

Abdelkader Fassi Fehri  
Mohammed V University Souissi, Rabat

Marie-Thérèse Vinet  
Université de Sherbrooke, Canada

### **Distribution of Number and Classifier in Arabic and Chinese and parametrization**

It has been observed that Ls which exhibit Number (Nb) marking on Ns (tend to) lack classifiers (Cl), or vice versa (cf. e.g. Greenberg (1972)). This has led some authors to postulate a systematic complementarity distribution between Nb and Cl, and treat the two markings as discriminately complementary (cf. e.g. T'sou (1976), Chierchia (1998), Borer (2002)). Second, the 'count-mass' dichotomy has been thought in the literature to be the core classification, and been associated with a discriminative distribution of Cl and Nb, respectively, so that e.g. Chinese Ns are solely thought of as mass, whereas English Ns can be count (in addition to mass, count being the marked member). Third, (under)specification in Nb/Cl values plays an important role cross-linguistically, hence yielding complementarity effects, but also providing bases for identifying parametric variation (see e.g. Bresnan (2001)).

In this contribution, we provide evidence from Arabic and Chinese that a principled complementarity between Cl and Nb (or Pl, plural) cannot be motivated. We also show that N classification cannot be built only on a count/non-count (or mass) binary distinction. Such a binary classification obviates the role played by the 'singulative'/non-singulative (or 'individuating') dichotomy (in addition to collective) in typically characterizing Pl morphology in e.g. Chinese and other classifier Ls. It also obviates the role played by 'singulativity' in characterizing integral atomic units (or individuals), or integral kinds, as opposed to non-integral entities such as masses. Integrity and

parthood (or partitioning) will be shown to be the necessary relevant ingredients for defining a classificatory system based on two attribute-value feature pairs: [ $\alpha$  atomic] and [ $\beta$  singulative]. The system characterizes four N classes traditionally documented in Ls: individuals (I), kinds (K), masses (M), and collectives or groups (G). Is and Gs are both [+ atom], Ms and Gs [- sing], Ks and Is [+ sing], Ms and Ks [ $\emptyset$  atom], etc. Pl may mark G in Chinese, and K in English, but neither of the two in Arabic. Taxonomic Pls are found with Ks and Ms, and set Pls with Is and Gs in Arabic and English. The CI/PI (under)specification is 'top-down' oriented in English and (somewhat differently) in Arabic, but it is oriented 'bottom-up' in Chinese. Thus overt PI specification in English and Arabic may induce CI specification, without the need for the latter to be overt (hence the apparent complementarity). On the other hand, CI is overtly manifest in Chinese, but no PI is expressed (in the normal case).

The paper is organized as follows. In the first part, we examine the essential properties of various uses (and/or kinds) of CIs in Arabic and Chinese, as well as varieties of Nb (and PI) uses. We analyze some clear non-complementary distribution cases of CI and Nb, as well as some apparent complementary cases. In the second part, we investigate how and why the feature classificatory system must be rebuilt on atomicity and singulativity values, rather than the traditional 'count/mass' ontology. We examine the role played by unspecification in Nb and CI interaction, and how cross-linguistic variation in this interaction can be parametrized.

---

## 1. Kinds of classifiers and their distribution

Although the term 'classifier' is used quite extensively in the literature, it has no clearly established content, function, or form, nor has it a clear connection to categorization. Allan (1977: 13) takes CIs to denote "some salient perceived or imputed characteristics of the entity to which an associated noun refers". CI occurrence is claimed to be restricted to classifier constructions "which require the presence of a particular kind of morpheme, the choice of which is dictated by the semantic characteristics of the referent of the head of the noun phrase" (Aikhenvald 2000: 13). CI then qualifies as the morpho-syntactic mark (or realization) of the 'perceived' semantic characterization. CI forms often depend on categories (with which they are associated) so that (a) nominal, numeral, and D (article and deictic) CIs (including N class and gender), (b) genitive (possessive) CIs, and (c) verbal (predicate, aspectual) CIs can be distinguishingly identified (cf. Aikhenvald (2000: 9-13) for discussion). In this article we concentrate of some kinds of CIs found in the nominal domain, and investigate what could be their (more precise) content, function, and form, and what role is played by Nb (or Pl) in this phenomenology.

### 1.1. *The count/mass distinction: a misnomer*

An important property of Chinese is that all nouns, whether the counterparts of English count or mass Ns, need a CI or a measure phrase to be counted. Thus in Mandarin Chinese, counting individuals is mediated through the use of CIs, such as *zhi* in (1a) and (1b). Num(eral) such as *yi* (one) or *san* (three) precedes the CI N sequence. As illustrated by the

ungrammaticality of (1c), Num cannot be used without occurring with the mediating CI:

- (1) a. wo kanjian yi zhi gou  
 1s see a/one CI dog  
 I see a dog
- b. wo kanjian san zhi gou  
 1s see three CI dog  
 I see three dogs
- c. \* wo kanjian san gou  
 1s see three dog  
 I see three dogs

Note that CI can also occur with quantifiers such as *ji* "few", or *mei* "every", as indicated in (2):

- (2) a. ji ge pingguo  
 few CI apple  
 few apples
- b. mei ge ren  
 every CI man  
 every man

Furthermore, the CI can occur solely without the Num, as in (3):<sup>1</sup>

---

<sup>1</sup> CIs can also be preceded by a Dem(onstrative), with or without a numeral:

- (i) (zhei) san ben shu  
 (this) three CI book  
 (these) three books
- (ii) zhei ben shu  
 this CI book  
 this book

When the CI is necessary, then Q, Num, D co-occur with a CI.

- (3) wo xiang mai ben shu  
 1s want buy Cl book  
 I want to buy a book.

On the other hand, mass Ns can be counted through the use of measure or container words, or sorts, as exemplified in (4) and (5):

- (4) a. yi bei kafei  
 one Cl coffee  
 a cup of coffee  
 b. san ping jiu  
 three Cl wine  
 three bottles of wine  
 c. yi pian mianbao  
 one Cl bread  
 a slice of bread
- (5) ta kanjian san zhong you  
 he saw three sort oil  
 He saw three kinds of oil/ three oils.

These measure words are also taken to be Cls, and the postulated important function of both types of Cls is to *make Ns countable*.

Cls are often divided, however, into two groups. Cheng & Sybesma (1999; = CS), inspired by the references cited there, claim that (a) Cls either "create a unit of measure" (they are then called *mass-classifiers* or *massifiers*), or (b) Cls simply 'name the unit' in which the entity denoted by N naturally occurs (they are then *count-classifiers*). CS take it that Chinese makes a distinction between Ns that are inherently (lexically) partitioned into countable units

(count Ns) and those that are not inherently partitioned (mass Ns). The role of the mass-CI is then to 'create' units by which the amount of the mass is measured/counted. Count-classifiers, on the other hand, are associated with already partitioned (countable) units. Following Croft (1994: 162), CS (p. 516) appeal to an 'individualizing function' of count-CIs, which enables them to "extract ... distinguished, ... discrete occurrences" (Iljic (1994: 104)), or to be "a mark of individuation, or singularization" Paris (1981: 69).<sup>2</sup> The count/mass distinction, they claim, although inherently present in Ns, is grammatically reflected only at the CI level (in addition to the role it plays in other parts of Chinese grammar).<sup>3</sup>

Note, however, that the 'counting' function which unifies count-CIs and mass-CIs is still puzzling and obscure. If massifiers operate 'partition', as it is

<sup>2</sup> CS misleadingly equate this singularization function with that of D, which identifies one instance of what is denoted by N (p. 517).

