meaning of the BP (the same applies to the Syrian doublet *lsaan-aat* 'tongues' ∼/'əlsine 'languages'). The fact that a BP can feed morphological pluralization clearly shows that BPs have or at least can have the same status as 'singular' (that is, unmarked) nouns when it comes to plural suffixation. This is the opposite of what one would expect of inflected word forms, and represents strong evidence for the lexicality of BPs in a specific morphological sense.[12]

It is important to ask how generally BPs can feed further pluralization. In classical Arabic, BPs formed out of other BPs deserved their own label: *jamʕ-u al-jamʕ-i* 'the plural of the plural', cited by Wright (1967 [1896]: 231–2), who illustrates it with a wealth of exemples such as *kalb-un* 'a dog' > *'aklub-un* 'dogs' > *'akaalib-u* 'dogs'; *zahr-un* 'a flower' > *'azhaar-un* 'flowers' > *' azaa-hiir-un* 'flowers'; and even a treble formation, *firq-at-un* 'a band, a party, a sect' > *firaq-un* > *'afraaq-un* > *'afaariiq-u*. I could find no hint of a systematic interpretive opposition between these forms; at most, the formal difference may have been put to use to disambiguate between count/individualized versus mass/undifferentiated readings (cf. the discussion in Section 7.5 below). The dialect descriptions I could consult do not mention the application of stem-changing operations to their output; but they mention something close to it, namely the combination of affixation with stem changes. Fischer and Jastrow (1980) state that it is not common for a BP to receive the suffix of SPs in spoken dialects—but the phenomenon is attested. They give the following examples:

(7.10) | Singular | Plural | Gloss |
| --- | --- | --- |
| ṭarii' | ṭər'-aat | 'road' (Damascus) |
| 'a ṭr | 'uṭur-aat | 'road' (Cairo) |
| fraaš | fruuš-aat | 'bed' (Tunisi) |
| ṣaahib | ṣhaab-iin | 'master' (Damascus) |
| muul | mwaal-iin | 'master' (Rabat) |

Cf. also Saudi *kilma* ∼ *kalimaat* 'word/speech ∼ words/speeches' (Omar 1975: 46). In classical Arabic, broken pluralization was certainly compatible with the suffixes of SPs. Ratcliffe (1998: 88–90) discusses cases like *halq-at-un* 'ring, circle', with alternative plurals *halaq-un* (BP) or *halaq-aat-un*, that is, a SP formed on a stem expanded by *-a-*. As he notes, this stem expansion is

---

[12] BPs could also be suffixed by dual suffixes: *jamal un* 'male camel', BP *jimaal un*, dual *jimaal aani* 'two herds of male camels' (Wright 1967 [1896]: 191; Ojeda 1992: 322). However, here the interpretation is that of a dual of a semantically singular term, namely camels as one herd.

the only exponent of pluralization for some nouns like *xirq-at-un* 'rag',
pl. *xiraq-un*; this means that the *halaq-aat-un* type effectively cumulates the
changes of a BP and of a SP. Fischer (1980: 61) has more examples, such as
jamal-un 'camel' > BP *jimaal-un* > BP + SP *jimaal-aat-un*. The compatibility
of BPs with SP suffixation is a special case of a more general phenomenon.
Already in classical Arabic, some ablaut patterns occasionally or even system-
atically co-occurred with suffixes: *-at-*, effectively a collective (singular) for-
mation predominant with animates (*kaafir-un* 'unbeliever', pl. *kafar-at-un*;
dubb-un 'bear', pl. *dibab-at-un*; *turs-un* 'shield', pl. *tiras-at-un*), but also *-aan-*
(*suur-un* 'wall', pl. *siir-aan-un*; *fa'r-un* 'mouse', pl. *fi'r-aan-un*); cf. Wright
(1967 [1896]: 207–9, 216–18).[13]

No matter how restricted this mixed BP + SP pattern may now be (Holes
2004: 162 speaks of 'a tiny number of exceptions' to the mutual exclusiveness
of BP and SP formation), it raises a fundamental question: how to distinguish
changes in the stem shape as a primary exponent of pluralization (as in BPs)
from stem allomorphy as an adjustment concomitant with pluralization, that
is as a secondary exponent. If BPs are defined as non-suffixal and SPs as
suffixal but involving no change in the stem, then all the forms just discussed
are double plurals. But it seems more likely to view them as single, not double,
pluralizations, involving stem allomorphy in tandem with affixation: BPs
that happen to be suffixed. So, BPs cannot be strictly *defined* as non-suffixal,
but only as involving a change in the shape of the singular stem (this is in
effect the traditional notion). But then stem allomorphy can affect SPs as
well, as we can see not only from the exceptional *sana* ~ *sin-uun* 'year ~ years'
in MWA, but also from dialectal data (cf. the Egyptian examples in (7.7b),
Section 7.2.1.3).[14]

What all this shows is that in order to judge what counts as double
pluralization, one must know precisely what counts as single pluralization.
Taking every case of co-occurrence between stem change and affix as a double
plural is oversimplistic. The review of empirical properties suggesting that
BPs are lexical thus leads straight to the theoretical question on the precise

---

[13] Ratcliffe argues that the *a* added to the stem in SPs of the form *CuCaCaat* or *CiCaCaat* (from a
singular *CvCC*) is the primary exponent of plurality, and the suffix is secondary. This view receives
support from Fischer's (1980: 74) observation that this same *a* insertion also appears in Hebrew and
Aramaic and is probably an inherited Semitic feature. At least in these cases, then, the overlap of
broken and sound plural morphology cannot be reduced to a phonological readjustment of the
stem form.

[14] The co occurrence of suffixation and stem ablaut would seem to be particularly common in
Cypriot Maronite, as in *kampx* ~ *kumpxat* 'wheat' ~ 'wheats', *ʕaneb* ~ *ʕanubat* 'grape ~ grapes', *freš* ~
faršat 'bed ~ beds', *raxl* ~ *xulat* 'saddle ~ saddles', *maʕallem* ~ *muʕalmin* 'teacher ~ teachers'
(Tsiapera 1969: 49 50).

value of stem allomorphy as a morphological exponent. I will defer addressing that question to Section 7.4, after the evidence against the lexicality of BPs is examined.

7.3.2 *Why BPs are inflectional*

Having seen the evidence suggesting that BPs are lexical, we now turn to the evidence that they are not. This falls under four headings: the BP form does not determine a special semantics, it does not by itself determine a special agreement, it applies to agreeing adjectives where plurality is entirely a matter of grammar, and it does not by itself favour semantic lexicalization and reanalysis.

7.3.2.1 *Lack of a semantic correlate*   For all their lexical properties, BPs are unambiguously on a par with SPs when it comes to their distribution. Here a common misunderstanding must be dispelled: BPs are not mass plurals. The BP $\sim$ SP divide does *not* correspond to a partition between mass and count plurals, if only because most singulars only have one of the two: 'there simply cannot be such a semantic distinction, because sound and broken plurals are distributed through different sets of singulars and hence do not contrast' (Ratcliffe 1998: 70). The source for this misconception is presumably an unhappily worded pronouncement by Wright (1967 [1896]: 233):

As regards their meaning, the plurales fracti differ entirely from the sound plurals; for the latter denote several *distinct* individuals of a genus, the former a number of individuals viewed *collectively*, the idea of individuality being wholly suppressed. For example, ʕabd uuna are *slaves* (*servi*), i.e. several individuals who are slaves, ʕabiid un slaves collectively (*servitium* or *servitus*). . . . The plurales fracti are conse quently, strictly speaking, *singulars* with a *collective* signification, and often approach in their nature to *abstract nouns*. Hence, too, they are all of the *feminine* gender, and can be used as masc. only by the constructio ad sensum.

Yet Wright was (of course) well aware that BPs can be counted (the most individualizing construction), as is shown not only by several examples throughout his grammar (e.g. *nisaa'-un ʕašrun/ʕašru nisaa'-in* 'ten women' (p. 254); *θalaθat-u rijaal-in* 'three men' (p. 255)), but especially by his statement on p. 234 of the second volume that 'the genitive [governed by the numeral] must, in every possible case, be that of the *broken* plural', tempered on the following page by adding that 'perhaps it is superfluous to remark that the pluralis sanus is used in cases where no broken plural exists'. I will return in Section 7.5 to the relation between BPs and collectives; for now, what matters is to establish that BPs are not in any way coextensive with mass or

'collective' plurals. In fact, they have no special meaning. Again, this does not mean that they cannot have a lexicalized reading, but that it is not being a BP that turns a plural into a collective or mass. And since BPs have no semantic peculiarity as a morphological class, they are not lexical plurals in the sense of having a morphology that expresses a special reading.

7.3.2.2 *Excursus: BPs and deflected agreement*   Traditionally, the strongest argument for granting BPs a status distinct from SPs is that they trigger deflected agreement. This is the phenomenon, referred to in the above quote from Wright (1967 [1896]), whereby a plural noun triggers agreement in the feminine singular. In MWA, feminine singular is the norm for agreeing modifiers of non-human nouns, regardless of whether the noun is a BP like *kutub* or a SP like *sayyaaraat*:

(7.11)   Singular   Plural   (MWA; Belnap 1999: 171)
       a. kitaab   jadiid   kutub   jadiid-a
          book.MASC.SG   new.MASC.SG   book.MASC.PL   new-FEM.SG
          'new book'   'new books'

       b. sayyaar-a   jadiid-a   sayyaar-aat   jadiid-a
          car-FEM.SG   new-FEM.SG   car-FEM.PL   new-FEM.SG
          'new car'   'new cars'

Nouns denoting humans trigger instead 'strict' agreement in number (as well as in gender). The SP–BP distinction is irrelevant, as seen in the following example, where a BP noun (from *waziir* 'minister') is followed by a BP adjective and then by a quantifier in SP form:

(7.12) wuzaraa'   judud   kaθiir-un   (MWA; Belnap 1999: 170)
       minister.PL   new.PL   many.MASC.PL-INDEF
       'many new ministers'

BPs, then, play no role in determining the distribution of deflected agreement in MWA.[15] Turning to the spoken vernaculars, the work of of Belnap (1993, 1999), Belnap and Shabaneh (1992) and Brustad (2000: 52–62), has shown that the crucial variable for the distribution of deflected agreement is individuation rather than humanness. A Syrian example from Cowell (1964: 424; cited in Holes 2004: 204) illustrates how the same form *kutub* 'books' may

---

[15] Badawi *et al.* (2004: 102 4) seem to imply a special status for BPs in some of their formulations ('Non human broken plurals are with few exceptions treated as grammatically fem. sg.'; p. 102), but they too state that the decisive factor is whether the head noun is human, not its morphological form.

trigger plural agreement when its reference is specific and feminine singular when it is generic:

(7.13) a. l-kətˤb   maa   bihəmmuu
            the-book.PL   NEG   interest.3.PL.him
            'the books don't interest him'

       b. l-kətˤb   maa   bəthəmmo
            the-book.PL   NEG   interest.3.SG.FEM.him
            'books don't interest him'

Several semantic and discourse-related factors come into play in determining whether a plural noun's referent is conceptualized as a mass of indistinguishable parts or a denumerable collection of identifiable elements, and the picture is further complicated by syntactic factors (closeness of the target to its controller, category of the target) and by dialectal variation (see especially Brustad 2000). But again, the BP morphology does not play a role.

This leaves us with classical Arabic. Here, Beeston (1975) and especially the quantitative analyses of Belnap and Shabaneh (1992) and Belnap (1999) have shown that the deflected pattern of feminine singular agreement for non-human head nouns was an innovation that gradually spread from the tenth century onwards. After this innovation became established, as we have seen, deflected agreement has been triggered by inanimate or non-human nouns, not by BPs as such. But even before the innovation, what determined whether a plural agreed in the feminine singular was its place along a scale of individuation, plural agreement being most likely for humans and least for inanimates. This depended on the conceptualization: collectively interpreted plurals like *rusul* 'apostles' or *rijaal* 'men' could well trigger deflected agreement despite having human referents. In this context, it is true that 'sound plurals tend to occur with plural agreement more than broken plurals' (Belnap 1999: 181), but it is a tendency. How far this was from the determinism of a grammatical rule can be appreciated by looking at Belnap's (1999: 178–9) statistical results on deflected agreement in ancient texts up to and including the Koran. In this latter text, 37 out of 48 inanimate BPs (77 per cent) and 13 out of 17 inanimate SPs (76 per cent) trigger deflected agreement. In a previous seventh-century text, the corresponding percentages are (always for inanimates) 46 per cent versus 33 per cent. For pre-Islamic texts alone, the statistical probability of deflected agreement are: 0.725 for inanimate BPs and 0.630 for inanimate SPs, 0.501 for animal BPs and 0.361 for animal SPs, and 0.178 for human BPs and 0.158 for human SPs. While it is true that BPs are always more likely than SPs to trigger deflected agreement, the difference is

too small to attribute the choice of agreement pattern to the morphology of the head noun. Even in old Arabic, being a BP was not a sufficient condition for deflected agreement, not even for inanimates; nor was it necessary, given the high proportion of inanimate SPs that could trigger it. In fact, as Belnap notes, the pre-Islamic pattern is based on individuation like that of the modern Cairo dialect. In sum, Arabic grammar does not, and never did, make deflected agreement a consequence of the BP form.[16]

7.3.2.3   *Generality and uniformity of application*   The fact that adjectives can have BPs as well as nouns constitutes strong evidence against the lexicality of BPs. Adjectives are agreement targets, not controllers. Their number value is entirely a contextual property, deterministically grammar-driven. This is true even for deflected agreement, because the choice between that and strict agreement depends on the conceptualization of the head noun and not on that of the agreeing elements. But some adjectives are pluralized as BPs; therefore, being a BP by itself cannot signal lexicality.[17]

Notice that BP adjectives differ from unpredictable allomorphs, like for instance the comparative *better* for *good*. The suppletive stem of this allomorph is lexical in the sense of being listed as a form. But *judud* 'new.PL' in (7.13) above is not unpredictable in the same way: it instantiates the pattern *CuCuC* from a singular *CaCiiC* (*jadiid*), following a correspondence exemplified by the nouns *sariir* 'bedstead, seat' or *madiin-at-* 'city' (scheme III.2 in Wright's classification). So, the formation of this adjectival form is as regular or as irregular as BP formation generally, depending on one's stance; certainly it is not an isolated case.

The partial predictability of (some) BP patterns can in fact be viewed as an argument against the lexicality of BPs as a whole class. I have already mentioned the analyses by McCarthy and Prince (1990), Ratcliffe (1998) and Kihm (2003), which have shown how a significant part of BP formation can be reduced to the application of general rules. But even before these formal

---

[16] Once it is clear that deflected agreement is a tendency but not a grammatical consequence of the BP form, the question of the cause for this preference can be raised. There seems little doubt that the non plural form of BPs and the ambiguity between abstract and concrete collective reading (cf. *service* in English) played an important part, as Belnap adds (1999: 180). But to say that 'the BP class . . . tends to take feminine singular agreement' may be read as implying that their morphological form triggered deflected agreement independent of the semantic characterization, and this is the opposite of what the rest of that paper says. It is this putative *grammatical* source for deflected agreement that I am arguing against.

[17] The fact that BPs are much more common with nouns than with adjectives might be construed as evidence that this format sits more comfortably with agreement controllers. In fact, this depends on the fact that many nominals with adjectival function are derived from verbs (participles) or nouns (relational adjectives).

approaches had uncovered the prosodic regularities of BP formation, Fischer (1980) had grouped together SPs and the predictable or partly predictable patterns of BP formation into a class of inflectional plurals, opposed to the rest of BPs which make up the class of lexical plurals. In so far as at least a part of BPs are predictable by rule, one cannot analyse BP formation as a whole as a lexical pluralization strategy. The conclusion must rather be that stem reshaping is a morphological option of Arabic alongside affixation, which may or may not lead to lexical plurality.

7.3.2.4 *Reanalysis of BP forms*   Finally, even the privileged relation between BPs and lexicalized nouns can be construed as an argument that BPs are not lexical by themselves. We have seen in Section 7.3.1.5 that the BP *bilaad* 'cities' can be reanalysed as a singular meaning 'country' and re-pluralized as *buldaan*. On the face of it, this shows that BP formation can take a plural form out of its original inflectional paradigm, 'lexicalizing' it. But we also saw that this is part of a larger pattern whereby the input for broken pluralization comes from the reanalysis of the output of a stem-changing operation. The participle *kaatib* 'writing, *scribens*' has the BP *kuttaab* in its lexicalized meaning 'writer, *scriptor*' (otherwise, what is synchronically a participle with pattern *CaaCiC* gets a SP: *kaatib-uun* 'writing ones', *scribentes*). The same applies, for instance, to *kaafir* 'unbeliever', which owes its SP *kaafir-uun* to its form as a participle ('non-believing') but also has the BPs *kuffaar*, *kafara*, *kifaar* as a noun (according to Wehr's dictionary). This shows that the reanalysis of 'cities' as 'country' is not determined by BP formation, but can generally happen to forms generated by other stem-changing (or 'Level I') operations. It is not BP formation that leads to lexicalization, understood as the reanalysis of a grammatical form of a word as a related but distinct word; what happens instead is that a Level I-derived stem may get lexicalized, and this can involve BPs.

7.4 Derived stems in an inflectional paradigm

The status of BPs is ambiguous because lexicality is an ambiguous notion. Some, but not all, BP forms are listed in the mental lexicon in so far as they cannot be automatically produced by deterministic rules. This has little import on the whole question; mere formal listedness, perhaps in the form of a restriction to a deterministic rule ('turn pattern *A* into *B* for roots *x, y, z*'), does not immediately relate to those grammatical and morphological properties that accompany the intermingling of grammatical and lexical

knowledge. We can disregard sheer unpredictability, then, along with the fact that it is not absolute for all BPs.

What remains falls neatly into two categories: the evidence for lexicality refers to the morphology of BPs, the evidence against refers to their meaning and categorial distribution:

(7.14)   THE CASE FOR BPs AS LEXICAL PLURALS:

    (i) no principles of Arabic grammar ensure that a given singular has a unique BP

    (ii) BPs do not in principle block SPs

    (iii) different forms of plural can express different interpretations

    (iv) the BP stem can serve as base for adjectival and verbal derivation

    (v) some BPs, but no SPs, can be re-pluralized

    (vi) BPs are compatible with plural suffixes (possibly those of SPs)

    THE CASE AGAINST BPs AS LEXICAL PLURALS:

    (vii) BPs have no special meaning

    (viii) adjectives can have BPs

    (ix) all words derived on Level I, not just BPs, can feed BP-creation if lexicalized

What suggests lexicality concerns the morphology of BPs, namely their lack of disjunctivity (properties (i)–(iii)) and their ability to function as bases for inflection (properties (iv)–(vi)). Even property (iii), which refers to interpretation, centres on the paradigmatic relation of BPs with each other. Semantic differentiation is a natural product of polymorphism; what makes BPs different from SPs with respect to property (iii) is that a given singular may have more than one BP, and this is a morphological fact.

By contrast, the properties speaking against lexicality are independent of the morphological form of BPs. This type of pluralization has no special semantic value, so obviously it is available for plural agreement on adjectives, which is purely grammar-driven. Besides, the instances when a BP feeds further BP formation is a subcase of a more general pattern whereby the output of a stem-changing derivation is reanalysed. In sum, BPs have none of the semantic properties that typically accompany plurality as a lexicalized property, fused with the meaning of the lexeme itself.

The neat distinction between derivational-type properties of BP as morphological forms and inflectional-type properties of BPs as an abstract category suggests the solution: the forms themselves are the product of stem-forming processes, which are not intrinsically plural. But these forms, while the

product of word formation, serve as exponents of a perfectly regular category of *inflectional* number.

7.4.1 *The status of plurality in SPs and BPs*

The inflectional side of BPs is the easier to justify. Quite simply, Arabic *kutub* 'books' and *kibaar* 'big.ᴘʟ' do not differ from their counterparts in other languages provided with inflectional number, as abstract representations encoding lexical information and grammatical features. In particular, the abstract representation of two such BPs differs in nothing from that of SP plurals. BPs occupy no special place in the number system of Arabic—so much so, that, as we have seen, the distinction between BPs and SPs is less clear-cut than it would appear at a superficial inspection, because the properties of being deterministically predictable, having a suffix, and being formed by a pattern of stem-internal changes, fail to neatly partition the set of Arabic plurals into two mutually exclusive classes.

Attributing the forms, and just the forms, which realize noun stems as BPs to derivation (in the sense of word formation, or rather stem formation) amounts to accepting the traditional view that BPs are 'lexical'—with some qualifications. As I noted in Section 7.3.2.3, Fischer (1980) divided Arabic plurals between 'lexical' BPs, whose form is not predictable by deterministic rule, and 'inflectional', comprising SPs and the regular BP ablaut patterns (quadriliterals with vowels *a-aa-i*/*a-aa-ii*, *fuʕal* and *fiʕal*, plus the SP *CVCa-Caat* derived from *CVCCat* by inserting an *a* after the second *C*). While a distinction along these lines is justified in a philological perspective, the interpretation I am offering differs because derivational stem-forming status is not linked to listedness. What sets BPs on a different level from SPs *qua* morphological objects is not that they are lexicalized in the sense of being listed (one of the two senses of 'lexical' elucidated by Aronoff 1994), but that they realize an inflectional property on a stem by creating an alternant of it. The ways to do so are more or less regimented, and some singular–plural match is in fact regular and predictable; but the making of a new stem is essentially different from spelling out number through a suffix (with or without stem allomorphy). As is well known, the suffixes are exponents of plurality, but the stem forms of BPs by themselves are not; and even the abstract operations that underlie the formation of such BP forms (say, addition of a VX to the initial CVX, as per McCarthy and Prince 1990 and Ratcliffe 1998) are not by themselves expressions of plurality. The same formal pattern of stem extension that underlies BP formation is also at work in other areas of Arabic morphology, like diminutive formation and the formation of verbal stems:

(7.15) <CVX>   > <CVCVX>   (Ratcliffe 1998: 33)
        <yak>tubu > <yukat>tibu 'write', stem I > 'cause to write' stem II
        <yak>tubu > <yukaa>tibu 'write', stem I > 'correspond' stem III
        <kal>bun   > <kilaa>bun   'dog', sg. > pl.
        <kal>bun   > <kulay>bun   'dog', sg. > sg. diminutive

Although the addition of moraic units on the timing tier is formally an
instance of suffixing, these abstract suffixes are pre-defined strategies to turn
a stem into a different version of it; they have no 'content' in the sense in
which [plural] is the 'content' of English -*s*. Kihm (2003) has particularly
emphasized that BPs express plurality in a purely differential way, by their
formal opposition to singulars: there is no plural morpheme on BPs, no
matter how abstract. In this purely morphological sense, plurality has co-
alesced with the stem. This was also the intuition behind the approach to BPs
outlined in Noyer (1997: lviii–lxi).[18]

   SPs, on the other hand, are regular in two ways: the addition of a timing
unit to the stem is realized by a noun-invariant exponent (-*uun* for mascu-
lines and -*aat* for feminines), and this suffix unambiguously spells out
plurality. What is usually seen as the crucial regularity of SPs with respect to
BPs, namely the identity of the stem in singular and plural, is less important—
and indeed we have seen that cases of stem allomorphy with SP make this a
dubious criterion. The formal and interpretive invariance of the SP endings,
by contrast, allow us to see them, or better, the abstract lengthening realized
by them, as a morphological object endowed with form *and* meaning. With-
out being exactly suffixes in the traditional sense, the SP endings approximate
suffixes because their target is morphologically defined as the singular stem,
not prosodically defined as a heavy rhyme. In this I follow Ratcliffe (1998: 54),
who reformulates the traditional distinction of stem- and word-level morph-
ology into two levels of word derivation: 'Level I (base or stem) rules reference
a phonological word or any phonologically defined part of a word. Level II
(stem or word) rules reference a morphologically defined stem.' What makes
BPs different from SPs, I contend, is *entirely* due to their morphological

---

[18] Noyer (1997) claimed that BPs are stem inherent plurals, and I follow him here; but he argued
that the basis for this was that BPs *never* allow for suffixes. This is not true, as we have seen in Section
7.3.1.4; nor is it true that BP 'acts as a special type of agreement class, associated with feminine singular
agreement' (Noyer 1997: lix), as we have seen in Section 7.3.2.2. In fact, the attested facts suit Noyer's
analysis better, because in his system a feature changing *rule* like ablaut should not bleed a feature
inserting *affix*; English strong verbs, for instance, can be compatible with a suffix (*sol d*, *brok en*).
I follow his view that BPs express an inflectional property through a lexical choice; but, like Kihm
(2003), I do not take BPs to be plural exponents at all, inherent or otherwise. They are just stem
alternants that do not 'mean' plural, different e.g. from the uniquely plural *geese*; what they share
with SPs is the abstract syntactic representation, which contains the same plural number feature (and
hence can trigger strict agreement, contrary to Noyer's assumptions).

property of being a product of Level I, while SPs are created on Level II by affixing a stem (the same as the singular) created on Level I.

