

Counting Chad
Arnold M. Zwicky, Stanford University

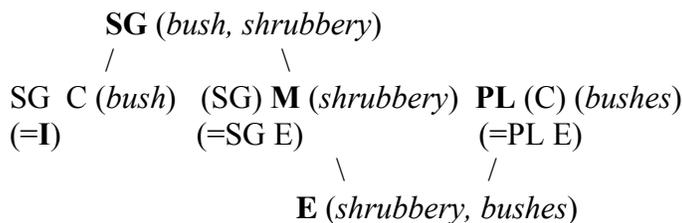
1. Count/Mass (C/M) distinction in English (familiar observations):
 - (A) It's a classification of **lexemes as types**, not of their tokens.
 - (B) The classification concerns **syntactic behavior**, though closely associated with semantics.
 - (C) Indeed, the syntactic classification concerns **a collection of rules**, not just one
 - (D) There is a **default association** between the syntactic classification and semantics: **taCisM**
(things are C, stuff is M)
 - (E) taCsiM **competes** with a variety of other principles.
 - (F) taCsiM is not an objective principle, but instead takes into account the **salience of significance** of a thing vs. stuff analysis of the referent for ordinary human dealing with it – especially for those dealings with it that are conventional within a culture.
 - (G) There are systematic ways of converting M nouns to C and vice versa

2. Re (A):

- 2.1. ROSE 'rose plant' is C, even when used to refer to collections of rose plants grown as dense hedges, covering trellises, or espaliered thickly over a wall (*You have a lot of roses/*rose growing out there*); POTATO 'potato plant' is also C, even though potato plants are often grown so that their vines make a dense mat (*The field of potatoes/*potato made a solid expanse of green*).
- 2.2. TARRAGON 'tarragon plant' and IVY are M, even when used to refer to individual plants grown in separate pots (*You have a lot of tarragon/*tarragons/ivy/*ivies growing those pots*).

3. Re (B):

3.1. Five syntactically significant classifications:



- 3.2.
 - a. **SG** (M or C), as in subject-verb agreement, selection of *this/these* and *that/those*
 - b. **PL** (C), as in: *many/(a) few *bush/bushes/*shrubbery*
 - c. (SG) **M**, as in: *much/(a) little *bush/*bushes/shrubbery*
 - d. **I** (= SG C), as in: *a/one/each/every *bush/*bushes/shrubbery*
 - e. **E** (= SG M v PL C), as in:
 - bare N: *A lot of/ Lots of *bush/bushes/shrubbery will burn easily.*
 - There should be *bush/bushes/shrubbery galore/aplenty in the desert*
 - All *bush/bushes/shrubbery in the desert will be fragrant*

4. Re (C): C/M is parallel to Main/Aux verbs in English (Aux verbs make a class with respect to Verb Phrase Ellipsis, Subject-Auxiliary Inversion, *not* placement, negative inflection), and to paradigm class distinctions in the morphologies of many language (1st declension nouns in Latin make a class by lacking *-s* in the Nom Sg, having the Gen Sg and Nom Pl in *-ae*, having *-i:s* in the Dat/Abl Pl).

5. Re (D):

5.1. Because of taCsiM, the names *count* and *mass* are pretty good ones.

5.2. But the “Bloomfieldian dogma” is that “form-classes are semantically arbitrary” (Wierzbicka 1988:500, hereafter W); contrasts like those between C BEAN PEA LENTIL and M CORN RICE BARLEY have been taken to support semantic arbitrariness.

5.3. Still, there’s a lot of mileage in taCsiM, which predicts vast numbers of C/M assignments: e.g. names of plants that grow as easily distinguishable individuals, like ROSE THISTLE TULIP ELM, are C; names of plants that grow in dense mats or clusters, like CLOVER VINCA IVY BAMBOO, are M. taCsiM predicts that ICE PLANT is M, since it is grown in expanses as a cover for dunes or banks along highways, and it is for many speakers/writers, including the author(s) of a website “How to control ice plant”:

So what do you do if you have a slope covered with ice plant that you would like to restore to something more interesting? Experts say that it’s best not to remove the ice plant by pulling it up... Ice plant can be easily killed with Roundup... and also for many people who aren’t writing about plants specifically, as in this passage from Ramachandran & Blakeslee, *Phantoms in the Brain* (p. 21):

...he was thrown from the driver’s seat onto the ice plant bordering the freeway.