<sup>3</sup> CS, as well as Paris (1981), observe that 'count-classifiers' and 'mass-classifiers' are distinct in two important ways, namely the optional occurrence of *de* (a marker of phrasal boundary inside the DP) with 'mass-classifiers' versus its absence with 'count-classifiers' (cf. (i)) and the modification of the 'mass-classifier' head with a small number of adjectives (*xiao* (small), *da* (big)) versus the lack of adjectival modification on 'count-classifier' heads (cf. (iib)) :

- (i) a. *yi wei laoshi*  
       one CI teacher  
       a teacher  
    b. *\*yi wei de laoshi* (Paris 1981)  
    c. *yi kuang (de) pingguo* (Paris 1981)  
       one CI-basket DE apples  
       a basket of apples  
 (ii) a. *yi da zhang zhi* (Cheng and Sybesma 1999 :516)  
       one big CI-sheet paper  
       one large sheet of paper  
    b. *\*yi da zhi gou*  
       one big CI dog

claimed, why already partitioned Ns are not countable? Worse, why are Ns designated as 'count' in the first place? If partitioning is a distinguishing property between two classes of Ns, as CS argue, it does not appear to be sufficient for counting, hence the necessary mediation of some Cl. Partitioning is then *not* counting. Following Fassi Fehri (2003b), we take CS partitioning to be *singulativity*. Partitioned Ns are then marked as [+ sing].

Furthermore, as CS (p. 515) claim, massifiers occur not only with 'mass' Ns, but also with 'count' Ns. The examples in (6) illustrate this distribution:

(6) a. yi qun ren

one Cl person

a group of persons

b. yi jin pingguo

one pound apple

a pound of apples

If the role of mass-Cls is to 'partition', as they claim, how can they be associated with already 'partitioned' Ns (so-called 'count' Ns)? If even inherently partitioned Ns are in need of Cls to be counted (be they interpreted as individual units, or as parts of potentially partitioned entities), then the 'counting' function (whatever that is) has to be associated with both types of Cls. It is a property of neither 'count Ns' or 'count Cls'. Likewise, the individualizing/singularizing function which is supposed to be behind the countability of so-called 'count' Ns remains obscure at best, if not completely misleading. Indeed, if complementary distribution between the two types of Cls is discriminatory, as is documented in the literature, and exemplified by (7), then that could be an indication that the two Cls share a common identity, yet

to be identified, which enables both to 'create' (rather than 'name' in CS informal terms) the entities to be counted (be they individuals or measures):

(7) a. \*si liang ge qiche

four Cl Cl car

four cars

b. \*yi bei ge kafei

one cup Cl coffee

a cup of coffee

Cls then serve to count individuals, to measure a portion of a substance, a portion of a collective, or a part of a partitive, as in the following examples, respectively:

(8) a. yi ge pingzi

one Cl bottle

b. yi ping jiu

one Cl/bottle wine

a bottle of wine

c. yi gua zhuzi

one Cl pearl

a string of pearls

d. yi di yanlei

one Cl/drop tear

a tear

Their diversity in categorizing objects according to shape, animacy, size, or other perceptual or physical dimensions does not obviate their main role in



---

counting (and/or quantifying; cf. e.g. Paris (1981), Krifka (1995), Cheng and Sybesma (1999), among others).<sup>4</sup>

### 1.2. A new classification of classifiers

If both types of CIs are adding some information to that carried by (bare) Ns which relates to countability, then the distinction of CIs (examined so far), according to whether they are count-CIs or mass-CIs (or even worse massifiers) cannot be adequate. The classification needed has to take into account the type of information contributed by the CI, but also by the N it applies to. If some Chinese bare Ns are interpreted as either singular or plural, and other not (as it is the case), then that point to an inherently based distinction of Ns, according to whether they are potentially 'partitionable' or not, basically [ $\pm$  partition], mistakenly taken to be equivalent to 'count' vs. 'mass'.<sup>5</sup> We identify these Ns that can be potentially partitioned and be interpreted as Pl or Sg as *kinds* (Ks), as proposed in Fassi Fehri (2003b). Ks  $\emptyset \pm$  denote potentially an unspecified number of integral wholes or singularities, and can be either Sg or Pl. They have the property of being 'singulative', as construed there. The unmarked N with respect to singulativity is thought of as being *mass* M. M has no semantic potential of denoting singulativities (or integral

---

<sup>4</sup> Paris (1981:73) mentions that in the spoken language the 'neuter' classifier *ge* is coming to be used more and more at the expense of other classifiers. Even though manuals indicate '*yi ben shu*' (a book), students from Beijing University, she observes, will rather use the form '*yi ge shu*' (a book). This distribution supports the hypothesis that classifiers are more connected to quantity than to quality or categorization.

<sup>5</sup> This view annihilates Chierchia's (1998) proposal that all Chinese Ns are mass (for similar ideas, see also Krifka (1995) and Borer (2002), among others).

partitions). The contrast at the N level is not then between 'count' and 'mass', but rather between K and M. Both K and M need a 'count' CI to be counted.

What the 'counting' CI does with singulative Ns (or integral partitions) is to make precise whether there is only one integral unit (to be counted), a single *atom* named an *individual* I, or unspecified/non-atomic number of units, a kind K. Both realizations of CIs associated with inherently built singularities are found (in Arabic and English, respectively). They apply to positively marked singulativities (at the N level), and make them either *positively atomic* (I-CI), or *unspecified* for atomicity (K-CI). Depending on the marking (as I or K), the partitioned entity will turn out to be 'countable' or not. These two choices of atomization values associated with CI marks are exemplified in Arabic and English respectively:

- (9) a. samak "fish; kind of fish" → samak-at "fish-I; an individual fish"  
 b. apple → apple-s "apple-K"

In these examples, specified atomization as [+ atom] through the I-CI (or 'individualizer') is associated with an inherent singularity (or an integral 'partition') in the case of Arabic, whereas it is associated with unspecified atomic CI (a K-CI, or a 'kindifier') in English (applying to inherent singularities as well). The Chinese singulativity appears close to the Arabic one, in that no (grammatical) mark is needed to form K, which is the unmarked case. I-CI are then needed to form Is, as in Chinese (3) or (1a) above, much more like what happens with Arabic (6a).

Note now that M-CIs (the 'true' massifiers) are found in Ls. Their role is to remove (or unspecify) the inherently positive singulative value found in K (its minimal counterpart), hence converting (a 'lexical') K to (a 'grammatical') M.

This is illustrated in Colloquial Moroccan Arabic (MA), where M can be overtly (and grammatically) derived from K:

(10) bğer "cattle, cows" → bəgr-i "cow-M; beef"

Other potential CIs to be found in this system are 'collective' or 'group' CIs (G-CIs), which are typically documented in Chinese.<sup>6</sup>

CIs which have been examined so far select nominal entities which involve the singulative and atomic features and rearrange/specify their value. The outcomes may or may not be countable as integral atoms. Only if the CI is [+atom], as in the case of I-CI (or G-CI) can it be countable. K-CI and M-CI do not output countable units. Call the successful counting CI A-CI [+atom] CI. CIs which do not output countable entities (Ks and Ms) are non-A-CI (i.e. [- or Øatom] CIs. Since atomicity appears not to be relevant for the opposition of the pair, we might better term them S-CI (singulative CIs) for convenience. Ns can be classified inherently as A-Ns or non-A-Ns or S-Ns, just like CIs are. S-Ns, namely Ms or Ks, can be made countable through CIs of different nature and range. These CIs count 'sorts', 'taxonomies', 'portions' of substances, etc. They tend to be found cross-linguistically, unlike A-CIs and S-CIs. They are phrasal (and lexical), and they can be easily translated from one language to another:

<sup>6</sup> Iljic (2001: 13) indicates that *-men* cannot be considered a plural suffix since it does not enter into a systematic opposition with the singular (see section 2.2 below) :