7.4.2 *Derivation, inflection, and disjunctivity*

Most importantly, Levels I and II do not coincide with derivation and inflection. If the term 'morphology' refers to the rules, principles, and elements employed in the formation of word forms, the separationist approach I am following (after Beard 1995) regards morphology as the realization of abstract information; the derivation–inflection opposition concerns this information, not the morphology that spells it out (Aronoff 1994: 15). The notion of BPs, I submit, is morphological in this sense. Number is a syntactic (morphosyntactic) category which may or may not be lexicalized, but that does not have to do with the distinction between BPs and SPs, nor more generally with the Levels at which these two plural types are formed. These are layers in the formation of a word: the operations of Level I create a morphological object (a stem, specifically a nominal stem), while those on Level II add on to it. But one and the same morphosyntactic category may find expression through the workings of either of these sets of operations, or indeed both: a BP like *kutub* 'books' is entirely derived on Level I, a SP like *kaafiruun* 'unbelievers' entirely on Level II, and a 'mixed' form like *halaq-aat-un* 'circles' (from *halq-at-un*) on both levels.

This brings us to the central issue of disjunctivity and the paradigmatic relation between plurals. Adding a VX string to a prosodically circumscribed <CVX> does not 'mean' plural; it is an established strategy of Arabic morphology for creating a stem out of another stem. Here, the relevant sense of 'lexical' is the one opposed to 'grammatical' ('lexemic' in Aronoff 1994: 19). Two alternative BPs, say *zuhuur* and *'azhur* 'flowers' are both lexical in this sense; they are full words, not grammatical morphemes, and they cannot be analysed into a lexical base and a grammatical morpheme. This ensures that there is no deterministic competition between them, as there is instead between grammatical morphemes. In principle, then, stem alternants escape disjunctivity, but only in principle. The choice between one form and another is not as free as that between *cat* and *dog*, because the stem alternants are word forms for a given lexeme, not two distinct lexemes. This distinguishes Arabic BPs from Italian -*a* plurals, for instance, which I have analysed as autonomous lexemes in Chapter 5. Hence, the grammatical system tolerates multiple plurals, but absolute equivalence between BP alternants is still disfavoured on functional grounds. This derives the state of affairs prevalent in spoken dialects, where multiple BPs are not exceptional but not particularly frequent either, and alternants tend to differentiate themselves on semantic or stylistic grounds.

An example of stem allomorphy from a different language may be useful. In Italian, the past participle of *vedere* 'to see' is *visto* 'seen'; but the alternant *veduto* is also available. The two are not in free variation (*veduto* is generally dispreferred in contemporary Italian), but the choice is not driven by grammar, as it is instead for the ungrammatical **chieduto* 'asked' instead of *chiesto*, from *chiedere* 'to ask'. The difference between the unavailable **chieduto* and the available *veduto* must be stated somehow; my claim is that the alternants *veduto* and *visto* stand in the same mutual relation as alternative BPs do in Arabic. The widespread but unsystematic tendency to differentiate grammatically equivalent doublets shows up in Italian too. The difference between *veduto* and *visto* is a matter of style, but the alternants *riflesso* and *riflettuto* distinguish two senses of the verb *riflettere* 'to reflect': *riflettuto* is the past participle of 'to reflect' in the sense of 'to think', while *riflesso* only means 'reflected' in the optical sense (and is then zero-nominalized with the meaning 'reflex').[19]

Being grammatically equivalent realizations of the same lexeme, the alternative BPs naturally tend to differentiate themselves on semantic, stylistic, or dialectal grounds (cf. Section 7.2.3 above). But this natural drift towards a functionally unique realization of plural is not the strict disjunctivity of grammatical morphemes, because it involves stems, that is sound forms of lexemes, and not grammatical morphemes.

7.4.3 *BPs as stems*

Aronoff (1994: 31, 39–44, 57–9) defines a stem as one of the realizations of a lexeme that acts as the base for inflectional morphology. While being part of the 'concrete' morphological spell-out component, stems are nonetheless abstract concepts in two senses: first, a stem as a type (e.g. the Latin third stem) may have tokens with irreducibly different forms (e.g. the third stem *rupt-* from 'to break' is built on the root of the lexeme, while *amat-* from 'to love' is derived from the present stem by *-t* suffixation); second, since a stem is a word form abstracted away from inflectional exponents, it may well fail to correspond to any existing word form for a given lexeme; for instance, *rupt-* never surfaces by itself, but it underlies forms like *rupt-um* 'broken. PAST.PARTICIPLE.NEUT.SG'.

Each instance of a BP pattern on a noun is a stem token. The schemes that derive BPs from the singulars are as many ways to form a stem from another.

---

[19] Latin verbal stems too, on which Aronoff (1994) based his proposal, display some variation in the shape of the stems: for instance, the perfect stem of *parcēre* 'to spare, be lenient to', is normally *peperci* but *parsi* is also attested, and is normal in Plautus (Sommer 1902: 613, §374). Perhaps strong weak alternations such as *proved* ∼ *proven*, or *backte* ∼ *buk* for the German preterite of *backen* 'to bake' are further examples. The uncertainty in usage between forms like *shrank* and *shrunk* mentioned by Pinker (1999: 77) certainly suggests a parallel.

Some of these strategies are regular and even productive, others unpredictable; but the definition of Aronovian stem accommodates both. It may well be that BP formation has a stable invariant format at a suitably abstract (prosodic) level, but this does not impinge on the analysis of BP forms as stem tokens. As an Aronovian stem, a given BP form is a purely morphological ('morphomic') entity, not a sign: the pattern *CiCaaC* of *kilaab* 'dogs' is a stem form that comes to functionally express plural because it is different from the *CaCC* form *kalb* 'dog'. There are regularities, as the systematically plural *CuCuC* (or the singular *CuCC*), but overall there is no plural stem as a category of Arabic grammar; what happens is instead that each noun which does not feed SP formation makes up a plural form by constructing a stem alternant according to one of a range of strategies.

In this light, the fact that BPs allow for further suffixation in the formation of plural, and can even feed further lexeme-creating derivation, is even more telling than the lack of inflectional blocking between BP alternants. This would be unexpected if the stem reshaping rule by itself 'meant' plurality, because in that case the frequence of suffixal BPs (especially when the suffixes are those of SP) would imply that forms already realizing plural take on morphologically useless affixes with embarrassing frequency. The use of BPs as bases for adjectival and verbal derivation would be even harder to justify, if the BP form incorporated an abstract plural morpheme. In sum, then, a realizational and separationist view of morphology enables us to see what is lexical in BPs and what is not. Their morphology is, because they are stem forms and hence very different from grammatical morphemes. The number information they express, on the other hand, is not.

7.5 Number, collectives, and the semantics of BPs

The conclusion we have reached is that BPs are lexical for how they realize plurality (namely through lexical stems), not for what they realize. This means that plurality as an inflectional category, which is what they realize, has no special properties in Arabic. I contend this is the right interpretation; but it is only part of the story. Alongside number, nouns can also be specified in Arabic in terms of countability, distinctive identity, and collectiveness. These additional dimensions of information interact with number because they create additional paradigmatic relations between noun forms. In particular, a singular 'collective' (e.g. 'livestock') can function as a notional plural for a singular count noun (e.g. 'a cow'). This is where BPs become relevant.

BPs are stem forms serving as plurals in an inflectional paradigm; what are generally known as collective nouns are instead grammatically singular,

related to count singulars by singulative derivation. Although the native grammatical tradition has carefully distinguished them as two conceptually distinct classes, they are morphologically one and the same, realized as Level I stem alternants, and with the same (singular) inflection. As a result, the distinction between stem-changing plurals and singular collectives is sometimes blurred; for instance, the pattern *CuCC* is typical of singular collectives and not of plurals (so that Ratcliffe 1998: 102 argues that *rukb-un* is not properly the BP of *raakib-un* 'rider'), but it surfaces to form the grammatical plural for adjectives of colour or defect, like *humr-un* from *'ahmar-u* 'red', or *ṣumm-un* from *'aṣamm-u* 'deaf' (Fleisch 1990: 479). And of course BPs also behave like weakly individuated collectives when they trigger deflected agreement. In sum, BPs act as inflectional plurals but they can have the morphosyntax of collective singulars, which is the basis for Wright's misleading generalization cited above in Section 7.3.2.1. This is an important part of their perceived 'lexicality', and requires discussing the relation between part-structure conceptualization and the Arabic number system.

7.5.1 *Number and the morphological expressions of individuality*

The main morphological means for packaging the reference domain of a noun into individual wholes is the singulative (*'ism l-wahda* 'noun of individuality', *nomen unitatis*). This typically involves the feminine suffix *-at-* (*-a* or a derivative in the dialects):

(7.16) COLLECTIVE   SINGULATIVE
      damʕ-un 'tears'   damʕ-at-un 'a tear'   (classical Arabic; Fischer 1972: 49)
      hadiid-un 'iron' hadiid-at-un 'a piece of iron'   (classical Arabic; ibid.)
      xux 'peach(es)'   xux-a 'a peach'   (Moroccan; Harrell 1962: 80)
      ḥṭeb 'fire wood' ḥṭeb-a 'a piece of fire wood'   (Moroccan; ibid.)
      bagar 'cattle'   bgar-a 'a cow'   (Gulf Arabic; Qafisheh 1977: 99)
      'akil 'food'   'akl-a 'a meal'   (Gulf Arabic; ibid.)

For those used to interpreting singular as 'one' and plural as 'many', the singulatives in the right column need no explanation. They are simply count nouns, and can of course be pluralized—with a SP, because they are morphologically derived: Syrian *dəbbaane* 'a fly', *tlətt dəbbaanaat* 'three flies' (Cowell 1964: 297), or MWA *baqar* 'livestock', *baqarat* 'cow', *baqaraat* 'cows'. What needs clarification is the first column: these nouns are grammatically singular and non-count. Just like English has mass nouns that denote continuous substances like *water* and factually discrete domains like *footwear*, these Arabic nouns conceptualize as masses substances and aggregates alike;

hence glosses such as 'peach(es)', which cover both a concrete mass reading ('peach-stuff') and a more abstract kind reading ('the peach fruit'). Grammatical descriptions typically present these bases as special 'collective' plurals of their respective singulative. This is justified for pedagogical purposes, and may in fact reflect the *Sprachgefühl* of native speakers; but for an understanding of the morphosemantics of Arabic number it is important to recognize that they are singular nouns outside the inflectional paradigm of singulatives, as we saw in Section 3.5.1 for Maltese ('structurally an Arabic dialect': Holes 2004: 2). These uncountable nouns are generally masculine, while singulatives, being derived through *-at-*, are invariably feminine. Besides, base mass nouns sometimes, if not always, have a BP of their own (I will return to this in Section 7.5.3).

The point that *-at* serves to derive a new noun with a count reading, and not to spell out the singular of the base form, is reinforced by the observation that it has the same packaging function with verbal nouns, which are not their plurals in any sense. The result of this derivation is called *'ism l-marra* 'noun of time', *nomen vicis*:

(7.17) Verb   Verbal noun   Instance noun
       ʕaṭaṣ 'to sneeze' ʕaṭəṣ 'sneezing'   ʕaṭṣa 'a sneeze'
       baas 'to kiss'   boos 'kissing'   boose 'a kiss'
       saafar 'to travel' safar 'travelling' safra 'a trip' (Syrian; Cowell 1964: 302)

In this case the criterion for granularity is an event in which the verbal predicate holds.

The singulative derivation in *-at* is for non-humans; another one, restricted to human ethnics (plus *jinn* 'demon(s)'), employs the suffix of the *nisba* relational adjectives, namely *-iyy* (MWA *-ii*): *yuunaan* 'Greeks', *yuunaanii* 'a Greek'; *badw* 'Bedouins', *badawii* 'a Bedouin'. Here the base is syntactically plural: *al ʕarab as-suuriyiin* 'the Syrian [pl.] Arabs' (Cowell 1964: 301). This might suggest that the singulative is simply its singular. But in grammatical terms, as opposed to functional, this is wrong: some *nisba* singulatives have their own plural, like *yuunaaniiyyuun* 'Greeks', and some notionally plural ethnic bases have their own (broken) plural, like *turk* 'Turk', pl. *atraak* (singulative *turkii*). Once more, grammatical and notional number are not isomorphic.

Another way to express semantic packaging is through classifiers. Examples from the Cairo dialect are *hittit lahm* 'a chunk of meat', lit. 'piece meat', and *hittit hadiid* 'a piece of iron', lit. 'piece iron' (Holes 2004: 171). Greenberg (1974: 26) reports that, in the Omani-Zanzibar dialect, numerals require a unit noun in *-at* if that is available, but occasionally there is a choice between

the two constructions: *θalaaθ baqraat* 'three cows' is synonymous with *θalaaθit rwaas baqar* 'three heads [of] cattle'. The classifier is in fact a lexical noun acting as unit counter, and regularly *plural* after the numeral, in an interesting parallel with Irish (cf. Chapter 6). In short, *-at* and classifiers are alternative morphological expressions of semantic packaging, but neither acts as the singular in an inflectional number opposition.

In fact, another use of *-at* shows that the discretizing function is not even tied to notional number, but strictly to unity. This suffix can attach to a singular denoting a human being, and derive a singular noun denoting the class of such beings: *muslim-un* 'Muslim' > *muslimat-un* 'Muslim world', *jammaal-un* 'camel driver' > *jammaalat-un* 'camel drivers'; it attaches especially often to *nisba* relational adjectives with singular reference, as in *ṣuufiyy-un* 'Sufi, mystic' > *ṣuufiyyat-un* 'Sufis' (Fischer 1972: 52, Fleisch 1990: 307). In this case what is 'one' is the class, so that the same suffix that derives a notional singular by segmenting a mass into wholes derives a notional plural by grouping the wholes into a class.

7.5.2 *Collectives*

Nouns that denote a collection as a single unit, like *'ibil* 'camel herd, camels', are no more noteworthy in Arabic than they are in any other language. What is peculiar about Arabic is that it has a class of non-count collective nouns, referring to collections but also to substances, which are systematically related to their corresponding count noun through the singulative *-at* suffixation as shown in (7.16) above: *baqar* 'cattle' ~ *baqarat* 'a cow'. The traditional view, expounded in Fleisch (1990: 302–6), holds that these bases primarily name a kind. For substance nouns, like 'iron', the use of the same word to name the whole kind (*iron oxidizes easily*) and any of its instance parts (*the iron in this bar*) perfectly parallels mass nouns in English. But Arabic also uses the same type of mass noun for entities typically experienced as aggregates of perceptually undifferentiated tokens: plants, fruits, and insects, but also animals like ducks, cattle, pigeons, owls, worms, fish, mice, birds; and objects like eggs, pearls, feathers, clouds, waves, and as possibly the only artefact, bricks.

These are all textbook examples of concepts that lack identity properties, that is, criteria for deciding the identity of one object with respect to any other (cf. Section 4.5.2). Unlike the Irish counting plurals of Chapter 6, which encode unity but not identity properties, and thus serve as standards for counting, Arabic collectives cannot be counted. Rather, the elements denoted by, for example, *zahr* 'flower(s)' are undifferentiated in the sense in which portions of water are: we can distinguish one portion form the other, but only thanks to their spatio-temporal location and not for any intrinsic property. Lack of identity is therefore the perceptual basis for the conceptualization as a

mass. Correspondingly, the denotation of Arabic collectives has the structure of a free-join semilattice (following Zabbal 2002: 102), which comprises the minimal parts and all their sums, up to the total sum that corresponds with the whole kind. This broadly agrees with the semantic characterizations of collectives expressed in the literature, gathered together in Ojeda (1992*b*: 306–7) and Zabbal (2002: 82). Collectives are transnumeral as kind-names are (cf. Sections 2.8.5 and 4.5), as shown by their occurrence as complements to classifiers and units of measure: cf. Saudi *kiilu leemuun* 'a kilo lemon(s)' (Omar 1975: 175).

The label of kind-names highlights the identity between the name of the kind and that of its instances, but it can be misleading. For Ojeda (1992*b*: 312), collectives denote 'the coarsest partition of a kind', that is, in his terms, the maximal sum. But this conflicts with the fact that they can also denote subsets smaller than the maximal sum, as Zabbal (2002: 84–5) demonstrates with examples like the following:[20]

(7.18) a. al-baqar-u   ta-ṣʕud-u   'ila l-jabal-i   (Lebanese)
            the-cow-NOM   3.FEM-climb. IMPF-SG   to   the-mountain -GEN
            'the cattle are climbing the mountain'

        b. nazal-a   l-hamaam-u   'alaa l-saṭh-i
            alighted.PERF-3.SG.MASC   the-pigeon-NOM   on   the-roof-GEN
            'the pigeons alighted on the roof'

I also collected this example, with unambiguously non-generic reference:

(7.19) a. il-beḍ   yallii beʕtni   yahum   miš   ṭaza   (Syrian)
            the-egg which sold.2.me them   NEG fresh

        b. l-biḍ   lli   'iʕtili   ma   kan-š   ṭrii   (Moroccan)
            the-egg which sold.2.me NEG were-NEG fresh
            'the eggs you sold me were not fresh'

We see here the essential difference between these Arabic expressions and kind-referring singular count nouns in English-like languages. One could say *the bear is resourceful*, but not *the bear surrounded the jam factory*, where the group-predicate and the episodic tense enforce reference to a collection of individual bears. The closest a singular count noun can get to this reading is the folk-tale conceptualization of a kind as embodied in one of its instances, as in *the bear and the fox went hunting together*. Arabic collectives on the other hand can refer to instances with the name of their kind, even with group-predicates: Zabbal (2002: 95) reports the Lebanese example *ta-jamaʕ-u*

[20] Zabbal (2002) identifies as Lebanese the Arabic of his main informant. Although this does not affect the argument, his examples are much closer to MWA than to regional vernacular, so they are not Lebanese in the same sense as examples from dialect monographs are Syrian or Egyptian.

l-baqar-u qurba l-nahr-i 'the cattle gathered near the river', where *baqar* refers to a specific collection of animals.[21] In sum, the interpretation of collectives varies between reference to the whole kind and to sums of its instantiations— just like that of mass terms. Zabbal (2002: 110–11) concludes that the genericity of many instances of collectives is brought about by the definite determiner, not by the form of the noun, and I think his conclusion is correct.

However, collectives differ from mass nouns in a crucial respect: they can be arguments to predicates which require an articulation into individual wholes (*'inna* is a sentence-initial particle that plays no role in the translation):

(7.20) a. 'inna-ni   'a-jmaʕ-a   n-naml-u   (Lebanese; Zabbal 2002: 98)
             PRT-1.SG 1-counted-SG the-ant-NOM
             'I counted the ants'[22]

(7.21) a. hseb   l-hut   wahda b wahda   (Moroccan)
        b. yayd   s-samak   wahda b wahda   (Egyptian)
        c. ʕadd   s-samak   wahde (b) wahde   (Syrian)
             counted   the-fish   one by one
             'he counted the fish one by one'[23]

The informants I consulted made it very clear that substituting a singular singulative to the collective 'fish' in (7.21) would definitely not fit the interpretation. This means that collectives, while grammatically singular, can also conceptualize the kind as a *plurality* made up of individual wholes that are accessible to semantic operations (in the terms of Moltmann 1997: 87–91). In fact, a non-collective plural would fit the context too. This is illustrated by the following example, where 'to distinguish between' can take either a collective or a BP as argument:

(7.22) laa 'astaṭiiʕ-u t-temiiṣ   bayna   hazaa   l-zahr
             NEG can.1   the-distinction between   this. SG   the- flower
             /tilka   l-'azhaar (Egyptian)
             /that. FEM.SG   the- flower.PL
             'I cannot distinguish between these flowers'

---

[21] If being collective is not a sufficient condition to denote a kind, it is not necessary either, for singulatives can do so too. Zabbal (2002: 108) shows the singulative *namla* 'an ant' (from the collective *naml*) in the sentence 'the ant exist everywhere'. My own observations confirm this. A Syrian speaker found both the collective *samak* 'fish' and the singulative *samake* 'a fish' acceptable in 'this fish is extinct'; a Moroccan one accepted *nḥal* 'bee(s)' and *naḥl a* in 'bees make honey', and two Egyptian speakers accepted the same pair in 'this bee is common in this country'. See also Holes (2004: 171).

[22] I reproduce Zabbal's example and gloss here, which have the nominative *naml u* instead of the expected accusative *naml a*.

[23] 'One by one' ensures that 'to count' has the intended reading: cf. *he counted the money* (*one by one)*.

(Notice the feminine singular agreement of the determiner of *ʾazhaar*, to which I will return in Section 7.5.4 below.) This aligns collectives with countable plurals, since we saw in Section 4.2.2 that mass plurals are incompatible with predicates like 'count' or 'distinguish'. Zabbal (2002) must be credited for making this important point; but then collectives cannot be equivalent to nouns like *furniture*, as he contends, (pp. 101–2), because these, not being count, do not allow access to the individual parts of their reference: **he counted the furniture*, **I cannot distinguish the furniture*.

This leaves us with a contradictory conclusion: collectives are mass because they cannot be counted and can refer to the whole kind or to any of its subparts; but they are like count terms because they allow predicates to refer to the individual wholes in their reference.

7.5.3 *Collectives as non-count plurals*

The fact that Arabic collectives like 'ants', 'fish', or 'flowers' can be arguments to distributive predicates (examples (7.20), (7.21), and (7.22)) shows how oversimplistic it would be to assimilate them to mass plurals. The evidence is that they can conceptualize their reference as made up of wholes, not as an unstructured mass. This distinguishes them from English nouns like *furniture* or *police*: contrast (7.21) with **he labelled the furniture one by one*, **she avoided the police one by one*. The problem then is why they do not tolerate numerals.

This mismatch between ostensibly discrete conceptualization and countability results from the interplay between the properties of collective nouns and the syntax of numerical modification in Arabic. There is no need to delve into the details of a very complex system, with a great deal of dialectal variation (see Holes 2004: 212–16 for a concise overview). All we need to note to start with is that Arabic unit numerals (3–10) govern the plural of count nouns (genitive when case distinctions are preserved). This requirement however is not absolute, and the exceptions strongly suggest a grammatical rather than a semantic motivation. Besides the use of classifiers, illustrated in the preceding section, a numeral may quantify otherwise uncountable collectives like *baqar-un* 'cattle' by means of the preposition *min* 'out of', or even directly, as is exceptionally the case with collective nouns of nationality like *ʾameerkaan* 'American(s)':[24]

(7.23) a. xams-u   min   l-baqar-i   (classical Arabic; Bloch 1978: 417)
           five-NOM   of   the-cattle-GEN
           'five cows'

---

[24] The feminine ending on 'five' in (7.24b) illustrates the gender feature on these numerals, which regularly contrasts with that of the governed noun. See Fleisch (1990: 509 11) for a brief historical discussion.

b. xams-e   'ameerkaan   (Syrian, Damascus dialect; Bloch 1978: 417)
   five-FEM.SG   Americans
   'five Americans'

Wright (1967 [1896]: 237, vol. 2) reports two more examples from the classical language of this lexically restricted construction. It seems then that the non-countability of collectives is not so much caused by their meaning, as by the grammar of the numeral construction.

The crucial factor, I suggest, is the same which played a decisive role in the analysis of Irish special plurals: numerals requires a syntactic encoding of countability, that is, of a uniform partition of reference into fixed-sized atoms. In the syntactic analysis I follow, based on Borer (2005: 109–19), the division of the reference domain into units and the quantification over these units have two distinct syntactic loci: the first is Classifier Phrase, just above NP, headed by a division operator encoded through a word (a classifier) or through grammatical number features; the second is a 'quantity phrase' (Borer dubs it 'Quantity Phrase' or '#P', to avoid the ambiguities of the term 'Number Phrase'), hosting elements that quantify over the units defined by the lower division operator (see Sections 4.2.1 and 6.6). This is schematized in (7.24), repeated from (4.2) in Chapter 4:

(7.24)

```
                        DP
                   ╱          ╲
                 D              #P
          referential index   ╱    ╲
                          <e>#        Classifier P
                        quantity    ╱          ╲
                                <e>DIV            NP
                               division
```
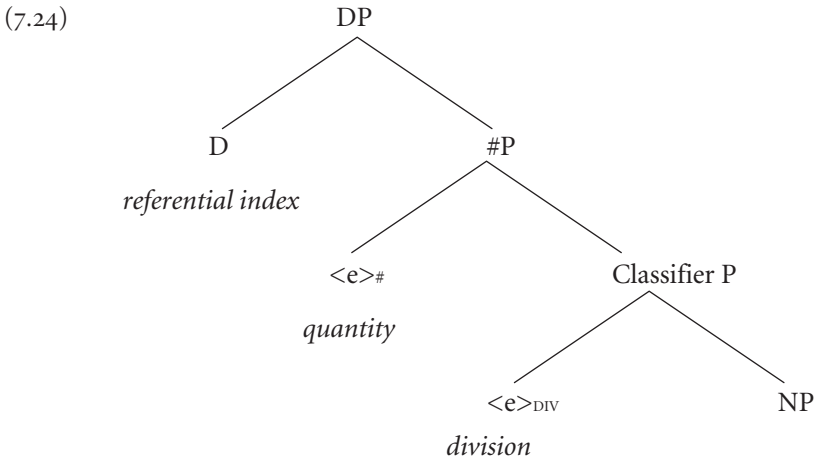
Obviously, counting requires a uniform standard of what is to be enumerated, so cardinal numbers in #Phrase require a defined value for the head of Classifier Phrase: the reference must be partitioned into units of the cardinality corresponding to the numeral. In Borer's (2005) words, 'while counters assign range to $<e>_\#$, they may only assign range to it if range has already been assigned to $<e>_{DIV}$' (p. 118).