5.4. W proposes, in fact, to maintain taCsiM thoroughly, denying that there is any arbitrariness at all in C/M assignment; beans are just bigger and easier to distinguish from one another than grains of rice are, for instance. For W, any remaining differences in C/M assignment (between dialects or languages) are differences in **the way people conceptualize things**. I reject this proposal, which requires positing differences in “conceptualization” when there is absolutely no independent evidence for it (American and British M LETTUCE ‘salad green’, but American M LETTUCE ‘lettuce plant’ vs. British C LETTUCE ‘lettuce plant’).

6. Re (E): Competing principles have to do with linguistic form, either entirely (e.g. 6.1) or in part (e.g. 6.2).

6.1. IH, Inheritance from Head: a N-N compound inherits the C/M classification of its head N. FOXGLOVE and MONKSHOOD are C, though their synonyms DIGITALIS and ACONITE, respectively, are M. (Similarly, the directly converted plant names HONESTY and IMPATIENCE are M.) IH predicts that ICE PLANT should be C, and it is for many speakers/writers, including the editors of the *Sunset New Western Garden Book* (1979 edition):

A seacoast trademark, dazzlingly brilliant ice plants (these are *Lampranthus*) thrive on sandy soil. (p. 157)

IH also predicts that E-MAIL/EMAIL should be M, and it is for many people.

6.2. IPP, Inheritance from Principal Product: a N denoting a plant inherits its C/M classification from a homophonous N denoting its principal product. IPP predicts that the culinary herb plant names TARRAGON and SAGE are M, that the medicinal herb plant names DIGITALIS

and ACONITE are M, that the names of plants grown for salad greens, like KALE SPINACH ROMAINE are M, that the names of plants grown for their large edible roots, like CARROT BEET POTATO, are C, that the names of plants grown for their showy flowers, like ROSE TULIP LILY, are C, etc.

6.3. Principles like IH and IPP – there are a number of others – together with taCsiM provide a motivation (often, multiple motivations) for the C/M classification of English nouns in general; I agree with W that these assignments aren't arbitrary, but unlike W I appeal to more than taCsiM, even taCsiM generously interpreted as following from “conceptualizations”.

7. Re (F):

7.1. Plants like box and privet, which occur in nature as easily separable individuals, are referred to by M nouns, because their conventional use is in overlapping or mingled hedges; but rose plants, which can be so grown but are not mostly so grown – it's not what they're for – are referred to by a C noun. And then there's IPP, which predicts, inter alia, that leafy green vegetable plant names like SPINACH and LETTUCE, whose referents occur in nature as easily separable plants comprising easily separable leaves, are M, because their conventional use is in mélanges, in salads or cooked greens.

7.2. If the referent of an M noun has individual parts that are not usually salient but can be – for RICE CONFETTI SAND vs. WINE WIND OIL, for example – then many speakers can treat the noun as semantically countable: *We counted the rice/confetti* ‘We counted the grains of rice/pieces of confetti’. But it's still syntactically M: *We counted rice/confetti for hours* (with bare N). Such facts argue against the idea that M nouns are “conceptualized” as indivisible substances.

8. Re (G):

8.1. M to C: ‘kind or type’ (*We sampled three German beers*), ‘serving or unit’ (*Bring us three beers – Anchor Steam*)

8.2. C to M: *There was a lot of dog on the road, ugh*

9. Two new observations:

(H) C/M is **not an opposition of contradictories**.

(I) Double C/M classification is **motivated**.