(i) \*jiu ge xueshengmen  
 nine CI student-men  
 nine students

(11) a. *ʔalaat-u ruʔuus-i ɡanam-in*

three heads-gen sheep-gen

three heads of sheep

b. *ʔalaat-u qita<sup>c</sup>-i xaʃab*

three pieces wood

three pieces of wood

To distinguish these counting CI from A-CI, let us designate the atoms they create for counting as Ts (for taxonomies, portions, measures), and the CIs as T-CI. We think that both A-CI and T-CI, unlike S-CI, create atoms or units of some sort, although presumably of different ontological nature, and both outputs of these CIs can be counted. The co-occurrence of A-CI and T-CI is then expected to be excluded, given that both operate the relevant atomization needed for counting, as illustrated in Chinese (7) above, where CI conflict arises, compared to English (12), where no such a conflict arises:

(12) three kinds of apples

Differences between our A-CI and T-CI have been amply noted in the literature, although misleadingly attributed to the 'count-mass' CI distinction. Thus CS (1999, 1998) note that even though 'count-CIs' and 'mass-CIs' are both nominals, only 'count-CIs' form a closed set, with elements functioning solely as CIs (e.g. *tiao* for counting long, flexible and narrow objects such as rivers, towels, trousers, streets, etc., *duo* for counting flowers, *liang* for counting different types of vehicles, etc.):

(13) a. *yi tiao he*

one CI river

a river

- b. san duo hua  
 three Cl flower  
 three flowers
- c. si liang qiche  
 four Cl car  
 four cars

'Mass-CIs', however, can occur as CIs as well as independent Ns (e.g. *ping* "bottle" as a unit for measuring wine, and *ping* "a bottle" as a container):

- (14) a. yi bei kafei  
 one Cl coffee  
 a cup of coffee
- b. san ping jiu  
 three Cl wine  
 three bottles of wine
- c. yi pian mianbao  
 one Cl bead  
 a slice of bread

It has been noticed that contrary to 'count-CIs', 'mass-CIs' do not 'categorize'. They provide a unit of measure for the entities they are associated with. This explains why 'count-CIs' cannot usually be translated easily into English or other languages, whereas 'mass-CIs' can be. It is also observed that 'mass-CIs' are found in most languages, whereas much fewer languages have 'count-CIs' (so-called 'classifier languages').

To summarize, we have shown that at least three types of CIs with distinct behaviours and distributions can be identified: A-CIs which form atomic entities, S-CIs which are not atomic, but are either [ $\pm$  sing], and T-CIs which

form taxonomies (or measures). Of the three types, both A-CI and T-CI outputs are countable. The counterpart to M is then K (rather than 'count'). K is 'partitioned' in the sense that it denotes 'singulativities', whereas M does not. Furthermore, 'partitioning' or 'singulativity' has been shown to be *not sufficient* for counting, contrary to the wide spread belief.<sup>7</sup>

## 2. Kinds of Number

Nb/PI marking interacts with CIs in interesting ways, as widely noticed in the literature, but forms of this interaction need to be more precisely characterized. This subsection is dedicated to identifying various kinds of PI (Nb) uses typically in Arabic and Mandarin Chinese, and to assess naturally expected kinds of interaction. One use of PI might be termed a *multiplier*. It multiplies individuals or sorts/taxonomies, in the sense that it generates a referential expression from the NP/CIP, by mapping the nominal expression onto a *multitude* of its realizations. It applies equally to A-CIs (or A-Ns), and to T-CIs, hence forming A-PI and T-PI, respectively. A second use of PI is found in the context of so-called the 'plural of abundance' or 'the plural of the plural' (or 'double plural'), well-documented in traditional grammars. The main feature of this kind of PI is that it can apply to Ms or Ks without operating a *nominal reference shift* (e.g. from 'substance' to 'sort'). It can be seen as operating on an already formed referential expression (or nominal terms), and having a modificational role of emphasizing the referent's quantity or its high amount. It can be thought of as a mass Nb marker (Wiese (1999)). Call it *M-PI*. What is typical of M-PI is that it does not apply to countable units, but to amounts, Ks or

---

<sup>7</sup> In fact, it is *not necessary* either, if collectives of Gs are non-singulative, as we propose later on.

Ms (or the class S, for which the A feature is irrelevant). Finally, a third type of PI is best thought of as a *Cl*, a K-Cl or a G-Cl, as we will argue.

### 2.1. Arabic

Consider first some instances of PI types in Arabic. The following examples instantiate a multiplier A-PI (with Is and Gs) in (15) and (16), and a multiplier T-PI (with Ms and Ks) in (17) and (18), respectively:

(15) samak-at "fish-unit; a fish" → samak-aat "fish-unit-pl; a multitude of integral fish units"

(16) firqat "team" → firaq "teams"

(17) zayt "oil" → zuyuut "oils; sorts of oil"

(18) samak "fish" → ?asmaak "sorts of fish; a multitude of fish"

In all these cases, we take the PI to be a multiplier, inputting an atomic individual or sort, and outputting a multitude of these entities.<sup>8</sup>

Let us examine now the diversity of PIs found in Arabic (Standard Arabic) and Colloquial Moroccan Arabic (MA), which manifests the realization of M-PIs. In Arabic (19c & d), the 'double PI' functions as M-PI, whereas other PIs are multipliers:

---

<sup>8</sup> Borer (2002) takes the PI in examples like e.g. (14) to be a 'divider' of M (the equivalent to a 'partitioner' in CS terms). This then led her to think that the PI there is a *Cl* which creates countable units, a view that she generalizes to all PI cases. But this view cannot be maintained given that e.g. the Arabic Sg *zayt* 'oil' can have the reading of "one sort of oil", rather than "oil" (see Fassi Fehri (2003b) for discussion). Consequently, "oils" is the PI of "an oil" rather than of "oil" in English, which then makes it a multiplier (of sorts), and not a nominal reference shifter from M to non-M (or A).

- (19) a. qawl "saying" → ʔaqwaal "sayings"  
 b. qawl-at "saying-I; a saying" → qawl-aat "saying-I-PI; a multitude of sayings, utterances"  
 c. ʔaqwaal "sayings" → ʔaqaawiiil "a lot of sayings"  
 d. farq "difference" → furuuq "differences" → furuuq-aat "a lot of differences/ various sorts of differences"

Note that the M-PI in (19d) may have a reading of 'double plural', i.e. PI of sorts of differences, where differences has to be itself PI. In MA, similar patterns are found, interacting also with CIs:

- (20) a. bettix "melon K" → bettix-at "melon-I; an integral piece of melon"  
 b. bettix-at "a melon" → bettix-aat "(many integral pieces of) melons"  
 c. bettix "melon" → btatex "a lot of melons (many sorts of melon)"  
 d. šḥam "fat" → šaḥm-ah "fat-I; a piece of fat" → šaḥm-aat "(many) pieces of fat"  
 e. šḥam "fat" → šḥum "sorts of fat; a lot of fat" → šḥum-aat "a lot of fat; many sorts of fat"

It is to be noted that as far as we can tell, the PI is never used in Arabic as a CI forming a K, as in the case of English 'apples', where the PI is not necessarily read as a multiplier. This latter PI translates into Arabic by a Sg K 'tuffaaḥ' ("apple"), rather than by any form of the PI types provided above.

In addition to the ample literature found in Arabic traditional grammars on these PI distinctions (although they are not identified, organized, or characterized the way we have done it here), contrasts are reported by Wiese (ibid), pointing to the existence of a multiplier/M PI distinction in Persian and Chinese, as illustrated in (21) and (22), respectively:



- (21) a. ab water "water"  
 b. ab-ha water-pl "plenty of water"  
 c. ab-i water-sg "some water"
- (22) a. háizi child "a child; children"  
 b. háizi-men child-pl "several / a lot of children"

These various types of PI indicate clearly that the functions of PI are far from being limited to classification, and hence the interactions between PI/Nb and CI have to be more carefully investigated.

