

Countability and Nominal Ambiguity – Mining Polysemy in BECL

Tibor Kiss

Ruhr-Universitaet Bochum
tibor@linguistics.rub.de

F. Jeffry Pelletier

University of Alberta
francisp@ualberta.ca

Halima Husić

Ruhr-Universitaet Bochum
husic@linguistics.rub.de

Johanna M. Poppek

Ruhr-Universitaet Bochum
poppek@linguistics.rub.de

While a large number of works on the count/mass distinction is agnostic towards the role and relevance of nominal ambiguity, recently, a growing body of work has considered the implications of lexical ambiguity to the count/mass distinction, as is witnessed in Payne and Huddleston (2002: 334ff.), Borer (2005), Falkum (2010), Pelletier (2012), Kiss, Pelletier, and Stadtfeld (2014), Kiss et al. (2016), and Kiss et al. (2017). On the other hand, several influential papers on the (formal) semantics of count and mass nouns do not cover lexical ambiguity at all (cf. Link (1983), Chierchia (1998), and Rothstein (2010)).

In the present paper, we would like to distinguish three types of lexical ambiguity, two of which are not really interesting when the count/mass distinction is considered. The third one, however, we consider highly relevant. The reason is as follows:

The first two types of lexical ambiguity are homonymy (1a, b), where there is no relation between two (or more) senses, and polysemy (2a, b), where there is a relation between the two senses that can be expressed at different levels of generality (the relation can indeed be subdivided, and certain divisions are more productive than others).

- (1) a. *chicken#1* (the flesh of a chicken used for food) (mass)
b. *chicken#3* (a person who lacks confidence, is irresolute and wishy-washy) (count)
- (2) a. *chicken#1* (*the flesh of a chicken used for food*) (mass)
b. *chicken#2* (a domestic fowl bred for flesh or eggs; believed to have been developed from the red jungle fowl) (count)

In both cases, the respective classification of the noun-sense pair being count or mass is unproblematic as long as the different senses are treated separately.

In the final case, the count/mass distinction, however, is the very locus of the polysemy. This is interesting insofar as it may tell us something about the (semantic) nature of the count/mass distinction in itself.

Homonymy will presumably not shed light on the count/mass distinction because it is a defining feature of homonymy that we find a coupling of unrelated meanings to the same form. As an illustration, consider Payne and Huddleston's (2002: 334f.) examples of count/non-count polysemy (which includes pairs, which we would call homonymous): it includes cases of support verb construction (3a), as well as the contrast between a nominal and a participial meaning (3b), among others.

- (3) a. This proposal has three advantages. / They took advantage of us.
b. It is certainly a fine building. / There's plenty of building going on.

At this point, it should be stressed that the very concept of a sense is by no means fail-safe. To begin with, senses of open word classes do not usually come neatly packaged and separated from other senses. What is more, it becomes necessary to rely on lexical resources providing such senses. As Kilgarriff (1997) has pointed out, this is not without pitfalls. Consider the following illustration in light of the question whether a *Universal Grinder* should be considered part of an analysis of the count/mass distinction. Consider the definitions of the word **moose** in the Oxford Dictionary of Current English (ODCE), the American Oxford Dictionary (AOD), and the Canadian Oxford Dictionary (COD). The ODCE restricts itself to one sense of the word: “a ruminant mammal, the largest living deer ... with very large antlers”. The AOD adds a second – clearly homonymous – sense, not found in the other two: “a major benevolent fraternal order”. The COD, however, replaces the American second with: “[t]he flesh of the moose as food”, a sense which is again not found in the other two dictionaries. Of course, an analysis of the count/mass distinction, which would flatly rely on the senses provided by the three dictionaries, would perhaps come to the conclusion that **moose** in an example like *We had moose for dinner.* exemplifies polysemy at the very level of the count/mass distinction – a conclusion that should not be drawn easily.

In the following, we would like to illustrate how we addressed the issue of polysemy in BECL (Kiss, Pelletier, and Stadtfeld 2014, Kiss et al. 2016). With regard to the definition of senses, BECL assumes a pragmatic approach: the data analyzed in BECL has been derived from frequently occurring nouns in the Open American National Corpus (www.anc.org), for which sense definitions are provided in WordNet (Fellbaum 1998). It should be noted here that even this pragmatic reliance on WordNet does not necessarily mean that all senses come neatly cut and packaged.

