Lexemes as listed structures: Evidence from Italian irregular plurals

Paolo Acquaviva
University College Dublin
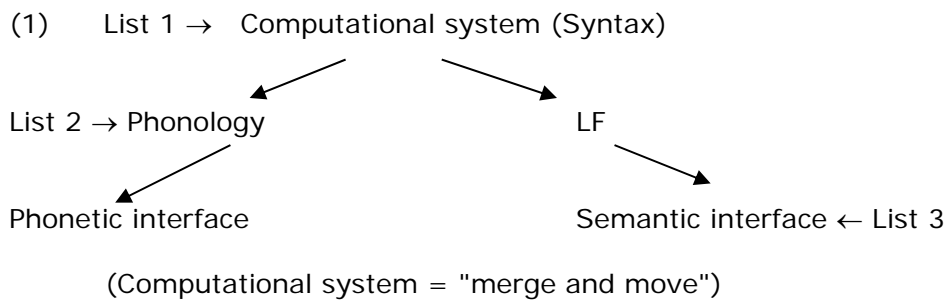Paolo.acquaviva@ucd.ie

Ce qui avait commencé comme une étude assez spécifique sur les pluriels irréguliers de l'italien aboutit, quelques années plus tard, à une enquête beaucoup plus étendue sur la pluralité comme déterminant du sens lexical, avec la monographie Lexical Plurals *(2008). La contribution qui suit représente une étape importante dans ce parcours. Elle fait suite à un article (publié en italien en 2002 dans* Lingue e linguaggio*) qui soutenait l'extranéité de ces pluriels irréguliers au paradigme flexionnel, à l'appui de nombre de considérations de différents ordres. L'auteur reprend ici les plus importants de ces arguments, mais cette fois dans un cadre théorique précis (celui de la Distributed Morphology). Ceci permet d'inscrire l'analyse de ces formes dans le contexte d'une question théorique centrale : étant donné une construction grammaticale, comment un lexème peut-il se distinguer d'un autre par un élément flexionnel tel que le pluriel ? Comment délimiter ce qui est 'lexical' de ce qui ne l'est pas, dans une approche constructionnelle ? L'analyse qui suit, qui sera développée et mise en relation avec d'autres phénomènes de pluralité lexicale, propose que la pluralité de ces noms est inséparable du choix du genre féminin. Il s'agit, donc, de pluralia tantum, où le nombre est décidé par le genre, de dernier étant une propriété du lexème, tout en restant un élément dans la structure qui forme un mot.*

1. Distributed Morphology: "distributing" the lexicon

The label "distributed", with which Halle and Marantz (1993) expressed the salient theoretical innovation of their model of morphology, encapsulates an intuition with far-reaching implications: there is no one lexicon providing "units" (listed or otherwise) for syntax, semantics and morphology. Syntax manipulates abstract features without regard for what may count as lexical in either of the two senses elucidated by Aronoff (1994), that is, "listed/idiosyncratic" and "substantive/lexemic"; its output is then mapped to morphology, whose output is in turn mapped to phonology, without any reference to lexical building blocks or bases receiving contextual (grammatical) specifications. Rather, the lexicon is distributed across different dimensions of listedness. This factorization of listed information emerges clearly from the model of grammar defended in Marantz (1997: 204):

(1)   List 1 → Computational system (Syntax)

List 2 → Phonology   LF

Phonetic interface   Semantic interface ← List 3

        (Computational system = "merge and move")

Instead of viewing syntax as the locus for the arrangement and manipulation of lexical items, Marantz sees it as an abstract computational system (along the lines of Chomsky 1995), whose units (the syntactic atoms) are not lexical items in any sense, and indeed would not seem to be products of any linguistic component:

(2)   "List 1 or the "narrow lexicon," most directly replaces the Lexicon as it provides the units that the syntax operates with. This list 1 contains the atomic roots of the language and the atomic bundles of grammatical features." (*ibidem*: 203)

Quite distinct from this dimension are the purely morphological and the purely interpretive dimensions of listedness:

(3)   "List 2 or the "Vocabulary," provides the phonological forms for the terminal nodes from the syntax … The Vocabulary includes the connections between sets of grammatical features and phonological features, and thus determines the connections between terminal nodes from the syntax and their phonological realization." (*ibidem*: 204)