## 2.2. Chinese

PI marking in Chinese does not appear to be as inflectionally productive as it is in e.g. Arabic, English, or French. Thus a bare N may express Nb without any morphological mark dedicated to it, and hence be interpreted as Sg or Pl, as already observed in (22a). Moreover, K Ns in Chinese are also expressed through a bare N. Note further that (in)definiteness is not marked either on the bare N. These distributions are summarized in (23):

- (23) a. wo kanjian gou le  
 1s see dog Asp  
 I see (a/the) dog(s)
- b. wo xihuan gou  
 1s love dog  
 I love dogs

However, expressed Pls/Nbs are found, and they carry additional closely connected features. For example, the suffix -men occurs on pronominal forms and on certain nouns, mostly those with animacy and humanness, as in (24).

Iljic (1994, 2001) argues that this suffix is a 'collective' marker, rather than a normal plural:

(24) a. wo-men

1s-MEN

we/us

b. faoshi-men

professor-MEN

professors

c. \*yizi-men

chair-MEN

chairs

The marker (yi)xie, on the other hand, expresses indefinite plurality, as in (25):

(25) yi xie shu

one XIE book

a few books

Third, classifiers may reduplicate to express a universal quantification, as illustrated in (26):

(26) ta ge ge xuesheng dou rende

3s Cl Cl student all know

(S)he knows all the students

Let us discuss in more detail the role of these markers. First, the suffix *-men* on Ns has been analyzed as either a PI (Li and Thompson 1981, Li 1999), or a collective (Iljic 1994, 2001, Cheng & Sybesma 1999, Norman 1988, Chao 1968, Lü 1947). But although *-men* implies plurality, it specifies something more. It indicates that the members of a particular group should be considered together as a unit, and it "... marks a subjective location: several individuals

are grouped together relative to the speaker or some other subjective origin" Ilijic (1994: 91). In this case, it is best treated as *collective* or a 'grouper' G, referring to a whole.<sup>9</sup> The use of collective forms with *-men* concerns the way of viewing the members of a group.<sup>10</sup> As mentioned by Ilijic (2001), *-men* is regularly attached to pronouns, and only sporadically to Ns. But its optional use with Ns is not random. When *-men* is used in narrative contexts, it marks a subjective group, and it bears an affective interpretation, limited to humanness or interpersonal relation words expressions (terms of address, kinship, etc.).

Another argument demonstrating that *-men* is a collective, rather than a regular PI, is the fact that it can suffix to proper Ns, as in (27), taken from (Ilijic 1994):

(27) Xiao Qiangmen

Xiao Qiang-MEN

Xiao Qiang's group

The suffix then serves to identify a group relative to a certain person.<sup>11</sup> This type of collective reading on proper nouns is far from unique. It can also be found in other languages, namely Tok Pisin (cf. Mühlhäusler 1981: 43) and

---

<sup>9</sup> Norman (1988: 121) explains that this suffix was originally a compound nominal meaning 'every person'.

<sup>10</sup> The term *collective* has been used in the literature in a variety of ways (Corbett 2000: 117-120, Kemmer 1993: 92-94, Greenberg 1972: 19-25) and this situation has generated a lot of confusion. The internal plural value of collective markers is usually left imprecise, and the term acquires a different content, depending on the language described. Typically, there is confusion between K and G, as we construe them. In our conception, only G qualifies as 'collective'.

<sup>11</sup> Ilijic (2001) indicates that this is the use of well-read speakers of Chinese.

Berbice Dutch (Kouwenberg 1994: 238).<sup>12</sup> Corbett (2000:102) discusses associative plural forms added to proper nouns, nouns for kin terms, titles and occupations in Hungarian.

A further argument for the G nature of *-men* is that it is incompatible with simultaneous counting, as demonstrated in (28):

- (28) a. san ge xuesheng  
           three CI students  
       b.\*san ge xuenshengmen  
           three CI students-MEN

This situation can be explained if the suffix is taken to be forming a *whole*, which is perceived as a *singleton* or an atom, whereas Num selects a *multitude* of atoms. G formation is not a multiplying function, but rather a unifying function, which is incompatible with counting. Such a situation recalls that of Pls forming Ks (as in English generics). The following examples in Chinese may be cases of K Pls which exclude counting:<sup>13</sup>

---

<sup>12</sup> In Tok Pisin, the plural form can appear with the form *ol* (from *al/* preceding N, as in (ia). The associative or 'group' plural in this grammar is given through the inversion of the N with *ol*, as in (ib):

- (i) a. ol man  
           3p man  
           men  
       b. pater ol  
           priest 3p  
           the priest and his flock

In Berbice Dutch, Kouwenberg (1994: 238) notes that "the plural suffix *-apu* may also be used to mark an associative plural, as in *eke papapu* (my father and people of his generation)".

<sup>13</sup> It is worth comparing this state of fact compared to the incompatibility of numerals with a mass noun like *hair*, usually taken as an unmarked collective form in English. As can be observed below, only the sort reading (pieces/kinds) is available for the plural (ic), or the distributive (id):

- (29) a. che-liang  
       car-CI  
       vehicles
- b. \* san che-liang  
       three vehicles
- c. shu-ben (shuben)  
       book-CI  
       books
- d. \* wu shu-ben  
       five books
- e. xin-jian (xinjian)  
       letter-CI  
       correspondence (letters)
- f. \* shi xin-jian  
       ten correspondence

K or collective nouns in Chinese are generally formed by attaching a CI as a suffix to the related N, as illustrated above, and this form is incompatible with a preceding numeral. It appears then that 'grouping' and 'partitioning' are two notions which are in contradiction in grammar. The suffix *-men* should therefore be treated as a *classifier*, a G-CI, not a multiplier, like a normal plural. It then parallels the K PI in English, or the K PI in the examples (29).<sup>14</sup>

- 
- (i) a. \* I have seen three hair(s) on the floor.  
       b. \* They have three fine heads of hair.  
       c. I have seen three pieces/kinds of hair on the floor.  
       d. They each had a fine head of hair.

<sup>14</sup> However, as shown in (ib), numerals do appear with nouns suffixed with *-men* when these nouns are constituents in apposition, as frequently observed with pronominal forms in (ia). The examples are from Zhang & Sang (1986: 492), after Ijic (2001: 24) :

The suffix *-men* can co-occur with some CIs, such as *qun* "group", which has a collective G reading, but not a K CI, such as *zhong* and *lei*.<sup>15</sup>

- (30) a. *zhei qun haizi-men*  
 dem CI child-MEN  
 this group of children (Yu 1957: 3, after Iljic 2001:21)
- b. \* *zhei zhong haizi-men*  
 dem type child-MEN  
 this type of children
- c. \* *yi lei laoshimen* (Iljic 2001: 21)  
 one CI teachers-MEN

- 
- (i) a. *ta-men si ge*  
 them four CI  
 the four of them
- b. *zhe jia you gemenr san ge*  
 this family have brother+MEN+R three CI  
 In the family, there are three brothers (, cited in Iljic 2001).

<sup>15</sup> Another property of *-men* is that it is always interpreted as a specific definite, and never as a generic, hence the contrasts in (i) and (ii) (Rygaloff 1973):

- (i) a. \* *you ren-men*  
 have man-MEN  
 There are men
- b. *you ren*  
 have man  
 There is (are) somebody (some people).
- (ii) a. \* *ta-men shi laoshimen*  
 they-MEN be teachers  
 They are teachers
- b. *ta-men shi laoshi*  
 they-MEN be teachers  
 They are teachers.

