Building a Reference Lexicon for Countability in English

Tibor Kiss, Francis Jeffry Pelletier, Tobias Stadtfeld

Ruhr-Universität Bochum, Simon Fraser University, Ruhr-Universität Bochum
Bochum Germany, Vancouver Canada, Bochum Germany
tibor@linguistics.rub.de, jeffpell@sfu.ca, stadtfeld@linguistics.rub.de

Abstract

The present paper describes the construction of a resource to determine the lexical preference class of a large number of English nouns ($\approx$ 14,000) with respect to the distinction between mass and count interpretations. In constructing the lexicon, we have employed a questionnaire-based approach based on existing resources such as the Open ANC (`http://www.anc.org`) and WordNet (Miller, 1995). The questionnaire requires annotators to answer six questions about a noun-sense pair. Depending on the answers, a given noun-sense pair can be assigned to fine-grained noun classes, spanning the area between count and mass. The reference lexicon contains almost 14,000 noun-sense pairs. An initial data set of 1,000 has been annotated together by four native speakers, while the remaining 12,800 noun-sense pairs have been annotated in parallel by two annotators each. We can confirm the general feasibility of the approach by reporting satisfactory values between 0.694 and 0.755 in inter-annotator agreement using Krippendorff's α.

Keywords: countability, mass/count distinction, questionnaire-based lexical resource

1. Introduction

While the distinction between count and mass expressions has played a role in linguistics since (Jespersen, 1924), its analyses are usually based on a small set of nouns. What is more, several rules – which are really only rules of thumb – have been stated in the literature to allow the determination of the lexical countability of nouns. Thus (Bunt, 1979), among many others, describes the compliance of a noun with the indefinite determiner as indicating countability (*a car*), while a noun in singular in an amount expression is to be considered uncountable (*much water*). Even writers of more comprehensive works such as (Pelletier and Schubert, 1989) use only eleven examples of mass nouns and eight for count nouns. In addition, (Allan, 1980) has noted more than thirty years ago that there seem to be different "degrees of countability" between nouns. Again, this is not acknowledged, nor disputed, nor reflected in the vast majority of analyses of the count-mass distinction.

Scaling up to a larger set of nouns is neither easily done nor amenable to an unambiguous conclusion that will answer all questions surrounding the topic. For example, certain nouns and types of contexts are particularly susceptible to grinding (Pelletier, 1975), allowing an assumed countable noun (1-a) to be reinterpreted as a mass noun (1-b).

(1) a. This recipe requires an apple. (count)
b. Put more apple into the parfait. (mass)

Conversely, the universal packager (Jackendoff, 1991) in (2-a) and universal sorter (Bunt, 1985) in (2-b) allow a prototypical uncountable noun to appear in the scope of the indefinite determiner or even as a plural, therefore also contradicting the aforementioned rules of thumb.

(2) a. He ordered a beer. (a container filled with beer)
b. The waiter recommended two different wines. (two sorts of wine)

We are well aware that attributing countability to the level of the lemma (or word) might prove to be problematic. An assignment of countability classes to individual senses in individual contexts could lead to even deeper insights. While this is in fact our long-term goal, in devising the present resource, we chose a compromise between a lemma-based approach, and a purely sense-based approach. To implement the lemma-based approach, it is necessary to exclude the aforementioned (and other) contextual influences on the countability reading of a noun (or noun sense) during annotation. Hence, we prepared several thousand nouns for annotation (see Section 2.) and defined restrictive test patterns (see Section 3.) which each noun has to undergo to determine its syntactic and semantic features. In combination, these features explicitly define the countability of a sense of a noun and can be used to implicitly define a notion of lexical countability of a noun. To ensure the overall quality of annotations, we conducted an extensive inter-annotator reliability study (see Section 4.). In Section 5., we present the observed distribution of common countability classes in the data.

As we have already indicated, we see the current resource as a bootstrap for further and finer-grained data-based analyses of the count-mass distinction. By tracking nouns in a large corpus, we hope to answer long-standing questions about, e.g., whether grinding (Pelletier, 1975) is a matter of coercion/construction of one sense into another, or a matter of choosing between existing senses of a polysemous word (Cheng et al., 2008; Djalali et al., 2011; Pelletier, 2012). As a next step, the present countability lexicon forms part of a token-based cluster analysis to determine how lexical contexts and individual senses influence class assignments. We thus hope to provide a sound empirical basis for approaches such as (Borer, 2005; Frisson and Frazier, 2005; Chierchia, 2010; Rothstein, 2010; Pelletier, 2012; Katz and Zamparelli, 2012).

