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 'individuative') 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: [α atomic] and [β 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 [Ø 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 Cl/Pl (under)specification is 'top-down' oriented in English and (somewhat differently) in Arabic, but it is oriented 'bottom-up' in Chinese. Thus overt Pl specification in English and Arabic may induce Cl specification, without the need for the latter to be overt (hence the apparent complementarity). On the other hand, Cl is overtly manifest in Chinese, but no Pl 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 Cls in Arabic and Chinese, as well as varieties of Nb (and Pl) uses. We analyze some clear non-complementary distribution cases of Cl 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 Cl interaction, and how cross-linguistic variation in this interaction can be parametrized.

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).[4]

1.2. A new classification of classifiers

If both types of Cls are adding some information to that carried by (bare) Ns which relates to countability, then the distinction of Cls (examined so far), according to whether they are count-Cls or mass-Cls (or even worse massifiers) cannot be adequate. The classification needed has to take into account the type of information contributed by the Cl, 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 [± partition], mistakenly taken to be equivalent to 'count' vs. 'mass'.[5] 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 Ø±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

[4] 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.

[5] 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' Cl to be counted.

What the 'counting' Cl 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 Cls 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-Cl), or *unspecified* for atomicity (K-Cl). 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 Cl 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-Cl (or 'individualizer') is associated with an inherent singularity (or an integral 'partition') in the case of Arabic, whereas it is associated with unspecified atomic Cl (a K-Cl, 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-Cl 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-Cls (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) bgər "cattle, cows" → bəgr-i "cow-M; beef"

Other potential Cls to be found in this system are 'collective' or 'group' Cls (G-Cls), which are typically documented in Chinese.[6]

Cls 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 Cl is [+ atom], as in the case of I-Cl (or G-Cl) can it be countable. K-Cl and M-Cl do not output countable units. Call the successful counting Cl A-Cl [+ atom] Cl. Cls which do not output countable entities (Ks and Ms) are non-A-Cl (i.e. [- or Øatom] Cls. Since atomicity appears not to be relevant for the opposition of the pair, we might better term them S-Cl (singulative Cls) for convenience. Ns can be classified inherently as A-Ns or non-A-Ns or S-Ns, just like Cls are. S-Ns, namely Ms or Ks, can be made countable through Cls of different nature and range. These Cls count 'sorts', 'taxonomies', 'portions' of substances, etc. They tend to be found cross-linguistically, unlike A-Cls and S-Cls. They are phrasal (and lexical), and they can be easily translated from one language to another:

[6] 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 Cl student-men
nine students

(11) a. ṯalaaṯat-u ruʔuus-i ġanam-in
three heads-gen sheep-gen
three heads of sheep
b. ṯalaaṯ-u qiṭa^{c}-i xašab
three pieces wood
three pieces of wood

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

(12) three kinds of apples

Differences between our A-Cl and T-Cl have been amply noted in the literature, although misleadingly attributed to the 'count-mass' Cl distinction. Thus CS (1999, 1998) note that even though 'count-Cls' and 'mass-Cls' are both nominals, only 'count-Cls' form a closed set, with elements functioning solely as Cls (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 Cl river
a river

b. san duo hua
three Cl flower
three flowers

c. si liang qiche
four Cl car
four cars

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

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

form taxonomies (or measures). Of the three types, both A-Cl and T-Cl 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.[7]

2. Kinds of Number

Nb/Pl marking interacts with Cls 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 Pl (Nb) uses typically in Arabic and Mandarin Chinese, and to assess naturally expected kinds of interaction. One use of Pl might be termed a *multiplier.* It multiplies individuals or sorts/taxonomies, in the sense that it generates a referential expression from the NP/ClP, by mapping the nominal expression onto a *multitude* of its realizations. It applies equally to A-Cls (or A-Ns), and to T-Cls, hence forming A-Pl and T-Pl, respectively. A second use of Pl 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 Pl 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-Pl.* What is typical of M-Pl is that it does not apply to countable units, but to amounts, Ks or

[7] 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 Pl 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 Pl types in Arabic. The following examples instantiate a multiplier A-Pl (with Is and Gs) in (15) and (16), and a multiplier T-Pl (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" → ʔasmaak "sorts of fish; a multitude of fish"

In all these cases, we take the Pl to be a multiplier, inputing an atomic individual or sort, and outputting a multitude of these entities.[8]

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

[8] Borer (2002) takes the Pl 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 Pl there is a Cl which creates countable units, a view that she generalizes to all Pl 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 Pl 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-Pl; a multitude of sayings, utterances"

c. ʔaqwaal "sayings" → ʔaqaawiil "a lot of sayings"

d. farq "difference" →furuuq "differences" →furuuq-aat "a lot of differences/ various sorts of differences"

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

(20) a. bettix "melon K" →bettix-at "melon-I; an integral piece of melon"

b. bəttix-at "a melon" →bəttix-aat "(many integral pieces of) melons"

c. bəttix "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 Pl is never used in Arabic as a Cl forming a K, as in the case of English 'apples', where the Pl is not necessarily read as a multiplier. This latter Pl translates into Arabic by a Sg K 'tuffaaḥ' ("apple"), rather than by any form of the Pl types provided above.

In addition to the ample literature found in Arabic traditional grammars on these Pl 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 multipler/M Pl 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 Pl indicate clearly that the functions of Pl are far from being limited to classification, and hence the interactions between Pl/Nb and Cl have to be more carefully investigated.

2.2. Chinese

Pl 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. laoshi-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 Pl (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
Iljic (1994: 91). In this
referring to a whole.[9] T
of viewing the membe
regularly attached to p
use with Ns is not rando
subjective group, and it
or interpersonal relation

Another argument
regular Pl, is the fact th
(Iljic 1994):

(27) Xiao Qiangmen
Xiao Qiang-MEN
Xiao Qiang's group

The suffix then serves to
type of collective reading
found in other languages

[9] Norman (1988: 121) explains
person'.
[10] The term *collective* has bee
117-120, Kemmer 1993: 92-94,
confusion. The internal plural va
acquires a different content, dep
between K and G, as we constru
[11] Iljic (2001) indicates that this is

are grouped together relative to the speaker or some other subjective origin" Iljic (1994: 91). In this case, it is best treated as *collective* or a 'grouper' G, referring to a whole.[9] The use of collective forms with *-men* concerns the way of viewing the members of a group.[10] As mentioned by Iljic (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 Pl, is the fact that it can suffix to proper Ns, as in (27), taken from (Iljic 1994):

(27) Xiao Qiangmen
Xiao Qiang-MEN
Xiao Qiang's group

The suffix then serves to identify a group relative to a certain person.[11] 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

[9] Norman (1988: 121) explains that this suffix was originally a compound nominal meaning 'every person'.

[10] 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'.

[11] Iljic (2001) indicates that this is the use of well-read speakers of Chinese.