This definiteness contrasts with that of definite generics in English as in *Dogs bark*, in that *-men* is necessarily specific, as already mentioned.

This may suggest that *-men* can form Ks, in addition to Gs. But as Ilic (2001: 22) notes "*-men* does not refer to types or groups established on a purely qualitative basis (abstract classes), but to concrete groups of people, delimited in specific situations", and it crucially "pertains to the grammatical category of person" (Ilic 1994 :91).

The other marker of plurality in Chinese is *yi(xie)*, for indefinite plurality, as already mentioned. *Xie* cannot be preceded by any other numeral but *yi* (31):

- (31) a. *yi xie shu*  
           one XIE book  
           books
- b. \* *san xie shu*  
           three XIE book

The situation recalls that of *waahəd* 'one, some' in MA, which is in fact a quantifier more than a Num, a sort of 'some', which contributes indefiniteness.<sup>16</sup> *Yi* can then be seen as contributing indefiniteness, and *xie* plurality. Support for this treatment comes from the fact that it is possible to express definiteness with *-xie* by using the demonstrative *zhe* (Wu 1999); the outcome is still interpreted as plural:

- (32) *zhexie xuesheng dou xihuan wo*  
           these student all like me  
           All of these students like me

<sup>16</sup> In e.g. (i), *waahəd* contributes indefiniteness, even though it is used with a definite article:

(i) *waahəd r-rajel ja*  
       some the-man came  
       Some man came.

*Yi-xie* does not categorize semantically any particular set of Ns, and hence does not appear to be a CI. Furthermore, it can co-occur with other CIs, e.g. *ge*, in (33a). With the complex form *haoxie* "a good number of", various CIs are possible even though they may be left out (cf. 33b, c; examples from Iljic (1994: 101)):

- (33) a. You xie ge ren bu hui shuo hua [please align glosses adequately]  
 Have XIE CI man NEG can/know speak word  
 There are some people who cannot speak (very well)
- b. haoxie (jian) fangzi  
 a good few CI house/room  
 a good few rooms
- c. haoxie (ben) shu  
 a good few CI book  
 quite a few books

But Iljic (*ibid*) treats *yi-xie* as collective too, although it appears to be a regular indefinite PI. If that were so, then the PI markers examined so far qualify only as CIs. The question is then why they cannot be used as either Mul-PI (multiplier-PI) or Mass-PI.

A further form of plurality arises through CI reduplication. In Chinese, CI reduplication induces universal quantification, which can be viewed as a form of plurality. As is well-known, universal quantification is often used as plurality expression (cf. English *you* vs. *you all*; Melanesian Pidgin English *all man*, *all-a man* "the men" (Hall 1943)). Thus CI reduplication can yield a universal reading in (34), or a distributive reading in (35). Paris (1981: 70) points out that the reading is different if the reduplicated classifier is preceded by *yi* "one".



Only in the presence of *yi*, as in (35), can the reading be iterative or distributive:

(34) ta ge ge xuesheng dou rende

3s CI CI student all know

(S)he knows all the students

(35) ... Kancha-le yi tiao tiao xin hangxian (Paris 1981)

...prospect Asp one CI CI new maritime line

have prospected one after the other

Note that the universal reading in (34) is ruled out with an indefinite. In a study on *dou* quantification, Wu (1999: 113) explains that the role of *dou* is "to distribute the property of a predicate over an NP", not to measure. An indefinite cannot be quantified by *dou*:

(36) \* yi ben shu Zhangsan dou du-le

one CI book Z all read Asp

*Dou* is unable to quantify a weak NP. Only strong NPs (in Barwise & Cooper's (1981) sense) can be quantified by *dou*, a universal quantifier in (37a), or a definite in (37b):<sup>17</sup>

<sup>17</sup> It is to be noted that *dou* expresses an 'individualized property of the NP, as in (i), or it rather signifies 'a totality without exception', as in (ii) (Huijun Zhou, p.c.):

(i) a. ben ben shu dou chucuo

CI CI book DOU make mistake

There are mistakes in every book.

b. ta men men gongke dou de you

3s CI CI course DOU get excellent

He got an excellent mark in every course

c. jian jian shangyi dou you dong

CI CI garment DOU have hole

There is a hole in every article of clothing.

- 
- (37) a. mei-ge xuesheng dou chuxi-le huiyi  
       every student all attend-Asp meeting  
       Every student has (all) attended the meeting.
- b. zhexie xuesheng dou chuxi-le huiyi  
       these student all attend-Asp meeting  
       These students have all attended the meeting

### 3. Non-complementarity of Classifier and Number

It has been observed by a number of scholars that languages with overtly express nominal CIs tend to lack a Nb category (cf. e.g. Greenberg (1972)). The question is then whether CI and Nb are mutually exclusive, or in exclusive complementary distribution ECD, and hence constitute two manifestations of the same parameter, or ECD is only apparent, if not illusory. Tsou (1976), for example, claims that nominal CIs and Pl morphemes in natural languages are in complementary distribution. In the same vein, Borer (2002) adopts the complementary distribution view: "Not only is it the case that classifier languages do not (appear to) have plural inflection, but languages which mark plural do not appear to have classifier inflection. In other words, it would appear that classifier inflection and plural inflection are in complementary distribution". CI and Pl inflection are then supposed to play a parallel role in the grammar (cf. also Chierchia (1998)). We will show that there are firm bases to

- 
- (ii) a. tuan tuan huoyan ba wo weizhu  
       CI CI fire BA 1s surround  
       All the fire surrounded me
- b. duo duo xianhua xiang yang kai  
       CI CI flower flower towards sun  
       All the flowers came into bloom towards the sun

think that the ECD view is both empirically illusory, and conceptually inadequate.

A first clear case of the inadequacy of ECD is provided by Arabic cases where CI and PI co-occur. This is exemplified by Arabic (19b) and MA (20b & d) above. In this kind of examples, the PI is a plural of individuals, an I-PI. K undergoes an overt morphological derivation to form I, and PI applies to the latter to form the I-PI. The following Arabic example instantiates the two derivational processes which occur together:

(38) tamr "date K" → tamr-at "an individual date" → tamr-aat "a set of individual dates"

Here the suffix -at forms the individual unit from the general kind, and the I-PI is formed through internal vowel lengthening (-aa basically).

That CI and PI are not incompatible is also supported by the fact that the K-PI is available as well; it then takes the form of a 'broken' PI:

(39) tamr "date K" → tumuur "kinds/sorts of date"

As pointed out in Fassi Fehri (2003 b), PI formation in these cases is classifier/ontology preserving, in the sense that it does not operate, for example, a classification/reference shift, from a kind/sort to a unit. It is rather a multiplier of 'already' available units, which are of different nature. They then lead to different PI interpretation: I-PI and T-PI, respectively.

A second sort of problematic cases for ECD are cases where CIs co-occur in the same construction. That is, a CI construction is not necessarily limited to one CI. Various instances of this case are found. For example, the collective

marker *qun* can co-occur with the collective marker *-men* in Chinese, as in (30a) above, repeated here as (40), for convenience: \_

- (40) *zhei qun haizi-men*  
 dem CI child-MEN  
 this group of children

Likewise, CI may reduplicate, as observed with respect to (26) or (34) above. The following example from Cheng & Sybesma (1999) instantiates this phenomenon:

- (41) *tiao tiao daolu tong Beijing*  
 CI CI road connect Beijing  
 All roads lead to Beijing.

If CIs themselves are not mutually exclusive in the same construction (except when they are semantically incompatible), then there is no reason to think that PI, even when read as CI, is incompatible with another CI. The distributions encountered above are then expected.