In an annotation task, four speakers of Canadian English were provided with four syntactic and two further semantic tests for pairs of word forms and senses. They had to determine whether a) the word with the specific sense may occur in the syntactic contexts provided by the four syntactic tests, and b) depending on the outcome of the first two syntactic tests had to answer two semantic questions about the resulting constructions. We have collected the independent annotations and classified them so that four major, and 18 minor countability classes could be identified. A combination of a word and a sense was included in the classification if it received identical annotations from at least two independent annotators. The four major classes are: a) regular count noun-sense pairs, b) regular mass noun-sense pairs, c) noun-sense pairs that are count and mass, and d) noun-sense pairs that are neither count nor mass.

Cases of homonymy can be identified by inspecting all the noun-forms that end up in different classes, according to this procedure. As an illustration, consider two sense definitions of the word *accommodation* (WordNet sometimes provides examples, which are included below).

- (4) a. *accommodation#3*: in the theories of Jean Piaget: the modification of internal representations in order to accommodate a changing knowledge of reality (BECL 528, regular mass)
- b. *accommodation#2*: a settlement of differences; “they reached an accommodation with Japan” (BECL 235, regular count)

Of course, knowing that *accommodation* can be both a count and a mass noun does not reveal much, once it is clear that we are talking about two semantically rather different words sharing the same form, i.e. homonyms.

Polysemy of course is different, as is illustrated by the examples below:

- (5) a. *certainty#2*: something that is certain; "his victory is a certainty" (BECL 235, regular count)
- b. *certainty#1*: the state of being certain; "his certainty reassured the others" (BECL 528, regular mass)
- (6) a. *charity#1*: a foundation created to promote the public good (not for assistance to any particular individuals) (BECL 235, regular count)
- b. *charity#2*: a kindly and lenient attitude toward people (BECL 528, regular mass)
- (7) a. *classification#2*: a group of people or things arranged by class or category (BECL 235, regular count)
- b. *classification#3*: the basic cognitive process of arranging into classes or categories (BECL 528, regular mass)

The cases of polysemy provided in (5), (6), and (7) show instances of a general process of individuation, so that the count sense can always be derived from the mass sense in the form of the Appropriate Individuation Function (AIF) in (8):

- (8) Appropriate Individuation Function: If X is a noun with a mass sense [[X]], then X with a count sense is to be interpreted as an individuated instance for which [[X]] is true.

The AIF in (8), which surely needs a more formal characterization, including all its relevant parameters, is a generalization of polysemy rules like the *Universal Sorter* and the *Universal Packager*. An interesting question emerges from such pairs that can presumably be discussed more lucidly by considering nouns with concrete interpretations, such as *beer*:

- (9) a. I don't like *beer*. (mass)
- b. She offered me another *beer*. (count)

With regard to such contrasts, Payne and Huddleston (2002:336) say: "Names of drinks are primarily [mass], but systematically allow count interpretations where the noun denotes a serving of the drink ... This is one of the regular cases where the secondary sense ... is predictable and need not be listed."

Although this is implicit in Payne and Huddleston's remark, it should be made explicit that such pairs are not only regular (and hence in no need to be listed), but also that they are less interesting from the perspective of the count/mass distinction, because the two senses are not only clearly different from each other, but are very obviously mass, and count, respectively. The noun in (9a) refers to a substance, while the noun in (9b) refers to a portion of a substance. Leaving all ontological questions aside, it is neither a surprise that the second noun – denoting an individuated entity – is a count noun, nor that the first one is not. In this respect, we are not actually gaining anything from including senses into an analysis from the count/mass distinction, and thus polysemy in the sense made explicit above does not really form a challenge for current semantic theories of the count/mass distinction, such as Chierchia (1998) and Rothstein (2010).¹ Of course, the data supports the assumption that a rule-based mechanism such as AIF must be included in an analysis of the count/mass distinction (cf. also Poppek et al. 2018).

As an intermediate conclusion, one could state that here we find issues of polysemy but that the polysemy does not target the count/mass distinction in itself.

The final class of examples we want to consider strongly differ from the ones discussed so far. They are members of BECL class 726 that already shows some peculiarity with regard to

¹ We cannot currently answer the question whether *another beer* is a naturally or semantically atomic entity.

the syntactic contexts in which they may appear and also with regard to the semantic properties. Noun-senses that belong to this class have the following characteristics:

1. In a construction with more X[SG] the insertion of this noun-sense is grammatical, and the comparison in such a sentence is not based on number, which means that this noun is not an object mass noun.
2. In a construction with more X[PL] the insertion of this noun-sense is grammatical, and the resulting construction is semantically not equivalent to one with a classifier, which means that the plural marking does not imply a hidden classifier.
3. The noun-sense[SG] can be combined with the indefinite article as the syntactic subject of a definition.
4. The noun-sense[SG] can occur without a determiner as the syntactic subject of a definition.