(4)   "The final Lexicon replacement in [1] is List 3 or the "Encyclopedia"—the list of special meanings. The Encyclopedia lists the special meanings of particular roots, relative to the syntactic context of the roots, within local domains." (*ibidem*: 204)

Different aspects of listedness are thus distributed over different components of grammar. The most conspicuous absentee in such a picture is the lexeme, the concept that represents Aronoff's (1994) second sense of "lexical": a possibly abstract unit mapped into one or more word-forms when in a syntactic context, and encoding intrinsic (i.e. not context-dependent) grammatical, interpretive and formal properties of a word.[1] This is not to say that there exist no listemes apart from the atomic sound/meaning mappings (the Vocabulary in List 2): the point is rather that there are no other *morphological* listemes. Elements of List 3 may well be associated with complex

---

[1] The need for a distinction between lexematic and inflectional morphology has been argued for extensively by Beard (1995).

morphosyntactic structures; however, this does not affect the way in which these structures are built up:

(5)   "What you see is what you get; i.e., if the morphophonology justifies decomposition into a complex structure of terminal nodes, the syntax must create this structure and the structure must be interpreted in the regular way for such constructions (with of course the possibility that roots in the construction might have special meanings in the context of (elements of) the construction)." (*ibidem*: 212)

The final qualification in (5) clarifies that roots, as opposed to affixes, are the real locus of idiosyncratic interpretation, even when what gets a non-compositional interpretation is a whole complex structure:

(6)   "Things with special meanings are roots". (*ibidem*: 212-213)

The clear division between roots and grammatical formatives might seem to parallel Beard's (1995) separation between L-(exical) derivation and I-(nflectional) derivation. The roots of Distributed Morphology, however, are not lexemes: they are the un-analyzable core of complex morphological objects, radically underspecified for grammatical information — including categorial information. Harley and Noyer (1998) develop Marantz's (1997) insight and argue that the spell-out of syntactic nodes, that is the insertion of morphological material in place of abstract features, obeys different laws for root-nodes and other nodes: insertion is non-deterministic for the former (which they call L-nodes), because the grammar does not choose between alternative roots, but it is deterministic for the latter (or F-nodes), because the choice of the appropriate Vocabulary item is there driven by the Subset Principle (Halle 1997), a formulation of the Paninian principle to the effect that more richly specified forms take precedence over less richly specified ones.

Whatever the merits or demerits of this approach to morphology (which I personally find convincing on the whole), it certainly is coherent. Against this theoretical background, which I will argue neither for or against, an irregular pattern of Italian plural nouns assumes a deep significance: the data that I am about to discuss provide, in my opinion, strong support for the notion of lexeme as a morphological listeme. Specifically, I will argue that there exist listemes of morphological (not semantic) nature, which do not belong to any of the three lists depicted in (1) above. Because this conclusion emerges only thanks to the assumptions summarized in (1), it cannot be construed as an argument against the whole picture, but specifically against the idea that atomic morphemes (List 2) are the only morphological listemes.

2. Empirical challenge: Italian plural nouns in *-a* as "lexical" plurals

The nominal morphology of Italian, like that of Spanish (Harris 1991, Aronoff 1994), shows a clear distinction between gender and inflectional classes. Regardless of the principled relation between the two types of categories, nouns which are not invariable fall into one of four inflectional classes, defined by the word-final vowel in singular and plural:

(7)   *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'

This is not the place to discuss the various regularities and irregularities concealed by such a basic outline, especially in connection with the relation between each class and one of the two genders (for which issue see especially Dressler and Thornton 1996). Of direct relevance is the very large class *-o / -i*, which extends beyond nouns to adjectives

and pronouns, and which regularly realizes the masculine gender with the single exception of *mano, mani* 'hand, hands'. A small subset of this large class, numbering about a dozen items in common use (the exact number depends on individual usage), displays a very unusual property: in addition to the regular plural in *-i*, an alternative plural form is available ending in *-a*, which however triggers feminine (plural) agreement. Each of these nouns has therefore three forms, not two: one masculine singular ending in *-o*, one masculine plural ending in *-i*, and one feminine plural ending in *-a*. Unsurprisingly, this last form is marked, in more senses than one, and its meaning does not exactly coincide with that of the regular plural. Descriptive grammars traditionally refer to these as collective plurals, a label which seems intuitively justified by the following examples:

(8)   *braccio* 'arm'; *bracci* 'arms (of objects)'; *braccia* 'arms (parts of the body)'
        (also: *ciglia* 'eyelashes', *sopracciglia* 'eyebrows', *corna* 'horns', *dita* 'fingers', *ginocchia* 'knees', *ossa* 'bones')
        cervello 'brain'; *cervelli* 'plurality of organs/heads'; *cervella* 'brains (single organ)'
        (also: *budella* 'intestines')
        fondamento 'ground, base'; *fondamenti* 'grounds, bases'; *fondamenta* 'building foundations'
        (also, *lenzuola* 'bed sheets')
        membro 'member'; *membri* 'members' (plurality of individuals); *membra* 'limbs'
        muro 'wall'; *muri* 'walls (plurality of walls)'; *mura* 'walls (walled perimeter)'
        urlo 'shout'; *urli* 'shouts'; *urla* 'shouts'

The label of collective is certainly appropriate for terms like *fondamenta* 'building foudations' and *mura* 'walls (circle of walls)', and also with *cervella* 'brains' and *budella* 'intestines', which resemble mass nouns (they cannot be preceded by numerals). Oppositions like *lenzuoli / lenzuola* 'bed sheets' are much less clear, although many speakers (not all) consequently use the *-a* form to refer to a single bed-linen set (a pair of sheets) for one bed. In the case of body parts, paired or otherwise, the plural in *-a* by no means refers to the naturally given set: *dita* 'fingers', for example, refers to any non-singleton set, not just a ten-membered one, and the fingers need not be that of a single individual. As for the pair *urli / urla*, it is even less clear what role, if any, a collective interpretation may play. But the real problem for a collective interpretation of the ending *-a* is provided by a slightly different type of nouns, which lack the regular plural in *-i* altogether:

(9)   *paio* 'pair'; *paia*, *paii 'pairs'
        riso 'laughter'; *risa*, *risi 'peals of laughter'

        uovo 'egg'; *uova*, *uovi 'eggs'
        (also *centinaia* 'hundreds', *migliaia* 'thousands', *miglia* 'miles')

Although semantic factors are clearly at work (as evidenced by the nouns referring to measurement units), here there is simply no question of a collective interpretation for the forms in *-a*.

I have analysed this remarkable class in some detail elsewhere (Acquaviva 2002), and I will only summarise here the conclusions reached in that work. Semantically, the interpretive common denominator of all these nouns is their weak individuation: elements of the plural set are conceptualized as instances of un-differentiated or weakly differentiated tokens, without distinctive individual properties. This characterization is especially clear with units of measurement, but it also encompasses members of natural aggregates, quasi-mass nouns and collectives proper. The most relevant result, however, concerns the morphological, not semantic, status of the irregular plural forms. There are many reasons for rejecting the traditional notion of "marked inflectional class" (or irregular realization of plural), and for viewing this type of plural as belonging to word-

formation rather than inflection. In a nutshell, *braccia* is really an inherently plural lexeme derived from *braccio*. Its inherent number, its exceptional ending, its non-individualizing semantics are all facets of the same lexeme-forming process.

(10)   LEXEME A   LEXEME A'

         braccio, bracci (masc)   *braccia* (fem, pl, + idiosyncratic information)

No fewer than six arguments converge towards the conclusion that a form like *braccia* is a lexical, not inflectional, plural:[2]
– The form in -*a* fails to block the regular plural in -*i*. More precisely, the expected blocking obtains only for the few nouns in (9); the clear majority displays no blocking, as shown in (8), and even granting that these are *all* individual exceptions, the fact remains that blocking of the -*i* form is simply not a concomitant property of irregular plurals in -*a.* This (obvious) fact alone is sufficient to prove that *braccia* and *bracci* (which are not contextual variants) are not alternative realizations of the same morphosyntactic property [plural]. But the traditional label of "marked" plurals either amounts to precisely that claim, or else must be understood as something like (10).
– The ending -*a* is nowhere else an exponent of plural in Italian morphology. There is no independent evidence that this form encodes inflectional information at all. Notice that all the nouns in this class are non-suffixed, as there are no suffixes that share this irregular ending (e.g. *ditine* 'small fingers' FEM.PL., regularly ends in -*e*, as opposed to a hypothetical **ditina*).
– Gender is systematically an inherent property of nouns, which does not depend on the syntactic context. If *braccia* (fem.) was the plural of *braccio* (masc.), the switch in gender would be totally accidental, unconnected with any other property of this noun.
– Moreover, it is not strictly true that feminine and plural are concomitant properties of these nouns. In Italian, two coordinate singular NPs trigger plural agreement on the verb; when this involves a past participle and the auxiliary *essere*, gender becomes visible as well, and the question arises which gender the verb realizes if the NPs have conflicting values for this feature. If at least one of the conjuncts is masculine, agreement is (predictably) masculine:

(11)   [la mela e il formaggio] che sono stati mangiati /*state mangiate
            FEM.SG MASC.SG.   MASC.PL. / *FEM.PL
            'the apple and the cheese that have been eaten'

The point is, however, that the same masculine agreement is obligatory when a feminine NP like *la mela* is conjoined to *l'uovo* 'the egg', which is masculine as singular but feminine when plural (so much so that it does not even have a masculine plural at all; cf. (9) above). As (12) shows, the two singular NPs trigger plural number on the verb, but *masculine* gender:

(12)   [la mela e l'uovo] che sono stati mangiati/*state mangiate
            FEM.SG MASC.SG.   MASC.PL. /   *FEM.PL
            'the apple and the egg that have been eaten'

Indeed, even two nouns which are both feminine in the plural fail to trigger feminine agreement on the plural verb:

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

---

[2] All of these arguments also show the clear difference between Italian irregular plurals and Romanian "neuter" plurals (Mallinson, 1988, Farkas, 1990) which too change gender according to number, but behave differently in almost every other respect. Acquaviva (2002) contains some discussion on this contrast.

It really seems that the correlation between feminine and plural is not an inflectional property that may be transmitted by syntactic agreement: the complex subject in (13) is not feminine, although it is plural as a whole.

– The correspondence of plural and feminine also breaks down, and in a revealingly systematic way, when a singular pronoun refers back to a plural NP headed by one of the relevant nouns:

(14)　　　　le uova costano 50 centesimi l'una　　　　　/*l'uno
　　　　　　FEM.PL　　　　　　　　FEM.SG　/ *MASC.SG.
　　　　　　'the eggs cost 50 cents each'

Thanks to its distributive interpretation, a singular pronoun can refer to a plural NP in this example. As can be seen, what counts is semantic number, as opposed to grammatical: the feminine on the pronoun is made necessary by the plural predication (the predicate's truth conditions must be satisfied by a non-singleton set of eggs), even though *una* is morphosyntactically singular. When no such distributive interpretation is available, as in (15), a singular pronoun referring to *due uova* sounds decidedly odd, or downright unacceptable:

(15)　　　　volevo due uova, ma ne ho *?trovata una /*trovato uno
　　　　　　　FEM.PL　　　　　　　　　FEM.SG / *MASC.SG.
　　　　　　　'I wanted two eggs, but I found one'

If *uova* was just an irregular inflectional form, the connection between gender and semantic plurality would be totally accidental. If feminine gender is instead a morphological reflex of the same lexeme-creating process which turns a nominal base into a *plurale tantum*, the behaviour of gender an number in pronominal agreement becomes understandable.

– Finally, it is well known that plurals in *-a* differ somewhat in meaning from their regular counterparts in *-i* (when these are available), but in a way that cannot be traced back to a single interpretive property — certainly not "collectiveness". Nouns such as *mura* 'set of walls' have indeed a collective interpretation, but the identification and delimitation of the constituent elements of their referents is not clear (unlike, for instance, in *committee*). Nouns like *budella* 'intestines' likewise refer to objects conceptualized as complex aggregates without a non-arbitrary delimitation of constituent parts; they actually come close to mass nouns (and cannot be counted: *due budella*), except that they cannot be subdivided under preservation of their identity (while a spoonful of water remains water, a fragment of mucosa from the intestine is not *budella*). Nouns for natural aggregates (e.g. *braccia* 'arms', *corna* 'horns', *dita* 'fingers') are countable and consist of individuals, but their reference is not limited to naturally-occurring sets. As noted, plurals like *uova* 'eggs' or *paia* 'pairs' have no collective sense at all (only the singular is a collective concept, and that just for units of measurement). The *-a* ending, in sum, does not encode a special type of plural meaning. Rather, the special shades of plural meaning these forms undeniably have (except in a handful of cases) vary with the choice of the word; that is to say, on a lexical basis.