2. Data Preparation and Presentation during Annotation

Since we wanted to capture possible alternations depending on individual senses, we extracted all nouns occurring at least ten times within the Open American National Corpus (Open ANC), for which sense definitions are provided in WordNet.

During the annotation process, annotators were told to give their judgement on a variety of tests (see Secton 3.) for each sense of a noun. The data was presented to them in a spreadsheet. Here, pairs of nouns and individual disambiguated sense description taken from WordNet (examples in Table 1) were displayed along with drop-down menus providing possible answers for each test pattern. The presentation of lexical items itself was randomized. Different senses of lexical items, however, were kept grouped during presentation to simplify the task of differentiating between multiple senses of a noun.

The initial training phase comprised of a group annotation of 1,000 noun-sense pairs to allow annotators to become familiar with the annotation process, and also to identify possible problems. In the second stage, the four annotators received pairwise randomized files for annotation, so that each noun was annotated by two annotators, but that the annotation did not happen at the same time. Discussion among annotators concerning specific test outcomes was forbidden at this stage. We thus ensured that a large subset of senses was annotated by at least two annotators, and that the annotators did not contaminate one another's judgments.

3. The Questionnaire

To reduce the influence of various context variables on the countability of a noun, we have defined restrictive test patterns. Among possible interferences, we find advertisement-writer readings, as illustrated in (3) or interpretations resulting from grinding, as illustrated in (4).

(3) John bought more car for less money than Bill.

(4) John bought more car in respect to overall volume than Bill.

The patterns that we have set up exclude the interpretations arising in (3) and (4). Without such extended contexts, however, a noun like *car*, as listed in its first sense in WordNet, leads to ungrammaticality if inserted into the test pattern of Test I, as is illustrated in (5).

(5) *John bought more car than Bill.

Therefore, any alteration – except where variables are provided – of the test patterns presented below was disallowed. In the following, we represent noun-sense pairs by the combination of the noun and a number prefixed by a # sign. So *car#1* refers to the noun *car* with its first sense given in WordNet.

3.1. Test I

The pattern of the first test is restricted to the following:

The test given in Fig. 1 consists of two steps. Given a noun to be annotated, the annotator choses an appropriate verb

NP_1 VERB more NOUN[sg] than NP_2.

Figure 1: Pattern Test I

and then has to decide first, whether the insertion of a noun into the frame provided in Fig. 1 results in a grammatical or ungrammatical sentence, or the insertion is not applicable for that noun. Inserting *fruitcake#2* and *lingerie#1* into the pattern provided in Fig. 1 leads to grammatical sentences, as is illustrated in (6) and (7), and also by the value *yes* in Table 1. In the case of (6), NP_1 is replaced by *the boy*, and NP_2 by *the girl*, and similarly in (7) using *Nicole* and *Lisa*. The variable VERB is replaced by *ate* and *owns*, respectively.

(6) The boy ate more fruitcake than the girl.

(7) Nicole owns more lingerie than Lisa.

As was already mentioned, *car#1* cannot be inserted into the pattern, and hence would lead to an outcome of *no* (cf. also Table 1). Pluralia tantum (i.e. *plural-only nouns*, i.e., nouns without a singular form) such as *goggles* cannot be inserted into the first test pattern. An insertion would thus lead to the value "not applicable" (*na* in the tables).[1]

If the insertion of the noun resulted in a grammatical sentence in the first step of Test I, the annotator has to decide whether the comparison in the constructed sentence is based on the number of entities (e.g., pieces in the case of *lingerie#1*), or on a different kind of measurement (e.g., mass/volume in the case of *fruitcake#2*). This second step, which leads to the outcome of Test I.2, as indicated in Table 1, employs the experimental results provided in (Barner and Snedeker, 2005; Bale and Barner, 2009) on modes of measurement.

