

Varieties of mass/count interpretation of hybrid nouns

Kurt Erbach¹ & Yasutada Sudo²

¹Heinrich-Heine-Universität Düsseldorf & ²University College London

Divide and Count Workshop, DGfS 2023

Formal semantics of mass/count

The dominant view of mass/count in the current literature

- **COUNT** semantically encodes **individuation**
- **MASS** is compatible with (loose) individuation

We can't simply say **MASS** = not individuated, because of **object mass nouns** (alt. fake mass nouns), e.g. *furniture, equipment*

Many different theories of **MASS**, e.g.

- Bale & Barner (2009); **COUNT** = individuation; **MASS** = no meaning
- Chierchia (1998, 2010, 2015, 2021): **COUNT** = stable atoms; **MASS** = vague atoms
- Landman (2016): **COUNT** = disjoint, individuated 'base'; **MASS** = overlapping 'base'

Flexible individuation

Based on nouns like *wall* and *fence*, it is suggested that the notion of individuation that **COUNT** encodes must be sufficiently 'flexible' (Sutton & Filip 2016; Rothstein 2010, 2017)

- (1) The bedroom has a white wall and a wooden floor.
- (2) The bedroom has six walls.

E.g. Rothstein proposes that **COUNT** has a context-sensitive denotation:

$$\llbracket \mathbf{COUNT} \sqrt{N} \rrbracket^c = \{ x \mid x \in \text{atom}(\llbracket \sqrt{N} \rrbracket^c, c) \}$$

- For *wall*, it's easy to find contexts that non-trivially differ with respect to atoms
- For *boy*, it's much harder for extra linguistic reasons



The decompositional approach

A common feature of previous formal semantic theories of mass/count is **decomposition**

Nouns are syntactically decomposed. Roughly:

$$N_{\text{count}} = \mathbf{COUNT} + \sqrt{N} \qquad N_{\text{mass}} = \mathbf{MASS} + \sqrt{N}$$

These sub-components are assigned model-theoretic denotations:

- $\llbracket \mathbf{COUNT} \rrbracket$ introduces flexible individuation
- $\llbracket \mathbf{MASS} \rrbracket$ is analysed differently by different theories
- $\llbracket \mathbf{PL} \rrbracket = *$ -operator (closure under $\sqcup$)

Limits of the decompositional approach

Assuming **compositionality** we want **COUNT** and **MASS** to have constant denotations that can apply to all nouns

But this has led to 'generalisation to the most bleached case':

- **COUNT** universally encodes flexible individuation (due to *wall, fence*)
- **MASS** universally has very weak meaning (due to object mass nouns)

We argue against the decompositional approach based on variation among **hybrid nouns** (alt. flexible nouns, dual-life nouns) in Midwest American English

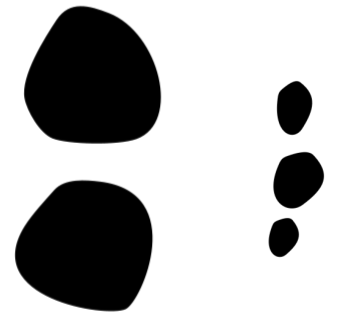
Hybrid nouns

Hybrid nouns can be used as mass or count, e.g. *stone, rope, apple, chocolate*

Mass/count often has a clear interpretative effect for hybrid nouns

E.g., Barner & Snedeker's (2005) Comparative Task (see also Bale & Barner 2018)

- (3) Who has more stone_{mass}? ...> **volume-based comparison**
- (4) Who has more stones_{count}? ...> **number-based comparison**



We'll point out that there's more to the semantics of hybrid nouns

Our observations and claims

Key observations: Variety of count hybrid nouns

1. Some count hybrid nouns are not so flexible with respect to individuation (e.g. *apple*)
2. Some plural count hybrid nouns are compatible with no individuation (e.g. *potatoes*)

Proposal: Renounce decomposition in terms of **COUNT** and **MASS**, and characterise interpretations at the level of whole nouns