As an illustration, consider the following noun-sense pairs (all from BECL 726):

- (8)
- a. *accumulation*#3: the act of accumulating
 - b. *ambition*#2: a strong drive for success
 - c. *burial*#2: concealing something under the ground
 - d. *cake*#3: baked goods made from or based on a mixture of flour, sugar, eggs, and fat
 - e. *casserole*#1: food cooked and served in a casserole
 - f. *confirmation*#1: additional proof that something that was believed (some fact or hypothesis or theory) is correct; "fossils provided further confirmation of the evolutionary theory"
 - g. *contradiction*#3: the speech act of contradicting someone; "he spoke as if he thought his claims were immune to contradiction"
 - h. *exploration*#1: to travel for the purpose of discovery
 - i. *perversion*#3: the action of perverting something (turning it to a wrong use); "it was a perversion of justice"
 - j. *revival*#1: bringing again into activity and prominence; "the revival of trade"; "a revival of a neglected play by Moliere"; "the Gothic revival in architecture"
 - k. *swamp*#1: low land that is seasonally flooded; has more woody plants than a marsh and better drainage than a bog

What is interesting about these noun sense pairs is that they show up in syntactic contexts for mass and count nouns without an apparent difference in meaning (hence only one relevant sense is provided). They are thus true instances of so-called dual-life nouns (cf. also Kiss et al. 2017).

They may occur in a comparative construction with *more* in plural form, as well as in the singular form. Thus *more swamp* (mass) would imply a greater/bigger swamp while *more swamps* (count) would imply more individuated swamps. The former comparison is based on volume, the latter on the number of entities. What makes this case of polysemy special is that the count and mass properties are unified in one single sense. Furthermore, such noun-senses may appear together with an indefinite determiner, as well as without a determiner at all, i.e. bare.

Appearing in the context *more*[PL] and the combination with an indefinite determiner are hallmarks of countability, while the context *more*[SG] and the lack of any determiner suggest

a classification as a mass noun. This observation is complemented in that the first semantic test suggests a classification as a mass noun, while the second semantic test clearly indicates that the count interpretation is not combined with a meaning shift inherent in AIF.

So, finally, we have arrived at a case of polysemy, which refers directly to the count/mass distinction. If the count/mass distinction is a semantic one, then a systematic investigation of nouns of this type (true dual-life nouns) may allow deeper insights in the nature of this distinction.

References:

Borer, H. (2005). *Structuring Sense. Vol. I: In Name Only*. Oxford: Oxford University Press.

Chierchia, G. (1998). Plurality of mass nouns and the notion of 'semantic parameter'. *Events and Grammar*.

Falkum, Ingrid Lossius (2010). Systematic polysemy and the count-mass distinction. *UCL Working Papers in Linguistics*.

Fellbaum, C. (ed.) (1998) *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.

Kiss, T., Pelletier, F.J., Husic, H., Poppek, J. M. and Simunic, R.N. (2016). A Sense-based Lexicon of Count and Mass Expressions: The Bochum English Countability Lexicon. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation LREC 2016*. Portoroz, Slovenia.

Kiss, T., Pelletier, F. J., Husic, H. and Poppek, J. M. (2017). Issues of Mass and Count: Dealing with "Dual Life" Nouns. In *Proceedings of the 6th Joint Conference on Lexical and Computational Semantics (ACL-*Sem 2017)*. 189-198. Vancouver, Canada

Link, G. (1983). The Logical Analysis of Plurals and Mass Terms: A Lattice Theoretic Approach. In R. Bauerle, C. Schwartze and A. von Stechow (Eds.), *Meaning, Use and Interpretation of Language*. Berlin: De Gruyter.

Payne, J. and Huddleston, R. (2012). Nouns and Noun Phrases. In R. Huddleston and G.K. Pullum (eds.), *The Cambridge Grammar of the English Language (5th ed.)* (pp. 323-523). Cambridge: Cambridge University Press.

Poppek, J. M., Kiss, T., and Pelletier, F.J. (2018). Kinds, Containers, Instances: Mass Nouns and Plurality. To appear in *Proceedings of Grammar and Corpora 2016*. Heidelberg: Heidelberg University Publishing.

Rothstein, S. (2010). Counting and the mass count distinction. *Journal of Semantics*, 27 (3), 343-397.