This is where Arabic collectives differ from count plurals. They do not 'assign range' to $<e>_{\text{DIV}}$, that is they do not grammatically encode the interpretation of the corresponding singular, because there is no corresponding singular. As we have seen, collectives are already singular for syntactic purposes—like substance mass nouns, which are the same morphological class. Count plurals (whether BPs or SPs) define their reference on the basis of a standard of unity, encoded on the head of Classifier Phrase and corresponding to a count singular. Arabic collectives, by contrast, are singulars that name pluralities directly, and so there is no syntactic expression of what they are collections of. Their meaning makes the individuals in the denotation accessible to semantic operations, as we have seen, but as grammatical entities collectives have nothing to do with singular reference. Only the addition of an extra morpheme (the singulative affix) allows their reference to be articulated into fixed units as a matter of grammatical encoding and not of world knowledge. Therefore, they cannot be directly counted, even though a minimally different structure like 'five of N' as in (7.24a) is grammatical.

In effect, Arabic collectives name semantically plural kinds (cf. Section 4.5.2), viewed through their multiple instantiations rather than as unitary abstract concepts. They can be used for contextually salient sets rather than for the whole set in existence, as (7.18) and (7.19) illustrate, but even then they are better paraphrased as 'cattle/pigeon/egg-property in multiple instances', rather than as 'cows/pigeons/eggs', which can denote any collection down to two- or indeed one-membered ones. The vagueness inherent in this characterization corresponds to the label 'pluriel indeterminé', which does not foreground the single members and yet differs enough from a mass to be able to trigger plural number on an agreeing adjective, as in the Koranic *'as-sahaab-a θ-θiqaal-a* 'THE.cloud(s).ACC THE.heavy.ACC.PL' (Fleisch 1990: 309; Koran XIII, 12).[25] As a conceptual characterization of this interpretation, I would like to suggest lack of identity rather than lack of unity: the elements are viewed as many, but not individuated enough to fix their identity, for instance, by letting one correspond to *a*, another to *b*, and a third to *c*. A plurality of indistinguishable elements does not differ very much from a mass, since the distinction is a matter of conceptualization and is not regimented in grammatical categories. Brustad (2000: 24–5) has argued for the linguistic significance in Arabic of a similarly information-based continuum based on the notion of individuation, resulting from the convergence of

[25] Fischer (1980: 78) observes that the grammatical treatment of collectives as syntactic plurals was especially common in pre classic and Koranic Arabic, where BPs could freely denote either the kind, with singular agreement, or its instances, with plural.

semantic factors (definiteness, specificity, distributive quantification) with textual and pragmatic salience. It seems clear that the part-structural notions of unity and identity are an important ingredient, if probably not the only one, of individuation understood in this sense.

7.5.4 *BPs, collectives, and part-structure conceptualization*

BPs are not collective plurals, because the properties of being a BP or a collective lie on different planes. Collectives are notionally plural but syntactically singular terms, defined as a class by their non-countable interpretation, functionally opposed neither to singulars nor to plurals, but to singulatives. BPs are instead defined as a morphological class, functionally opposed to singulars. The fact that both are stem forms derived on Level I brings about some overlapping between the two concepts in cases where a form may have an ambiguous status—although the real overlap is minimal, since collectives are masculine while BPs in their non-individuating reading trigger feminine singular agreement (at least in classical Arabic: Fischer 1980: 78).[26]

   BPs, *qua* plurals, can have and often do have variously lexicalized readings. In some cases, they pluralize collectives themselves:

(7.25) Collective singular Collective plural (Syrian; Cowell 1964: 369)

samak	'fish'	ʾasmaak	'(many or various) fish'
dəbbaan	'flies'	dababiin	'(many or various) flies'
moož	'wave'	ʾamwaaž	'(many or extensive) waves'
ɣalaṭ	'error'	ʾaɣlaaṭ	'(various kinds or instances of) error'

The meaning of this pluralization ranges from null, that is being a mere formal alternant of the singular, to marking great numerical or non-numerical size (as in 'extensive waves'), the latter two fusing into one for mass nouns proper; of course the sense 'many types of X' is also in principle available. The precise reading varies with the word, and indeed some of these BPs are countable: Cowell (1964: 369) mentions the Syrian *wraaʾ* 'leaves', plural of the collective *waraʾ*, and *wruud* 'flowers, roses', from the collective *ward*. These BPs are countable (*xams ʾwruud* 'five roses', *tlətt ʾwraaʾ* 'three leaves'), and coexist with a singulative, the plural of which is also (obviously) countable. Other nouns have a plural but not a singular collective: such is the case of *sagaayer* 'cigarettes', which is count like the singulative *sigaara* 'a cigarette' (pl. *sigaaraat* 'cigarettes'), but has no singular corresponding to the left-hand column in

---

[26] Zabbal (2002: 2) shows that deflected agreement with BPs signals a group interpretation rather than a distributive one. Notice, however, that BPs with deflected agreement can be antecedent to

(7.26). In sum, not all collectives have a plural, not all plurals of collectives are mass, and not all plural collectives have a singular in the first place. If we add that not all singulative singulars have a corresponding plural, we can see that almost everything in this four-way contrast is subject to lexical idiosyncrasy. The only regularities are the ending and the meaning of singulatives (-*a* ∼ -*aat* and count, respectively), which form an inflectional paradigm. What we have here is the intersection of a number opposition with stem-forming derivations, variously available and interpreted. BPs can be said to be lexical because they are stem-forming derivations, not because they have a particular meaning.

7.5.5 *Number, kinds, and part structure*

We can now return to the observations of Cowell (1964: 370) cited in Section 4.5.1. In Cowell's words, 'Some nouns which in the singular designate a substance in general, or as a sample of its kind, have plurals (in -*aat*) designating a certain batch or indefinite quantity of that substance'. Crucially, these plurals remain mass and do not denote pieces. When analysed along the lines suggested in Section 4.5, the concrete–abstract opposition is related to the part-structural conceptualization expressed by number, where the singular refers to the kind as a unitary concept abstracting away from its instances, while the plural foregrounds the manifold nature of its instances. In Cowell's (1964: 370) examples, this concrete instance-reading is expressed by SPs: *zeet* *ᵓz-zeetuun* 'olive oil' (sg., as a name for a commodity) ∼ *zeet-aat ᵓs-sammaan* 'the grocer's oil' (SP, referring to a particular instance); *mayyet ᵓl-bahr* 'sea water' (sg., the substance) ∼ *mayy-aat ᵓl-bahr* 'the sea's water' (SP, the mass of water contained in the sea). These SPs are all mass and uncountable; so, not only can BPs of mass singulars be countable as in (7.25), but SPs can also be uncountable when pluralizing the same sort of mass noun. The independence of morphology from conceptualization could not be clearer.

The abstract kind reading leads to the same conclusion. It is the interpretation of the singulars *zeet* 'oil' and *mayyet* 'water' in the foregoing examples; but this reading can also be expressed by a BP with feminine singular deflected agreement, here exemplified by *dbuba* and *'adbab*, plurals of *dubb* 'bear':

---

reciprocals, which strictly demand a reference articulated in wholes:

(i) a. n naas   maa katfəhamš   baʕdha   (Moroccan; Brustad 2000: 55)
       the people NEG understand.3. FEM.SG eačḥ. other.FEM
       'people don't understand each other'

   b. əl 'asdiqa   bətruuh   la ʕand baʕd on   (Syrian; Brustad 2000: 59)
       the fṙiend.PL go.3.FEM.SG to at   eačḥ.other
       'friends go to each other's places'

This once more shows how misleading it is to speak of BPs (even with deflected agreement) as mass.

(7.26)   a.   d-dbuba   mənqarḍa   (Moroccan)
                the-bear.PL   extinct.FEM.SG
                'bears are extinct'

          b.   əl-'adbab   munqarida   (Syrian)
                the-bear.PL   extinct.FEM.SG
                'bears are extinct'

Deflected agreement also expresses a generic interpretation for *l-kətᵊb* 'books' in (7.13b) above; cf. also Holes (2004: 171). Like all the evidence considered in this section, this shows at the same time the variety of morphological (and syntactic) resources employed by Arabic to express different conceptualizations along the scale of individuation, and also the lack of a unique, grammaticalized match between conceptualization and morphological expression. Lexicalized number is common in Arabic and has a wealth of morphological manifestations; but it does not correspond to any one grammatical class—in particular, not to BPs.

7.6 Conclusion: BPs and lexical plurality

It is obvious that BPs are somehow lexical, yet their lexicality is hard to pin down, even though they have been intensively studied. The main reason is that the evidence bearing on this issue is complex and downright contradictory. The main goal of this chapter has been to disentangle the many strands of empirical evidence, and to distinguish morphological realization, inflectional information, and lexical semantics.

BPs are a way to fill a cell in an inflectional opposition by deriving a new version of the base stem. In this case, inflection borrows the tools of word formation. BPs are therefore genuinely lexical, but only as forms of words, not words; more precisely, they are stem forms, not lexemes like Italian *-a* plurals. Their lack of disjunctivity is not due to the fact that words do not block each other, but to the fact that stem forms do not block each other as meaningful affixes do. It is crucial, in my account, that BP forms are not morphemes spelling out number. Instead, they are but one facet of the generalized use of stem forms in the grammar of Arabic (and Afroasiatic), across the divide between inflection and lexeme formation. From the theoretical viewpoint, the main conclusion of this chapter is that the stem is a substantive theoretical concept, justified on empirical grounds.

Their form aside, BPs are entirely within the inflectional number system. They do not make up a grammatically defined class, like Irish counting plurals, because from the point of view of their meaning they are like any

other plural—which is why BPs extend to agreeing adjectives. Their forms do not spell out plurality, and what inflectional endings are there are the same as those of the singular; yet they are not singulars or collectives: they are just stem forms, derived from singulars and functionally opposed to them.

This is not to say that BPs are immune from semantically lexicalized number. On the contrary, they enter into the complex web of part-structural conceptualizations expressed in Arabic nominal morphology, and their form makes them particularly close to collectives. But they do not constitute a class on semantic grounds. In fact, BPs cut across contrasts in terms of countability, kind- and instance-readings, and individuation, even though all these are notions needed to explain their semantic range. In short, semantic lexicalization of number is prominent in Arabic, but it is not the meaning of BPs. For these reasons, the main lesson to derive from this analysis of BPs is how independent, but always *inter*-dependent, morphology, syntax, and semantics can be with each other.

8

The system of Breton plural nouns

8.1 Introduction

For the student of plural, Breton is something of a classic. The number oppositions defined by this language are not particularly exotic; it is the relation between morphological and semantic categories that gives this system its fascinating complexity and its exceptional value as a window on the relation between number and part-structure conceptualization. The exponence of plural on Breton nouns is by itself a rich and complex system. In addition, semantic properties such as animacy, countability, boundedness, and cohesion, all impinge on the choice of plural form for a given noun. The choice itself is not always deterministic, and many nouns have more than one possible plural. Most significantly, nouns inflected for plural often appear word-internally as input to nominal, adjectival, or verbal affixation; and, notoriously, they can even be input to pluralization itself.

Unsurprisingly, such a complex system has been the object of much study.[1] In particular, Stump (1989, 1990) has shown that the Breton facts are incompatible with the Split Morphology model advocated by Anderson (1986) and Perlmutter (1988), whereby word-creating morphology is separated from the morphology which spells out the grammatical relations between words in a syntactic structure. In direct contradiction to this view, Breton makes use of the same morphological tools (plural exponence) both to make up words and to spell out their contextual relations. Unlike in languages where plural is a derivational category, plural in Breton is both unambiguously inflectional and unambiguously non-inflectional. What is distinctive is that this applies to plural as a whole, that is, in principle to all exponents. Although some exponents appear more often than others in a non-inflectional capacity, there is no morphological class of lexical or collective plurals, any more than one of purely inflectional ones. The whole category, then, is in principle

---

[1] See Trépos (1956, 1968: 68 77), Hemon (1975: 29 45), Anderson (1982, 1986), Stump (1989, 1990). For recent overviews, see Ternes (1992) and Stephens (1993: 365 8) and the references there contained. I have also relied on the more recent dialectal descriptions in Humphreys (1995) and Wmffre (1998).

available both for spelling out syntactic plurality on noun stems and for creating the noun stems themselves.

This might be taken as a knock-out argument against any theoretically significant distinction between inflection and lexeme formation, and so against the notion of lexeme or lexical base itself (a position upheld in Distributed Morphology, notably by Embick and Noyer 2001; Marantz 1997, 2003). In fact, what Breton shows is not so much that inflection and derivation are one, as that they can involve the same morphological realization, a substantially different proposition which underpins the lexeme-based approaches of Aronoff (1994) and Beard (1995). In this latter approach, which clearly separates grammatical operations from their realization, lexeme formation and inflection are two different *uses* of morphology, which remain distinct even when they share the same realization (see Stump 1998: 18–22, and, in a non-lexemic approach, Borer 2005: 51–8).

In this chapter, I will argue that this is the correct interpretation of Breton plurals. Breton is an extreme illustration of this distinctness, because the same plural morphology can be totally inflectional (obligatory, disjunctive, transparent, word-final) or totally non-inflectional (lexically restricted, non-deterministic, idiosyncratic, word-internal). However, it would be wrong to assume that plurality in Breton can freely switch back and forth between an inflectional and a non-inflectional function, as one might think on the basis of a few examples. In fact, many, probably most instances of pluralization are deterministic and semantically transparent, and certainly not all plural nouns can be further affixed (very few, in particular, support double pluralization). Why should the word-internal use of plurality be so restricted, indeed why should it be restricted at all, if it was a property of Breton grammar? My claim is that the peculiarities of Breton plurals are best understood by relating them to lexical plurals in other languages, not by positing a language-specific (and category-specific) ability to attach to complete inflected words (Stump 1990), nor generally by denying any distinction between inflection and word (stem, lexeme) formation. Plurality can have a non-inflectional function, because morphological form is distinct from its function; but this non-inflectional use is lexeme-specific, because it consists in making up lexemes. What happens in Breton is not qualitatively different from other instances of lexicalized plurality. It is the extent to which it happens that is exceptional, and that because of the properties of Breton grammar.

To see this, it will be necessary to view Breton plural nouns in the context of the system of number inflection, specifically in connection with the other means to express part-structure information. Section 8.2 focuses on the former goal, introducing the core facts and their theoretical implications,

on the basis of the studies of Anderson (1986) and Stump (1989, 1990). The next two sections examine in detail the quite complex relation of plurality with the morphological expression of part-structure information, first (Section 8.3) by relating it to singulatives and duals, and then (Section 8.4) focusing on plurality itself with respect to 'collectiveness'. We will see that, while Breton noun morphology is strongly sensitive to part-structure conceptualization, there is no direct match between type of exponence and any particular plural construal. This leads to the conclusion stated in Section 8.5: the frequent use of plurality as a lexical formant is a facet of the general tendency of the language to express part structure information through nominal morphology; this cuts across all varieties of plural exponents, because it is the category itself that can get lexicalized, not one affix or another (or one exponence class, like broken plurals in Arabic). Breton thus brings out with unusual clarity how part structure information, which is part of lexical semantics, can be encoded through morphological number, which is part of grammar.

8.2 Breton plurals between inflection and word formation

Building on previous studies, this section will recapitulate the main facts about Breton plurals, placing them in the context of the number system and showing in what ways they call into question the distinction between inflection and word formation.

8.2.1 *Plural nouns within the Breton number system*

Number is in Breton a morphosyntactic category just like in other Indo-European languages. Syntactic agreement opposes singular and plural, and this opposition cuts across the paradigms of nouns, pronouns, prepositions, and verbs (examples from Hemon 1995):[2]

---

[2] Breton does not have a standardized transcription system in general use, not just for its strong dialectal differentiation but also for the ideological value attached to competing conventions (Ternes 1992: 381 4). For the purposes of this chapter, I will follow the orthography of the sources, including the phonological transcriptions of Ternes (1970) and Humphreys (1995). Inconsistencies like the use of an accent on the suffix *ou* are due to this choice. I depart from the sources in two respects: I have generalized the spelling *ier*, *ien* for what Trépos (1956) writes *yer*, *yen*, and I have added hyphens to clarify the affixal structure. To follow the data, it must also be noted that Breton, like the other Celtic languages, has a system of initial consonant mutations triggered by a number of morphosyntactic contexts; for instance, the initial *p* of *per* 'pears' is voiced in *kals a ber* 'many pears', cited below in Section 8.2.2. Other instances of mutation occur in word internal position: the *d* of the plural suffix in *merc'h ed* 'girls' is regularly devoiced before the deverbal affix *a* in *merc'h et a er* 'womanizer' (from (3) below). The dual prefixes *di /daou* likewise trigger lenition, turning *brec'h* 'arm' into *di vrec'h* and, less transparently, *glin* 'knee' into *daou lin*. Finally, *c'h* represents a glottal or pharingeal (more rarely, velar) fricative.

(8.1)	SINGULAR	PLURAL	GLOSS
a.	lenn	lenn-ou	'lake ∼ lakes'
b.	te	c'hwi	'thou (SG) ∼ you (PL)'
	eñ	int	'he ∼ they'
	hemañ	ar re-mañ	'this one (m) ∼ these ones'
	va	hon	'my ∼ our'
	e	o	'his ∼ their'
c.	ennañ	enno	'in.3.SG ∼ in.3.PL'
	gantañ	ganto	'with.3.SG ∼ with.3.PL'
d.	skriv	skrivont	'write.PRES ∼ write.PRES.3.PL'
	skrivas	skrivjont	'write.PRETERITE ∼ write.PRETERITE.3.PL'

There is no optionality in syntactic number agreement, nor is there any sign of 'general number' (cf. Section 2.5). The interpretation of singular and plural is the same as in other inflected languages; in particular, plural marking does not imply any collective nuance. As a syntactic category, number in Breton is perfectly regular, with a distinctly non-exotic interpretation, and realized with the generality and obligatoriness that are the hallmark of inflectional morphology.

In nouns, however, this straightforward opposition finds a morphological expression that is anything but straightforward. The singular has no ending at all; the plural has instead a rich array of possible forms, involving affixes, stem-internal changes, combinations of the two, and several cases of suppletion. Even more than the wealth of forms, what makes this a complex system is that the singular–plural opposition interacts with other oppositions, first and foremost that between collective and singulative. As in Arabic, nouns that refer to masses or collections can be turned into individual-denoting count nouns by a feminine suffix, in this case *-enn*. The resulting singulative can in turn be pluralized:[3]

(8.2)	COLLECTIVE	SINGULATIVE	(Trépos 1956: 252, 254, 235)
	stered 'stars'	stered-enn, stered-enn-ou 'a star, stars'	
	sili, silied 'eels'	sili-enn, sili-enn-ou 'an eel, eels'	
	ero/erv 'furrows'	erv-enn, erv-enn-ou 'a furrow, furrows'	

Leaving aside the details, to which I will return, the intersection of singular–plural and singulative–collective oppositions implies that the expression of number on nouns is inextricably linked to the expression of properties relative to individualization.

---

[3] The term 'singulative' was coined by Johan Caspar Zeuss (1871: 294) to describe precisely this derivation in the Britannic branch of Celtic (see Cuzzolin 1998).

In addition, a restricted group of nouns forms a semantically motivated class of duals, by prefixing to the singular a form of the numeral 'two' that agrees in gender. This is not a third grammatical number value on a par with singular and plural, not just because it only applies to a small lexically restricted class (which makes it a 'minor number' in the typology of Corbett 2000: 96–101), but also because such forms can themselves feed plural suffixation, as detailed in Section 8.3.2 below. In fact, Breton duals are best seen as lexicalized paucals, special forms for nouns denoting natural pairs. Indeed, some among them are just morphologically regular plurals, which are dual because of their meaning alone: for example, the formally regular plural *bot-ou* means 'two shoes', 'paired shoes' and not generically 'shoes'. Duality is thus a special interpretation of plurality rooted in the lexical semantics of some nouns. Like individuation, it is not an additional value of the number opposition, but a distinct opposition, which interacts with number both in meaning and, as we will see, in form.

It is in the distribution of exponents that Breton plurals most clearly differ from prototypically inflectional plurals. I will briefly recapitulate the main non-inflectional properties of the system here, since some of them have already been mentioned (Sections 2.7.3, 2.8.3, and 3.4.1), and others will be discussed later on in this chapter:

- the relation between a given singular and its plural is far from deterministic; multiple alternative pluralizations are relatively common (semantically differentiated or otherwise), in contrast with the strict disjunctive blocking typical of inflection;
- pluralization is sensitive to lexical semantics in a way that parallels derivational rather than inflectional morphology; some exponents are sensitive to the animacy of the bases (-*ed* and -*ou* are the default suffixes for animates and inanimates, respectively); in other cases, one and the same noun may select one among two or more plural alternants, depending on whether the noun is conceptualized as referring to discrete individuals, to indistinguishable units, or to a mass;
- some plural suffixes have a special sense for certain choices of nouns (some plurals are semantically duals; the collective -*ach* can be pejorative, the plural -*ou* diminutive); and
- plural affixes can appear inside some derivational, or at least non-inflectional, material (verbal, nominal, and adjectival affixes, as well as diminutives) and also inside pluralization itself; this gives rise to doubly pluralized nouns, like *merc'h-ed-ou* 'girl.PL.PL'.

It is especially this last aspect that has brought Breton to the fore of the theoretical debate. For ease of reference, I repeat the examples introduced in Section 3.4.1.2:

(8.3) a.

Gloss	Singular	Plural	Plural of dimin.	(Stump 1990: 105)
'bird'	labous	labous-ed	labous-ed-ig-où	
'cat'	kazh	kizh-ier	kizh-ier-ig-où	

b.

Gloss	Singular	Plural	Derived verb	(Stump 1990: 107)
'part'	darn	darn-où	darn-aou-iñ 'to distribute'	
'bird'	evn	evn-ed	evn-et-a 'to hunt for birds'	

c.

Gloss	Singular	Plural	Derived adjective (Stump 1990: 108)
'rock'	maen	mein	mein-ek 'rocky'

d.

Gloss	Singular	Plural	Der. agentive noun (Stump 1990: 113)
'apple'	aval	aval-où	aval-aou-er 'apple-hunter (hedgehog)'
'girl'	merc'h	merc'h-ed	merc'h-et-aer 'womanizer'

e.

Gloss	Singular	Plural	Double plural	(Stump 1990: 114)
'child'	bugel	bugal-e	bugal-e-où	
'girl'	merc'h	merc'h-ed	merc'h-ed-où	

What is plural in the last column of these examples is not an entire word but only a part of it, which has no syntactic role: *aval-aou-er* 'hedgehog', for instance, is singular despite the embedded plural *aval-ou*. Plural morphology, in these cases, forms stems and not words.

8.2.2 *Inflection and derivation*

As Stump (1989, 1990) pointed out, these facts call into question the very conceptual distinction between derivation and inflection, that is, between that part of morphology which creates a word as distinct from all other words, and that which specifies a grammatical form among other forms that a single word can assume. His crucial observation is that the plural morphemes that appear word-internally as lexical formants (like *-ou* in *aval-ou-er*) are the same ones which function as inflectional plurals when modifying entire words. The forms listed under 'plural' in (8.3) are for all intents and purposes the inflectional plurals of the relevant nouns, whether they are suffixed or realize plurality through stem revowelling, as *mein* 'stones', or through an obsolete and unproductive stem extension pattern, as *bugal-e* 'children'. The point is crucial, because Anderson (1986) (who must be credited for noticing the significance of these data) had argued that the only reason some plurals can appear word-internally in Breton is that they are 'collectives', notionally

but not morphologically plural: monomorphemic nouns that refer to plural-
ities or masses, like *bili* 'gravel', *dilhad* 'clothes' or *per* 'pears'. Anderson's
suggestion is promising, because it relates the word-forming function of
Breton plurals to the value of plurality as a lexical property of some stems.
I will develop this very insight in what follows. But as it stands, the claim that
word-internal plurals are not inflectional plurals is empirically falsified.
Stump (1989: 264–5) gives three arguments showing that collectives are plural
both morphologically and syntactically:

(i) It is not just monomorphemic collectives that can feed pluralization,
    diminutive affixation or word-class-changing derivation; the same
    processes also target plurals formed by regular suffixation: cf. *darn-
    où*, *evn-ed*, *kizh-ier*, *bugal-e*, and *merc'h-ed* in (8.3).

(ii) There is no syntactic difference between derived plurals and collectives,
     which both count as plural for agreement purposes:

(8.4) N′ eo ket mad ar bili-se;   re   vihan int (Stump 1989: 264;
                                                              Trépos 1968: 66)

      NEG is NEG good the gravel.THAT; too small are
      'that gravel is no good, it is [lit. they are] too small'

The key word here is the plural *int*, as opposed to the singular *eo*, referring to
bili 'gravel'.[4]

(iii) The singulative *-enn* attaches both to derived plurals and to basic
      collectives:

(8.5) | GLOSS | SINGULAR | PLURAL | SINGULATIVE *(sg)* |
      |-------|----------|--------|-------------------|
      | 'starling' | tred | tridi | tridi-enn 'a starling' |
      | 'fish' | pesk | pesk-ed | pesk-ed-enn 'a fish' |
      | 'pears' | — | per | per-enn 'a pear' |

This seems to show that monomorphemic stems like *per* have the grammat-
ical properties of every other plural.