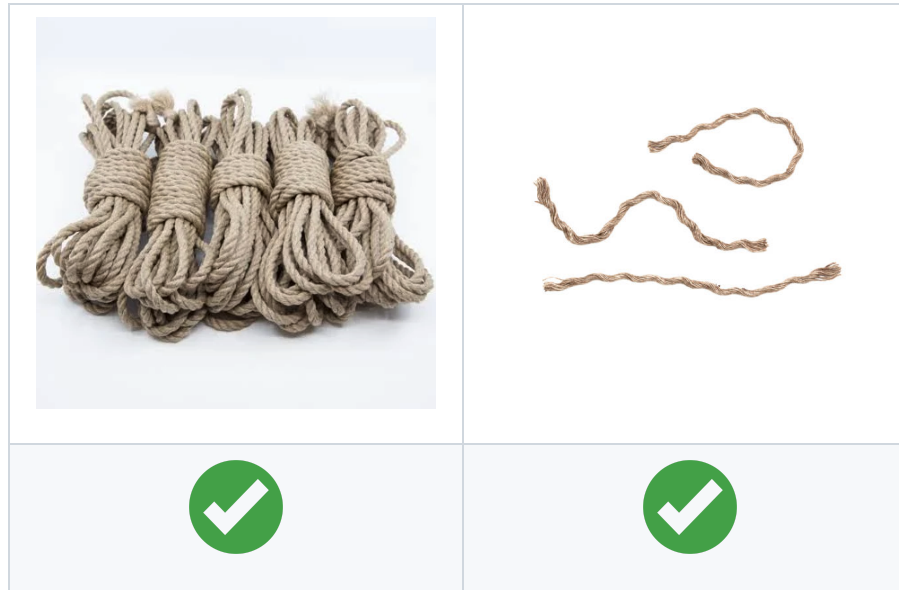
- N_{count} may encode an idiosyncratic manner of individuation $\Rightarrow$ 1.
- $N_{\text{pl.count}}$ and N_{mass} have trivial interpretations, only if the other one is infrequent $\Rightarrow$ 2.

Varieties of individuation

Flexible individuation: *rope*

*Rope*_{count} can describe all sorts of perceptually countable instances, similarly to *wall*

(5) We've got some ropes here.

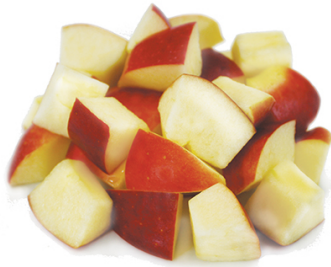


(6) Let's cut one of the ropes into smaller ropes.

Specific individuation: *apple*

*Apple*_{count} can only describe whole apples, not sliced or diced instances

(7) We have some apples here.

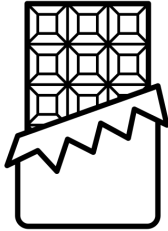
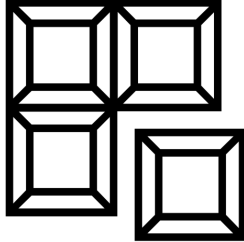
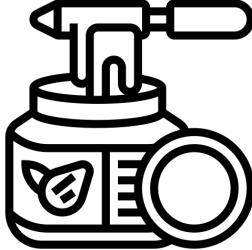
			
			

(8) #Let's cut one apple into smaller apples.

Specific individuation: *chocolate*

*Chocolate*_{count} can describe a chocolate truffle, but not a bar or an arbitrary piece

(9) I will give you a chocolate.

(10) #Let's break a chocolate into smaller chocolates.