That CIs are not mutually exclusive can be supported by more subtle evidence coming from comparing Arabic, English, and French. Consider the following contrast in Arabic and English (closer to that found in (11) above):

- (42) a. *ṭalaṭat-u fuṣuṣ-i taḥm-in*  
 b. three cloves garlic

---

<sup>16</sup> Here *-men* appears to operate more like a 'kindifier' than a 'grouper', close to what happens in the English translation, where "children" is a K (formed by the PI), and "group" is a counterpart to the Chinese G CI. However, *-men* has more specifications which make it rather a collective, as already explained.

(43) a. *ṭalaṭat-u ruʔuus-i ḡanam-in*

b. three heads of sheep

In both sets of examples, the measuring words (functioning as CIs) are in the PI form, whereas the head N is in the Sg form. But there is reason to think that the Ns are of different classes, the first being M, and the second K. For example, in French, the difference is manifested through PI marking in the K case, and its absence in the M case:

(44) a. *trois gousses d'ail*

b. *trois têtes de moutons* (\* *mouton*)

Consider another contrast which points to a similar conclusion. Taking into account the difference in forming K and I (including I-PI) in Arabic, English, and French, we expect K Ns (in their Sg form) in Arabic, or their PI form in English or French, to be compatible with "units", a measure word/CI, when denote integral units, but not necessarily with other N classes. The expectation is borne out:

(45) a. *ṭalaat-u waḥadaat-i tuffaaḥ-in*

b. three units of apples (\* apple)

c. *trois unités de poissons* (\* poisson)

(46) \* *ṭalaat-u waḥadaat-i tuffaaḥ-aat-in*

three units apple-I-PI

three units of individual apples

The ungrammaticality of (46) can be understood as barring the use of basically the same CI twice (on the unit and the N). The exclusive use of the PI form in English and French is necessary to make the K/M distinction visible.

Conceptually, the inadequacy of ECD is even more striking. If PI is of three types at least, as we have seen, and only one of them actively interacts distributively with CIs, then it is hard to see how strong ECD of CI and (any) PI can be maintained.

In examples like (15) and (16) above, PI has been shown to be a multiplier of individuals or groups, in (17) of sorts. In all cases, PI has to input an atom (individual, group or sort), and multiply it, outputting either an I-PI or a T-PI, as we have explained earlier. In the multiplier configuration, the complementary distribution is hardly conceivable on conceptual grounds, given that CI and PI are playing clearly separate roles in the grammar.

Second, as in constructions (19c & d) or (20 d & e) above, PI may play the role of *modifier* of the quantity of a PI (already formed). Thus there is no uniqueness in making use of PI inflection. In e.g. (19d), repeated here as (47) for convenience, a 'broken' PI form (the unmarked form with respect to quantity of plurality) combines with a 'sound' form, marked by the suffix –aat (PI fem), which emphasizes the high amount of quantity:

(47) a. farq "difference" → furuuq "differences"

b. furuuq "differences" → furuuq-aat "a lot of differences"

As observed above, similar patterns of these M-PIs are found in MA, which typically interact with CIs, to yield various interpretations, as in (20d & e), repeated here as (48):

(48) a. šḥam "fat" → šaḥm-ah "fat-I; a piece of fat" → šaḥm-aat "(many) pieces of fat"

b. šḥum "sorts of fat; a lot of fat" → šḥum-aat "a lot of fat; many sorts of fat"

We conclude therefore that ECD is simply wrong.

---

Note, however, that we do not deny the existence of apparent complementarity, or incompatibility of co-occurrence in some cases. In this section, we have illustrated cases arising from semantic incompatibility, e.g. the incompatibility of partitioning/counting and grouping at the same time within the same noun phrase, or cases of redundancy. In the next section, we turn to investigate what we think is the essential source of absence of co-occurrence of PI and CI, namely underspecification.

#### **4. Classifier variation and parametrization**

In this section, we restate the essential properties of the four feature-value N/CI ontology that we claim to be more empirically and conceptually adequate than the traditional binary 'count/mass' dichotomy. We also investigate the essential source behind Greenberg's (and others) intuition that languages which overtly express nominal CIs appear to lack a Nb category (or vice versa), even though, as we have argued, no form of ECD can be maintained. The role of underspecification of Nb and CI in conditioning the apparent existence of two types of languages: 'CI Ls' and 'Nb Ls'. Second, we propose an orientation of underspecification which we claim to be the source of variation.

##### *4.1. A two featured classification*

In dealing with classification, we have appealed to a binary featured classification ('atomic' and 'singulative') which takes positive and negative (and/or unspecified) values, instead of a single featured classification ('count/non-count'). We will provide more evidence here that this binary based

ontological classification is empirically more adequate, and that counting is not a classifier operation.

Let us recapitulate the main properties of the mereological classification proposed in Fassi Fehri (2003b). Recall that Ms, Ks, Gs, and Is are taken to be the 'natural' classes of Ns, with various cross-classifying semantico-syntactic properties. They are then organized through a minimal *mereology* which makes essential use of a *N-part* relation and a *N-integral* property of parts/wholes, thus capturing 'atomicity' and 'integrity/singulativity' of parts and wholes. A [+ atom] is an entity which has no (improper) N-part (which could make the whole-part relation symmetric and reflexive). I and G are [+ atom], having no (improper) N-parts. K and M, on the other hand, may have N-parts, but are not necessarily specified as such. Let us then mark this vagueness in specification as [∅atom].<sup>19</sup> A [+ sing] ('singulative' or 'individuative') feature is a property of an entity that is 'N-integral'.<sup>20</sup> Ks and Is are [+ sing] because they have integrity: if they have parts, these parts must be N-integral (i.e. cannot be

---

<sup>19</sup> The most quasi-standard formal properties of (proper) parts are the following:

- (i) there is no part x such that x is a (proper) part of x (irreflexivity);
- (ii) if x is part of y, then y is not part of x (antisymmetry);
- (iii) if x is part of y, and y is part of z, then x is part of z (transitivity).

The mereology proposed appeals to proper parts, as well as to (improper) parts.

<sup>20</sup> Singulativity may be equivalent to what is usually called 'individuation', a property of singularities. We have chosen a new term to avoid confusion found in the literature, with respect to the extension of the term 'individual' to K and G, and to be able to build a cross-classifying conceptual system which reflects the relation between both I and K (which are singulative), as distinct typically from Gs, which are 'collectives' (or [- sing]). Note that we have not chosen 'collective' as the marked option instead, because that option appears to make the wrong predictions in terms of the markings found in the Ls considered, where collectives or Gs are not (positively) marked, and they do not appear to exhibit any derivational link with what could be their negative counterparts. The question deserves, however, intensive investigation.



divided), be they proper or not. If K has one or more than an N-integral part, then the part-whole relation will be improper for the first case, and proper for the second. For I, the N-integral part is only improper. What is essential here is that N-integrity is a property of the part (and/or the whole). In simple words, Is and Ks cannot be divided or form a proper part-whole relation. It is the integrity/singulativity property that unify Is and Ks, which are otherwise different in terms of the atom property. Ms are [- sing]: they do not name N-integral wholes or parts. Gs are also conceptually [- sing] too: they may apply to a plurality of wholes, including integral wholes, but the latter are not N-integral. A 'team' may be composed of 'John' and 'Mary', but the whole and the parts are not in an N-part-whole relation, nor are they N-integral whole-parts. The system can be articulated as follows:

(49) N classification

- a. I = [+ atom, + sing]
- b. K = [ $\emptyset$ atom, + sing]
- c. G = [+ atom, - sing]
- d. M = [ $\emptyset$ atom, - sing]

The system accounts for shared and distinct properties assessed above through the cross-classifying attribute-value pairs. For example, the [+ atom] feature value of I and G describes their shared properties with respect to distributive Qs, negative 'kind of' behaviour, absence of reverse number predication, distributivity and cumulativity, taxonomic readings, etc.<sup>21</sup> Ks and Ms, on the other hand, being [ $\emptyset$  atom], have common properties with respect to distributive reference, 'kind of' modification, etc. Likewise, other

---

<sup>21</sup> See Fassi Fehri (2003b) typically for the reverse number predication property, as well as other properties mentioned here.

quantificational or predicational properties depend on the [+sing] feature value. We assume that this system is made available by UG to any language, and we expect these distinctions to be manifested either in lexical classes (or roots), or in grammatical f-words (or f-features). If this is true, then the question is how much of this system is 'inherently lexical' (or found in the lexicon of roots), or functional (carried by f-words or features).