   Stump's arguments have uneven force. The first is valid but it is less
cogent than it looks. As we will see in greater detail in Section 8.3.3.1, it is
true that plurals can feed further derivation or pluralization whether they are
monomorphemic stems (like *mein* in *mein-ek*) or regularly affixed inflectional
forms (like *aval-aou* in *aval-aou-er*), but the process is lexically restricted,

---

[4] 'Gravel' is the gloss given by Stump; Ywan Wmffre (pers. comm.) suggests 'pebbles' as a closer
translation, in so far as *bili* differs from *grouan* 'gravel' (Trépos translates: 'ces galets ne sont pas bons,
ils sont trop petits'). The point, in any event, is that *bili* triggers plural agreement. Ternes (1970: 187)
has more examples of monomorphemic collectives anaphorically resumpted by plural pronouns.

especially for double pluralization. Word formation can indeed target regularly inflected plurals, but whether this happens or not still depends on the choice of the plural noun and is therefore lexically conditioned—although not along the lines proposed by Anderson (1986). Stump's third argument is effectively neutralized by the observation that the singulative attaches not just to syntactic plurals but to singular nouns too, as we will see in Section 8.3.1. The second argument is unassailable, however, and is confirmed by the observation that numerals and *ped* 'how many?' require the singular and are ungrammatical when followed by a derived plural as well as by a monomorphemic collective plural:[5]

(8.6)   a. peder   ber-enn/*ber
            four   pear-SINGULAT.SG/pear.PL
            'four pears'

        b. ped   marc'h   ho   peus?   (Trépos 1956: 18)
            how.many   horse.SG   2.PL   have.2.PL
            'how many horses do you have?'

Correspondingly, collectives are instead grammatical after determiners requiring the plural: *kalz a ber* 'many pears', *neubeud a ber* 'few pears' (where *per* displays initial sonorization). There is no room for doubt, then, that the monomorphemic, stem-internal plural forms are grammatically equivalent to, or are the same grammatical word as, inflectional suffixed plurals. This does not mean, of course, that every plural exponent has an equal likelihood of being employed as a derivational formative; there are patterns and tendencies, some of which will be relevant later on. But it would be empirically wrong to state that the phenomena exemplified in (8.3) concern only idiosyncratic forms, or 'collectives' as opposed to grammatical plurals. Breton puts truly inflectional forms to truly derivational use.

8.2.3 *Inflection-as-derivation is lexical plurality*

For Stump (1990), these phenomena show that the processes which create words can use as building blocks fully inflected words as well as units smaller than words. Derivation and inflection can apply, in Stump's terminology, to

---

   [5] *Peus* in (8.6b) is a form of the verb *endevout* or *kaout* 'to have'. Breton is unique among (modern) Celtic languages in expressing possession by a verb distinct from 'to be'; in fact, the verb employed in this function is transparently related to 'to be', since it consists of a pronominal prefix and a form of *eus*, which is a suppletive stem of *bezañ* 'to be'. Hemon (1995: 252) and Press (1986: 142) give the second person plural form as *hoc'h eus*, with an invariant stem *eus* instead of the allomorph *peus* (which instead appears in Ternes 1992: 425).

both a word and a root (the part shared by all the word forms in a paradigm), with derivation turning either into a root, and inflection into a word.[6]

In a sense, this analysis cannot be faulted. That complex words can be constructed by using already complete words is exactly what the Breton data show. But to take this as evidence that inflection can feed word formation is at least slightly misleading; what gets used for word formation in Breton are not generally inflected forms, but specifically plural nouns. The word forms that make up verbal paradigms, for instance, are never suffixed by derivational or inflectional morphology. If what explains *merc'h-ed-où* 'girls' (structure [girl.PL.PL]) or *merc'h-et-aer* 'womanizer' ([woman.PL.er]) is the ability to derive an inflected word from an already inflected word, one should ask why there are no things like double past tense on verbs, for instance (which would not be semantically odder than double plurals). What is more, there would be no reason why the same should not extend to grammatical words like auxiliary verbs or pronouns and inflected prepositions, leading to hypothetical formations like **gant-añ-o* 'with-3.SG-PL'.

Actually, Breton *does* offer sporadic examples of inflectional categories other than plural inside a word: Trépos (1956: 262) mentions isolated dialectal verbal forms which seem to display something like emphatic reduplication of an inflectional marker, like *me a vantefafe* 'I would throw (up in the air)' (conditional) for the regular *me a vantefe*.[7] He also reports (p. 268) that the singulative (pluralized) can be added to the pronoun *hini* 'one', to form *hini-enn-ou* 'a few ones'. It is true, then, that inflection and derivation are in Breton more mutually permeable than in other languages. But these sporadic cases of inflected words feeding further morphological processes come nowhere near the extension of nominal plurals, and nothing makes this expected in Stump's account.

The alternative I propose should be clear by now. Breton, much more than other languages, uses for word formation a category whose chief domain of application is inflectional morphology, namely, nominal plurality. Using grammatical means for word formation is not a contradiction in terms, in so far as the semantic content of a grammatical category can be defined independently of its grammatical and syntactic function. The inflectional category involved in Breton word formation is chiefly plurality, because this

---

[6] The analysis is refined and expanded in the theory of Paradigm Function Morphology of Stump (2001), but without new evidence from Breton.

[7] This apparent reduplication probably reduces to a morphophonologically conditioned phenomenon. Yvon Gourmelon (pers. comm.) informs me that the conditional marker *f* is (dialectally) repeated after the *ta* of emphatic forms, but only if the preceding stem ends in a vowel: for example, the regular and emphatic forms of the vowel final stem *ya* 'to go' are respectively *ne a ya fe* and *ne a ya f ta fe* 'I would go'.

category can encode part-structure information which is part of the meaning of the lexical base. This happens specifically on nouns and not, for instance, on pronouns, because nouns are lexical and pronouns are grammatical, and the use of plural morphology as a word formant is an instance of lexical plurality. To simply claim that inflection in Breton can re-enter the word formation process is not so much wrong as misleading, because the plural that appears in the body of a word is a special type of inflectional marker—one used as a lexical formant. The exceptional patterns of pluralization in Breton are thus related to lexicalized uses of plural in languages that, unlike Breton, never allow further affixation of inflected words. Only by placing the Breton facts in the context of lexical plurality can we see this cross-linguistic connection.

8.3 The grammatical relevance of part structure

To substantiate the claim that the peculiarities of Breton plural nouns derive from the lexicalization of plurality, we must make explicit in what ways plurality impacts lexical semantics. That Breton plural nouns can mean much more than simply 'many *x*' is well known, but it has not played any significant role in the morphological analysis of this category, except for the (wrong) contention that monomorphemic collectives like *per* 'pears' are not grammatical plurals. This section will therefore clarify the relation between the meaning of plurals and that of duals and singulatives, which are the other two morphological categories that express part structure conceptualization. Besides achieving a better understanding of the place of plural in Breton grammar, this analysis will show how important the conceptualization in terms of unity and identity (boundedness, discreteness, re-identifiability) is for Breton noun morphology. The traditional descriptive label 'collective plurals' has a real significance, and underlies the use of plurality as a formant of the lexical stem's interpretation. Just as importantly, however, it does not match any one morphological type: this discrepancy between interpretive and morphological categories lies at the heart of the Breton plural system.

8.3.1 *Singulatives*

Like the closely related Welsh, old Breton had two singulative endings -*in* and -*en*, one for each gender value (Hemon 1975: 39); middle and modern Breton only retain the feminine -*enn*, so that singulatives are invariably feminine, as in Arabic. The fact itself of determining the gender of a noun indicates that -*enn* is a derivational suffix, and that the derived singulative and the base it is suffixed to are distinct nouns and not inflectional forms of the same noun.

This is confirmed by the fact that singulatives can be pluralized, as exemplified in (8.2) above. In fact, the pluralization of singulatives is probably the most regular instance of nominal pluralization in Breton: the plural marker is always -*ou* (-*ed* for humans), and the meaning is always 'many individuals' (McKenna 1988: 224; Wmffre 1998: 14).[8]

In principle, a morphologically distinct word may be associated with a cell of another word's paradigm; that is what happens with suppletive plurals like the Russian *čelovek* ~ *ljudi* 'person ~ people'; or an inflectional paradigm may have some cells based on stem X and others on a stem X_i derived from X, as in Arabic broken plurals. In its most straightforward use, exemplified in (8.2) and (8.5) above, the Breton singulative does exactly that: -*enn* attaches to a syntactically plural base denoting aggregates and derives a singular noun denoting the atoms of those aggregates: plural *stered* 'stars' (stem X)> singulative *stered-enn* 'a star' (stem X_i). Unsurprisingly, in many instances this has caused the singulative to oust the original singular, as *stered-enn* has replaced the obsolete *ster* (cf. Trépos 1956: 236; Hemon 1975: 40).

It would be wrong, however, to reduce the singulative to a word-formation means for deriving singulars out of plurals. That is the situation in modern Welsh, where the masculine -*yn* and the feminine -*en* turn unsuffixed plural bases into suffixed singulars (King 1993: 48; Cuzzolin 1998: 124–5). In Breton, by contrast, singulatives are demonstrably indifferent to the grammatical number of the bases they attach to: what counts is that the entities denoted by the base are not individual, in the technical sense of being neither discrete nor identifiable. Trépos (1968: 67) reports that some substance nouns which feed singulative derivation are syntactically singular (cf. also Hemon 1975: 40): these include *dero* 'oak(-wood)', *ed* 'wheat', *glao* 'rain', *gwiniz* 'wheat', *kolo* 'straw', *plouz* 'straw', *traez* 'sand'. *Gwiniz*, for instance, triggers singular agreement, and Trépos expressly rules out the plural *int* in place of the singular *eo*:[9]

(8.7) gwelet em   eus   ho   kwiniz, kar   eo (Trépos 1968: 67)
       seen   vpt.1.sg have 2.pl wheat, nice is
       'I have seen your wheat, it [sg.] is beautiful'

---

[8] Breton also expresses the analogue of a singulative derivation by the use of classifier like elements, like *penn* 'head' or *pez* 'piece': *moc'h* 'pigs' ~ *pennmoc'h* 'a pig', *dilhad* 'clothes' ~ *pez dilhad* 'a garment' (Trépos 1956: 236), /*pu:r*/'leeks' ~/*pen*'*pu:r*/'a leek' (Humphreys 1995: 255). These might well involve compounding rather than classifiers, however, at least to judge by the decontextualized examples in the descriptions (cf. also Hemon 1995: 23).

[9] vpt in the gloss of this example stands for 'verb particle', a grammatical morpheme placed in front of inflected verbs (except for some constructions and for 'to have') when they follow a clause initial element other than subject or object.

The meaning indicated for the corresponding singulatives is instructive (Trépos 1968: 67). *Glav-enn* (with regular phonological change from *glao*), *plouz-enn*, and *ed-enn* mean 'a raindrop', 'a straw', 'a wheat stalk'; but *derv-enn* means, or at least can mean according to Trépos, 'an oak wood' (i.e. a forest); *gwiniz-enn* 'a wheat field', *traez-enn* 'a beach', and *kolo-enn* 'a beehive' (actually 'straw skep'). So, in a derived noun *X-enn*, the precise sense of 'X-unit' varies with the word; it apparently varies also within the same word, because Trépos mentions a textual example (1956: 269) where *traez-enn* means 'a grain of sand' rather than 'a beach' (actually the example is plural). But singulatives are also formed on *count* singulars. Trépos (1956: 268–9) reports alternations between underived singular and singulative singular such as *botez* ~ *botez-enn* 'shoe', *lod* ~ *lod-enn* 'part', *karreg* ~ *karreg-enn* 'boulder, rock', *kiez* ~ *kiez-enn/ kioz-enn* 'bitch', and more; he explains that the singulative here 'rend l'objet plus proche, plus matériel, plus tangible' [makes the object closer, more material, more tangible], illustrating this with the following contrasts:

(8.8)   a.   eur   giez   eo   (Trépos 1956: 268)
          a   bitch   is
          'it is a bitch'

         b.   eur   gioz-enn   vad   'neus   Yan[n]
          a   bitch-SINGULAT   that   have.3.SG.MASC   Yann
          'its a good bitch that Yann owns'

(8.9)   a.   peb   hini   'no   e   lod[10]
          each   one   have.FUT.3.SG.MASC   his   part
          'everyone will have his part/share'

         b.   bras-oc'h   eo   e   lod-en[n]
          large-COMPAR   is   his   part-SINGULAT
          'his part/share is larger'

In (8.8), *kiez* is a predicate nominal, while *kioz-enn* refers to a particular animal; in (8.9), *lod* refers to a part as an abstract subdivision ('lorsque le partage n'est pas encore fait', 'when the partition has not been made yet'); *lod-enn* refers instead to a situated object. Similarly, *to* 'roof' is used in compounds like *plouz-to*, *kolo-to*, both 'roofing (thatching) straw', while *to-enn* denotes the actual roof (ibid.). It appears then that the derivation in *-enn* produces count nouns interpreted not just as bounded, but as bounded individual objects endowed with unity and identity. The function of singulatives, then, consists not only in picking discrete entities out of

---

[10]   What Trépos represents as '*no* here apparently corresponds to the full form *en devo*.

undifferentiated masses, but also in turning abstract object types into identifiable objects. This clarifies the function of singulative formation from plurals like *stered* 'stars' or *gwer* 'glass, glasses (drinking vessels)'. Yvon Gourmelon (pers. comm.) has provided some enlightening examples of the contrast between these plurals and those of the corresponding singulatives: while *stered* refers to the stars in the sky, *stered-enn-ou* can refer to a few individual stars, or to the brass stars on military epaulettes, or to the stars printed on a bottle label. *Gwer* refers to glasses (as a kind); *gwer-enn-ou* may denote some glasses on a table. What changes is not primarily the status as mass or count, but more precisely the identifiability of the objects in question, which can be individually pointed to.[11]

Not only is *-enn* indifferent to the number of a nominal base; it does not even need a nominal base at all, but can nominalize adjectives and verbs: *koant* 'beautiful' > *koant-enn* 'a belle', *baz* 'shallow' > *baz-enn* 'a shoal, sandbank'; *drailh-a* 'to break' > *drailh-enn* 'a fragment', *prezeg* 'to preach' > *prezeg-enn* 'a sermon'. Trépos (1956: 270) notes that the singulative parallels in this function the diminutive, which can turn verbs, adverbs, or adjectives into nouns with individual interpretation: for instance, with the *-an* diminutive suffix, *huneg* 'sleepy' > *huneg-an* 'dormouse, the sleepy one' (this use of the diminutive is a characteristic of Britannic languages; cf Pedersen 1913: 57–8). Besides, the Breton *-enn* occasionally has a diminutive function itself: *kalon-enn* 'heart-shaped pendant' (from *kalon* 'heart'), *lagad-enn* 'eyelet' (from *lagad* 'eye') (Trépos 1956: 270; see also Stump 1998: 19 and the comparative study of Stump 1993). Notice that these bases are count nouns, so *-enn* cannot have a packaging function.[12]

What all this shows is that the Breton singulative is a nominalizing derivation whose distribution and function are determined by the semantics of the

---

[11] The interpretation of *enn* as an individualizing, not just packaging, operator, makes a falsifiable prediction: no noun denoting an abstract property (like 'beauty' or 'wisdom') should be formed by singulative affixation in Breton. This is because abstract property nominalizations are unique concepts, quite different from the re identifiable referents that properties are true of. Hence, they cannot be 'individuated' (i.e. provided with distinctive identity) through singulative affixation. Note that this applies specifically to nominalized properties, not generally to abstracts. This is clarified by an example due to Yvon Gourmelon (pers. comm.): from *kred* 'faith, trust', the singulative *kred enn* 'belief' is derived, which is still abstract. However, it refers to the contents of believing, not to the property of having faith, as is apparent from an example like *kred enn ou an lliz* 'the dogmas of the church'. My ignorance of the language prevents me from claiming that the prediction is borne out, but none of the sources I have consulted features a single abstract noun in *enn*.

[12] On the relations between singulative and evaluative semantics, see Jurafsky (1996). Cuzzolin (1998) cogently argues that the Welsh suffix *yn* had the two functions already in middle Welsh as a reflex of the semantics of the base: diminutive with count nouns, singulative for mass nouns (rarely) and especially for nouns of aggregates with discrete but indistinguishable parts. This important result highlights the tight connection between this derivation and the part structure semantics of the lexeme.

input lexeme, in particular by the part structure of its denotation. It has an individualizing function, which turns several types of bases into singular nouns denoting bounded and identifiable entities. Applied to nouns (singular or plural) denoting masses or mass-like collections of indistinguishable elements, this derivation produces nouns referring to bounded atomic entities variously related to the base: members of collections, atomic parts of granular masses, detached pieces of matter, but also objects made up of a material, or bounded extensions of a mass, like a field or a beach. When applied to nouns (singular) denoting bounded entities that cannot be re-identified as the same (abstract kinds, equivalence classes like the constituents of an abstract partition), the singulative derivation produces nouns referring to the individual instances endowed with identity properties (situated objects, which can be picked out from among others). When applied to property-referring stems, it produces nouns denoting entities with that property. No grammatical principle prevents -*enn* from attaching even to singular nouns denoting already individual entities; but in that case, the individualizing function would be redundant and the suffix can only have a diminutive reading, subject to lexical, dialectal, and presumably idiolectal variation.[13] From the point of view of exponence, the singulative derivation nearly coincides with the domain of application of the suffix -*enn*, but not totally, because of the lexically restricted classifier-like constructions mentioned in note 8. It is therefore best seen as an abstract derivation, expressed through -*enn* suffixation except in lexically specified cases, with a semantically and syntactically constant output (feminine singular nouns interpreted as bounded individuals). The grammatical diversity of its possible inputs conceals an underlying semantic unity, which cannot be brought out without reference to part structure.

8.3.2 *Duals*

As mentioned in Section 8.2.1, duals are a marked class restricted to nouns for natural pairs, mostly twin body parts and matching items of clothing. Morphologically, dual formation consists in prefixing to the singular the numeral 'two' (*daou* masc., *diou* fem.), with phonological readjustments; here are some examples from Hemon (1975: 42) and Trépos (1968: 70):

---

[13] This does not cover *hini enn ou* 'a few ones', based on the pronoun *hini* 'one' (Section 8.2.3). Even there, however, *enn* is reported to force a specific reading of the pronoun, highlighting the distinct individuality of each of the elements in its denotation. Here *enn* does not select a lexical meaning at all, because the pronoun is a grammatical morpheme. This type of affixation is clearly exceptional, since the singular of the form discussed by Trépos (1956: 268) is 'inusité'.

(8.10)   Singular   Dual
           Masculine   lagad 'eye'   daou-lagad
                         dorn 'hand'   daou-arn
                         glin 'knee'   daou-lin
           Feminine   brec'h 'arm'   di-vrec'h
                         skouarn 'ear'   diou-skouarn
                         morzed 'thigh'   di-vorzed

While the English *two ears* can refer to any two-membered set of ears, no matter who each ear belongs to, the reference of *diou-skouarn* ranges over naturally occurring twin organs. This makes the denotation dependent on world knowledge about the natural cohesion of referents, as is typical for lexically restricted 'minor numbers' (Corbett 2000); the restriction to naturally occurring pairs does not arise compositionally from the form 'two N'. If the facts stopped there, the Breton dual would only be an example of lexicalized morphological number; a particular match of form and meaning.

Crucially, however, form and meaning do not actually match. Exactly the same interpretive facts hold for other nouns that are formally plurals (Trépos 1956: 228, 249, 265; Trépos 1968: 70; Ternes 1970: 206–8; Press 1986: 71; Humphreys 1995: 259–60). This applies to nouns that include *bot-ou*, 'shoes, clogs' (sg. *bot-ez*), *loer-ou* 'tights, socks' (sg. *loer*), *treid* 'feet' (sg. *troad*). Particularly instructive is the case of the plurals *ot-ou*, *brag-ou*, *brek-ez*, all of which mean 'one pair of trousers'. The meaning of these plurals is clearly lexicalized, and it coincides with that of formal duals. So, duality is really an abstract category, realized through different morphological means according to the lexeme. Not only is the prefixed 'two' not necessary for duals; at least in some dialects, it is not sufficient either. Humphreys (1995: 254), in his description of the dialect of Bothoa, distinguishes proper duals from plurals that are dual in form only, and notes e.g. /d'əwgəd/, composed with *daou-*, means just 'eyes', whether they belong to one or more persons, or could apply to 'eyes' in the soup (which are not paired). Ternes (1970: 207) reports a similar if not identical evolution in the dialect of Groix. Five nouns there have *di-* as part of the *singular* stem, which, when pluralized, gives rise to a dual reading, as in /diskwaj/∼/diskwaj-ew/'shoulder ∼ pair of shoulders'. Whatever the dialectal diffusion of this bleaching of the dual prefix (which Trépos does not mention), it shows that the Breton dual is a category expressed through morphological means but primarily defined by its semantics.[14]

[14] An interesting observation of Ternes (1970: 210) about the Groix dialect brings out the primarily semantic nature of duality. To emphasize duality, the numeral 'two' may be used. Like all numerals, 'two' governs the singular, not the plural. But as we have seen, some nouns are plurals in form and dual in meaning, like in this dialect /bron ew/ '(two) breasts'. When governed by the numeral 'two', such

What unifies what I will call plural-duals (e.g. *bot-ou*) and prefix-duals (e.g. *daou-lagad*) is that both articulate the reference domain into individual units, each of which is 'one' for the whole-properties given by natural cohesion, but 'twin' because it must comprise two (or exceptionally more) mutually indistinguishable elements. There are two levels of oneness: the basic entities that make up each pair (an ear, an eye, a glove, . . .) are one because there cannot be a two-membered set unless each of its members counts as one; and the pairs themselves, which are one because of the functional cohesion holding together every two ears, eyes, etc. These pairs are not only whole, that is have unity properties, but they are also distinguishable from each other and identifiable, that is have identity properties. The members of each pair lack instead identity properties, being conceptualized as interchangeable. Since the denotation of duals is based on pair-sized units, its formal structure is akin to that of singulars: *treid* 'feet', for instance, is true of each element of a set {a, b, c, . . . }, where the single elements are not single feet but natural pairs of feet (cf. Desbordes 1983: 27–8; Ternes 1970: 191; Favereau 1997: 45). It should not be too surprising, then, that duals can be syntactically singular: Trépos (1956: 249) cites numerous textual examples like *eur maneg-ou* 'a pair of gloves' (lit. 'a gloves'), or even *peb a luned-ou* 'to each his pair of spectacles'; cf. also McKenna (1988: 221) and Humphreys (1995: 261).[15]

The relevance of all this for plurality is that duals interact with pluralization in a way that affects its conceptualization. Because this conceptualization is only available to certain words as a function of their lexical semantics, it is a lexical property. Therefore, the interaction of plural and dual concerns the lexical use of plurality. The starting point is the observation that duals can be pluralized, even those that are already formally plural:

(8.11)   a. Singular   Prefix-dual   Prefix-dual + plural
            morzed 'thigh' di-vorzed   di-vorzid-i
            brec'h 'arm'   di-vrec'h   di-vrec'h-iou
            dorn 'hand'   daou-arn   daou-arn-ou   (Trépos 1956: 227)

nouns must therefore be singular: /diw vron/, lit. 'two breast'. By contrast, duals that are prefixed with *di* are syntactically singular and therefore can be governed by 'two', giving rise to doubly dualized forms like /diw di skwarn / 'two ears' (this is a real dual, not one of the five words where *di* is just part of the stem). So, for grammatical purposes /bronewl/ is just a plural and /di skwarn/ is just a singular; duality is an entirely semantic matter. For the record, Trépos (1956: 82) reports that the usual dual for 'breasts' is the prefixed *di vronn*, while the formally plural *bronn ou* 'serait grossier, parce que trop évocateur'.

  [15] All of these examples are formally plurals, not prefixed duals of the form 'two N'. This appears to be systematic: Humphreys (1995: 261) states that in the Bothoa dialect formally plural duals normally take the indefinite article, which is only singular, while this is rare for prefixed duals; as far as anaphora and verbal agreement are concerned, however, formally plural duals count as plurals (cf. also Hemon 1975: 42 3).

b. Sᴍɢᴜʟᴀʀ | Pʟᴜʀᴀʟ-ᴅᴜᴀʟ | Pʟᴜʀᴀʟ-ᴅᴜᴀʟ + ᴘʟᴜʀᴀʟ

b. Sɪɴɢᴜʟᴀʀ	Pʟᴜʀᴀʟ-ᴅᴜᴀʟ	Pʟᴜʀᴀʟ-ᴅᴜᴀʟ + ᴘʟᴜʀᴀʟ
troad 'foot'	treid	treid-ou
loer 'tights/socks leg'	loer-ou	loer-ei-er
bot-ez 'shoe'	bot-ou	bot-ei-er
—	ot-ou 'trousers'	ot-ei-er (Trépos 1956: 227–8)

This important paradigm requires a number of comments. First, all descriptions agree on this pattern of pluralization, over and above dialectal differences. This does not imply that Trépos's specific examples are all attested across the dialectal spectrum, neither as forms nor as lexemes, but it does imply that all dialects can pluralize the dual (both prefix- and plural-duals). Second, the *-ei-er* endings of the last three forms is a regular phonological development of *-ou-ier* in the main western dialect group, corresponding to *-euier* of the eastern Vannes dialect (cf. Hemon 1975: 43). Third, Trépos does not indicate a singular for *ot-ou* 'trousers', which may well be a *plurale tantum*, but that is irrelevant. The point is that both prefix- and plural-duals can be pluralized as a class, even though this leads to a double pluralization (structure sᴛᴇᴍ+ᴘʟ+ᴘʟ). From the point of view of morphology, then, these formations show that the pluralization of duals is not sensitive to the difference between prefixed singulars (*di-vrec'h* 'arms') and plurals lexicalized as duals (*treid* 'feet').