Interim summary

As expected, $apple_{\text{count}}$ and $chocolate_{\text{count}}$ cannot describe unindividuated instances

But they also cannot describe arbitrary pieces, even if clearly perceptually individuated

☞ These count nouns encode **specific individuation** (unlike *wall* and *rope*)

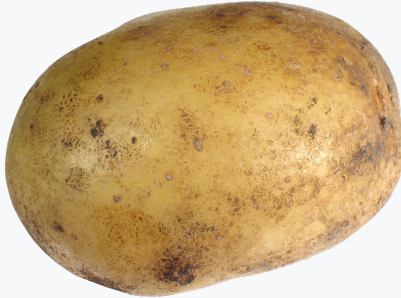
- $\llbracket \mathbf{apple}_{\text{sg.count}} \rrbracket = \{x | x \text{ is a whole apple}\}$
- $\llbracket \mathbf{chocolate}_{\text{sg.count}} \rrbracket = \{x | x \text{ is a chocolate truffle}\}$
- $\llbracket \mathbf{wall}_{\text{sg.count}} \rrbracket^c = \{x | x \text{ is one wall in } c\}$
- $\llbracket \mathbf{rope}_{\text{sg.count}} \rrbracket^c = \{x | x \text{ is a perceptually salient chunk of rope in } c\}$

(The encoded specific individuation is probably not completely random but can be idiomatic like *chocolate*)

Plural count hybrid nouns

An apple vs. a potato

A potato is similar to an apple

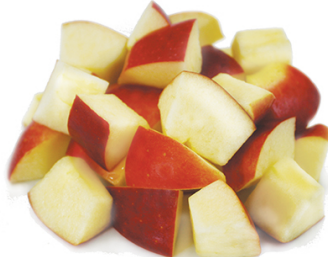
(11) I'll give you an apple.			(12) I'll give you a potato.	
				
				

(In child-directed speech, reference to non-whole instances might be possible)

Apples

Given how *an apple* behaves, *apples* can be understood in terms of the $*$ -operator

(13) I'll give you some apples.

(14) Who ate more apples?

... number-based comparison

Potatoes

Potatoes is not simply sums of singular potatoes: $\llbracket \text{potatoes} \rrbracket \neq *(\llbracket \text{potato}_{\text{sg.count}} \rrbracket)$

(14) I'll give you some potatoes.



(15) Let's cut one of these potatoes into smaller potatoes.

(16) Who ate more potatoes? $\leadsto$ **both volume and number comparisons possible**

- **Flexible individuation, *-plural** (*rope-class*)
rope, paper, brick, wire, cable, string, hair, (sub)sandwich, fibre, oat, plank, board, pipe, steak, talk, meeting, exercise
- **Specific individuation, *-plural** (*apple-class*)
apple, chocolate, lemon, banana, eggplant, artichoke, hamburger, pizza, cake, beard, chicken, duck, mango, candy, song, show, movie, bone
- **Specific individuation, mass-plural** (*potato-class*)
potato, strawberry, carrot, leaf, french fries, chickpea, pea, potato chip, lentil, tomato, noodle, blueberry, pebble
- **Flexible individuation, mass-plural** (*cloud-class*)
cloud, wind, detail

Towards an analysis

Key observations

Two observations about count hybrid nouns:

1. **Specific individuation**

$N_{\text{sg.count}}$ may encode a specific manner of individuation (e.g. *apple, chocolate*)

2. **Plural count nouns with mass denotations**

$N_{\text{pl.count}}$ may have a mass-like denotations (e.g. *potatoes, clouds*)

1. Specific individuation

All singular count nouns encode individuation, but some encode specific manners

☞ A generic individuation function for 'Spelke objects' will be too permissive; a Natural Unit function will be too restrictive

Idea: The interpretation of $N_{\text{sg.count}}$ is not compositionally derived from $\sqrt{N}$ and **COUNT**

- $\llbracket \text{apple}_{\text{count}} \rrbracket = \{x | x \text{ is a whole apple}\}$
- $\llbracket \text{potato}_{\text{count}} \rrbracket = \{x | x \text{ is a whole potato}\}$
- $\llbracket \text{chocolate}_{\text{count}} \rrbracket = \{x | x \text{ is a chocolate truffle}\}$
- $\llbracket \text{wall}_{\text{sg.count}} \rrbracket^c = \{x | x \text{ is one wall in } c\}$
- $\llbracket \text{rope}_{\text{sg.count}} \rrbracket^c = \{x | x \text{ is a perceptually salient chunk of rope in } c\}$

2. Plural count nouns with mass denotations

Potatoes can describe all sorts of instances, unlike *apples*

NB: *potato*_{mass} is also available

(17) There's some (mashed) potato on your shirt.