Mainly for these reasons, I have proposed (Acquaviva 2002) to remove these plurals from the inflectional system altogether, and to view them as inherently plural lexemes (like *scissors* or *cattle*) derived from bases that refer to weakly individualised inanimate concepts. The latter semantic property is a necessary, not sufficient, condition, and it would be wrong to see it as *the* semantic side of the lexeme-creating process described in (10). The interpretive value of these derivations is rather inherent plurality. This interacts with the meaning of the root to give rise to the "special" shades of plurality we have just considered, and with the morphosyntactic features of gender and number to give rise to grammatically [feminine, plural] nouns. I will now focus on this morphological aspect of the derivation.

3. What are "inherent properties" properties of?

(10) is an informal statement about the paradigmatic relation holding between three (types of) word-forms, which claims that only two of them (*braccio* and *bracci*) are variants of the same noun in two morphosyntactic contexts, [singular] and [plural] respectively. What of the third, *braccia*?

It should be clear from the preceding discussion that *braccia* is not a positional alternant of *bracci* in a different morphosyntactic context. The two plural forms are not automatically selected on a contextual basis, but are instead in free competition with each other, just like *cows* and *cattle*. This intuitively correct assumption must be made more precise: since there is no evidence to the contrary, I assume that the syntactic inputs realized by *bracci* and *braccia* only differ in the morphosyntactic feature of gender (masculine and feminine respectively), and in no other *grammatical* feature; specifically, no special information about plurality, collective or otherwise, is hidden in the syntactic structure that surrounds the noun. The point is important, because it is could be easily falsified (if, for example, one of the two forms were confined to clauses with a certain type of plural predicate[3]) and because it will presently shown to be crucial for the main argument.

Another assumption that must be made explicit is that the regular and the irregular plural forms invariably share one and the same root, at least as far as the form is concerned. No morphonological alternation, no minor difference of any kind justifies treating the root of *braccia* or *ossa* as different from that of *bracci/braccio* or *ossi/osso*. This applies to *all* plurals in *-a*: it is a systematic regularity that sharply distinguishes this kind of inherent plurals from formally unrelated pairs like *cow/cattle*. This systematic formal identity of the roots does not rule out, in principle, a distinction in terms of abstract information, or indeed morphological features like gender. It does however put the burden of proof on those who want to argue for a systematic doubling of homophonous roots, like *bracc-₁* [masc] and *bracc-₂* [fem].

The importance of these two assumptions becomes clear when the nouns in question are analyzed in Distributed-Morphological terms, along the guidelines of Marantz 1997, Harley and Noyer 1998, 1999. In this perspective, nouns are the morphological realization of a complex syntactic output, which involves a category-neutral root dominated by projections headed by grammatical heads. The best-studied case is provided by verbal nominalizations, roots which take their verbal character (loose terminology for thematic and aspectual properties) from a head "small v" and their nominal character from a higher Det head:

(16)　a　　$[_{vP}$ v [Root]]]
　　　　b　　$[_{DP}$ Det [Root]]]

Determiners, however, are really correlates of nouns, rather that constitutive elements. The insight underlying category-neutral roots can be sharpened by viewing nouns as roots embedded under a head [n], or "small n", locus of the grammatical information intrinsically associated with nominal character.

(17)　　　　　$([_{DP}$ Det) $[_{nP}$ n [Root]] (])

Following an assumption which has become widespread after Ritter 1993, 1995, I will regard the category of number as being external to the "lexical core" which for Ritter was just N but here is represented by the nexus [n[Root]]; and inside this core, I will identify [n] as the locus of gender, following Kihm's idea (2001) that to turn a root into a noun means giving it a place in the paradigmatic system of gender.[4] In sum, both regular

---

[3] See Winter (2002) for a recent new proposal on different types of plurals and their contextual restrictions.