At first sight, the meaning of pluralized duals should be straightforward: if a dual means a natural pair, its plural should refer to sets of such pairs. This is not what happens, however. Nouns whose dual is formally a simple plural can refer, when pluralized again, to sets of individual elements, not necessarily arranged in pairs: this is confirmed by all sufficiently detailed descriptions, whether they are comparative or single-dialect studies (Ternes 1970: 200; Hemon 1975: 44; McKenna 1988, §406b; Humphreys 1995: 260).[16] For Trépos (1956: 265), this category of double plural expresses 'soit l'idée de « plusieurs paires de … », soit simplement l'idée de « un grand nombre de … » '. The following sentences, which I owe to the courtesy of Yvon Gourmelon (pers. comm.), show that a double plural like *bot-ei-er* (corresponding to *bot-oi-où* in other dialects) is an appropriate antecedent for 'one by one', and must therefore make the single shoes, not just the pairs, accessible to semantic interpretation. To obtain the reading 'one pair after the other', the pair unit must be explicitly mentioned, as in (8.12b):

---

[16] Crediting Greg Stump as his source, Corbett (2000: 36) states that a pluralized plural dual like *manigeier* 'gloves' does not have to denote pairs and can refer, for example, to three gloves. This last point contradicts Trépos' (1956: 265) claim that non-paired readings imply a large number, but Breton varies so much dialectally that it may well be correct.

(8.12) a. sellet   en   doa   ouzh   ar   botei-er   an   eil
           seen   VPT.3.SG have.PAST.3.SG   at   the shoe-PL   a   one
           war-lerc'h   eben
           on-back   other
           'he examined the shoes one by one'

        b. sellet   en   doa   ouzh   ar   botei-er   an   eil
           seen   VPT.3.SG   have.PAST.3.SG   at   the   shoe-PL   a   one
           re   war-lerc'h   ar   re-all
           pair   on-back   the   other
           'he examined the shoes one pair after the other'

The addition of a second plural morpheme to these plural-duals, then, eliminates the restriction to natural pairs. Of course the resulting forms can still refer to natural pairs, just like the English *shoes* can, but not as a necessary entailment. To express an explicitly non-cohesive reading, where the elements in the denotation do not form natural groupings, one can pluralize the singulative; as Steve Hewitt points out to me, *maneg-enn-ou* 'glove-SINGULAT-PL' means 'individual gloves not in pairs', where it is an acceptable form.[17]

Prefix-duals apparently cannot express this non-pair reading, which makes sense since they are affixed with 'two'. This does not mean that they are extraneous to the 'paired' ~ 'non-paired' opposition, however. Rather, the non-paired reading is associated (on a lexically restricted basis) with the plural form *without* a prefix, resulting in a four-way paradigm where duality and plurality cross-classify:

(8.13) | SINGULAR | PLURAL | TWO+SINGULAR | TWO+PLURAL |
        |----------|--------|--------------|------------|
        | | | | (Trépos 1956: 227) |
        | brec'h 'arm' | brec'h-iou | di-vrec'h | di-vrec'h-iou |
        | lagad 'eye' | lagad-ou | daou-lagad | daou-lagad-ou |
        | skouarn 'ear' | skouarn-iou | di-skouarn | di-skouarn-iou |
        | dorn 'hand' | dorn-ou | daou-arn | daou-arn-ou |

The singular and plural duals in the third and fourth column refer respectively to paired objects and sets of paired objects; the non-prefixed singulars in

[17] Trépos (1956: 266) reports that in some dialects the double plurals in *ei er* (from *ou er*) denote pluralities of atoms, while those in *ou you* denote pluralities of pairs: *bot oi ou* features in the example 'put on your shoes, children!', while *bot ei er* is appropriate for a context like 'go out and clean the shoes', where it 'évoque simplement l'idée de chaussures en vrac' ('simply suggests the idea of loose shoes'). Such contrasts, however, are exceptional, at least in the present day (thanks to Steve Hewitt for clarifying this point).

the first column denote of course single objects. The non-prefixed plurals in the second column refer instead to sets of the corresponding elements without a conceptualization in pairs, the same function of double plurals like *bot-ei-er* 'shoes' in most dialects. Again, the conceptual distinctions remain the same across morphologically different structures: the expressions of pluralities of pairs are different from those of pluralities of atoms, whether the latter correspond to STEM+DUAL+PL as in (8.11) (including where this means effectively STEM+PL+PL) or to a simple STEM+PL, as in (8.13). There is one important difference, however: the nouns in the class of prefix-duals only refer to body parts, as opposed to items of clothing that derive their cohesion from that of body parts. Duality has for body parts a particularly strong semantic motivation, and so when it is removed, they are liable to shifting their lexical meaning. According to Trépos (1956: 227, 265), *lagad-ou* can refer to 'eyes' in the soup, *skouarn-ou* to the lugs of a vessel, *dorn-ou* to handles of pots, and *brec'h-iou* to the 'arms' by which horses riding in circle are tied to a central pole ('les bras d'un manège'; cf. Humphreys 1995: 258–9 for more recent dialectal examples). This shift is not obligatory, but only a possible by-product of the conceptualization without cohesive duality. In this way, a morphologically straightforward plural is paired to the semantically most peripheral sense of the lexeme. Revealingly, exactly the same holds for the simple plurals of nouns like 'hands' or 'wings' in Biblical Hebrew (Kautzsch 1980 [1910]: 243, §87.3), where, as in modern Hebrew (cf. Section 2.6.4), the dual is a lexical number rooted in the semantics of certain nouns, as opposed to a grammatical category marking nouns and pronouns (see Plank 1996 on the similar distribution of dual in Maltese).

In sum, the complexity of Breton duals stems from the lack of isomorphism between morphological realization and a basically simple abstract pattern, where an opposition between singulars and plurals intersects an opposition between two ways of defining what counts as 'one':

(8.14)   atomic individuals   sets of atomic individuals
          two-membered individuals   sets of two-membered individuals

These semantic oppositions do not map to a corresponding set of morpho-logical oppositions. The conceptualization of 'one' as 'twin' is expressed either by dual prefixes or by a lexicalized reading of plurality, depending on the noun. These bases can in turn be pluralized, but then only the prefixed bases (prefix-duals) have a compositional meaning 'many pairs'. Where duality is instead a lexicalized plural reading (plural-duals), as in *bot-ou* 'shoes', a further pluralization generally has the effect of shifting the part-structure conceptu-alization again, so that *bot-ei-er* means 'loose shoes', moving diagonally from

the bottom-left to the top-right corner of (8.14). The applicability of these morphological operations varies across the lexicon as well as across dialects, so that the Breton dual ends up being very different indeed from a neat inflectional system, where form and meaning oppositions are isomorphic and regular. We see in this subdomain a clear illustration of the mismatch between the system of exponents and the semantic oppositions they spell out. This very loose fit pervades the use of Breton nominal number; missing this connection between duals and lexical plurality would be a serious analytical shortcoming.

The use of plural forms in duals is lexical, not only for its listedness but also because it affects lexical semantics, by determining what counts as 'one' entity satisfying the predicate, as well as by altering the conceptualization and leading to a shift in descriptive content, as in 'hand' > 'handles'. But this function as a lexical formant is compatible with the plural being grammatically inflectional, belonging to a regular and productive type and determining syntactic agreement. Duals are thus a clear example of how inflectional plurality can affect the lexical semantics of a noun. Our next step is to show that the lexeme-forming function of plural number on duals is a special case of a much more general phenomenon.

8.3.3 *Plural and part structure*

Usually, a plural noun can refer to any collection of entities satisfying the noun's lexical predicate. Plural-duals differ because they only refer to some of these collections, determined on the basis of the meaning of the lexeme and of encyclopaedic knowledge. We will now see that they are just a prominent subclass of a more general phenomenon, whereby grammatical plurality qualifies a noun's lexical semantics.

8.3.3.1 *Duals are lexicalized plurals* If dual was a purely grammatical category, its denotation would contain all and only the sums with two members, irrespective of their referents. Instead, the Breton dual denotes only certain pairs: those of symmetric body parts forming cohesive natural sets.[18] What is primary is that the referents are cohesive natural aggregates,

---

[18] Ojeda (1995) proposes exactly this encyclopedia based restriction for the denotation of Italian *a* plurals: a sub mereology restricted to natural pairs and their sums (also their sums, since these are semantically plurals and not collective singulars like 'a natural pair of X'). From a formal point of view, the criterion of 'oneness' based on natural pairs is exactly like that based on sharing the same herd or the same owner in Papago (Ojeda 1998), for they are all ways to define equivalence classes (cf. Section 4.4.2). What is noteworthy about Breton is that pluralizing a plural dual does not necessarily lead to a compositional interpretation as sums of natural pairs, as Ojeda's model would predict. Instead, each morphological form defines its own criterion for belonging to an equivalence class.

not that they have two members; it is because the most salient such natural aggregates are pairs that the lexicalized noun forms referring to them have a dual interpretation. In fact, they can also denote natural aggregates with more than two members. Ternes (1970: 200–1) shows that the dialect of Groix expresses by the identical morphology the three-way distinction of single part, natural set, and many natural sets, whether the natural set has two elements like thumbs, or more than two, like paws or furniture legs:

(8.15) a. /mød/   'thumb'
            /mød-ew/   'thumbs of one person'
            /mød-uw-i:r/   'thumbs of several persons'

        b. /pat/   'paw, leg of a piece of furniture'
            /pat-ew/   'paws of one animal, legs of one piece of furniture'
            /pat-uw-i:r/   'paws of several animals, legs of several
                             pieces of furniture'

Ternes reports the same three-way distinction for /deãt/ 'tooth' ∼ /dəntew/ 'denture' ∼ /dəntuwi:r/ 'dentures', where the natural set 'denture' has more than two members. Similarly, in this dialect the lexicalized plural /bronew/, from the singular /bron/, may have either the dual reading 'breasts' or the non-dual 'teats of an animal'.[19] The same contrast between a single plural with a cohesive paucal reading and a double plural with a non-cohesive reading emerges with some of Trépos' (1956: 225) examples of double plurals: *dilhad* 'suite of clothes' ('l'ensemble des pièces qui habillent une personne'), *dilhaj-ou* 'clothes' in suites or loose (as in 'a dealer of old clothes'). In short, duals are a special case of morphological paucals, that is lexicalized plurals referring to cohesive natural aggregates.

    If duals are paucals, then the corresponding non-dual plural forms are non-paucals: there appears a link between non-dual reading and large size. On this basis, Trépos (1956: 265–7) argued that *bot-ei-er* has the non-paired reading 'loose shoes' because it refers to *large* collections of non-specifiable size ('pluriels indéterminés'; p. 265). The same can be said about a plural like *dour-ei-er* 'water-PL-PL' 'large amounts of water, flooding'. This link between size and lack of individuation calls to mind the distinction between plurals of abundance and paucity in classical Arabic (cf. Section 7.5.4), and in this case too the key notion is not size but part structure: double plurals refer to a mass of non-whole parts, or to a uniform multitude of non-distinguishable elements. Some of the semantic contrasts are glossed as follows (Trépos 1956: 233, 266):

---

[19] Hemon (1975: 42) also mentions the prefixed form *di guestad* 'ribs' from *kestad* 'ribs', but I don't know if it is (or was) a paucal referring to the set of ribs, or a dual referring to the twin ribcages.

(8.16)   a.   dour   'water'
              dour-you   'streams'
              dour-ei-er   'flooding, water streaming after heavy rains'

         b.   park   'field'
              park-ou   'fields'
              park-ei-er   'countryside, field landscape'

Park-ou fits a context like *park-ou ma breur* 'my brother's fields', while *park-ey-er* is appropriate for *galoupad ar park-ey-er* 'to run across the fields', said of hunters (p. 266). Here single and double plural correspond to a count and mass plural reading respectively. The form *tud-ou*, formed on the suppletive plural *tud* 'people', seems to have a similar function: *tud* in *dindan treid-ou an dud* 'under the feet of people' (p. 228; note the double plural *treid-ou* 'feet-pairs', normally *treit-ou*) contrasts with the more generic double plural *tud-ou* in *an holl dud-ou* 'all the peoples (of the earth)' (p. 224). In fact, the sequence -*ei-er* (PL+PL) can be directly added to a singular to turn it into a mass plural, whether or not there is a single (count) plural available. That is Trépos' analysis for *glao-eier* 'heavy rains' (from *glao* 'rain'), and even *bragez-eier* 'trousers', which is formed on the plural-dual *brag-ez* 'a pair of trousers' and is therefore a triple plural. Hemon (1995: 24) reports also the contrast between *koad-où* 'woods' and *koad-ei-er* 'woods, forests in general', and between the plural mass noun *ed-où* 'flour' and *ed-ei-er* 'wheats, cereals'.

   There are grounds for being cautious about these contrasts. Two native speakers separately told me that *park-ei-er* is the usual (or only) plural of *park* in their dialect; Wmffre (1998: 14) states that it is the normal form in central Breton, with *park-ou* confined to petrified place names and the fixed phrase *dreuz park-ou* 'across fields' (as opposed to by road). It is also interesting that most descriptions illustrate this phenomenon by the very pair *park-ou* ∼ *park-ei-er*, although Trépos (1956) lists many more forms (ten on p. 232 alone). In fact, several of these forms were not accepted by my informants, and Trépos himself acknowledged that his data came from different dialects (although he generalized his conclusions). It seems clear that re-pluralization is by no means a regular process, and the contrasts in (8.16) are in no way representative of a systematic set of morphosemantic alternations. Breton morphology cannot just turn count plurals into mass double plurals with the same regularity it turns singulars into plurals. This is essential for a correct appreciation of the status of plural in Breton as a whole, and I will return to this issue in Section 8.5. With these cautionary notes in place, the fact remains that pluralization *can* be employed to alter the part-structure conceptualization of a lexeme, as a lexical if not a grammatical process.

8.3.3.2 *Inflectional plural with lexicalized meaning*   Once we extend our view beyond double plurals, there is plentiful evidence that the exponents of inflectional plurality can convey lexicalized readings as mass or kind. A detailed description would have to rehearse most of the data in Trépos' (1956) monograph, so I will just concentrate on the suffix *-ou*, which is most clearly inflectional. Abundant evidence shows that it is the unmarked plural marker for inanimate nouns. It is quantitatively the most common marker, being used for 74.4 per cent of noun bases in the Bothoa sample of Humphreys (1995: 240), and accounting for 55 per cent of all plural suffixes in central Breton according to Wmffre (1998: 14). It is the only plural for singulative bases in *-enn*, except when they refer to humans (which is relatively rare); in fact it can apply even for human referents, as in the exceptional singulative *hini-enn-ou* 'some people' (cf. Section 8.2.3 above). It appears in many loanwords, like *velosiped-ou* 'bycicles' (Trépos 1956: 39; today rather *bissiklet-où*, Steve Hewitt informs me), /*distr'aksión-əw*/ 'distractions', /*af'iš-əw*/'placards' (French *affiches*) (Humphreys 1995: 241). Perhaps most tellingly, it is employed on a pre-nominal modifier that exceptionally marks plural agreement, namely *pikol* 'large': *piko-lou rec'hell* 'large rocks' (Trépos 1956: 81). Unlike other exponents which can express plurality but often have some sort of collective overtone, *-ou* is as close as one could get to a default, not only morphologically but also semantically.

For all that, *-ou* (as a plural exponent) can have a lexicalized reading on some lexemes.[20] It features, for instance, in a derivation that turns verbal bases into plural nouns:

(8.17)   kouez-a 'to fall'   > kouez-ach-ou 'fallout, debris' (Trépos 1956: 50)
         dislonk-a 'to vomit' > dislonk-ach-ou 'vomit, puke'
         skub-a 'to sweep'   > skub-ach-ou 'sweepings, dirt'

The derived nouns have the collective suffix *-ach* followed by the plural *-ou*; another dialectal example is *koed-aj-eu* 'woodwork' (McKenna 1988: 224), based on *koed* 'wood'. Trépos adds that the form without *-ou* does not seems to be attested. The match between Breton *skub-ach-ou*, French *balayures*, and English *sweepings*, is therefore not just morphological but also semantic, and it is not accidental. It exemplifies the use of plural to conceptualize a noun's referent as a material mass (cf. Section 4.5.1). Steve Hewitt (pers. comm.) points out another parallel. Spoken Breton distinguishes between *bezan en*

---

[20] I add 'as a plural exponent', because *ou* also has (or rather had) an affective, non plural reading on proper names, as in *Jak ou* 'little Jacques' (Yvon Gourmelon, pers. comm.); this is often reflected in place names of the form *ker X ou* 'village X' (Trépos 1956: 119). Evidently, *ou* cannot be an intrinsic ally inflectional plural affix, because it may not be plural at all.

vakañs 'to be on holiday' and *e-pad ar vakañs-ou* 'during the holidays'. Unlike the singular in the first example, the plural *vakañs-ou* in the second unambiguously refers to an occurrence with an extension in time, as required by 'during'. This no more refers to an aggregate of discrete holidays than its English translation does. Other instances of lexicalized *-ou* plurals are those that can feed singulative affixation. This happens very rarely, to judge by the examples in the literature. If the singulative suffixation targeted plurals, this scarcity would be surprising; if what counts is instead the part-structural reading of the base, the rarity (but not the impossibility) of *-enn* suffixation to a plural in *-ou* simply reflects the fact that *-ou* plurals generally denote aggregates of whole individuals, which cannot be further individualized. The exceptions mentioned by Trépos (1956: 34, 35, 227) are *louz-aou-enn* 'weed', *goul-aou-enn* 'light', *gwerelaou-enn* 'morning star', *gêl-aou-enn* 'blood-sucker', *loer-ou-enn* 'single sock, tight leg'. Each of these plural bases is demonstrably lexicalized. *Louz-(a)ou* is a non-compositional mass plural, also meaning 'medicine', not 'medicines' (thanks to Steve Hewitt for discussion on this point); *goulou* means 'light' and is in fact a singular (Hemon 1975: 43) of which *gwerelaou* is a formal variant (Trépos 1956: 34); *gêl-aou* has been replaced as a plural by *gêl-aou-enn-ed*, so that *gêlaouenn* is synchronically monomorphemic (1956: 227); and *loer-ou* is a plural-dual. These forms do not refer to pluralities of individual entities but to complexes of (non-individual) parts. They are, in a word, 'collectives'. With this, we have reached the crucial question as to the grammatical status of this category within the Breton plural system.

8.4 'Collectives' and plural morphology

In languages with a collective–singulative opposition, the label 'collectives' refers to noun forms that are plural by syntactic and morphological criteria, but differ from regular count plurals in the conceptualization of their reference domain. Loosely speaking, they are true not of sets of distinct individuals but of manifold masses, whose elements contrast with whole individuals in various possible ways, traceable to lack of unity and/or of identity. Typical examples from Breton are *bili* 'gravel', *gwer* 'glasses, glassware', *gwez* 'trees', *kelien* 'flies', *merien* 'ants', *nez* 'lentils', *stered* 'stars'. There is no need to emphasize again the semantic coherence of such non-aggregative plurals, nor the impressive regularity with which the same descriptive concepts (small animals, plants, fruits, granular masses, objects experienced in sets) occur whenever a language morphologizes this category. 'Collectives' understood as a semantically defined class of plurals certainly play a role in Breton

grammar. The question is whether they also form a morphological class; my answer, based on three types of evidence, is that they do not.

8.4.1 *'Collectives' as bases for singulatives are not a morphological class*

Since singulatives, as we have seen in Section 8.3.1, also target singulars and even non-nominal bases, collectives cannot be *defined* as the bases to singulative derivation; at most, as those bases which are also grammatically plural. But these plural bases are not homogeneous morphologically. The main class is made up of monomorphemic stems like *per* 'pears', on which plurality is intrinsic to the lexical stem. Such is also the case for (almost) all collectives in Welsh, like *coed* 'trees' ∼ *coed-en* 'a tree'. But in Breton, singulatives can also target noun forms that, unlike *per*, are overtly marked as plural, through suppletion, ablaut, suffix, or a combination of the last two (the form *penn-deñved* is from Wmffre 1998: 15):

(8.18)	Singular	Plural	Singulative (sg.)	(Trépos 1956: 236–42)
a.	marc'h 'horse'	kezeg	penn-kezeg/loan kezeg	(suppletion)
b.	dañvad 'sheep'	deñved	penn-deñved/deñved-enn	(ablaut)
c.	bran 'crow'	brin-i	brin-i-enn	(ablaut+suffix)
d.	loer 'sock'	loer-ou	loer-ou-enn	(suffix)

This fact was indeed one of Stump's (1989) reasons for claiming that collectives have the same grammatical status as suffixed plurals. As noted in Section 8.2.2 above, being a possible input for singulative derivation does not actually prove that a form is plural; but if we admit, as we must, that the forms in the second column in (8.18) are grammatically plural, then these data do show that collectives need not be monomorphemic bare stems like *per*. Each of the examples in (8.18), which Trépos cites from the *Atlas Linguistique de Basse-Bretagne* (1924–1943), would deserve careful discussion. I will just clarify that *penn-kezeg* and *loan kezeg* are 'compound singulatives', typically used for names of animals (cf. note 8 above); that the singulative based on the plural can alternate with one based on the singular ('sheep', for instance, is *deñved-enn* in some points of the Atlas and *dañvad-enn* in others); and, above all, that singulatives built on morphological plurals historically tend to replace the unmarked singulars (as *stered-enn* 'star' did for *ster*; cf. Section 8.3.1 above). This tendency has an obvious functional basis, in so far as the unmarked singular and the derived singulative singular end up being synonymous. The main point in this connection, however, is that the singulative can definitely attach to a bimorphemic base STEM+PL as well as to a monomorphemic inherently plural STEM$_{PL}$. Therefore, if collectives are defined as grammatically

plural bases for singulative derivation, these bases can be mono- or bi-morphemic. When collectives are understood in this way, therefore, they cannot be equated with monomorphemic plural stems, nor with any one particular morphological expression of plural.

8.4.2 *'Collectives' as plural stems are not a morphological class*

It is not only as bases for singulatives that Breton plurals have a non-aggregative reading. The well-known instances of plurals inside derivational affixes, illustrated in (8.3) above, clearly have the non-referential kind of interpretation typical of plural as inherent inflection (cf. Booij 1994, 1996, and Section 3.4.1). Some relevant examples are repeated here as (8.19):

(8.19) evn 'bird'   evn-ed 'birds'   evn-et-a 'to hunt for birds'
       maen 'rock'   mein 'rocks'   mein-ek 'rocky'
       aval 'apple' aval-où 'apples' aval-aou-er 'apple-hunter (hedgehog)'
       merc'h 'girl' merc'h-ed 'girls' merc'h-et-aer 'womanizer'

Trépos (1956: 121–5) makes some very interesting points about the value of such word-internal plurals. The form indubitably has a plural interpretation, in the light of pairs like *sili-enn-a* 'slip between the hands' (like a single eel, which is the meaning of the singulative *sili-enn*) versus *sili-aou-a* 'to fish for eels' (with the plural *sili-aou*).[21] The plural that appears inside derivation is clearly lexicalized, as opposed to triggered by the syntactic context, and has thus the same status of duals, which we have analysed as a special type of lexical (paucal) plurals. Indeed, duals appear as word formants too, and the dual *daou-lin* 'two knees' (from *glin* 'knee') can provide a dramatic illustration: *daou-lin-a* means 'to kneel down'; but a near-synonym exists, *penn-daou-lin-a*, based on the singulative of *daou-lin*, and Trépos reports that it is typically used for men, who kneel by putting down one knee only (p. 124). So, lexicalized plural (including dual) inside derivation definitely has a semantic effect. Trépos' last point in this connection concerns a semantic property of central importance. A verb like *labous-et-a*, based on *labous-ed* 'birds', means 'to hunt/shoot birds'; but it must denote multiple bird-shooting events, not a single event where many birds are killed. This is interesting, because it shows how the plurality of the verb-internal noun is necessarily fused with the lexical semantics of the verb, modifying its actionality (a lexical property) and not just qualifying its internal argument. Notice that this is not predicted under

---

[21] Trépos (1956: 121) adds that while *marc'h a*, built on the singular *marc'h* 'stallion', means 'to be on heat (of mares), to seek the stallion', *merc'h et a*, built on the plural *merc'h ed* 'girls', 'suggère des aventures multiples'.

an analysis like that of Stump (1990), which simply holds that plurals brought about by inflectional morphology can re-enter word formation. If that was all there is to say about plurals inside derivation, one would expect the verb *labous-et-a* to support both a many-birds and a many-events reading, contrary to fact, because that is the ambiguity caused by a plural argument. These word-internal plurals, then, are truly plural and are also truly lexical.