10. Re (H):

10.1. For syntactic category distinctions like Main/Auxiliary and for paradigm class distinctions like the English regular *d*-past in *seemed* vs. the irregular *t*-past in *slept*, the overwhelming preference is for opposite categorizations (every lexeme has either one classification or the other, but not both), but some lexemes can be categorized both ways, in the language of a single speaker:

a. British possessive HAVE as both Main V (*Do you have any wool?*) and Auxiliary V (*Have you any wool?*);

b. DREAM as both *d*-past (*dreamed*) and *t*-past (*dreamt*).

10.2. Compare the classification of E-MAIL/EMAIL, SPAM, and CHAD, as both M and C for many current speakers/writers, as in these electronic examples:

a. [Posting to a mailing list] Yes, conversational! I say blame it on **email** [M]! I'm 34. Started emailing in 1992. Outside of **email** [M] I am sort of a grammatical purist... My **emails** [C] are *filled* w/ comma splices & run-on sentences... I have always said that I have no linguistic spine... This spinelessness extends to **email** [M].

b. [Posting to a mailing list] [NN]'s mail was a **spam** [C]. **Spam** [M] is junk... Publicity by **email** [M] to unknown persons is **SPAM** [M]... If anyone sends an **email** [C] to [site name] saying that [NN] had sent a **spam** [C] to all research scholars...

c. [Guardian web site, "Election glossary" of 11/6/00] The presence or otherwise of **chad** [M], in those states which use the punchcard ballot system, is a key factor in the current legal disputes over recounting. There are [sic] a whole family of **chads** [C]... Pregnant chad: **A chad** [C] which is bulging or pierced...

d. Add to these: Steve Curl editorial cartoon of 12/15/00, involving two characters, both chewing stuff from bags labeled "Genuine FLORIDA DIMPLED **CHADS** [C]", one of them saying:

DESPITE THE ELECTION FINALE,
MEDIA PUNDITS PRESS ON...
IT WAS ONE WAY TO RECYCLE
BALLOT EFFLUENT. THIS
CAME TO BE KNOWN AS
"CHEWING **CHAD** [M]".

10.3. Current state of my knowledge about C/M variability for these nouns:

a. E-MAIL/EMAIL is M only for many speakers (I am one, and Emma Pease is another), but doubly classified by many others (see above). So far as I know, there are no speakers for whom it is C only. Everyone recognizes/produces *You have (some) e-mail*.

b. SPAM is M only for many speakers (I am one, and Emma Pease is another), doubly classified by some (see above), and C only for others (on at least one Indiana University website, an article dated 7/18/95).

c. CHAD seems mostly to be doubly classified, though there might be people with C only (the earliest citations (currently known), found by Gerald Cohen in U.S. telegraphy patent applications from the late 30s/early 40s, are entirely C) and people with M only (my first experience with the noun was in instructions from my local (California) voting district, and they involved a clearly M use).

10.4. Comparison to ICE PLANT, which seems to be either C or M for any particular speaker.

10.5. This variation is not a matter of alternative grammars (or of alternative "conceptualizations" determining alternative grammars), but of alternatives within one grammar, like the choice of a complement clause marked by *that* vs. an unmarked complement clause. Shifts occur within a discourse, even within a single sentence (cf. Bresnan & Deo 200 on variation in the paradigms of BE in British English).

10.6. The variation seems to be lexeme-specific. The classification of CHAD and the classification of SPAM seem to have nothing to do with one another. Despite their close semantic relationship, SPAM and E-MAIL/EMAIL don't necessarily have the same classification(s) (though it might be true that anyone with a C use of SPAM also has a C use of E-MAIL/EMAIL). In any case, there doesn't seem to be a lexeme-general principle or constraint here.