Proposal: What differentiates *potato* and *apple* is the relative frequency of N_{mass}

- Grimm & Wahlang's (2021) corpus data (350 million words) with Allan's (1980) diagnostic for countability
- *Potato* is mostly marked count (84.3%) vs. *apple* (61.4%)
- 10 most skewed hybrid nouns: *pickle, dumpling, sprout, carrot, bouquet, leaf, pea, potato, beet, egg*

Frequency effects and RSA with Lexical Uncertainty

Idea: When N_{mass} is relatively infrequent, $N_{\text{plural.count}}$ can be used to mean the same thing

We use a Rational Speech Act (RSA) model with Lexical Uncertainty (Bergen, Levy & Goodman 2016, Scontras & Goodman 2017)

- $N_{\text{pl.count}}$ has two potential denotations, **PLURAL**: $\ast(\llbracket N_{\text{sg.count}} \rrbracket)$ vs. **MASS**: $\llbracket N_{\text{mass}} \rrbracket$
- When $N_{\text{pl.count}}$ is used, the listener probabilistically guesses what the intended interpretation is, based on cooperativity
- The relative frequency of N_{mass} (implemented as speaker cost) and the frequency of intended meaning affect the reasoning
- If N_{mass} is not costly, cooperative speaker avoids ambiguous N_{plural} to mean **MASS**
- Details in Appendix

Further directions

1. Symmetric frequency effects of mass/count

- If N_{mass} is infrequent, N_{plural} can mean the same thing  *potatoes*
- If N_{plural} is infrequent, N_{mass} can mean the same thing  object mass nouns

2. We expect cross-linguistic/dialectal variation and language change over time

- Frequencies change for many reasons
- Very specific idiosyncratic meanings of singular count nouns can fail to be acquired

"Spearman, and others [...], carried out **many ingenious researches** using mental tests and guided by his 'two factor' hypothesis."

(Godfrey Thomson, 1947, "Charles Spearman, 1863–1945", doi.org/10.1098/rsbm.1947.0006)

Summary

Summary

Key observations: Count hybrid nouns show

1. Interpretive variation with respect to individuation

- **Flexible individuation:** *rope, string, paper*
- **Specific individuation:** *apple, chocolate, potato*

2. Unexpected mass meaning in the plural for some: *potatoes, clouds, lentils*

Proposal

- Singular count nouns may encode compositionally opaque individuations
- The denotations of plural count nouns are computed via competition with mass

The observations pose challenges for the traditional compositional approach

Appendix

RSA with Lexical Uncertainty for **PLURAL** vs. **MASS**

RSA with Lexical Uncertainty

We assume that N_{plural} has two potential interpretations, **PLURAL** and **MASS**

We implement this ambiguity in terms of a parameter p on $\llbracket \cdot \rrbracket$ (Bergen, Levy & Goodman 2016, Scontras & Goodman 2017)

$$\llbracket N_{\text{plural}} \rrbracket^p = *(\llbracket N_{\text{sg.count}} \rrbracket)$$

$$\llbracket N_{\text{plural}} \rrbracket^m = \llbracket N_{\text{mass}} \rrbracket$$

RSA with Lexical Uncertainty

- The literal listener L_0 infers what interpretation i is intended based on the utterance u and a parameter x on $\llbracket \cdot \rrbracket$

$$L_0(i|u, x) \propto P(i) \cdot \llbracket u \rrbracket^x(i)$$

- The n th speaker S_n picks out the best message for i given x , relative to L_{n-1}

$$S_n(u|i, x) \propto e^{\lambda \cdot (\log L_{n-1}(i|u, x) - \text{cost}(u))}$$

- Infrequent expressions are more costly
- The n th listener L_n decodes u relative to S_n

$$L_n(i, x|u) \propto S_n(u|i, x) \cdot P(i) \cdot P(x)$$

Simulation