Similar comments apply to plurals-inside-plurals. We have seen in Section 8.3.3.1 that the double pluralization involving *-ei-er* (from *-ou* + *-er*) is available as a more or less lexically restricted option across dialects. As we have seen, this instance of double pluralization is far from being a regular grammatical process, and when it is available, its interpretation can vary significantly. This typically lexical irregularity in distribution and interpretation also characterizes other instances of double pluralizations:

(8.20) SINGULAR	PLURAL	PLURAL + PLURAL (Trépos 1956: 223–30)
—	tud 'people'	tud-ou
—	dilhad 'clothes'	dilhaj-ou
houarn 'iron'	hern	hern-iou
louarn 'fox'	lern	lern-ed
bugel 'child'	bugal-e	bugal-e-ou
merc'h 'girl'	merc'h-ed	merc'h-ed-ou

This list should not give the impression that double pluralization is regular or even particularly widespread. Trépos emphasizes that the grammar allows this possibility, but he then locates each example in a particular point of the *Atlas Linguistique de Basse Bretagne*, and he notes for many other cases that an apparent double plural is in fact the only plural in use (like the English *children*, which is historically a double plural with the endings *-er* and *-n*). The diffusion of these concurrent plurals seems to be even more restricted today; for instance, Yvon Gourmelon reports that *merc'h-ed-ou* is only used as a fixed vocative in his dialect, with a pejorative nuance that has nothing to do with semantically transparent pluralization. Turning now to the morphology of these formations, the first two examples, as we have seen in Section 8.3.3.1, are mass or generic plurals built on monomorphemic plural stems. *Hern-iou* 'irons' and *lern-ed* 'foxes' are likewise based on a plural stem, so that the double plural has in fact a single suffix. The remaining examples feature instead pluralization of suffixal plurals, resulting in two consecutive plural suffixes. Again, 'collectives' fail to correspond to one morphological category—in particular, they fail to coincide with the listed forms of ablaut monomorphemic plurals. If being collective means admitting secondary pluralization, then collectives are just as morphologically diverse as plurals are. If it means

being a plural with a non-aggregative interpretation, then the plural bases that feed further inflectional or derivational affixation are collective; but these bases can be mono- or bi-morphemic, and do not correspond to one particular morphological expression of plural. Whether one starts from the morphological or from the semantic end, collectives in Breton do not define a morphological class.

8.4.3 *'Collectives' are not a subset of opaque plurals*

If semantic collectives do not coincide with a morphological class, perhaps they could include such a class as a subset. All plurals with a certain morphology would then have a collective interpretation but not the other way around; this would still amount to a significant match between form and meaning. Trépos (1956) makes just this claim. At various points in his book (pp. 53, 72, 264) he states that several morphological types carry a collective interpretation: suppletives like *kezeg* from *marc'h* 'horse', monomorphemic ablaut plurals like *kerreg* from *karreg* 'rock', and certain suffixal plurals that do not involve a productive affix, like those in *-i* (e.g. *kendirv-i* from *kenderv* 'cousin') or in *-er* (e.g. *kij-er*, regularly developed from *kiz-ier*, from the singular *kaz* 'cat'); the latter appear systematically in double plurals like *park-ei-er*. The clearest formulation is on p. 82, where the regular aggregative reading 'many discrete units' is explicitly linked to morphologically transparent composition:

Il y a des pluriels dans lesquels le singulier n'est pas immédiatement reconnaissable; ce sont les pluriels dans lesquels la voyelle a subi une métaphonie: pluriels internes, et pl. formés à l'aide des suffixes *i*, *er*, *en*; ce sont aussi les pluriels supplétifs. Ces pluriels donnent plus ou moins au bretonnant l'idée d'une collection dans laquelle il ne distingue que difficilement les unités.

[There are plurals in which the singular is not immediately visible; these are the plurals where the vowel has undergone ablaut: internal plurals and plurals suffixed with *i*, *er*, *en*; these are also suppletive plurals. These plurals give to the Breton speaker more or less the idea of a collection in which the units cannot be easily distinguished.] (Trépos 1956: 82)

However, there is evidence that the lack of morphological transparency does not correlate systematically with collective or non-distributive interpretation, not even as a sufficient condition (as opposed to necessary and sufficient). We have seen above in (8.12) that the plural *bot-ei-er* 'shoes' can be the antecedent of 'one by one'; moreover, it can be the argument of the verb 'to count', which likewise imposes a discretization of the reference domain (cf. **he counted the furnishings* and the discussion in Section 4.2.2):

(8.21) kontet   en   doa
        counted VPT.3.SG   have.PAST.3.SG
        ar   botei-er   (Yvon Gourmelon, pers. comm.)
        the   shoe-PL
        'he counted the shoes'

Another plural in *-er* is *breud-er* 'brothers' (Trépos 1956: 65). Its singular is *breur*, with a stem allomorphy that makes the relation between singular and plural form morphologically non-transparent; yet *breud-er* supports a distributive interpretation, as in the following example (where *bredeur* is the form of this plural in the Treger dialect):

(8.22) ma bredeur   so peb un auto gante (Steve Hewitt, pers. comm.)
        my brother.PL is each a   car with.3.PL
        'my brothers have a car each'

As we have just seen, Trépos includes the plurals in *-en* among those that refer to collections 'in which the units cannot be easily distinguished'. This suffix typically attaches to nouns referring to humans, but its distribution is lexically limited and contrasts with *-ed*, which is the default plural ending for humans and animates generally. In the light of this lexically restricted distribution and of Trépos' words cited above, one would expect a plural like *studier-ien* 'students' to be incompatible with an openly distributive context. This is not what we see (data from Steve Hewitt and Yvon Gourmelon):

(8.23) ar   studier-ien   a   oa   aet   araog hini   ha   hini/
        the   student-PL   VPT   be.PAST gone ahead one and one/
        a hini-enn-où/   an   eil   war-lerc'h egile
        one- SINGULAT-PL/ the   second   on-back   other
        'the students left one by one'

Note that some plurals in *-en* are indeed incompatible with a distributive context. Such is the case for nominalized adjectives, like *paour* 'poor', which becomes *peur-ien* 'poor.PL' when it heads a generic plural noun phrase:

(8.24) ar   bewr-ien   a   oa   aet   kuit
        the   poor-PL   VPT   be.PAST gone away
        (*a hini-enn-où)   (Yvon Gourmelon, pers. comm.)
        (one. SINGULAT-PL)
        'the poor left (*one by one)'

This particular example is instructive because the distributive modifier is impossible both in the Breton example and in its English gloss. The deviance

clearly has a semantic basis, namely the necessarily generic interpretation of nominalized adjectives in such definite plural phrases (contrast *the poor* with *the poor people*, which may instead refer to a specific collection). But this is not the case in (8.23), where *studier-ien* exhibits the same suffix. Hence, this suffix has no bearing on the distributive interpretation. The conclusion can be generalized to the other morphological types that Trépos claimed to carry a collective interpretation: as (8.21)–(8.23) show, morphological opacity is not a sufficient condition for non-distributive interpretation.

In sum, then, Breton collectives are not a morphological class. A collective meaning does not imply a particular form, nor does a particular form imply a collective meaning. Undeniably, the relation between plural form and plural meaning is far from random, and Trépos' remarks quoted above are obviously correct when read as pointing to a privileged relation between non-distributive interpretation and morphologically opaque pluralization. But this preferential relation does not have the status of a grammatical rule: the collective reading is not the meaning of any morphological form. Contrary to common assumptions, then, Breton has no *morphological* class of collective plurals. And yet, as we have seen in the last section, the collective–singulative opposition plays a central role in the distribution of plural morphology (including plural-duals), underpinning its use as a lexical formant. We must conclude that part-structural oppositions are relevant for the use of morphology, but they are not isomorphic to the oppositions defined by the morphological system itself. The exponence of plurality constitues a morphological subsystem which can be put to use to spell out the inflectional number category, or to qualify the part-structure information encoded on lexemes. As Anderson (1986) saw, the two functions are distinct; but as Stump (1989, 1990) saw, their morphology is not. Since one function is inflectional and the other is derivational, Breton plurals as a morphological category are neither intrinsically inflectional nor derivational. They very clearly illustrate the separation between the system of morphological exponence and the abstract categories it spells out.

8.5 Conclusion: the peculiarity of Breton plurals

The category of number on Breton nouns is unambiguously inflectional. What is noteworthy is that the morphological expression of number, specifically of plurality, is not limited to spelling out this morphosyntactic category, but can serve word-formation purposes too. In this use, plurality encodes context-independent semantic properties of lexemes (hence, not of purely functional elements like pronouns or auxiliary verbs); it applies to certain

nouns but not others, as a function of words and not of grammar; and its exponents are sometimes not disjunctively selected, and on occasion feed further morphological affixation.

Although plurality as an inflectional morphosyntactic property and as a lexical property are spelled out by one and the same morphological system, these two functions must be clearly distinguished. A failure to do so would lead one to think that Breton plural nouns are all lexical/derivational (since their morphology can be) or non-lexical/inflectional (since their morphosyntactic category is), both of which views are wrong. In particular, it would be wrong to claim that, when used as lexeme formants, these morphemes (affixal or otherwise) are 'not really plural': they are and remain plural, both because they have the same form as inflectional plural morphemes, and because they contribute semantic information that is distinct from that of singular noun forms. This information may be called plurality, provided it is clear that, as a characterization of the part structure of the reference domain of the lexeme, this term can mean 'non-discrete' (as in mass plurals like *dislonk-ach-ou* 'vomit') or 'made up of discrete but non-distinct elements' (as in *park-ei-er* 'fields, countryside'), or 'made up of elements organized in natural collections' (as in duals/paucals like *treid* 'feet', where individualization lies at the level of the collection and not of the single element), or 'kind' (as in *aval-aou-er* 'apple(s)-hunter, hedgehog'). Once we recognize the possibility that the very same plural exponents can spell out of grammatical information as well as stem-constituting part-structural information, Breton plurals cease to be so unique, because many other languages express these notions through plural morphology.

The unusual nature of Breton noun plurals, then, is not caused by the use of plurality as a lexical formant, but by its co-occurrence with two other peculiar traits. One is the prominence of part structure conceptualization for morphology. As we have seen, the conceptual contrast between individuals (bounded and identifiable) and non-individuals (lacking either properties) plays a pervasive role in the Breton number system: it determines the distribution of singulatives, which turn non-individual-referring nouns into individual-referring ones, over and above their syntactic number; it underpins the class of duals, which are in fact lexicalized paucals whose elements have unity but lack identity; and it is associated with several uses of plural to express lack of individuality. While noun morphology is so sensitive to part structure, syntax only knows a schematic singular ∼ plural contrast. When we speak of nominal plurality we speak of the intersection between the syntactic opposition in number and its encoding in noun morphology.

The other factor, probably the most conspicuous, is the extent of the mismatch between morphological resources and their function. Unlike Arabic, for instance, where only broken plurals can function as collective singulars because they are just stem variants, in Breton all exponents may in principle have a lexeme-forming function, albeit with obvious preferences and depending on the vocabulary of single dialects (ultimately, on historical accident). More striking still is the independence from positional requirements. Again, we have seen that that double plurals like *merc'h-ed-ou* 'girl.PL.PL' are far from common; but even so, the ease with which plural endings can appear word-internally is remarkable. What is significant in these cases is not so much the internal position of plural, but its independence from the syntactic function. This trait should not be seen in isolation, because the lack of isomorphism between the organization of exponents and of the categories they express is a theme that has surfaced elsewhere in this chapter, particularly in connection with duals; besides, recall that the singulative suffix has an evaluative function with certain nouns, something that the plural affix -*ou* can also have (see note 20).

In this context, it cannot be an accident that the most peculiar aspect of Breton grammar is precisely the organization of nominal plurality, where lexically encoded part structure interacts with syntactic number, and both functions are mapped into a complex system of forms. Apart from showing the connections with lexical plurals in other languages, the analysis here proposed thus relates the peculiarities of Breton plurals to independently given facts about Breton grammar. In so doing, it highlights the central theme of the whole book: lexical plurals at the centre of the complex map between lexical semantics and autonomous morphology.

9

Conclusion: Plurals and lexicality

9.1 Lexical and grammatical knowledge

Throughout this book, I have been using the expression 'lexical knowledge' advisedly, skirting attempts to define the concept of lexicon or lexical item. The reason is that the lexicon is necessarily a multi-dimensional concept. Even understood simplistically as a repertoire of meaningful linguistic building blocks, it would have to encompass various kinds of listemes, each equally justified at different levels of analysis: syntactic, morphological, semantic, psychological (see Section 2.2). For our investigation, there is no gain in isolating one dimension of analysis and decreeing that the units in that dimension are lexical items. The real issue is not the lexicon, but lexicality: the properties characterizing knowledge of substantive 'words', as opposed to constructional principles and to functional expressions with syncategorematic value. The study of plurality as an ingredient of nominality strongly suggests that what we know about words does not fall into a deterministic grammatical part and a totally idiosyncratic and non-linguistic lexical part. Rather, lexical knowledge is linguistic, not just encyclopaedic, and it is not disjoint from the knowledge encapsulated in the grammatical system but subsumes it in important respects.

The whole book has provided evidence for this claim. Imagine trying to summarize all the phenomena described so far by saying that plural nouns are either regular, transparent, and productive, or irregular, opaque, and listed. As I hope is clear by now, this would be a grotesque oversimplification. Although an idiosyncratic form or meaning often accompanies non-canonical uses of inflectional plurality, there is more to these phenomena than lack of inflectional regularity. The reason is that nouns, unlike pronouns, encapsulate more than purely grammatical knowledge. When plurality is inherent in a noun, it is part of what defines a speaker's knowledge of that word, in form and meaning, as distinct from other words fitting the same grammatical context. The evidence from lexical plurals suggests that this kind of knowledge has a substantive linguistic content. Interpretively, the lexicalization of

plurality amounts to a conceptualization as 'not-one'; morphologically, to an exponence of plurality that may be idiosyncratic, or independent of the syntactic context, or non-uniquely determined by it, or inseparably fused with the exponence of the word root, or a combination of these properties. The empirical range thus defined is vast but not boundless, and the connections between morphology and semantics are far from random. A revealing account of the range of semantic and morphological phenomena defining lexical plurality, therefore, requires a substantive view of what constitutes lexical knowledge as part of linguistic competence, and of the way it can (and cannot) interact with knowledge of the grammar as a computational system. The goal of this final chapter is to make explicit the conception of lexicality that emerges from our investigation as a whole. Sections 9.2 and 9.3 focus on the key concepts of lexeme and stem, which are instrumental for capturing the distinctive properties of lexical plurals on the paradigmatic and on the syntagmatic axis. Section 9.4 surveys the role of syntactic structure in determining aspects of lexical knowledge, and clarifies the close relation of my position with recent constructional approaches. Finally, Section 9.5 concludes the discussion with an overview of the results obtained—and those not obtained.

9.2 Lexemic plurals

A lexeme is here understood as the single abstract base instantiated by different inflected forms in different syntactic contexts. Lexemes are abstract concepts on the paradigmatic axis, provided with a syntactic category and, in the case of nouns and in some languages, with context-invariant diacritics like gender or class (modulo automatic diacritic-changing rules, as in the case of Romanian ambigenerics; cf. Chapter 5). By definition, the forms of a lexeme must have the same class diacritics, and therefore plural nouns with idiosyncratic gender or class instantiate distinct lexemes. We have seen many such examples, focusing in particular on Italian -*a* plurals in Chapter 5. Importantly, a pair like *uovo* ∼ *uova* 'egg (masc. sg.) ∼ eggs (fem. pl.)' also instantiates two distinct lexemes, even though these are the only possible forms for what is intuitively a single noun. But a lexeme is a technical, not an intuitive concept, and does not necessarily correspond to a semantic listeme, to an address for lexical retrieval, to an item in the mental lexicon, or to a 'sense'. In a morphosyntactic representation, a lexeme is realized by those elements that define a base for inflection: one lexical root (or more in the case of compounds), and zero or more morphemes; for nouns, this must

include the morphosyntactic locus for nominality and, when present, for class diacritics (see Section 9.4). In addition, every nominal lexeme carries a certain conceptualization of the part structure of its denotation. Nominal lexemes may be intrinsically plural, when the information inherent in them includes grammatical and semantic plurality. Being an integral part of the lexeme, plurality affects lexical semantics and encapsulates a conceptualization as 'not-one', which can take different forms along the axes of unity and identity (cf. Chapter 4). This often involves a special morphology, which might define a class of paucals (see Section 2.6.4 for Hebrew and Section 8.3.2 for Breton). The concept of lexeme thus captures the connection between morphological class diacritics, conceptualization, and morphosyntactic plurality.

Since class diacritics are constitutive of the identity of a lexeme, all plural doublets that differ in class diacritics instantiate distinct lexemes. They do not compete for the realization of the same inflectional content on the same lexeme, and so the grammar allows them to coexist, although usage typically differentiates their meaning on functional grounds. However, this says nothing about plural doublets that do not differ in class diacritics, like the German *Worte* 'words in a linguistic context' ∼ *Wörter* 'words in isolation' (both neuter), or the English *oxen* ∼ *oxes* and *pennies* ∼ *pence*. Lack of blocking, on its own, seems too weak a reason to view all such doublets as distinct lexemes, considering that they are grammatically identical and lexemes are primarily defined on grammatical grounds. I will tentatively hypothesize that alternants having the same extension, like *Worte* ∼ *Wörter*, are variants of the same lexeme, while plurals with different denotations, like *brothers* ∼ *brethren*, realize distinct lexemes. In sum, plurals with idiosyncratic class diacritics are lexical as lexemic; others may be lexemic too, but that depends on a more precise definition of lexeme than I can offer here.

9.3 Inherently plural stems

A lexeme is a paradigmatic concept, which results from abstracting a common lexical base over different grammatical contexts. Some plurals, however, are lexical (also) in a syntagmatic sense, when they express plurality through a certain shape of their core word form, rather than by morphemes attached to it. What is crucial is that they spell out plurality without contextually determined inflectional appendages. The notion of 'stem', in the sense of Aronoff (1994), identifies the inner part of a word form that excludes inflectional and context-determined affixes, without being necessarily atomic or listed. Nouns that express plurality by a certain choice of stem, rather than by a stem-external marker, are lexical plurals as a property of a morphological

form (the stem) and not of the abstract lexeme. The notion of stem is indispensable to pinpoint the sense in which Irish counting plurals and Arabic broken plurals are 'lexical' (Chapters 6 and 7), which is not the same sense in which plurals are lexical when they realize an inherently plural lexeme.

Plurals with stem-internal exponence often, but not always, display non-inflectional properties. In *geese* and *feet*, for instance, the fronting of the root vowel is a sub-regular adjustment affecting nouns of a certain shape with the regularity of an inflectional operation, just like -*s* affixation. There is no evidence for *geese* or *feet* as morphological objects independent of grammatical pluralization, and these forms are syntactically and semantically equivalent to regular plurals. In many other cases, however, stem-internal plurality has a privileged relation with non-canonical readings that are sensitive to the meaning of the lexical base. The link between word-internal exponence and non-canonical reading is particularly clear where a plural marker appears inside the base for inflection, as in derivational uses of plural like the Breton *sili-aou-a* 'to fish for eels' (from the plural *sili-aou* 'eels').

Above all, stem-inherent plurals can escape strict disjunctivity and allow alternative plural formations, especially when they too are stems; this emerged very clearly in the discussion of Irish and Arabic plurals. By definition, the stem of a noun includes that part which has a purely lexical content, as opposed to grammatical. But grammatical knowledge is what drives the choice between grammatical formatives, not between lexical words. When plurality is fused with the stem, it escapes this grammar-driven automatic competition (cf. Section 7.4.2 and note 18 in Chapter 7). Arabic stem-inherent plurals, and Breton ones to some extent, display precisely this mix: purely grammatical information, defining a cell in an inflectional paradigm, expressed by spelling out the lexical, non-grammatical, part of the word. If the expression of plural inside a stem had exactly the same status as a grammatical number affix in regular inflection, stem-inherent plurals should behave like any regularly inflected plural, without allowing competing plural alternants, without being favoured for non-canonical readings, and without being used in other senses but plural. Since this is not what happens, our analysis bolsters the case for the stem as a theoretical concept.

9.4 Lexical and constructional knowledge

Neither lexemes nor stems are pieces of a syntactic representation. Both however are associated with concrete morphosyntactic structures, made

up—I have assumed—of syntactic heads, spelled out according to the principles defining a postsyntactic morphological component. While not pursuing a syntactic analysis, I have made certain structural assumptions that have consequences for the concept of lexicality. This does not concern so much the noun-external domain, which I take to involve a DP-peripheral number head, as the immediate environment of the noun. As sketched out in Chapters 1, 4, and 6, I have followed Borer's (2005) view that the noun is immediately governed by a head encoding a division of the noun's reference (heading a 'Classifier Phrase'), in turn governed by a head expressing a quantity based on that division (heading a 'Quantity Phrase', or #P):

(9.1) $[_{DP}$ Det $\ldots$ $[_{\#P} <QUANTITY> [_{ClassifierP} <DIVISION> [_{NP}$ N $]]]]$

The projection #P hosts adjectival determiners like *many* or *much*, as well as numerals above 1 in English. Below that, Classifier Phrase may accomplish its dividing function through morphosyntactic number features that the noun inflects for, or through classifiers; some expressions incorporate both a counting and a dividing function, like the English indefinite article *a*, and lexicalize both heads (see Borer 2005; Chapter 4). I depart from Borer's assumptions in further decomposing the noun into a category-assigning 'little [n]' and a category-free root, following recent proposals in Distributed Morphology (cf. Marantz 1997, 2003; Harley and Noyer 2000; and Arad 2003):

(9.2) $[_{DP}$ Det $\ldots [_{\#P} <QUANTITY> [_{ClassifierP} <DIVISION> [_{nP}$ n $[$ Root $]]]]]$

Importantly, heads do not necessarily match discrete elements in the morphological spell-out, so (9.2) does not claim that all nouns end in an affix. What it states is that nouns, like all lexical categories, are not syntactic atoms but syntactic constructs, where nominality and the 'core' lexical root have distinct abstract loci. This decomposition proves empirically convenient to describe nouns that share the same root but differ in non-contextual grammatical information, like class diacritics. Noting that to be a noun, in certain languages, means to be assigned to a certain morphosyntactic class, Kihm (2001) indeed proposed that [n] hosts gender or noun class information, which remain fixed across syntactic contexts.

These structural assumptions allow a very natural syntactic interpretation of lexical plurality. When number is entirely a property of the syntactic context, it is encoded as the head of Classifier Phrase, directly above the noun-defining complex $[_{nP}$ n $[$ Root $]]$. When a particular choice of number is instead inherent in a noun, the relevant information is expressed on [n] itself. Number is then determined not generically by the choice of a noun, but

specifically by the choice of [n], which identifies a nominal lexeme. The observed tight connection between lexicalized plurality and gender or class value, encoded on [n] follows straightforwardly.

This structural interpretation has other positive consequences. The decomposition of nouns into [n] and [Root] makes it possible to say that plural doublets involve distinct nouns, even though they share one and the same root. Consider for example a doublet like the Italian *braccia* ∼ *bracci* 'arms (body parts, also measurement)' ∼ 'arms (any other sense)'. I analysed them as forms of distinct lexemes, only *braccia* being inherently plural and expressing a non-individual conceptualization. This information is clearly lexical, not inflectional; but if it was a property of unanalysed noun stems, we would have to posit two homophonous stems distinguished by their semantics. This synchronically accidental homophony, however, would systematically extend to all plural doublets differing in gender. Surely, the root looks the same because it is the same; the difference, in grammar as well as in conceptualization, is encoded on [n].

Another positive consequence is that, if lexicalized plurality is expressed in [n], it becomes understandable why so many lexical plurals lack a discrete number affix. Both the Italian plurals in -*a* and the Irish counting plurals differ from their regular inflectional counterparts precisely because their endings are not number suffixes. Arabic broken plurals too, of course, are lexical for their form and lack a morphological encoding for number (the point is due to Kihm 2003). And in Breton, lexicalized readings correlate with a lack of transparent plural morphology as a clear tendency, if not as a grammatical rule (see Section 8.4.3). Of course many other lexicalized plurals have a regular number exponent, but it is revealing that so many of them have 'short' forms. The reason is that, when plurality is encoded on [n], all the necessary morphological information is already within [[n] Root]; if a morphologically well-formed form is available to spell out this complex, it will suffice to realize the noun. In this case, the higher dividing operator heading Classifier Phrase ('Number' in previous approaches) is necessary for syntax, but not for morphology, and remains null if morphologically dispensable.