3.2. Test II

The pattern of the second test is similar to the configuration of the first one, but this time the noun in question is pluralized:

The pattern in Fig. 2 illustrates the first step of the test. Here the annotator has to determine whether the insertion of the

[1] Some types of nouns that are usually included into the plural-only group might not get an *na* here because their morphological form hides the fact that they are plural. Thus *poor, homeless, educated, handicapped, downtrodden, ...* [as they occur in the phrases *the poor, the educated, the handicapped,...*] might be thought to be singular by annotators, and thus when they apply the Test I.1 they might generate such acceptable sentences as *Mary has fed more poor than John has, John has helped more handicapped than Sally,...*. These are acceptable sentences, but the noun under consideration is plural, not the singular required by Test I.1. Once the annotator accepts these as acceptable, they move on to Test I.2 and should judge that this comparison is based on the number of poor or handicapped people. When they move to Test II.1, they should realize that the noun has to be plural already: forms such as **poors, *homelesses, ...* don't exist. And they should re-analyze their answers to Test I.1. But this was not monitored during the annotation process.

Noun	WordNet description	Test I.1	Test I.2	Test II.1	Test II.2	Test III.1	Test III.2
car#1	a motor vehicle with four wheels	no	na	yes	not equivalent	yes	no
fruitcake#1	a whimsically eccentric person	no	na	yes	not equivalent	yes	no
fruitcake#2	a rich cake containing dried fruit and nuts [...]	yes	not number	yes	not equivalent	yes	yes
lingerie#1	women's underwear and nightclothes	yes	number	na	na	no	yes
whiskey#1	a liquor made from fermented mash of grain	yes	not number	yes	equivalent	no	yes

Table 1: Examples of Test Outcomes.

NP_1 VERB more NOUN[pl] than NP_2.

Figure 2: Pattern Test II.1

noun into the test pattern results in grammaticality. Both senses of *fruitcake* and the first senses of *car* and *whiskey* are grammatical in Test II.1. *Lingerie* on the other hand receives the value *not applicable*, as it lacks a valid plural form. Some plural-only nouns (e.g. *clutches, goggles*) do show plural marking. Nevertheless they are ungrammatical in this pattern, due to incompatibility with the quantifier *more*; but other plural-only nouns (e.g. *poor, handicapped*) will pass this test. (See also the discussion in footnote 1 above.)

If the insertion of a noun is judged grammatical in Test II.1, annotators are required to construct a second sentence with respect to the pattern in Fig. 3:

NP_1 VERB more CLASSIFIER of NOUN[sg] than NP_2.

Figure 3: Pattern Test II.2

Annotators then have to judge whether the constructed sentences of Test II.1 and II.2 are semantically equivalent.

(8) a. He drank more **whiskeys** than her.
b. He drank more **kinds/glasses of whiskey** than her.

In (8), both sentences show the same meaning (value *equivalent*). Since they are equivalent, we conclude that the plural marking of the noun implies a hidden classifier.[2] In case of a noun such as *car*, however, annotators will come to the conclusion that the sentences are *not equivalent* and hence *car* does not implicate a hidden classifier when in plural or in the scope of an indefinite determiner.

3.3. Test III

Test III is also twofold. The pattern of the first step of this test is defined as follows:

[$_{NP}$IND.DET.+NOUN[sg]] is {SOME PROPERTY OF NOUN}

Figure 4: Pattern Test III.1

Annotators are again tasked to judge the grammaticality of adding an indefinite determiner to the noun in the test sentence. In a second step, annotators need to construct another sentence, identical to the given pattern in Fig. 4, but this time with the indefinite determiner omitted.

This test offers some synergies in regard to detecting a hidden classifier (see Test II). If an annotator is asked to construct a test sentence with an indefinite determiner for a noun such as *whiskey*, he or she will be forced to use a classifier in the necessary description part of the test.

(9) #A whiskey is a glass filled with whiskey.

(10) Whiskey is a drinkable liquid containing alcohol.

But using a classifier in the description section of the test is not allowed; so therefore (9) is not a valid test sentence.

4. Inter-Annotator Reliability Study

Four native speakers of Canadian English accomplished the annotation task. The total number of annotated noun-sense pairs comprises 13,804, where all four annotators have carried out the initial 1,000 annotations as part of the training phase. After the training phase, the remaining noun-sense pairs where randomized and assigned to two annotators each, so that each noun-sense pair received two parallel annotations. The first group annotated a total of 6,375 noun-sense pairs, the second group annotated 6,429 pairs. We found Krippendorff's α a suitable measure for inter-annotator agreement in the given task (Artstein and Poesio, 2008), and calculated the inter-annotator agreement for

[2]'Classifier' in the linguistic sense, where it is a word, phrase, or morpheme that in some way indicates a manner of "individuating" the noun in question. For some classifiers, an alternate term might be "measure phrase", in the sense that we have such phrases as *pound of butter*, where *pound of*, which is a measure phrase, works to individuate the relevant amounts of butter. Of course, just which hidden classifier is actually employed in any given case could depend on either or both of the linguistic context [*He tasted several whiskeys* could employ the *kinds of* classifier] and the non-linguistic context [where one is in a tasting room with several kinds of whiskey].