#### 4.2. Cross-linguistic instantiations

In looking at the Ls we have examined so far, and having in mind the classifier system proposed, we have illustrated some operations (and/ or 'derivations') that change or form a new ontological entity from another basic ontology by changing the value of one feature. The operations appear to be identified by the two features proposed. One of them is 'atomize' (or its reverse value 'de-atomize'). The other is 'singulativize' (or its reverse 'de-singulativize'). These operations are illustrated below:

(50) atomize

K → I

samak → samak-at

(51) de-atomize

I → K

apple, dog → apples, dogs

(52) singulativize

M → K

xašab "wood" → xašab "wood" (xašab-at "wood-unit")

(53) de-singulativize

$K \rightarrow M$

$bg\grave{e}r \rightarrow b\grave{e}gr\text{-}ii$

We will come back to a potential instantiation of (51). The classification system appears to instantiate these operations, which involve value changes of one of the two features, and only those. Complex operations which involve changes of the values of the two features at the same time are not found (or so it appears):

(53)  $*I \rightarrow M$

(54)  $*G \rightarrow M$

(55)  $*M \rightarrow I$

(56)  $*M \rightarrow G$

We have found no counterexamples to this observation with respect to (54) and (56). Apparent counterexamples to (53) and (55) can be instantiated by 'universal grinder' and 'universal packager' discussed by Pelletier and Schubert (1989) and Jackendoff (1991), as in the following constructions:

(57) There is dog everywhere.

(58) I had three coffees.

If the individual "dog" is 'converted' to a mass "dog" in (57), or the mass "coffee" converted to an individual "coffee" in (58), then these operations, if grammatical, may be potential counterexamples to the claim that only single changes are permitted. We have found, however, no instances of (overt) grammatical CIs that operate these changes. More research is needed, but let us assume for the time being that only single operations are grammaticalized, and let us qualify the classifiers involved in these formations as *ontological* classifiers.

In Arabic, instances of 'atomic' and 'singulative' ontological classifiers are found, complemented by inherently built-in ontological distinctions at the N level for both features. In Chinese, the only built-in ontological distinction at the N level appears to be 'singulative' (the K/M distinction), and 'atomic' is left unspecified. As a consequence, an 'atomic' N must be formed through a 'count-classifier', to enable a singulative N to become an individual I. A 'count-classifier' is then an 'atomizer', and as such an ontological CI. If so, then Chinese 'count Ns' are not countable, much more like K Ns are not in Arabic. But they are not Ms though. As for CS 'mass-classifiers', they are not necessarily ontological, since they apply equally to singulative and non-singulative Ns. They are necessary though for counting, and may introduce units of measure or taxonomies that can be counted, without affecting the ontology of the N, much more like what is happening with the Arabic or English examples. If in both cases CIs are postulated, then we may still have to distinguish the two cases, A-CIs being ontological (in our sense), and T-CIs are not.

If our story is on the right track, then the difference between Ls like Chinese and Arabic (or English) is that the former (called a 'classifier L') has only one built-in feature distinction in N (the K/M singulative distinction), but Arabic has also the 'atomic' distinction built-in (although sometimes manifested also by an atomic CI). The distinction between traditionally called 'CI Ls' and others can then be stated as follows, building on the availability of a singulative/atomic classificatory system:

(59) In 'CI Ls' only the singulative distinction is built-in at the N level, in 'non-CI Ls' the atomic feature is also built in.

As a consequence of (59), a CI is necessarily needed for atomization (as an I), or taxonomization (as a T) in the first type of Ls. As we have shown earlier, Nb is blind to the A/T distinction, so it cannot be a CI in the ontological sense (except in the cases mentioned).<sup>22</sup>

#### 4.4. Classifier and Number distributions

The difference between CI Ls and non-CI Ls reflected at the N level correlates with a massive use of CI in type 1, and the massive use of Nb in type 2, with apparent complementarity examined above, hence giving the impression that they belong to the same category. Given that they are not, as we have shown above, we still have to explain why PI tends to be rarely used with CI in Chinese, or that the CI use is neither necessary nor productive in English, or Arabic. In fact, these observations have to be made more precise, given that either CI or Nb (or both) are necessary in counting Ms or Ks:

- (60) a. three kinds of oil  
b. three categories of persons  
c. three heads of cattle  
d. three oils

Looking closely at the facts, we found that the kind of PI which is lacking in Chinese is a general (free) multiplier PI. PI morphemes do exist in Chinese, but their character is more specified than the more general character of Nb in 'Nb Ls'. General Nb is dedicated basically to forming sets of atoms from other sets of atoms, including singletons. If atomic CI and general Nb are valuing the atom feature, then they appear to have the same function, and hence

---

<sup>22</sup> If it is thought of as a very general CI, which distinguishes 'count' from 'non-count', then the

predicted to be complementary. But in fact they are not. The atomic CI is forming an ontological class of Ns, namely individuals, whereas the general Nb is forming sets, blind to ontological distinctions.

Ls like English which express general Nb tend to express no CI. Ls like Chinese which express CI tend to express no general Nb. Suppose this is generally true, but not true of the specific cases. Suppose also that the CI specification of the Nb specification can be deduced by default, either by (a) taking into account some 'natural' correlation between Nb and CI or vice versa, or (b) just default specification.

Consider again the example (3) above, repeated here as (61) for convenience:

(61) wo xiang mai ben shu  
 1s want buy CI book  
 I want to buy a book.

Here 'book' is interpreted as atomic (singular). But in comparison with (2a), repeated here as (62), the CI does not make the N necessary singular. It can be a PI, as the translation indicates:

(62) ji ge pingguo  
 few CI apple  
 few apples

We have then to make a distinction between the atomicity introduced by the CI, which is ontological and forms I integral units, and that introduced by a potential higher Nb, which may or may not be atomic, hence the contrast between (3) and (2b). If this is correct, then there is an unexpressed Nb dominating the CI in Chinese, which is unspecified with respect to atomicity,

---

distinction between the T reading and the A reading of PI/Nb still remains to be accounted for.

but which can be attributed a positive or a negative atomic value, depending on syntactic contexts. Note that *yi-* in (8) above may be taken as overtly realizing Nb, if *yi-* is the counterpart to "a" in English, rather than to "one", as proposed by Chen (2003). If so, then Nb and Cl are not complementary in Chinese, typically in the singular case. In this type of situation, the (under)specification appears to be *oriented from bottom to top* (or from Cl to Nb). That is, the atom feature value of Cl may affect that of Nb in the general case, but not vice versa. If the Cl is atomic, then Nb is atomic too, resulting in a singular interpretation by default. This is close to the situation of *samak-at-un* 'integral fish' in Arabic, where the interpretation is singular, by default.

