The cognitive basis of the mass-count distinction: evidence from bare nouns

Edit Doron & Ana Müller
The Hebrew University of Jerusalem & University of São Paulo

Current view in the literature

The linguistic mass-count distinction does not reflect the cognitive distinction between entities which homogeneous vs. discrete (atomic), hence it is arbitrary:

- There are mass nouns denoting discrete entities -- furniture, jewelry, clothing, e.g. a chair is furniture but part of a chair is not furniture (Chierchia 1998).
- There are count nouns denoting homogeneous entities -- fence, line, bouquet, e.g. two bouquets together may form one bouquet (Rothstein 2010).

Our view

The linguistic mass-count distinction is not arbitrary, but reflects a cognitive distinction, not the one above but between entities with unstable vs. stable units.

- Mass nouns denote entities with inherently unstable units; e.g. furniture is bona-fide mass, since what counts as a unit of furniture in a given context is vague, it could be the whole sitting room set or just one of its parts (This living room set is so much furniture!)
- Count nouns have stable units in each given context (Chierchia 2009); e.g. bouquet is bona-fide count, and considering parts of a bouquet to be bouquets changes the context (#This bouquet is so many bouquets!).

Since the point in the last bullet was already made in Chierchia 2009, we will concentrate on the bullet above it: the units of e.g. furniture are unstable.

But first we argue, on the basis of Karitiana, for the general point that the mass-count distinction is a cognitive rather than a formal linguistic distinction.

Karitiana

In Karitiana (a Tupi-Arikém language spoken in Rondônia, Brazil), the mass-count distinction is not encoded morphologically. There is no number morphology in the language (1):

(1)  Maria Ø-nakam’at gooj
      Maria 3-decl-caus-make-nfut boat
      'Maria built some boat(s).'

Yet the mass-count distinction is encoded distributionally. Count Nouns may be modified by numerals (2):

(2)  Myhin-t ‘ejepo Ø-naakat i’ot-Ø
      one-obl stone 3-decl-cop-nfut participle-fall-cop.agr
      'One stone fell.'
and plurality can also be encoded by modifiers:

(3) 'ejepo Ø-naaka-j i-'ot-ot-Ø            myhin-t myhin-t
    stone decl-cop-fut participle-fall-cop.agr one-obl one-obl
    'The stones will fall one at a time.'

whereas Mass Nouns cannot be modified by numerals (4):

(4) Myhin-t *(kilo-t)ouro na-aka-t i-'ot-ot-Ø
    one-obl kilo-obl gold decl-aux-nfut participle-fall-con.cop
    'One *(kilo of) gold fell.'

Thus in Karitiana, the mass-count distinction among nouns directly reflect the
individuability of units.

Count→Mass "universal grinder" coercion if found as well, e.g. tree ↔ wood below:

**Count**

(5) 'ep i-ti-pasagngã-t João
    tree 3-inverse-count-nft João
    ‘The trees, João is counting (them).’

**Mass**

(6) 'ep Ø-na-aka-t jepyry-t
    wood 3-decl-cop club-cop.agr
    ‘The club is of wood.’

(7) atykiiri Ø-naam-’a-t ej otom 'ep tyym
    then 3-decl-make-nft tomb cover wood too
    ‘Then they also make a cover of wood for the tomb’

(8) ambí 'ep
    house wood
    ‘wooden house’

**Hebrew**

Hebrew has number morphology

(9)a. xatul / xatul-im        b. tmun-a / tmun-ot
cat-masc / cat-masc-pl        picture-fem / picture-fem-pl

**Counting: one/two/several, with count but not mass nouns**

(10)a. yéled exad        b. yelad-im axad-im        c. šneý yelad-im
      boy one              boy-pl one-pl                two boy-pl
      ‘one boy’            ‘several boys’             ‘two boys’