In a sense, [n] has the same semantic function as the dividing operator that heads Borer's Classifier Phrase, namely that of expressing a division of the noun's reference. As discussed in Section 4.2, a part-structure conceptualization is part of what it is to be a noun, and so it seems natural to ascribe this function to [n]. In this way, [n] encapsulates an important part of what Borer (2005: 106) calls 'the conceptual properties' of noun stems, which she views as atomic and grammatically inert. If what is atomic and featureless are not

noun stems but only their roots, we obtain a finer-grained picture of how grammar combines with lexical knowledge to bring about nouns. Bare roots combine with category-assigning [n] heads, which fill in the grammatical and conceptual properties of what we recognize (by abstraction) as lexical nouns in isolation. However, these formal and conceptual properties are only fully determined in a syntactic context, and the head <DIVISION>, above [n], hosts the contextual part-structural information that is not lexeme-inherent. While this outer layer determines important aspects of a noun's interpretation, including often the specification as mass or count (Borer 2005: 132–5), it is the inner layer, with [n], that encodes the basic conceptualization defining nominal reference as such. This is part of the meaning of a noun, as *Aktionsart* is part of the meaning of a verb (cf. Rijkhoff 2002), and it underlies the countability preferences determined by the noun's lexical semantics, and often by its choice of morphosyntactic diacritics on [n] (cf. the examples from Swahili in Section 2.7.4 and from Kiowa or Turkana in Section 3.5.3). That is the justification for distinguishing two loci of part-structure determination, inside and outside the lexeme-defining domain. In the structure described in (9.2), I identify the former with [n], and the latter with <DIVISION>. Plurality is lexical when it is inside the construct that defines a noun in a syntactic context.

9.5 Concluding remarks

This investigation has addressed the question 'how much grammar is there in a lexical word?' by considering plurality as a component of lexical nouns. The answer, in a nutshell, is that plurality is an integral part of certain nouns, with a range of empirical consequences on the semantic and morphological level. Semantically, plurality affects the part-structural conceptualization that is an intrinsic characterization of nouns as entity-denoting expressions. This results in readings that qualify the referents as non-whole (parts of a mass, or non self-standing elements of a greater whole), and/or not identifiable (tokens, abstract standards of quantification, perceptually or functionally indistinguishable entities, tropes). So, the meaning of plural is not a function on the meaning of singular, and certainly does not entail countability. As a pendant to these semantic considerations, plurality can be lexicalized in a morphological sense, when it is an integral part of a lexeme, or when its realization is inseparably fused with the realization of a lexical stem. The two aspects of lexicalization are related on principled grounds.

What *cannot* happen in this scenario? Since lexical plurality has a specific semantic basis in part-structure conceptualization, we do not expect inherently plural nouns to be characterized in terms that have nothing to do with part structure; for instance, in terms of physical shape—even though physical shape determines the categorization of nouns in many classifier languages. This excludes a hypothetical noun denoting 'many round *x*s', corresponding to a singular denoting 'an *x*'. It also excludes nouns that are inherently plural and restricted to referents of a certain sex, unless this correlates with a part-structural conceptualization (for instance, as undifferentiated). When the choice of number discriminates between a reading as abstract kind and a reading as concrete instantiation, the singular should correspond to the former and the plural to the latter. Morphologically, if a language has gender or noun class, we expect plural doublets to be differentiated according to these features. We also expect multiple competing plurals (for the same singular) to be more common in languages where number shows other signs of being derivational, than in languages where it is clearly a regular inflectional process. If a language has both stem-inherent and stem-external (affixal) plural nouns, we expect most semantically lexicalized forms to be included in the former class. Finally, we do not expect anything resembling semantic or morphological lexicalization on plural pronouns. A language may have multiple pronominal series, possibly including a collective one ('many forming a group'), but there should not exist special plural pronouns with an exclusive non-count reading, or specially marked to refer to abstract standards of measurement, or to undifferentiated tokens. These conceptualizations are typical of plurality when it is part of lexical knowledge, not of plurality per se.

Insofar as these expectations are met, they provide falsifiable evidence that plurality is a component of grammatical knowledge which may co-determine lexical knowledge, thus showing that the two are distinct but interrelated. I have argued for this conclusion pursuing mainly scope in Part I and detail in Part II, but always at a relatively informal level. Once the main point is accepted, it will become possible to do better. For instance, by detailing the formal mereological properties of non-canonical plural readings, possibly putting to use the analytic techniques of formal ontology. Or by giving an explicit syntax and compositional semantics of the way roots combine with [n] and then with higher heads. Or by investigating a typologically significant sample, rather than a few case studies. Besides, it seems desirable to explore the connection between nominal part structure and verbal aspect, to clarify the relation between the concepts of lexeme and stem and the actual morphosyntactic representation, and to provide formalized and

predictive analyses of the morphology of stem-inherent and lexeme-inherent inflection. As it stands, the description I have offered has shown that lexical plurality is a coherent empirical domain with great theoretical interest. My hope is that it will stimulate research from different perspectives on this elusive topic.

References

Abd Rabbo, M. (1990) 'Sound Plural and Broken Plural Assignment in Classical Arabic', in M. Eid (ed.), *Perspectives on Arabic Linguistics I*. Amsterdam, Philadelphia: John Benjamins, 55 93.

Aboul Fetouh, H. (1969) *A Morphological Study of Egyptian Colloquial Arabic* (Janua Linguarum 33). The Hague, Paris: Mouton.

Acquaviva, P. (2002) 'Il plurale in a come derivazione lessicale', *Lingue e linguaggio* 2: 295 326.

——— (2004) 'Constraining Inherent Inflection: Number and Nominal Aspect', *Folia Linguistica* 38: 333 54.

——— (2005) 'The Morphosemantics of Transnumeral Nouns', in G. Booij, E. Guevara, A. Ralli, S. Sgroi, and S. Scalise (eds), *Morphology and Linguistic Typology: On Line Proceedings of the Fourth Mediterranean Morphology Meeting*. University of Bologna, 252 65.

——— (2006) 'Goidelic Inherent Plurals and the Morphosemantics of Number', *Lingua* 116: 1860 87.

Aikhenvald, A. (1998) 'Warekena', in D. Derbyshire and G. Pullum (eds), *Handbook of Amazonian Languages IV*. Berlin: Mouton de Gruyter, 225 439.

——— (2003) *Classifiers*. Oxford: Oxford University Press.

Aïm, E. (2003) 'Aspects de la flexion nominale en Somali', in G. Dal, B. Fradin, F. Kerleroux, N. Hathout, M. Plénat, and M. Roché (eds), *Silexicales 3: les unités morphologiques*. Villeneuve d'Ascq, Université de Lille, 11 19.

ALA (Academia de la Llingua Asturiana). (2000) *Normes ortográfiques y conxugación de verbos*. Uviéu [Oviedo]: Academia de la Llingua Asturiana.

Allan, K. (1977) 'Classifiers', *Language* 53: 285 311.

——— (1980) 'Nouns and Countability', *Language* 56: 541 67.

An Caighdeán Oifigiúl (1975) [1958] *Gramadach na Gaeilge agus Litriú na Gaeilge: An Caighdeán Oifigiúl*. Baile Átha Cliath [Dublin]: Oifig and an tSoláthair.

Anderson, G. (1999) 'Reduplicated Numerals in Salish', *International Journal of American Linguistics* 65(4): 407 81.

Anderson, S. (1982) 'Where's Morphology?', *Linguistic Inquiry* 13: 571 612.

——— (1986) 'Disjunctive Ordering in Inflectional Morphology', *Natural Language and Linguistic Theory* 4: 1 31.

——— (1992) *Amorphous Morphology*. Cambridge: Cambridge University Press.

Andrzejewski, B. W. (1960) 'The Categories of Number in Noun Forms in the Borana Dialect of Galla', *Africa* 30: 62 75.

Arad, M. (2003) 'Locality Constraints on the Interpretation of Roots: The Case of Hebrew Denominal Verbs', *Natural Language and Linguistic Theory* 21/4: 737 78.

ARONOFF, M. (1976) *Word Formation in Generative Grammar*. Cambridge, Mass.: MIT Press.

ARONOFF, M. (1994) *Morphology by Itself*. Cambridge, Mass.: MIT Press.

BACH, S. and J. SCHMITT JENSEN (1990) *Storre Italiensk grammatik*. Copenhagen: Munksgaard.

BACON, J. (1995) *Universals and Property Instances: The Alphabet of Being*. Oxford, Blackwell.

BADAWI, E S., M. CARTER, and A. GULLY (2004) *Modern Written Arabic: A Comprehensive Grammar*. London: Routledge.

BAERMAN, M. (2007) 'Morphological reversals', *Journal of Linguistics* 43(1): 33 61.

BAILEY, C. (1966) [1947] Titus Lucretius Caro, *De Rerum Natura Libri VI*, edited by C. Bailey. Oxford: Clarendon Press.

BAKER, M. (2003) *Lexical Categories*. Cambridge: Cambridge University Press.

BARKER, C. (1992) 'Group Terms in English: Representing Groups as Atoms', *Journal of Semantics* 9: 69 93.

———— (1999) 'Individuation and Quantification', *Linguistic Inquiry* 30: 683 91.

BEARD, R. (1982) 'Plural as a Lexical derivation (Word Formation)', *Glossa* 16: 133 48.

———— (1995) *Lexeme Morpheme Base Morphology: A General Theory of Inflection and Word Formation*. Albany, New York: State University of New York Press.

BEESTON, A. (1975) 'Some Features of Modern Standard Arabic', *Journal of Semitic Studies* 20: 62 8.

BELNAP, K. (1993) 'The Meaning of Deflected/Strict Agreement Variation in Cairene Arabic', in M. Eid and C. Holes (eds), *Perspectives on Arabic Linguistics V.* Amsterdam, Philadelphia: John Benjamins, 97 117.

———— (1999) 'A New Perspective on the History of Arabic Variation in Marking Agreement with Plural Heads', *Folia Linguistica* 33(2): 169 85.

———— and O. SHABANEH (1992) 'Variable Agreement and Nonhuman Plurals in Classical and Modern Standard Arabic', in E. Broselow and M. Eid (eds), *Perspectives on Arabic Linguistics IV.* Amsterdam, Philadelphia: John Benjamins, 245 62.

BERNSTEIN, J. (1993) Topics in the Syntax of Nominal Structure across Romance. Ph.D. Dissertation, City University of New York.

BIERMANN, A. (1982) 'Die Grammatische Kategorie Numerus', in H. J. Seiler and C. Lehmann (eds), *Apprehension: Das sprachliche Erfassen von Gegenständen. I: Bereich und Ordnung der Phänomene.* Tübingen: Gunter Narr, 229 43.

BIGALKE, R. (1996) *Abruzzese*. München: Lincom Europa.

BISANG, W. (1999) 'Classifiers in East and Southeast Asian Languages: Counting and Beyond', in J. Gvozdanović (ed.), *Numeral Types and Changes Worldwide*. Berlin: Mouton de Gruyter, 113 85.

BLOCH, A. (1978) 'Morphological doublets in Arabic dialects', in S. H. Al Aui (ed.), *Readings in Arabic Linguistics*. Bloomington: Indiana University, 405 29. Originally published in *Journal of Semitic Studies*, 1971, 16: 53 73.

BLOOMFIELD, L. (1933) *Language*. New York: Henry Holt.

Bobaljik, J. (2005) 'Itelmen diminutive plurals: a belated reply to Perlmutter 1998', *Yearbook of Morphology 2004.* Berlin: Springer, 317 19.

Bochner, H. (1984) 'Inflection within Derivation', *The Linguistic Review*, 3: 411 21.

Bodrogligeti, A. (2002) *Modern Literary Uzbek.* Munich: Lincom Europa.

Bonfante, G. (1973) 'Esiste il neutro in italiano?' in G. Bonfante, *Studii romeni.* Roma: Società accademica romena, 161 87.

Booij, G. (1994) 'Against Split Morphology', *Yearbook of Morphology 1993.* Dordrecht: Kluwer, 27 50.

—— (1996) 'Inherent versus Contextual Inflection and the Split Morphology Hypothesis', *Yearbook of Morphology 1995.* Dordrecht: Kluwer, 1 16.

—— (1999) 'Lexical Storage and Regular Processes', *Behavioral and Brain Sciences* 22: 1016.

Borer, H. (2005) *In Name Only.* Vol. 1 of *Structuring Sense.* Oxford: Oxford University Press.

Brennan, A. (1988) *Conditions of Identity.* Oxford: Clarendon Press.

Brugmann, K. (1900) *Griechische Grammatik.* Munich: Beck.

Brunet, J. (1978) *Grammaire critique de l'italien*, vol. 1. Saint Denis: Presses universitaires de Vincennes.

Brustad, K. (2000) *The Syntax of Spoken Arabic.* Washington, DC: Georgetown University Press.

Buccellati, G. (1976) 'The Case Against the Alleged Akkadian Plural Morpheme ĀNU, *Afroasiatic Linguistics* 3(2): 28 30.

Buckley, R. (2004) *Modern Literary Arabic: A Reference Grammar.* Beirut: Librairie du Liban.

Bunt, H. (1985) *Mass Terms and Model Theoretic Semantics.* Cambridge: Cambridge University Press.

Bybee, J. (1985) *Morphology: A Study of the Relation Between Meaning and Form.* Amsterdam, Philadelphia: John Benjamins.

Carlson, G. (1977) Reference to Kinds in English. Ph.D. Dissertation, University of Massachusetts at Amherst.

—— and F. Pelletier (1995) *The Generic Book.* Chicago: University of Chicago Press.

Carstairs, A. (1987) *Allomorphy in Inflection.* London: Croom Helm.

Carstairs McCarthy, A. (1994) 'Inflection Classes, Gender and the Principle of Contrast', *Language* 70(4): 737 88.

Carstens, V. (1993) 'On Nominal Morphology and the DP Structure', in S. Mchombo (ed.), *Theoretical Aspects of Bantu Grammar.* Palo Alto: Stanford Center for the Study of Language and Information, 151 80.

Cartwright, H. (1970), 'Quantities', *Philosophical Review* 79(1): 25 40.

Cesaro, A. (1939) *L'arabo parlato a Tripoli.* Milano: Mondadori.

Chapman, C. (1996) 'Perceptual Salience and Affix Order: Noun Plurals as Input to Word Formation', *Yearbook of Morphology 1995.* Dordrecht: Kluwer, 175 84.

Cheng, L. and R. Sybesma (1999) 'Bare and Not So bare Nouns and the Structure of NP', *Linguistic Inquiry* 30: 509 42.

Chierchia, G. (1998*a*) 'Plurality of Mass Nouns and the Notion of Semantic Parameter', in S. Rothstein (ed.), *Events and Grammar*. Dordrecht: Kluwer, 53 103.

(1998*b*) 'Reference to Kinds across Languages', *Natural Language Semantics* 6: 339 405.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge, Massachusetts: MIT Press.

(1995) *The Minimalist Program*. Cambridge, Massachusetts: MIT Press.

Clahsen, H. (1999) 'Lexical Entries and Rules of Language: A Multidisciplinary Study of German Inflection', *Behavioral and Brain Sciences* 22: 991 1013.

Contini Morava, E. (1999) 'Noun class and number in Swahili', in E. Contini Morava and Y. Tobin (eds), *Between Grammar and the Lexicon*. Amsterdam, Philadelphia: Benjamins, 3 29.

Corbett, G. (1991) *Gender*. Cambridge: Cambridge University Press.

(1993) 'The Head of Russian Numeral Expressions', in G. Corbett, N. Fraser, and S. MacGlashan (eds), *Heads in Grammatical Theory*. Cambridge: Cambridge University Press, 11 35.

(1996) 'Minor Number and Plurality Split', *Rivista di Linguistica* 8(1): 101 22.

(2000) *Number*. Cambridge: Cambridge University Press.

and R. Hayward (1987) 'Gender and Number in Bayso', *Lingua* 73: 1 28.

Cowell, M. (1964) *A Reference Grammar of Syrian Arabic (Based on the Dialect of Damascus)*. Washington, DC: Georgetown University Press.

Creissels, D. (2000) 'Typology', in B. Heine and D. Nurse (eds), *African Languages: An Introduction*. Cambridge: Cambridge University Press, 231 58.

Cuzzolin, P. (1998) 'Sull'origine del singolativo in celtico, con particolare riferimento al medio gallese', *Archivio Glottologico Italiano* 84: 121 49.

Dayal, V. (1992) 'The Singular Plural Distinction in Hindi Generics', *Proceedings of SALT II*, Ohio State University Working Papers in Linguistics.

De Bhaldraithe, T. (1977) [1953] *Gaeilge Chois Fairrge: an deilbhíocht*. Dublin: Dublin Institute for Advanced Studies.

Delbrück, B. (1893) *Vergleichende Syntax der Indogermanischen Sprachen: Erster Theil*. Strassburg: Trübner.

Déprez, V. (2005) 'Morphological Number, Semantic Number and Bare Nouns', *Lingua* 115: 857 83.

Desbordes, Y. (1983) *Petite grammaire du Breton moderne*. Lesneven: Hor yezh.

di Sciullo, A. M. and E. Williams (1987) *On the Definition of Word*. Cambridge, Mass.: MIT Press.

Dimmendaal, G. (1983) *The Turkana Language*. Dordrecht: Foris.

Donaldson, B. (1987) *Dutch Reference Grammar*. Leiden: Nijhoff.

Doyle, A. (1992) *Noun Derivation in Modern Irish*. Lublin: Redakcja Wydawnictw Katolickiego Uniwersytetu Lubelskiego.

Dressler, W. U. and A. Thornton (1996) 'Italian Nominal Inflection', *Wiener Linguistische Gazette* 57 59: 1 26.

Dryer, M. (1989) 'Plural Words', *Linguistics* 27: 865 95.

Duffield, N. (1995) *Particles and Projections in Irish Syntax*. Dordrecht: Kluwer.

Embick, D. and M. Halle (2004) 'On the Status of *Stems* in Morphological Theory', Ms., University of Pennsylvania and MIT.

    and R. Noyer (2001) 'Movement Operations After Syntax', *Linguistic Inquiry* 32(4): 555 95.

Eschenbach, C. (1993) 'Semantics of Number', *Journal of Semantics* 10: 1 31.

Farkas, D. (1990) 'Two Cases of Underspecification in Morphology', *Linguistic Inquiry* 21(4): 539 50.

    and H. de Swart (2003) *The Semantics of Incorporation*. Palo Alto: Stanford Center for the Study of Language and Information.

Favereau, F. (1997) *Grammaire du breton contemporain*. Morlaix: Skol Vreizh.

Fenech, E. (1996) 'Functions of the Dual Suffix in Maltese', *Rivista di Linguistica* 8(1): 89 99.

Fischer, W. (1972) *Grammatik des klassichen Arabisch*. Wiesbaden: Harrassowitz.

    (1980) 'Die Arabische Pluralbildung', *Zeitschrift für arabische Linguistik* 5: 70 88.

    and O. Jastrow (eds) (1980) *Handbuch der arabischen Dialekte*. Wiesbaden: Harrassowitz.

Fleisch, H. (1990) *Traité de philologie arabe* (2 vols; 2nd edn). Beirut: dar el Machreq.

Gabbay, D. and J. Moravcsik (1979) [1973] 'Sameness and Individuation', in F. Pelletier (ed.), *Mass Terms: Some Philosophical Problems*. Dordrecht: Reidel, 233 48. Originally published in *Journal of Philosophy* 70(13), 1973.

Gangemi, A., N. Guarino, C. Masolo, and A. Oltramari (2002) 'Understanding Top Level Ontological Distinctions', in A. Gómez Pérez, M. Grunninger, H. Stuck enschmidt, and M. Uschold (eds), *Proceedings of IJCAI 2001 (International Joint Conference of Artificial Intelligence) Workshop on Ontologies and Information Shar ing*. American Association for Artificial Intelligence Press, 26 33.

Geach, P. (1962) *Reference and Generality*. Ithaca, New York: Cornell University Press.

GGBC (1999) = Ó hAnluain, Liam. (ed.) *Graiméar Gaeilge na mBráithre Críostaí*, Baile Átha Cliath [Dublin]: an Gúm.

Giammarco, E. (1979) *Abruzzo*. Pisa: Pacini.

Gil, D. (1988) 'Georgian Reduplication and the Domain of Distributivity', *Linguistics* 26: 1039 65.

    (1996) 'Maltese "Collective Nouns": A Typological Perspective', *Rivista di Linguistica* 8(1): 53 87.

Goetze, A. (1946) 'The Akkadian Masculine Plural in *ānū / ī* and its Semitic Back ground, *Language* 22: 121 30.

Goidánich, P. G. (1967) *Grammatica italiana*. Bologna: Zanichelli.

Göksel, A. and C. Kerslake (2005) *Turkish Comprehensive Grammar*. London: Routledge.

Goodman, N. (1977) [1951] *The Structure of Appearance*. Dordrecht: Reidel.

Grafström, Å. (1968) *Etude sur la morphologie des plus anciennes chartes languedociennes*. Stockholm: Ålmqvist & Wiksell.

GREENBERG, J. (1963) 'Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements', in J. Greenberg (ed.), *Universals of Language.* Cambridge, Mass.: MIT Press, 73 113.

——— (1974) 'Numeral Classifiers and Substantival Number', in L. Heilmann (ed.), *Proceedings of the Eleventh International Congress of Linguists.* Bologna: il Mulino, 17 37.

GREENE, D. (1974) 'Distinctive Plural Forms in Old and Middle Irish', *Ériu* 25: 190 9.

GREENE, D. (1992) 'Celtic', in J. Gvozdanović (ed.), *Indo European Numerals.* Berlin: Mouton de Gruyter, 497 554.

GRIFFIN, N. (1977) *Relative Identity.* Oxford: Clarendon Press.

GRØNBECH, K. and J. KRUEGER (1976) *Introduction to Classical Mongolian.* Wiesbaden: Harrassowitz.

GUARINO, N. and C. WELTY (2000), 'Identity, Unity and Individuality: Towards a Formal Toolkit for Ontological Analysis', in W. Horn (ed.), *Proceedings of ECAI 2000 (European Conference on Artificial Intelligence).* Amsterdam: IOS Press, 219 23.

GUPTA, A. (1980) *The Logic of Common Nouns.* New Haven: Yale University Press.

HAASE, M. (2000) 'Reorganization of a Gender System: The Central Italian Neuters', in B. Unterbeck, M. Rissanen, T. Nevalainen, and M. Saari (eds), *Gender in Grammar and Cognition.* Berlin: Mouton de Gruyter, 223 36.

HAIMAN, J. (1988) 'Rhaeto Romance', in M. Harris and N. Vincent (eds), *The Romance Languages.* London: Croom Helm, 391 419.

HALL, R. (1956) 'Il plurale italiano in *a*: un duale mancato? *Italica* 33: 140 2.

——— (1983) *Proto Romance Morphology.* Amsterdam, Philadelphia: John Benjamins.

HALLE, M. and A. MARANTZ (1993) Distributed Morphology and the pieces of inflection, in K. Hale and S. J. Keyser (eds), *The View from Building 20.* Cambridge, Mass.: MIT Press, 111 76.

HARBOUR, D. (2003) 'The Kiowa Case for Feature Insertion', *Natural Language and Linguistic Theory* 21(3): 543 78.

——— (2007) *Morphosemantic Number.* Berlin: Springer.

HARLEY, H. and R. NOYER (2000) 'Formal versus Encyclopedic Properties of Vocabulary: Evidence from Nominalisations', in B. Peeters (ed.), *The Lexicon Encyclopedia Interface.* Amsterdam: Elsevier, 349 74.

——— and E. RITTER (2002) 'Person and Number in Pronouns: A Feature Geometric Analysis', *Language* 78: 482 526.

HARRELL, R. (1962) *A Short Reference Grammar of Moroccan Arabic.* Washington, DC: Georgetown University Press.

HARRIS, J. (1991) 'The Exponence of Gender in Spanish', *Linguistic Inquiry* 22: 27 62.

HEMON, R. (1975) *A Historical Morphology and Syntax of Breton.* Dublin: Institute for Advanced Studies.

——— (1995) *Breton Grammar* (10th edn., translated and revised by Michael Everson), Dublin: Everson Gunn Teo.

HEYCOCK, C. and R. ZAMPARELLI (2005) 'Friends and Colleagues: Plurality, Coordination and the Structure of DP', *Natural Language Semantics* 13(3): 201 70.

HIRSCH, E. (1982) *The Concept of Identity*. Oxford: Oxford University Press.

HOFMANN, J. B. and A. SZANTYR (1965) *Lateinische Laut und Formenlehre*, Vol. 2, München: Beck.

HOLES, C. (2004) *Modern Arabic: Structures, Functions, and Varieties* (2nd edn). Washington, DC: Georgetown University Press.

HOLMES, P. and I. HINCHLIFFE (2003) *Swedish Comprehensive Grammar*. London: Routledge.

HUGHES, A. (1994) 'Gaeilge Ulaidh', in K. McCone, D. McManus, C. Ó Háinle, N. Williams, and L. Breatnach (eds), *Stair na Gaeilge*. Maigh Nuad: Roinn na Sean Ghaeilge, Coláiste Phádraig, 611 60.

HUMPHREYS, H. L. (1995) *Phonologie et Morphosyntaxe du Parler Breton de Bothoa en Saint Nicholas du Pélem (Côtes d'Armor)*. Brest: Ar Skol Vrezoneg/Emgleo Breiz.

IDRISSI, A. (2000) 'On Berber Plurals', in J. Lecarme, J. Lowenstamm, and U. Shlonsky (eds), *Research in Afroasiatic Grammar: Papers from the Third Conference on Afroasiatic Languages*. Amsterdam, Philadelphia: John Benjamins, 101 24.

JACKENDOFF, R. (1991) 'Parts and Boundaries', *Cognition* 41: 9 45.

    (1997) *The Architecture of the Language Faculty*. Cambridge, Mass.: MIT Press.

    (2002) *Foundations of Language*. Oxford: Oxford University Press.