[4] Gender may possibly be encoded under [Num] instead, as Ritter (1993) proposed for Romance languages; but as will become clear, what is crucial is just that gender be associated with a head distinct from the root itself.

braccio / bracci and irregular *braccia* have the structure [n [Root]], abstracting away from number.

Recall now that the irregular *braccia* has inherent idiosyncratic information, part of which finds morphological expression in the features [fem, pl]. The question that presents itself is: assuming [fem] and [pl] to be inherent properties, just what are they properties of? This question may be rephrased as "how can inherent features be encoded in a representation where category and root are dissociated?" One possibility, of course, is to argue that they should not be dissociated. But I believe this is not the right direction, although no arguments are presented here to that effect. My point is rather that *assuming* this dissociation to be correct (i.e. assuming category-neutral roots), an interesting conclusion follows about the necessity of lexemes as distinct from roots.

Both *bracci* and *braccia* are plural forms sharing, by assumption, the same root; also by assumption, both are realizations of a morphosyntactic complex [n[Root]], where only [n] is a grammatical head expressing morphosyntactic information. This much suffices to rule out one possibility: inherent [fem] and [pl] cannot be just properties of the root, because if (as assumed) this root is the same for *bracci* and *braccia*, there is then no way to prevent the features [fem, pl] from appearing whenever the root is instantiated; in short, there would be no masculine forms for these nouns at all. For a similar reason, the root cannot lexically select [fem] in the context of [pl], no matter whether it is inherently genderless (18a) or masculine (18b):

(18) a   [ ] →[fem] / _____ [pl] (restricted to certain roots: *bracc-*, *dit-*, *oss-* ...)
       b   [masc] →[fem] / _____ [pl] (restricted to certain roots: *bracc-*, *dit-*, *oss-* ...)

(18) makes [fem] a necessary correlate of plural number, thus incorrectly blocking the regular masculine plural *bracci*. In addition, a mechanical feature-rewriting operation like that has no connection whatsoever with the various special properties surveyed in the previous section. In short, if the root is indeed the same in the masculine *bracci* and in the feminine *braccia*, it cannot be where the inherent information of the latter finds expression.

There remains [n]. It is certainly conceivable that the same root may combine with different values for the features expressed on this head, in particular for gender: that is what happens for pairs like *zio / zia* 'uncle / aunt' or *cugino / cugina* 'cousin (male) / cousin (female)'. The present case is more complex, however, because the feminine alternant *braccia* is also (by hypothesis) inherently specified for plural number, a category available for the masculine form *bracci* as well. In the general case, i.e. for regular NPs, it is today generally accepted[5] that number is encoded under a separate head above N, which in the present analysis means just above [n]: *[Num [n [Root]]]*. Syntactic head-raising, or a morphological rearrangement of this input, then derives the linear order Root-n-Num. Recall that in Distributed Morphology, morphosyntactic inputs (possibly rearranged by Fusion, Fission or Merger; see Halle and Marantz 1993, Harley and Noyer 1999) are spelt out through Vocabulary insertion: for the root, this happens non-deterministically (one root is as good as another), while for everything else the grammar deterministically chooses the one Vocabulary item that spells out most features of the input. As we will now see, the relation holding in Italian between the abstract representation of gender and number and its exponence make it impossible to encode the exceptionality of *braccia* on [n].

In Italian, as in many inflecting languages, gender and number are systematically fused together (Ritter 1993, Kihm, 2001): a single final vowel regularly provides exponence for both categories (or for neither, in invariable nouns like *cinema* 'cinema / cinemas'). In Distributed-Morphological terms, this means that the syntactic heads [Num] and [n] systematically undergo Fusion, as opposed to Merger (whose output is a complex head, not an simplex one). The Vocabulary is then consulted to interpret a single fused head without internal structure, endowed with gender and number features.

---

[5] See Ritter (1991), (1993), Cinque (1994).

The final *-i* of *bracci*, therefore, spells out both masculine gender and plural number (as well as the diacritic which identifies the noun as belonging to the *-o / -i* inflectional class). All that can be said about the alternative plural *braccia* is that its *-a* ending spells out feminine and plural. But this line of reasoning leads to a dead end: *both* plural endings are exponents for gender and number, and there is no way to state the fact that one is the regular and productive, and the other is highly idiosyncratic. Indeed, pairs like *zio /zia* 'uncle / aunt' only differ in gender (and ending), and neither element displays any irregularity. By forcing *bracci / braccia* in the same category as these perfectly regular doublets, we would again imply that all of the peculiarities of *-a* plurals listed in the previous section are accidental.