(11)a. yald-a axat        b. yelad-ot axad-ot        c. šteý yelad-ot
      girl-fem one.fem    girl-fem-pl one-fem.pl    two-fem girls-fem
      ‘one girl’          ‘several girls’           ‘two girls’
As in Karitiana, the mass-count distinction in Hebrew does not rely on morphology but on distribution (e.g. as we just saw, with numerals)

**Number morphology does not provide a reliable distinction between Mass Nouns (MNs) and Count Nouns (CNs), as many Mass Nouns have plural morphology.**

True, some MNs cannot be pluralized (14):

(14)a. *xamcan* *xamcan-im*  
*oxigen* oxigen-pl  
*gravel* gravel-pl  
*boc* *boc-im*  
*máy-im*  
‘one cotton’  

But others do not have singular form (15):

(15)a. *may*  
*b. 'atiq-ot* *'atiqa*  
*c. šimur-im* *šimur*  

But must nevertheless be viewed as mass nouns since they cannot be counted (16a), and do not show reciprocal ambiguity (16b) unlike plural count nouns (16c) (cf. Gillon 1992)

(16)a. *šney šmarim*  
*two yeasts*  

b. *qurey ha-akaviš ve-ha-atiqot domim* *ze la-ze*  
the spider webs and-the-ruins resemble each other  

non-ambiguous

(16)b. *ha-xatulim ve-ha-xatulot domim* *ze la-ze*  
the-cats and-the-cats.fem resemble each other  

ambiguous

Many MNs have both sing and pl forms:

(17)a. *xol* *xol-ot*  
b. *géšem* *gšam-im*  
*c. šéleg* *šlag-im*  

sand sand-pl  
*rain* rain-pl  
*snow* snow-pl  

These too are MNs, as e.g. they cannot be counted

(18)a. *šney xolot  b. *šney šlagim
two sands       two snows

A second distributional criterion for the mass-count distinction:
CNs can be distinguished by appearing with plural morphology when they appear with quantifiers such as harbe 'much', me’at 'a little', xóser 'lack of', ódef 'a surplus of' and classifiers such as kilo 'a kilo'.

Count Nouns
(19)a. kilo tapuxim/ agasim   b. harbe anavim   c. me’at zeytim
      kilo apples/ pears      much grapes      a little olives
      d. xóser adašim/agvaniyot e. ódef boțnim/duvdevanim/šezifim
         lack   lentils/ tomatoes    surplus peanuts/cherries/plums

(20)a.* kilo tapúax/ agas   b. * harbe anav    c. *me’at záyit
     kilo apple/ pear         much grape      a little olive
     d.* xóser adaša/agvaniya e. * ódef bóten/duvdevan/šezif
         lack   lentil/tomato    surplus peanut/cherry/plum

MNs, on the other hand, can appear in the singular.

Mass Nouns
(21)a. kilo šum       b. harbe peťruzílya   c. me’at nána
       kilo garlic       much parsley      a little mint
       d. xóser téred     e. ódef šu’it
          lack   spinach    surplus bean

Paradox
We find some nouns which are normally considered CNs, and are indeed all countable as shown in (22):

(22) xamiša gzarim / šney milonim / šiv’a bcalim/ asara tutim
      five carrots/ two melons    seven onions/ ten mulberries

šmona míšmišim/ ánanasim/ tírasim / kruvim / avațixim
      eight apricots / pineapples/ corncobs/ cabbages/ watermelons

yet are found with measure phrases in the singular (23), which is a MN characteristic:
(this is not with the "universal grinder" interpretation)

(23) kilo gêzer / milon/ bacal/ tut
     kilo carrot/ melon/ onion/ mulberry

ódef míšmiš / ánanas / tíras
surplus apricot/ pineapple / corn
If I got to the market, it is natural to say to me (a) but not (b)

(a) tiqni gézer / tut / mišmiš
   buy carrot / mulberry/ apricot
   'Buy carrots/ mulberries/ apricots.'