Let us turn now to Ls with productive general Nb morphology, such as English. The plural mark on "dogs" can be a kind K forming, and hence unspecified for atomicity (therefore not readily countable), or it can be a multiplier, in which case it is [- atom], and it presupposes the existence of atomic entities that can be multiplied. If so, then the positive value of the atomic feature of the the unexpressed Cl is forced by the selection of the multiplier Pl. Note that the same logic appears to apply to "oils", where Pl is a multiplier, and an atomic Cl is needed too for the sake of forming a negatively atomic set of oils. But there is an important difference between Cl use in Chinese and Pl use in English, to express atomicity. While Cl are sensitive to properties of Ns they apply to (basically the K/M distinction, or the specification of the singulative property), Pls appear to be usually blind to this distinction. However, because pluralized Ns even in English do count two types of entities (atomic and no-atomic entities), the unexpressed Cl must inherit a potential K/M specification to distinguish the two types of counting, much more like what happens in Chinese. If so, then we can hypothesize that the K/M distinction is

universally available at the N level, although it might be formed at other levels through morphology.

If singulativity is universally available at the N level, then CI use either selects the marked value of this feature to form individuals, *Is*, by making use of the positive atomic value, or it selects the unspecified value to form taxonomies. Since atomicity is specified at the CI level, there appears to be no need to re-specify it at the Nb level, at least in Chinese. Likewise, *Ls* like English which specify atomicity at the Nb level tend to do without specifying atomicity at the CI level. Arabic appears to be a mixed system, with the tendency to behave like English in the general case. Simplifying things to extreme cases, we can characterize the variation found as follows:

(63) Specify atom at the CI or the Nb level.

In Fassi Fehri (2003b), the first case is called bottom-to-top oriented *Ls* (with respect to atomicity), and the second case top-to bottom oriented. We think that a similar orientation is found in aspectuo-temporal systems, where systems that tend to specify Aspect do not specify Tense, and vice versa, although mixed systems are also found (Fassi Fehri (2003a, 2004)). Note that the choice stated in (63) does not imply that Nb *is* a CI (or vice versa), just as the orientation choice found in the aspectuo-temporal system does not imply that Aspect *is* Tense, or vice versa, although their overt expression tends to be complementary in many systems. But here too, mixed systems are frequent, clearly pointing out the misconception of classifying *Ls* into 'Tense *Ls*' and 'Aspect *Ls*', in which the two categories are mutually exclusive.



---

Summarizing, we claim that differences between Chinese and English are built around two dimensions that we have formulated in (59) and (63), repeated here for convenience in the form of (64) and (65):

(64) In Chinese (a 'Cl L'), only the [sing] distinction is built-in at the N level, in English (a 'non-Cl L'), the [atom] feature is also built in.

(65) Specify [atom] at the Cl level (for Chinese), or the Nb level (for English).

These differences are in fact manifestations of the classificatory system that we have adopted.

**References**

- Aikhenvald, A.: 2001, *Classifiers: A typology of noun categorization devices*. Oxford, New York: Oxford University Press.
- Allan, J.M.O.: 1977, Classifiers. *Language* **53**. 285-311.
- Barwise, J. and R. Cooper: 1981, Generalized Quantifiers and Natural Language. *Linguistics and Philosophy* **4**. 159-219.
- Borer, H.: 2002, *Structuring Sense. An Exo-Skeletal Trilogy*. USC ms. California.
- Bresnan, J.: 2001, *Lexical-Functional syntax*. Oxford: Blackwell.
- Chao, Y.R.: 1968, *A Grammar of spoken Chinese*. Berkeley: University of California Press.
- Chen, P.: 2003, Indefinite determiner introducing definite referent: a special use of *yi* 'one'+classifier in Chinese. *Lingua* **113**. 1169-1184.
- Cheng, L. & R. Syberma: 1999, Bare and Not-so-bare Nouns and the structure of NP. *Linguistic Inquiry* **30**. 509-542.
- Chierchia, G.: 1998, Reference to Kinds across languages. *Natural Language Semantics* **6.4**.
- Corbett, G.: 2000, *Number*. Cambridge : Cambridge University Press.
- Croft, W.: 1994, Semantic universals in classifier systems. *Word* **45**. 145-171.
- Fassi Fehri A.: 2003a, Arabic perfect and temporal adverbs. In A. Alexiadou, M. Rathert & A.von Stechow eds. *Perfect Explorations*. Berlin : De Gruyter.
- Fassi Fehri A.: 2003b, Nominal Classes and Parameters across Interfaces and Levels, with a particular reference to Arabic. *Linguistic Research* **8.2**. Rabat. IERA Publications.

- Fassi Fehri A.: 2004, Temporal/Aspectual Interaction and Variation across Arabic Heights. In J. Guéron & J. Lecarme eds. *The Syntax of Time*. 235-257. Cambridge, Mass: MIT Press.
- Greenberg, J.: 1972, Numeral classifiers and substantival number : Problems in the genesis of a linguistic type. *Proceedings of the Eleventh Congress of Linguists*. Bologna : Società editrice il Milano.
- Hall, R. A. Jr.: 1943, Melanesian Pidgin English. Baltimore : Linguistic Society of America.
- Iljic, R.: 1994, Quantification in Mandarin Chinese : Two markers of plurality. *Linguistics* 32. 91-116.
- Iljic, R.: 2001, The problem of the suffix *-men* in Chinese Grammar. *Journal of Chinese Linguistics* 29:1. 11-67.
- Jackendoff, R.: 1991, Parts and boundaries. *Cognition* 41. 9-45.
- Kemmer, S.: 1993, Verbal and nominal collectives. *Faits de langues* 2. Paris : édition Ophrys.
- Krifka, M.: 1995, Common Nouns in Chinese and in English. in G. Carlson and F.J. Pelletier (eds.) *The Generic Book*. Chicago : University of Chicago Press.
- Kouwenberg, S.: 1994, Berbice Dutch. In J. Arends, P. Muysken and N Smith (eds.) *Pidgins and creoles : an introduction*. 233-243.
- Li, H-Y. A.: 1999, Plurality in a Classifier language. *Journal of East Asian language* 8. 75-99.
- Li, C. N. and S.A. Thompson: 1981, *Mandarin Chinese: A functional Reference Grammar*. Berkeley, Los Angeles, London: University of California Press.
- Lü, S.: 1947, *Zhōngguó wénfǎ yàolüè (Outline of Chinese Grammar)* 3 vol., Shanghai: Shangwu yinshuguan.

- Mühlhäusler, P.: 1981, The development of the category of Number in Tok Pisin. In P. Muysken (eds.) *Generative Studies in Creole languages*. Dordrecht: Foris Publications.
- Norman, J.: 1988, *Chinese*. Cambridge: Cambridge University Press.
- Paris, M-C.: 1981, *Problèmes de syntaxe et de sémantique en linguistique chinoise*. Paris: Collège de France.
- Pelletier, F.J. & L.K. Schubert: 1989, Mass Expressions. In D. Gabbay & F. Guenther eds. *Handbook of Philosophical Logic*. Vol. 4. 327-407. Dordrecht: Kluwer.
- Po-Ching Y. and D. Remington: 1997, *Chinese: An Essential Grammar*. London and New York : Routledge.
- Rygaloff, A.: 1973, *Grammaire élémentaire du chinois*. Paris : PUF.
- T'sou, B.: 1976, The structure of nominal classifier systems. In P. Jenner, L. Thompson & S. Starosta eds. *Austroasiatic Studies II*. Honolulu: The University Press of Hawai.
- Wiese, H.: 1999, Two kinds of nominal 'singular' and 'plural' in natural languages: a semantic analysis. Paper presented at the Annual Meeting of the International Linguistic Association, New York. April 18-19.
- Wu, J.: 1999, *Syntax and semantics of quantification in Chinese*. Unpublished Ph.D dissertation. University of Maryland at College Park.