What is crucial is that [n] fuses with Number in the regular, productive case. Because of this undeniable fact, it is irrelevant whether the head [n] in the irregular *braccia* contains only the feature [fem] or, exceptionally, [pl] as well: gender and number end up fused in any case before morphological realization, and for this reason the irregularity of *braccia* cannot be encoded by a [n] head marked [fem, pl]. This is what militates against an account like that schematized in (19)-(20):

(19)   a   *braccio, bracci* ←→ Root + n [masc.] (+Number [sg/pl])
        b   *braccia* ←→ Root + n [fem. pl.] (+Number [pl])

(20)   *-a* ←→ n [fem, pl] (_____ *bracc-*, *dit-*, *oss-* …)

Notice that in (19a-b) the features on [n] are just disjoint, not in a subset relation; therefore, no blocking is (correctly) predicted between a realization rule like (20) and one that spells out as *-i* a fused (n-Num) input marked [masc, pl]. For this reason, and also because it is plausible that the exceptionality if *-a* as exponent for plural should be related to the exceptionality of having number encoded on [n] instead of [Num], I believe that something like (19)-(20) is in fact correct. But it is only part of the explanation; a necessary but not sufficient characterization of the way grammatical information is represented on irregular plurals. As we have seen, there is much more: *-a* plurals are profoundly irregular in a systematic way. What is missing in (19)-(20), or in any other account that reduces the phenomenon of irregular plurals to an unusual distribution of features, is the connection between grammatical and non-grammatical information.

In a way, we don't want to posit two distinct roots for regular and irregular plurals, because then the semantic similarity and the formal identity would be accidental; but we cannot locate all the differences on a grammatical element either, because then we would lose the explanation for the inherent character of this type of plurals. What I would like to propose is that all the peculiarities of a form like *braccia*, crucially including the purely morphological ones, must be stated neither on the root nor on the nominalizing [n], but on them both.

Lexemes ≠ syntactic atoms ≠ roots ≠ semantic listemes

Plurals in *-a* are revealing because they are provide evidence that they are non-atomic compounds of root+[n], but at the same time they enforce the conclusion that their irregularity is not encoded on either of their constituent parts; this would seem to make them atomic wholes. The only way to escape the contradiction is to recognize that they are simplex in one sense and complex in another. They are, clearly, syntactically complex structures of two elements, and as such associated with two Vocabulary items (the spell-out of the root and the *-a*). But this (morpho)syntactic structure is also a single entity, not only as an idiosyncratically interpreted idiom, belonging to List 3 in (1), but specifically as a morphological object:

(21)   *braccio, bracci* : construction of a N by rule (root + [n])
        braccia : pre-packaged [root+[n]], with a specific choice for the features of n.

Like a phrasal idiom, *braccia* matches a syntactically complex input with a single semantic listeme; this is related to, but distinct from (for some roots sharply distinct from) the compositional interpretation of the root and of the surrounding morphosyntactic features. However, as we have seen, the complex structure [root+[n]] also has a purely morphological status, distinct from its interpretive value. The gender and number that are inherent to this encyclopaedic entry (in terms of (4)) are not properties of the root or of [n] alone; my proposal is that they are properties of both taken together. But these are morphological categories; and even though grammatical plural is directly related to semantic plural, the two are not the same. The peculiar property of being qualified by morphosyntactic features that are normally associated to single (or fused) heads makes of a plural like *braccia* a morphological listeme, but one different from the only listemes that Distributed Morphology recognizes within morphology proper (cf. (1) and (3)), namely the Vocabulary items. Since these are sound-meaning mappings, and since *braccia* is made up of two such Vocabulary items (root + ending), it follows that morphology in the strict sense must encompass listemes of a different nature. As we have seen, what makes *braccia* different from *bracci* has partly to do with grammatically encoded information (features), and partly with semantic and encyclopaedic information that depends on the meaning of the root itself. This, I believe, is an appropriate characterization for the notion of lexeme. (21) should therefore be read as a rendition of (10), a statement of grammar about lexemes cast in the language of Distributed Morphology, which does not include them.