11. Re (I):

- 11.1. Sometimes it's a matter of competing predictions: for ICE PLANT, taCsiM predicts M, but IH predicts C; for the plant name LETTUCE, taCsiM predicts C, but IPP predicts M (via taCsiM, which predicts M for the principal product, salad greens). Insofar as the conflict is between semantic and purely formal principles, speakers tend to opt for one or another of the alternatives – apparently, on the basis of whichever model they heard first (as in children's acquisition of the C/M contrast – Gordon 1985, Hall 1994).
- 11.2. But for some other nouns, it's the application of taCsiM itself that's at issue. Some referents are *semantically bivalent* with respect to taCsiM; in our usual dealings with these referents, they are saliently **both stuff and** (collections of) things (rather than being “conceptualized” as **either stuff or** things). This is the case for the variable CHAD SPAM E-MAIL/EMAIL and also for the non-variable MAIL INFORMATION RICE BARLEY LENTIL BEAN.
- 11.3. For some semantically bivalent nouns, speakers seem to settle on a double C/M classification, so that E-MAIL (bivalent with respect to taCsiM, but M by IH), SPAM (bivalent with respect to taCsiM), and CHAD (exquisitely bivalent with respect to taCsiM, functioning both as mere confetti in a pile and as items that have to be separately counted) are doubly classified by many speakers, who then produce clearly C and clearly M occurrences of these nouns within a (short) single discourse.
- 11.4. What determines which semantically bivalent nouns are doubly classified and which are singly classified? Apparently, the age at which you experience the noun in use. Kids will use taCsiM to classify nouns, but actual input trumps evidence of thingness or stuffness, and the classifications are pretty much frozen thereafter; MAIL INFORMATION RICE BARLEY are M, LENTIL BEAN are C, that's the way other people talk, and that's just the way things are. Adults confronting nouns that are new to them seem to be more open to exploiting semantic bivalence in their classifications.

12. Side issue:

- 12.1. Non-linguists often report that *chad* or *spam* or *e-mail* is “plural” for them, or that they've heard it used as plural, as in the following quotation from William Safire's “On Language” column, *New York Times Magazine*, 12/10/00, section on the Words of the Year for 2000, *chad* (p. 68):
- ...according to Peter Graham, now university librarian at Syracuse, who served early in his career as a key-punch operator: “We had what we called a *chad* box underneath the key punch. We resisted calling it ‘confetti’ because the small bits of paper, when they caught on your clothes, would not dislodge.” Graham notes that the noun was then construed as plural, on the analogy of *chaff*, but today's ballot counters are referring to *chads*, construing the word *chad* as singular. But this is a confusion between plural and mass (CHAFF itself is mass, not plural); non-linguists know the technical vocabulary *singular/plural* but not *count/mass* and use the only term they know that has roughly the right meaning.
- 12.2. But there might well be genuinely zero-plural uses of CHAD (parallel to alternations between *reindeer* and *reindeers* as plurals of REINDEER, or possibly to zero-plural plural-only nouns like CATTLE and POLICE) as in this AP wire story from 11/28/00 (supplied by Lynne Murphy on the American Dialect Society mailing list):

Chad are the tiny pieces of paper that pop out of a ballot when a voter chooses a candidate.

(The crucial point is the plural verb form *are*.) It remains to be seen whether the person who wrote this example would also accept sentences like *There were many chad on the floor* and *Only two chad were left to count*. In any case, outside of the domain of names of animals hunted or fished for sport, zero plurals are quite rare, so that you wouldn't expect very many speakers to interpret sentences like *There will be a lot of chad on the floor* or *We ate all the chad* as involving a zero plural rather than a (singular) mass noun.

References

- Bresnan, Joan & Ashwini Deo. ms. 2000. On the grammar of variable agreement: *Be* in the *Survey of English Dialects* and stochastic Optimality Theory.
- Gordon, Meter. 1985. Evaluating the semantic categories hypothesis: The case of the count/mass distinction. *Cognition* 20.209-42.
- Hall, D. Geoffrey. 1994. How children learn common nouns and proper names: A review of the experimental evidence. John Macnamara & Gonzalo E. Reyes (eds.), *The logical foundations of cognition* (Oxford: Oxford Univ. Press, 1994), 212-40.
- Wierzbicka, Anna. 1985. "Oats" and "wheat": The fallacy or arbitrariness. John Haiman (ed.), (ed.) *Iconicity in syntax* (Amsterdam: John Benjamins, 1985), 311-42. Reprinted in Wierzbicka 1988: 499-560.
- 1988. *The semantics of grammar*. Amsterdam: John Benjamins.