Tests I.2, II.2, III.1 and III.2. As Table 2 indicates, the values range from 0.694 to 0.755, and thus exceed the threshold of 0.67, which is considered to be a minimum value for acceptable annotation agreement in the linguistic domain (Artstein and Poesio, 2008, pp. 576, 591).

	Test I.2	Test II.2	Test III.1	Test III.2
All senses	0.727	0.694	0.755	0.755
Excluding noms.	0.729	0.709	0.767	0.758

Table 2: Krippendorff's alpha values for all Annotators

As an interesting intermediate result, the inter-annotator study allowed us to identify nominalizations (e.g. *reflection, tracking, invention, destruction*) as a major source for disagreement. To show the influence of nominalizations, we have determined the overall inter-annotator agreement for two of the four annotators, as well as their agreement on nominalizations only, as shown in Table 3 below.

	Test I.2	Test II.2	Test III.1	Test III.2
All data	0.731	0.695	0.773	0.761
nominalizations	0.618	0.630	0.703	0.669

Table 3: Inter-annotator agreement for two annotators comprising all nouns, and nouns that are analyzed as nominalizations

The drop of agreement is conspicuous. It might be explained in part because the properties of nominalizations are much harder to access through such a questionnaire. In particular, *-tion*-nominalizations such as *examination* are generally considered to be ambiguous between a result-interpretation and a complex event (or argument structure) interpretation (Grimshaw, 1990; Borer, 2013). So, *collection* may describe the collection that is the result of the activity of collecting things, or it may describe the activity itself. Although Grimshaw and Borer have provided clues for the disambiguation of such nominalizations, the pertinent features cannot be easily combined with the test patterns presented here.

5. Distribution of Countability Classes

The six proposed tests are partly interdependent and allow the definition of 80 different classes (four possible outcomes for the first two tests, five possible outcomes for the third test). It should be clear, though, that the tests describe a latent property of the nouns in question, and hence, that test outcomes which put the very same noun into completely opposed classes are not likely to occur. In fact, we can identify eight common countability classes (plus some cross-categorial subclasses) from the 13,804 annotated senses.[3] These categories are identified, with some examples, in Table 5.

[3]Plus an "other", "unknown" class. This class is reasonably large – more than 8% of all the senses. There are a couple of reasons that a sense might fall into this class. First, many of these occur because of annotator disagreement, or lack of knowledge of the sense (often about scientific terms). A second reason is due to senses of words that seem to have no clear categorization, such as *open, fuss, tantrum* in the phrases *in the open, making a fuss, throwing a tantrum.*

Group	**noun-sense**	**WordNet meaning**
Object-Mass	$glassware_1$	an article of tableware made of glass
	$mail_1$	the bags of letters and packages that are transported by the postal service
	$perjury_1$	criminal offense of making false statements under oath
	$suicide_1$	the act of killing yourself
Substance-Mass	$anger_2$	the state of being angry
	$dogma_2$	a doctrine or code of beliefs accepted as authoritative
	$minimalism_1$	an art movement in sculpture and painting
	$mist_1$	a thin fog with condensation near the ground
Dual-Life ("Glut")	$absence_2$	failure to be present
	$insertion_1$	the act of putting one thing into another
	$refining_1$	the process of removing impurities
	$tissue_2$	a soft thin paper
Unspecified ("Gap")	$bias_2$	a line or cut across a fabric that is not at right angles to a side of the fabric
	$fate_1$	an event (or course of events) that will inevitably happen in the future
	$fate_3$	your overall circumstances or condition in life
	$tail_2$	the time of the last part of something