(b) tiqni tapú'ax/ agas/ tapuz/ agvaniya/ šezif
   buy apple/ pear/ orange/ tomato / plum
   # 'Buy an apple/a pear/ an orange/ a tomato/ a plum.'

Again, this is not the "universal grinder" interpretation, it is an additional interpretation:

(25)a. rov ha-gézer raqv
   most (of) the carrot(s) is/are rotten
   'Most of the carrot(s) is/are rotten.'

(b) rov ha-tapúax raqv
   most (of) the apple is rotten
   'Most of the apple is rotten.'

(26)a. yeš gézer b-a-tiq
   there (is) carrot in-the-bag
   'There is (a) carrot in the bag.'

(b) yeš tapúax b-a-tiq
   there (is) apple in-the-bag
   'There is (an) apple in the bag.'

(27) ani ohevet gézer / *tapúax / tapuxim
   I love carrot / apple / apples

(28) mic gézer / *tapúax / tapuxim
   juice carrot / apple / apples

(29) maraq bacal / *agvaniya / agvaniyot
   juice onion / tomato / tomatoes

We would like to claim that the nouns in (22)—(23) are MNs of the type which has been called fake by Chierchia 1989, 2009, like English MNs such as furniture. In Hebrew, carrot, melon, onion, mulberry, apricot, corn are all fake MNs. They have units which are individuated in many contexts, but we are not normally interested in them, since the normal context for the use of these terms is the preparation of food, where we are normally interested in parts of these fruits/vegetables (melon, onion), or in aggregates of them (mulberry, apricot). Since we may wish to count using different units within the same context, the units are not stable and thus these nouns are predicted to be MNs.

What Hebrew has that English doesn't is singulative morphology which allows the selection of particular units, and thus shifts these nouns into CNs.

In Hebrew many fake MNs have corresponding singulative CNs (sometimes of a different morphological form, e.g. different gender, from the MN, as in (30), and sometimes homonymous to the MN, as in (31)).

<table>
<thead>
<tr>
<th>fake MN</th>
<th>CN singulative</th>
<th>CN-pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>se'ar</td>
<td>sa'ar-ot</td>
<td>sa'ar-a</td>
</tr>
<tr>
<td>hair.masc</td>
<td>hair-fem</td>
<td>hairs</td>
</tr>
</tbody>
</table>
We do not want to say that all noun roots are basically mass, since we find both marking of CN and MN. Both markings are attested.

**MN is marked:**

(32) deriving a collective MN by the fem suffix was productive in Biblical Hebrew:

<table>
<thead>
<tr>
<th>CN</th>
<th>CN-pl</th>
<th>MN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a</strong></td>
<td><strong>b</strong></td>
<td><strong>c</strong></td>
</tr>
<tr>
<td>dag</td>
<td>dag-im</td>
<td>dag-a</td>
</tr>
<tr>
<td>fish.mas</td>
<td>fish-pl</td>
<td>fish-fem</td>
</tr>
<tr>
<td><strong>b</strong></td>
<td><strong>c</strong></td>
<td><strong>d</strong></td>
</tr>
<tr>
<td>ale</td>
<td>al-im</td>
<td>alv-a</td>
</tr>
<tr>
<td>leaf.masc</td>
<td>leaves</td>
<td>foliage-fem</td>
</tr>
<tr>
<td><strong>c</strong></td>
<td><strong>d</strong></td>
<td><strong>e</strong></td>
</tr>
<tr>
<td>afun</td>
<td>afun-im</td>
<td>afun-a</td>
</tr>
<tr>
<td>pea.masc</td>
<td>peas</td>
<td>pea-fem</td>
</tr>
<tr>
<td>oréax</td>
<td>orx-im</td>
<td>orx-a</td>
</tr>
<tr>
<td>traveller.masc</td>
<td>travellers</td>
<td>caravan-fem</td>
</tr>
<tr>
<td>gole</td>
<td>gol-im</td>
<td>gol-a</td>
</tr>
<tr>
<td>exiled.masc</td>
<td>exiled-pl</td>
<td>exile-fem</td>
</tr>
</tbody>
</table>