Clearly, the conclusion I reached is not compatible with all the assumptions of this framework. If my argument is valid, then (5-6) cannot be right: roots are not the only sources of idiosyncrasy, because no one element is the sole depositary of idiosyncratic information in *braccia* (neither the root nor [n] individually, but both together). Note also that the concept of a listed morphological structure which does not block an alternative realization for each of its constituent parts (i.e. the root and [n]) sits unconfortably with the view that the realization of grammatical properties deterministically selects one output. While the choice between *destroy* and *destruction* is a choice between two distinct syntactic environments for the same root (Marantz 1997), that between *braccia* and *bracci* involves the same environment (___[n]); the gender value changes, but as we have seen this is not a choice between equivalent variants, nor between equally regular ones (like *zio* and *zia*). All idiosyncratic aspects, formal, semantic and morphological, are exclusive to one member of the pair, but this undeniable fact is not determined by grammar in any way. While not being a counterargument, this state of affairs certainly highlights the limits of a conception of morphology that plays down the distinction, extensively argued for by Pinker (1999), between listed and rule-constructed "words". Finally, and most clearly, if my argument is valid, the three dimensions of listedness into which Distributed Morphology has distributed the lexicon are not enough. In addition to syntactic atoms, Vocabulary items, and semantic listemes, there exist listed matches of meaning, exponence and morphosyntactic properties larger than category-free roots; as their defining properties are lexemic, I see no reason to call them in any other way than lexemes.

References

ACQUAVIVA P. (2002), "Il plurale in *-a* come derivazione lessicale", *Lingue e linguaggio* 2, 295-326.

CHOMSKY N. (1995), *The Minimalist Program*. Cambridge, MA: MIT Press.

CINQUE G. (1994), "On the Evidence for Partial N-Movement in the Romance DP", in *Paths Towards Universal Grammar: Studies in Honor of Richard S. Kayne*, Cinque, Guglielmo, Koster, Jan, Pollock, Jean-Yves and Zanuttini, Raffaella (eds.), Washington D.C.: Georgetown University Press, 85-110.

DRESSLER W. & THORNTON A. (1996), "Italian nominal inflection", *Wiener linguistische Gazette*, 57-59, 1-26.

FARKAS, Donka (1990), "Two cases of underspecification in morphology", *Linguistic Inquiry*, 21, 4, 539-550.

HALLE M. (1997), ""Distributed Morphology: Impoverishment and Fission", in *PF: Papers at the Interface*. MIT Working Papers in Linguistics 30, Bruening, Benjamin, Kang, Yoonjung and McGinnis, Martha (eds.), 425-449.

HALLE M. & MARANTZ A. (1993). "Distributed Morphology and the Pieces of Inflection", in *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Hale, Kenneth and Keyser, Samuel J. (eds.), 111-176. Cambridge, MA: MIT Press.

HARLEY H. & NOYER R. (1998), "Licensing in the Non-Lexicalist Lexicon: Nominalizations, Vocabulary Items and the Encyclopaedia", in *Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect* (MIT Working Papers in Linguistics 32), Harley, Heidi (ed.), 119-137.

HARLEY H. & NOYER R. (1999), "Distributed Morphology", *Glot International* 4, 3-9.

HARRIS J. (1991), "The Exponence of Gender in Spanish", *Linguistic Inquiry* 22, 27-62.

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 Guglielmo Cinque and Richard Kayne (eds.), *The Oxford Handbook of Comparative Syntax*, 2005, Oxford: Oxford University press, 459-512.

MALLINSON G. (1988), "Rumanian", in Harris, Martin and Vincent, Nigel (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", University of Pennsylvania Working Papers in Linguistics 4.2, 201-225.

PINKER S. (1999), *Words and Rules*, London: Weidenfeld and Nicolson.

RITTER E. (1991), "Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew", in *Perspectives on Phrase Structure* (*Syntax and Semantics* 25), Rothstein, Susan (ed.). New York: Academic Press, 37-62.

RITTER E. (1993), "Where's Gender?", *Linguistic Inquiry* 24, 795-803.

Winter Y. (2002), "Atoms and Sets: A Characterization of Semantic Number", *Linguistic Inquiry* 33, 493-505.