Table 4: Examples of Four Classes of Noun-Senses

Some interesting observations can already be derived from Table 5: First, we observe that noun senses with an object-mass interpretation occur very rarely (0.58% in total), particularly when compared to substance-mass expressions (25.79% in total). Secondly, the present method allows the identification of dual life nouns. They behave like mass expressions according to Test I, but like count expressions according to Test II. (That is, they receive a *yes* answer to both of Test I.1 and II.1). This is further confirmed in that they may occur together with an indefinite determiner, but also without a determiner. Thirdly, we were able to identify a group of nouns that are *unspecified* for mass and count, according to Tests I and II. (They receive a *no* for both Tests I.1 and II.1.) It is tempting, thus, to think of dual-life nouns as being correspondingly *overspecified* for mass and count. Or in other, perhaps more picturesque terminology, the dual life nouns represent a *glut* of countability, while the unspec-

Countability Class	Example	Test I.1	Test I.2	Test II.1	Test II.2	Test III.1	Test III.2	Average freq.
Fully Countable	*car#1* *fruitcake#1*	no	na	yes	not equiv.	yes	no	59.67%
Fully Uncountable								
substance-mass	*doubt#1* *seawater#1*	yes	not number	na	na	no	yes	18.26%
object-mass	*lingerie#1*	yes	number	na	na	no	yes	0.30%
Uncountable (with sorter/packager plurals)								
substance-mass	*whiskey#1* *dye#1*	yes	not number	yes	equiv.	no	yes	4.07%
object-mass	*china#4*	yes	number	yes	not	yes	yes	0.07%
Dual Life ("Glut")								
substance-mass	*fruitcake#2* *elevation#1*	yes	not number	yes	not equiv.	yes	yes	3.46%
object-mass	*theft#1*	yes	number	yes	not equiv.	yes	yes	0.21%
Unspecified ("Gap")	*creep#4*	no	na	no	na	yes	no	1.56%
	wintertime#1	no	na	n	na	no	yes	
Plural Only								
compatible with quantifier	*expenses#1* *remains#1*	na	na	yes	na	na	na	0.27%
not compatible with quantifier	*clutches#1* *ancients#1*	na	na	no	na	na	na	0.04%
Proper (and Proper-Like) Nouns								
	prohibition#1	no	na	na	na	no	yes	2.27%
(only with definite determiner)	*www#1* *mafia#1*	no	na	na	na	no	no	0.88%
Unique Entities	*heyday#1*	no	na	na	na	yes	no	0.31%
Unknown	*open#4*							8.61%

Table 5: Observed Distribution of Countability

ified nouns represent a *gap* of countability.

A brief survey of the noun-sense meanings, some of which are represented in Table 4, that are in these different classes is revealing, and once again illustrates the point that discussions which focus on only a few paradigmatic nouns or noun senses will inevitably ignore relevant parts of the lexicon of a language. Table 4 shows groups of noun-sense pairs that fall into each of the four classes we have just discussed: object-mass, substance-mass, dual-life, and unspecified.

As Table 5 has revealed, dual life nouns occur much more often than object-mass expressions, but still much less often than fully countable or fully uncountable nouns. As a first step, we plan to further scrutinize these classes with the goal of identifying common semantic features among (subclasses of) the nouns in this class. However, we also consider it extremely likely that the very identification of dual life nouns is ill-suited at the type level, and we propose to proceed further at the token level. (A position argued for in (Allan, 1980; Pelletier, 2012).)

6. Discussion

The present resource consists of 13,804 noun-sense pairs that are annotated for their classes according to the six tests described in this paper. While investigations into the distinction between count and mass nouns typically make use of a small number of prototypical nouns, the present resource allows the in-depth analysis of a large set of nouns. What is more, the class assignments can also be used for the supervised classification of nouns into count and mass types, without having to assume a binary classification.

The 13,804 noun-sense pairs provide a gold standard for the analysis of count and mass expressions at the type level, but future research should expand the scope of the current resource by taking a token-level analysis into account. While supervised classification methods can be applied at the type level, under the assumption that each sense of the noun leads to the same class assignment, a token-level classification cannot rely on this information. Individual senses and assignments according to individual senses cannot be detected without the involvement of unsupervised methods. As the discussion in Section 5. about dual life and unspecified nouns has already shown, an analysis at the token-level may provide answers to questions that emerged from the present type-level analysis, and will thus be the next step in our analysis of countability.