(33) Argument Structure Nominal morphology (ASN) sometimes marks collective MN as well, which makes sense, since Grimshaw 1990 shows ASN have the distributional properties of MN:

<table>
<thead>
<tr>
<th>CN</th>
<th>CN-pl</th>
<th>MN derived by ASN suffix</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a</strong></td>
<td><strong>b</strong></td>
<td><strong>c</strong></td>
</tr>
<tr>
<td>béged</td>
<td>bgad-im</td>
<td>bigud</td>
</tr>
<tr>
<td>garment</td>
<td>garments</td>
<td>clothing-ASN</td>
</tr>
<tr>
<td><strong>b</strong></td>
<td><strong>c</strong></td>
<td></td>
</tr>
<tr>
<td>ná’al</td>
<td>na’al-áyim</td>
<td>han’ala</td>
</tr>
<tr>
<td>shoe</td>
<td>shoes (dual)</td>
<td>footwear-ASN</td>
</tr>
<tr>
<td><strong>c</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rehiš</td>
<td>rehiš-im</td>
<td>rihuš</td>
</tr>
<tr>
<td>piece of furniture</td>
<td>pieces of furniture</td>
<td>furniture -ASN</td>
</tr>
</tbody>
</table>

**CN is marked**

again by the fem. suffix, this time marking the singulative (nomen unitatis) (in Arabic and Neo-Aramaic this is productive). Thus it seems that the interpretation of the derivational fem. affix depends on whether the root is count or mass, which might argue for the root being classified as to the count-mass distinction.
Singulative marking is productive in Neo-Aramaic (Christian Barwar, Khan 2008:343):

(35)  fake MN    CN singulative    CN-pl
a  tu-tha        tu-tha-ta         tu-tha-tha    mulberries-pl    mulberry-fem    mulberries
b  tar-e        tar-e-tha         tar-e-tha      leaf-pl          leaf-fem        leaves
  dadokk-e      dadokk-tha         dadokk-tha    seed-pl          seed-fem        seeds

We have called the nouns in (30)—(31) fake MN. The singulative type-shift derives CNs (though the singulative here does not have a fem exponent):

(36)  fake MN    CN singulative    CN-pl
  gézer        gézer              gzar-im    carrot
  milon        milon              milon-im    melon
  léxem        léxem              lexam-im    loaf of bread
  réxev        réxev              rexav-im    vehicle
  néšeq        néšeq              nešaq-im    weapon
  weaponry     weapon             weapons
What are fake mass terms? These are predicates P such that given a context c which individuates units of P, it is nevertheless natural to view parts or aggregates of these units as units as well for P in the same context.

Examples: furniture, footwear, clothing, mail, silverware, means of transportation

These predicates have perceptible units which are natural, like a chair, a knife, a letter, a shoe, a sock. Yet in most contexts we are not interested in these units, but in aggregates of these units: a pair of shoes, a knife-fork-spoon set, a living-room set, the contents of my mailbox at a given time. Since we want all these to also count as units, within the same context, the right type of predicate would be a mass term.

Additional examples: carrot, melon, watermelon, mulberry, pea

Here too there are units individuated in many contexts, but we are not normally interested in them, since the normal context for the use of these terms is the preparation of food, where we are normally interested in parts of these fruits/vegetables (melon, watermelon), or in aggregates of them (mulberry, pea).

Fake mass terms are different from homogeneous count nouns (Mittwoch 1988, Rothstein 2010), such as wall, fence, table, mountain, line, sequence, bunch, which are count nouns. The units of these predicates are individuated in dramatically different ways in different contexts, but in a given context, whatever is part of a unit does *not* fall under the predicate, though it could in a different context.

References