JENSEN, F. (1976) *The Old Provencal Nouns and Adjective Declension*. Odense: Odense University Press.

JOSEPHS, L. (1975) *Palauan Reference Grammar*. Honolulu: The University Press of Hawaii.

JURAFSKY, D. (1996) 'Universal Tendencies in the Semantics of the Diminutive', *Language* 72(3): 533 78.

KAUTZSCH, E. (1980) [1910] *Gesenius' Hebrew Grammar as Edited and Enlarged by the Late E. Kautzsch, Translated by A.E. Cowley*. Oxford: Oxford University Press.

KETREZ, N. (2004) ' lAr Marked Nominals and Three Types of Plurality in Turkish', in J. Cihlar, A. Franklin, D. Kaiser, and I. Kimbara (eds), *Proceedings of the 39th Meeting of the Chicago Linguistics Society*, 176 92.

KIHM, A. (2001) 'Noun Classes, Gender, and the Lexicon Syntax Morphology Interfaces: A Comparative Study of Niger Congo and Romance Languages', CNRS, Paris. Published in G. Cinque and R. Kayne (eds), *The Oxford Handbook of Comparative Syntax*, 2005, Oxford: Oxford University Press, 459 512.

    (2003) 'Les pluriels internes de l'arabe: système et conséquences pour l'architecture de la grammaire', *Recherches Linguistiques de Vincennes*, 32: 109 56.

    (2005) 'The Case for One Case/One Gender in Romanian: A Tentative Account of Romanian Declension', CNRS, Université Paris 8.

KING, G. (1993) *Modern Welsh: A Comprehensive Grammar*. London: Routledge.

KOPTJEVSKAJA TAMM, M. and B. WÄLCHLI (2001) 'The Circum Baltic Languages: An Areal Typological Approach', in Ö. Dahl and M. Koptjevskaja Tamm (eds), *Circum Baltic Languages; Vol. 2: Grammar and Typology*. Amsterdam, Philadelphia: John Benjamins, 616 750.

KRECZ, C. (1986) 'Parts and Pieces', *Philosophy and Phenomenological Research* 46(3): 381 400.

KRIFKA, M. (1990) 'Four Thousand Ships Passed Through the Lock': Object Induced Measure Functions on Events', *Linguistics and Philosophy* 13: 487 520.

KRIFKA, M. (1995) 'Common Nouns: A Contrastive Analysis of Chinese and English', in G. Carlson and F. Pelletier (eds), *The Generic Book*. Chicago: University of Chicago Press, 398 411.

——, F. PELLETIER, G. CARLSON, A. TER MEULEN, G. CHIERCHIA, and G. LINK. (1995) 'Genericity: An Introduction', in G. Carlson and F. Pelletier (eds), *The Generic Book*. Chicago: University of Chicago Press, 1 124.

KÜHNER, R. and C. STEGMANN (1971) *Ausführliche Grammatik der Lateinischen Sprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.

LANDGRAF, G. (1906) 'Bemerkungen zum sog. poetischen Plural in der lateinischen Prosa', *Archiv für Lateinische Lexicographie und Grammatik* 14: 63 74.

LANDMAN, F. (1989) 'Groups I', 'Groups II', *Linguistics and Philosophy* 12, 559 605, 723 44.

—— (1991) *Structures for Semantics*. Dordrecht: Reidel.

LANGACKER, R. (1987) 'Nouns and Verbs', *Language* 63: 53 94.

LARSON, R. and G. SEGAL (1995) *Knowledge of Meaning: An Introduction to Semantic Theory*, Cambridge, Mass.: MIT Press.

LAUSBERG, H. (1963) *Romanische Sprachwissenschaft*. Berlin: de Gruyter.

LECARME, J. (2002) 'Gender "Polarity": Theoretical Aspects of Somali Nominal Morphology', in P. Boucher and M. Plénat (eds), *Many Morphologies*. Somerville, Mass.: Cascadilla Press, 109 41.

LEPSCHY, A. L. and G. LEPSCHY (1988) *The Italian Language Today*. London: Hutchinson.

LEWIS, C. S. (1960) *The Four Loves*. London: Geoffrey Bles.

LEWIS, G. (1967) *Turkish Grammar*. Oxford: Oxford University Press.

LINK, G. (1983) 'The Logical Analysis of Plurals and Mass Terms: A Lattice Theoretic Approach', in R. Bäuerle, C. Schwarze, and A. von Stechow (eds), *Meaning, Use and Interpretation of Language*. Berlin: de Gruyter, 303 23.

—— (1998) *Natural Language and Algebraic Semantics*. Palo Alto: Stanford Center for the Study of Language and Information.

LJAŠEVSKAJA, O. N. (2004) *Semantika russkogo čisla*. Moscow: Jazyki slavjanskoj kul'tury.

LÖBEL, E. (2000) 'Classifiers versus Genders and Noun Classes: A Case Study in Vietnamese', in B. Unterbeck, M. Rissanen, T. Nevalainen, and M. Saari (eds), *Gender in Grammar and Cognition*. Berlin: Mouton de Gruyter, 259 319.

LÖFSTEDT, E. (1928) *Syntactica: Erster Teil*, Lund: Gleerup.

LOWE, E. J. (1998) *The Possibility of Metaphysics*. Oxford: Oxford University Press.

LUMSDEN, J. (1992) 'Underspecification in Grammatical and Natural Gender.' *Linguistic Inquiry* 23: 469 86.

MAAS, P. (1902) 'Studien zum poetischen Plural bei den Römern', *Archiv für Lateinische Lexicographie und Grammatik* 12: 479 510.

MCCARTHY, J. and A. PRINCE (1990) 'Foot and Word in Prosodic Morphology: The Arabic Broken Plural', *Natural Language and Linguistic Theory* 8: 209 83.

McCawley, J. (1975) 'Lexicography and the Count Mass distinction', in C. Cogen, H. Thompson, G. Thurgood, K. Whistler, and J. Wright (eds), *Proceedings of the First Meeting of the Berkeley Linguistics Society*, 314 321.

Mac Congáil, N. (2004) *Irish Grammar Book*. Indreabhán: Cló Iar Chonnachta.

McDonald, C. (1998) 'Tropes and other Things', in S. Laurence and C. McDonald (eds), *Contemporary Readings in the Foundations of Metaphysics*. Oxford: Blackwell, 329 50.

McKenna, M. (1988) *A Handbook of Spoken Breton*. Tübingen: Niemeyer.

Maiden, Martin (1991) *Interactive Morphophonology: Metaphony in Italy*. London: Routledge.

(1995) *A Linguistic History of Italian*. London: Routledge.

(2004) 'When Lexemes Become Allomorphs: On the Genesis of Suppletion', *Folia Linguistica* 38: 227 56.

and C. Robustelli (2000) *A Reference Grammar of Modern Italian*. London: Arnold.

Mallinson, G. (1986) *Rumanian*. London: Croom Helm.

(1988) 'Rumanian', in M. Harris and N. Vincent (eds), *The Romance Languages*. London: Croom Helm, 391 419.

Marantz, A. (1997) 'No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon', *UPenn Working Papers in Linguistics 4.2*, Alexis Dimitriadis (ed.), 201 25.

(2003) 'Blocking', Ms., MIT.

Marcantonio, A. and A. M. Pretto (1988) 'Il nome', in L. Renzi and G. Salvi (eds), *Grande grammatica italiana di consultazione*, vol. 1. Bologna: il Mulino, 315 32.

Mathiot, M. (1983) 'Papago Semantics', in A. Ortiz (ed.), *Handbook of North American Indians*, Vol. 10. Washington, DC: Smithsonian Institution, 201 11.

Matthews, P. (1972) *Inflectional Morphology*. Cambridge: Cambridge University Press.

(1974) *Morphology*. Cambridge: Cambridge University Press.

Meirav, A. (2003) *Wholes, Sums and Unities*. Dordrecht: Kluwer.

Meisterfeld, R. (1998) *Numerus und Nominalaspekt*. Tübingen: Niemeyer.

Merlo, C. (1952) 'Ital. "le labbra, le braccia" e sim.', *Italica* 29: 229 34.

Mifsud, M. (1996) 'The Collective in Maltese', *Rivista di Linguistica* 8(1): 29 51.

Mithun, M. (1988) 'Lexical Categories and the Evolution of Number Marking', in M. Hammond and M. Noonan (eds), *Theoretical Morphology*. San Diego, New York: Academic Press, 211 34.

(1999) *The Languages of Native North America*. Cambridge: Cambridge University Press.

Moltmann, F. (1997) *Parts and Wholes in Semantics*. Oxford: Oxford University Press.

(2003) 'Nominalizing Quantifiers', Journal of Philosophical Logic 32: 445 81.

(2004) 'Properties and Kinds of Tropes: New Linguistics Facts and Old Philosophical Insights', *Mind* 113, 449, 1 41.

Moltmann, F. (2005) 'Two Kinds of Universals and Two Kinds of Collections', Ms., University of Stirling.

Moravcsik, E. (2003) 'A Semantic Analysis of Associative Plurals', *Studies in Lan guage* 27(3): 469 503.

Nakanishi, K. and S. Tomioka (2004) 'Japanese Plurals are Exceptional', *Journal of East Asian Linguistics* 13: 113 40.

Ní Chiosáin, M. (1991) Topics in the Phonology of Irish. Ph.D. Dissertation, Uni versity of Massachusetts at Amherst.

Noyer, R. (1997) *Features, Positions, and Affixes in Autonomous Morphological Structure*, New York: Garland.

Ó Baoill, D. (1999) *An teanga beo: Gaeilge Uladh.* Baile Átha Cliath [Dublin]: Institiúid Teangeolaíochta Éireann.

Ó Dochartaigh, C. (1992) 'The Irish Language', in D. MacAulay (ed.), *The Celtic Languages.* Cambridge: Cambridge University Press, 11 99.

Ó Dónaill, N. (1981)[1978] *Foclóir Gaeilge Béarla.* Baile Átha Cliath [Dublin]: Oifig an tSoláthair.

Ó Sé, D. (2000) *Gaeilge Corca Dhuibhne.* Baile Átha Cliath [Dublin]: Institiúid Teangeolaíochta Éireann.

Ó Siadhail, M. (1982) 'Cardinal Numbers in Modern Irish', *Ériu* 33: 99 107.

(1989) *Modern Irish: Grammatical Structure and Dialectal Variation.* Cambridge University Press: Cambridge.

Ojeda, A. (1992a) 'The "Mass Neuter" in Hispano Romance dialects', *Hispanic Linguistics* 5: 245 77.

(1992b) 'The Semantics of Number in Arabic', in C. Barker and D. Dowty (eds), *SALT II: Proceedings from the Second Conference on Semantics and Linguistic Theory.* Ohio State University, Working Papers in Linguistics 40, 303 25.

(1993) *Linguistic Individuals.* Stanford: Center for the Study of Language and Information.

(1995) 'The Semantics of the Italian Double Plural', *Journal of Semantics* 12: 213 37.

(1998) 'The Semantics of Collectives and Distributives in Papago', *Natural Language Semantics* 6: 245 70.

Omar, M. (1975) *Saudi Arabic Basic Course.* Washington, DC: Foreign Service Institute.

Oomen, A. (1981) 'Gender and Plurality in Rendille', *Afroasiatic Linguistics* 8(1) 35 75.

Owens, J. (1985) *A Grammar of Harar Oromo (Northeastern Ethiopia).* Hamburg: Helmut Buske.

Pedersen, H. (1913) *Vergleichende Grammatik der keltischen Sprachen*, vol. 2. Göttingen: Vandenhoek and Ruprecht.

Pelletier, F. (1979)[1975] 'Non Singular Reference: Some Preliminaries', in F. Pelle tier (ed.), *Mass Terms: Some Philosophical Problems.* Dordrecht: Reidel, 1 14. Ori ginally published in *Philosophia* 5, 1975.

and L. Schubert. (1989) 'Mass Expressions', in D. Gabbay and F. Guenthner (eds), *Handbook of Philosophical Logic. Vol. IV: Topics in the Philosophy of Language.* Dordrecht: Reidel, 327 407.

Penchoen, T. (1973) *Tamazight of the Ayt Ndhir.* Los Angeles: Undena Publications.

Peperkamp, S. (1995) 'Prosodic Constraints in the Derivational Morphology of Italian', *Yearbook of Morphology 1994.* Dordrecht: Kluwer, 207 44.

Perlmutter, D. (1988) 'The Split Morphology Hypothesis: Evidence from Yiddish', in M. Hammond and M. Noonan (eds), *Theoretical Morphology.* San Diego, New York: Academic Press, 79 100.

Pianesi, F. (2002) 'Review of F. Moltmann, *Parts and Wholes in Semantics*', *Linguistics and Philosophy* 25: 97 120.

Pinker, S. (1999) *Words and Rules.* London: Weidenfeld and Nicolson.

Plank, F. (1996) 'Domains of the Dual, in Maltese and in General', *Rivista di Linguistica* 8(1) 123 40.

Plungjan, V. (1997) 'Vremja i vremena: k voprosu o kategorii čisla', in N. D. Arutjunova and T. E. Janko (eds), *Logičeskij analiz jazyka: jazyk i vremja.* Moscow: Indrik, 158 69.

Polomé, E. (1967) *Swahili Language Handbook.* Washington, DC: Center for Applied Linguistics.

Press, I. (1986) *Breton.* Berlin: Mouton de Gruyter.

Pustejovsky, J. (1995) *The Generative Lexicon.* Cambridge, Mass.: MIT Press.

Qafisheh, H. (1977) *A Short Reference Grammar of Gulf Arabic.* Tucson: The University of Arizona Press.

Quine, W. V. O. (1960) *Word and Object.* Cambridge, Mass.: MIT Press.

Quirk, R., G. Leech, and J. Svartvik (1985) *A Comprehensive Grammar of the English Language.* London: Longman.

Rainer, F. (1996) 'Inflection Inside Derivation: Evidence from Spanish and Portuguese', *Yearbook of Morphology 1995.* Dordrecht: Kluwer, 83 91.

Ratcliffe, R. (1990) 'Arabic Broken Plurals: Arguments for a Two Fold Classification of Morphology', in M. Eid and J. McCarthy (eds), *Perspectives on Arabic Linguistics II.* Amsterdam, Philadelphia: John Benjamins, 94 119.

—— (1998) *The 'Broken' Plural Problem in Arabic and Comparative Semitic.* Amsterdam, Philadelphia: John Benjamins.

Regula, M. and J. Jernej (1975) *Grammatica italiana descrittiva*, (2nd edn). Bern: Francke.

Riente, L. (2000) 'Ladies First: The Pivotal Role of Gender in the Italian Nominal Inflection System', MA Research Paper, McGill University.

Rijkhoff, J. (1991) 'Nominal Aspect', *Journal of Semantics* 8: 291 309.

—— (2002) *The Noun Phrase.* Oxford: Oxford University Press.

Ritter, E. (1991) 'Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew', in S. Rothstein (ed.), *Syntax and Semantics, vol. 25: Perspectives on Phrase Structure.* San Diego, New York: Academic Press, 37 62.

—— (1993) 'Where's Gender?', *Linguistic Inquiry* 24: 795 803.

Rocchetti, A. (1968) 'Les pluriels doubles de l'italien: une interference de la sémantique et de la morphologie du nom', *Les langues modernes* 62: 351 9.

ROHLFS, G. (1966) *Grammatica della lingua italiana e dei suoi dialetti, vol. 1: fonetica*, Torino: Einaudi.

ROHLFS, G. (1968) *Grammatica della lingua italiana e dei suoi dialetti, vol. 2: morfo logia*, Torino: Einaudi.

ROUNDS, C. (2001) *Hungarian. An Essential Grammar*. London: Routledge.

SAEED, J. (1987) 'Central Somali A Grammatical Outline', *Afroasiatic Linguistics* 8(2) 1 43.

SANTANGELO, A. (1981) I plurali italiani del tipo 'le braccia', *Archivio glottologico italiano* 66: 95 153.

SAUERLAND, U. (2004) 'A New Semantics for Number', Ms., Universität Tübingen.

SCALISE, S. (1994) *Morfologia*. Bologna: il Mulino.

SCHÖN, I. (1971) *Neutrum und Kollektivum. Das Morphem a im Lateinischen und Romanischen.* Innsbruck: Institut für Sprachwissenschaft.

SCHUH, R. (1989) 'Gender and Number in Miya', in Z. Frajzyngier (ed.), *Current Progress in Chadic Linguistics.* Amsterdam, Philapdelphia: John Benjamins.

SCHWARZE, C. (1988) *Grammatik der italienischen Sprache.* Tübingen: Niemeyer.

SCHWARZSCHILD, R. (1996) *Pluralities.* Dordrecht: Kluwer.

SCHWARZWALD, O. (1991) 'Grammatical vs. lexical plural formation in Hebrew', *Folia Linguistica* 25: 577 608.

SCHWYTZER, E. and A. DEBRUNNER (1950) *Griechische Grammatik.* Munich: Beck.

SEILER, H. J. and C. LEHMANN (1982) *Apprehension: Das sprachliche Erfassen von Gegenständen. I: Bereich und Ordnung der Phänomene.* Tübingen: Gunter Narr.

SENFT, G. (1986) *Kilivila: The Language of the Trobriand Islanders.* Berlin: de Gruyter.

SERZISKO, F. (1982) 'Numerus/Genus Kongruenz und das Phänomen der Polarität am Beispiel einiger ostkuschitischer Sprachen', in H. J. Seiler and F. Stachowiak (eds), *Apprehension: das sprachliche Erfassen von Gegenständen. II: Die Techniken und ihr Zusammenhang in Einzelsprachen.* Tübingen: Gunter Narr, 179 200.

SHARIFAN, F. and A. LOTFI (2003) ' "Rices" and "waters": The Mass Count Distinction in Modern Persian', *Anthropological Linguistics* 45: 226 44.

SIMONS, P. (1987) *Parts: A Study in Ontology.* Oxford: Clarendon Press.

SMITH STARK, C. (1974) 'The Plurality Split', in M. La Galy, R. Fox, and A. Bruck (eds), *Papers from the Tenth Regional Meeting, Chicago Linguistics Society,* 657 71.

SOMMER, F. (1902) *Handbuch der lateinischen Laut und Formenlehre.* Heidelberg: Winter.

SONG, S. C. (1975) 'Rare Plural Marking and Ubiquitous Plural Marker in Korean', in R. Grossmann, J. San, and T. Vance (eds), *Papers from the 11th Regional Meeting of the Chicago Linguistic Society,* 536 46.

SPITZER, L. (1941) 'Feminizacion del neutro', *Revista de fililogía hispánica* 3: 339 71.

STANKIEWICZ, E. (1962) 'The Singular Plural Opposition in the Slavic Languages', *International Journal of Slavic Linguistics and Poetics* 5: 1 15.

STEMBERGER, J. and B. MACWHINNEY (1988) 'Are Inflected Forms Stored in the Lexicon?', in M. Hammond and M. Noonan (eds), *Theoretical Morphology.* San Diego, New York, Academic Press, 101 15.

STEPHENS, J. (1993) 'Breton', in M. Ball, with J. Fife (eds), *The Celtic Languages*, London: Routledge, 349 409.

STRAWSON, P. (1997) [1976] 'Entity and Identity', *Entity and Identity and Other Essays*, Oxford: Clarendon Press, 21 51. Originally published in H. Lewis (ed.), *Contemporary British Philosophy: Fourth Series*. London: Allen and Unwin.

STUMP, G. (1989) 'A Note on Breton Pluralization and the Elsewhere Condition', *Natural Language and Linguistic Theory* 7: 261 73.

    (1990) 'Breton Inflection and the Split Morphology Hypothesis', in R. Hendrick (ed.), *Syntax and Semantics 23: The Syntax of the Modern Celtic Languages*. San Diego, New York: Academic Press, 97 119.

    (1993) 'How Peculiar is Evaluative Morphology?', *Journal of Linguistics* 29(1), 1 36.

    (1998) 'Inflection', in A. Spencer and A. Zwicky (eds), *The Handbook of Morph ology*. Oxford: Blackwell, 13 43.

    (2001) *Inflection: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.

SZABOLCSI, A. and F. ZWARTS (1993) 'Weak Islands and Algebraic Semantics for Scope Taking', *Natural Language Semantics* 1: 235 84.

TAMM, M. (2004) 'Mass and Collection', in G. Booij, C. Lehmann, J. Mugdan, S. Skopeteas, in collaboration with W. Kesselheim (eds), *Morphologie/Morphology. Ein Internationales Handbuch zur Flexion und Wortbildung* [*An International Hand book on Inflection and Word Formation*], vol. 2. Berlin: Mouton de Gruyter, 1067 73.

TEKAVČIĆ, P. (1972) *Grammatica storica dell'italiano, vol. 2: morfosintassi*. Bologna: il Mulino.

TERNES, E. (1970) *Grammaire structurale du breton de l'Ile de Groix (dialecte occidental)*. Heidelberg: Winter.

    (1992) The Breton Language', in D. MacAulay (ed.), *The Celtic Languages*. Cambridge: Cambridge University Press, 371 452.

THACKSTON, W. (1994) *Koranic and Classical Arabic: An Elementary Grammar of the Language*. Bethesda, Maryland: IBEX Publishers.

THORNTON, A. (2001) 'Some Reflections on Gender and Inflectional Class Assignment in Italian', in C. Schaner Wolles, J. Rennison, and F. Neubarth (eds), *Naturally! Linguistic Studies in Honour of Wolfgang Ulrich Dressler Presented on the Occasion of his 60th Birthday*. Turin: Rosenberg e Sellier, 479 87.

    (2003*a*) 'L'assegnazione del genere in italiano', in F. Sánchez Miret (ed.), *Actas del XXIII Congreso Internacional de Lingüística y Filología Románica*, Vol. 1. Tübingen: Niemeyer, 467 81.

    (2003*b*) 'La rappresentazione dell'informazione morfologica nelle entrate lessi cali', in A. Bisetto, C. Iacobini, and A. Thornton (eds), *Scritti di morfologia in onore di Sergio Scalise in occasione del suo sessantesimo compleanno*. Rome: Caissa Italia editore, 203 21.

THURNEYSEN, R. (1980) [1946]. *A Grammar of Old Irish*. Dublin: Institute for Advanced Studies.

Tiersma, P. (1982) 'Local and General Markedness', *Language* 54: 832 49.

Tobin, Y. (1999) 'The Dual Number in Hebrew: Grammar, or Lexicon, or Both?', in E. Contini Morava and Y. Tobin (eds), *Between Grammar and the Lexicon*. Amsterdam, Philadelphia: Benjamins, 87 122.

Trépos, P. (1956) *Le pluriel breton*. Rennes: Imprimeries Réunies.
      (1968) *Grammaire bretonne*. Rennes: Imprimerie Simon.

Tsiapera, M. (1969) *A Descriptive Analysis of Cypriot Maronite Arabic* (Janua Linguarum 66). The Hague, Paris: Mouton.

Tucker, A. and M. Bryan (1966) *The Non Bantu Languages of North Eastern Africa*. Oxford: Oxford University Press.

Unterbeck, B. (1993) *Kollektion, Klassifikation und Transnumerus. Eine Typologische Studie zum Koreanischen*. Frankfurt am Main: Lang.

Väänänen, V. (1963) *Introduction au latin vulgaire*. Paris: Klincksieck.

van Marle, J. (1996) 'The Unity of Morphology: On the Interwovenness of the Derivational and Inflectional Dimension of the Word,' *Yearbook of Morphology 1995*. Dordrecht: Kluwer, 67 82.

von Soden, W. (1969) *Grundriss der Akkadischen Grammatik*. Roma: Pontificium Institutum Biblicum.

Wackernagel, J. (1926) *Vorlesungen über Syntax*. Basel: Birkhäuser.

Wade, T. (1992) *A Comprehensive Russian Grammar*. Oxford: Blackwell.

Waltke, B. K. and M. P. O'Connor (1990), *An Introduction to Biblical Hebrew Syntax*. Winona Lake: Eisenbrauns.

Wehr, H. (1976) *A Dictionary of Modern Written Arabic, Edited by J. Milton Cowan* (3rd edn), Ithaca New York: Spoken Language Services, Inc.

Wierzbicka, A. (1988) *The Semantics of Grammar*. Amsterdam, Philadelphia, John Benjamins.

Wiggins, D. (1980) *Sameness and Substance*. Oxford: Blackwell.

Windisch, R. (1973) *Genusprobleme im Romanischen*. Tübingen: Gunter Narr.

Winter, Y. (2002) 'Atoms and Sets: A Characterization of Semantic Number', *Linguistic Inquiry* 33(3) 493 505.

Wmffre, Y. (1998) *Central Breton*. Munich: Lincom Europa.

Wright, W. (1967) [1896]. *A Grammar of the Arabic Language*. Cambridge: Cambridge University Press.

Zabbal, Y. (2002) 'The Semantics of Number in the Arabic Noun Phrase'. M.A. Dissertation, University of Calgary.

Zamparelli, R. (2003) *Layers in the Determiner Phrase*. New York: Garland.

Zeuss, J. C. (1871) *Grammatica celtica* (2nd edn). Berlin: Weidmann.

Zwicky, A. and G. Pullum (1983) 'Cliticization vs. Inflection: English *n't*', *Language* 59: 502 13.

Name index

Index of languages and language families

Subject index

OXFORD STUDIES IN THEORETICAL LINGUISTICS