Furthermore, while we have constructed an annotation of the senses listed in WordNet for nouns that occur in our

data, we as yet have no firm data on how many of these various senses actually occurred in the OANC corpus. Our plan is to employ a clustering algorithm on the senses of these nouns, and look to find where the differing senses occur in the corpus. In this way we hope to provide an answer for various long-standing questions, such as how commonly mass vs. count senses occur in actual linguistic production, and what is the proportion of count vs. mass senses of the various individual nouns that our annotators have identified as having both mass and count senses. Among other outcomes, this last topic would provide one answer to the long-standing question raised in (Allan, 1980) concerning the actual employment of levels of countability of nouns.

7. Acknowledgements

We gratefully acknowledge that the Alexander von Humboldt Foundation and the Deutsche Forschungsgemeinschaft (KI-759/5) have supported the work reported here. We would like to thank Mathieu Dovan, Meghan Jeffrey, Lisa Shorten, and Fiona Wilson for their work in annotating the data.

8. References

Keith Allan. 1980. Nouns and countability. *Language*, 56:541–567.

Ron Artstein and Massimo Poesio. 2008. Inter-coder agreement for computational linguistics. *Computational Linguistics*, 34:555–596.

Alan Bale and David Barner. 2009. The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics*, 26:217–252.

David Barner and Jesse Snedeker. 2005. Quantity judgments and individuation: Evidence that mass nouns count. *Cognition*, 97:41–66.

Hagit Borer. 2005. *Structuring Sense. Vol. I: In Name Only*. Oxford UP, Oxford.

Hagit Borer. 2013. *Structuring Sense. Vol. III: Taking Form*. Oxford UP, Oxford.

Harry Bunt. 1979. Ensembles and the formal semantic properties of mass terms. In Francis Jeffry Pelletier, editor, *Mass Terms: Some Philosophical Problems*, pages 249–277. Reidel, Dordrecht.

Harry Bunt. 1985. *Mass Terms and Model Theoretic Semantics*. Cambridge UP, Cambridge. Originally a University of Amsterdam PhD Dissertation, 1981.

Lisa Cheng, Jenny Doetjes, and Rint Sybesma. 2008. How universal is the universal grinder? In *Linguistics in the Netherlands 2008*, pages 50–62, Amsterdam. John Benjamins.

Gennaro Chierchia. 2010. Mass nouns, vagueness and semantic variation. *Synthèse*, 174:99–149.

Alex Djalali, Scott Grimm, David Clausen, and Beth Levin. 2011. What can be ground? Noun type constructions, and the universal grinder. In *Proceedings of the Berkeley Linguistics Society*. University of California, Berkeley.

Steven Frisson and Lyn Frazier. 2005. Carving up word meaning: Portioning and grinding. *Journal of Memory and Language*, 53:277–291.

Jane Grimshaw. 1990. *Argument Structure*. MIT Press, Cambridge, MA.

Ray Jackendoff. 1991. Parts and boundaries. *Cognition*, 41:9–45.

Otto Jespersen. 1924. *The Philosophy of Grammar*. Allen & Urwin, London.

Graham Katz and Roberto Zamparelli. 2012. Quantifying count/mass elasticity. In J. Choi *et al*, editor, *Proceedings of the 29th West Coast Conference on Formal Linguistics*, pages 371–379, Somerville, MA. Cascadilla Proceedings Project.

George Miller. 1995. WordNet: A lexical database for English. *Communications of the ACM*, 38:39–41.

Francis Jeffry Pelletier and Lenhart Schubert. 1989. Mass expressions. In F. Guenthner and D. Gabbay, editors, *Handbook of Philosophical Logic, Vol. IV*, pages 327–407. Kluwer, Dordrecht. See also the updated entry of this entry in Vol. 10, pp. 249–336 of the 2nd edition.

Francis Jeffry Pelletier. 1975. Non-singular reference: Some preliminaries. *Philosophia*, 5:451–465. Reprinted in (Pelletier, 1979, pp. 1-14).

Francis Jeffry Pelletier, editor. 1979. *Mass Terms: Some Philosophical Problems*. Kluwer Academic Pub., Dordrecht.

Francis Jeffry Pelletier. 2012. Lexical nouns are neither mass nor count, but they are both mass and count. In Diane Massam, editor, *A Cross-Linguistic Exploration of the Count-Mass Distinction*, pages 9–26. Oxford UP, Oxford.

Susan Rothstein. 2010. Counting and the mass-count distinction. *Journal of Semantics*, 27:343–397.