Editorial Introduction: Background to the Count–Mass Distinction

Francis Jeffry Pelletier, Tibor Kiss, and Halima Husić

The Count–Mass Distinction – henceforth, abbreviated as CMD – is a morphological and syntactic distinction that can be made for noun phrases (“determined phrases”) in a language, and as such it varies from language to language. Indeed, it can be argued that there are some languages that do not manifest such a morpho-syntactic distinction. This distinction is also usually correlated with a conceptual or ontological distinction between “things” and “stuff”.

In English (a language that does manifest such a distinction), it is common to categorize the underlined NPs in the following sentences as being count NPs:

Jeremy owns a car. Natasha likes three boys. Juan ate many apples. Sarah read each book on the list.

And on the other hand, the NPs in the following sentences are commonly categorized as mass NPs.

Hans drinks only cold beer. Sally used red chalk to emphasize her point. Chlorine gas is poisonous. The recipe calls for garlic.

Among the generalizations drawn from such examples are that count NPs use numerals, allow indefinite determiners, can be pluralized, and admit certain “individuating” quantifiers. Mass NPs do not do this. One very apparent difference between the two types of NP is that the count ones presuppose that there are separate, individual cars, boys, apples, books; the mass ones do not, but rather suggest some indeterminate amount of a given sort of matter. (Note that this preliminary description of the CMD has not taken into account “abstract nouns”, where the notion of “matter” will have to be much different. This is touched upon by many of the contributions to this volume.) Some authors call this distinction between individuative NPs and non-individuative NPs a conceptual claim (that’s how we think about whatever the NP describes); other authors think of it as an ontological claim (that’s the way “reality” is organized). Either way seems to be a distinction between “things” and “stuff” (although there remains the issue of abstract nouns).
As we mentioned, this morpho-syntactic distinction is found within the class of NPs. But most writers also assign mass or count to the individual lexical items – words, for ease of reference – that are the head noun of the NP. This makes *car, boy, apple, book* and other words be classified as +count; *beer, chalk, gas, garlic* classified as +mass. And the individuation/non-individuation character of the NP is then also assigned to these words.

One class of puzzles in the area is this: although *cold beer* is a mass NP, and thus *beer* is a mass noun, it seems clear that there is nothing particularly strange or forced about the sentence *Hans drank three cold beers*. Given a presumption that the CMD is supposed to assign a unique one of +count or +mass to each noun, what shall be done with such apparent counterexamples? One strategy is to rule them “ungrammatical but interpretable”, and this can be carried out in different ways. A popular way is to say that *beer* (for example) is “really” a mass noun, but that it can be “coerced” into having a count meaning by certain processes, such as being “portioned” (a can or bottle or glass of beer) or being “sorted” (*Hans drank three sorts of beer*). The other way around might see the “basic” meaning of *chicken* be +count (“the common domestic fowl *Gallus gallus*”), but that it gets coerced into a mass meaning in such sentences as *Louisa cooked chicken for supper*.

The issues surrounding these (and other) “coercions” form a large part of the argumentation of very many papers in this volume. Other issues that emerge concern so-called *furniture*-nouns and *bouquet*-nouns. Although (in English) *furniture* is a mass noun, we in fact count pieces of furniture, rather than weigh the total or take the total volume, when determining who has more furniture. In the other direction, words like *bouquet, thing, fence* are count nouns, even though placing two bouquets in a bride’s hand does not make for a bride holding two bouquets but rather holding one large bouquet.

Another direction of research is in describing how different languages can differ in their versions of the CMD. Even in closely related languages, such as English and German, we find that direct translations will assign different count–mass statuses to the words. English *lightning* is called a mass noun; German *Blitz* is count. Is this a difference in conceptualization? It doesn’t seem to be a difference in “reality”. Further afield, there are “classifier languages” where, it is sometimes claimed, individual nouns are mass, and individuation takes place by means of classifiers applied to the noun. For an English example, *ice* is generally thought to be a mass noun, but *ice cube* (or *cube of ice*) uses *cube (of)* as a classifier. Some classifier languages, it is said, do this universally with all nouns.

All these topics are discussed in this volume, and the various surprising ways that they interact with one another are surveyed. A goal of this research area is to find some stable, comprehensive theory that finds a place for each type of word and an explanation for how different languages can diverge on
the ways they differentially treat these topics. The reader will decide how well this is carried out. But we can say with certainty that the papers in this volume make important progress toward this goal.

1. The Beginnings of the CMD

Almost all the topics that are of current interest in research concerning the CMD were mentioned at the very beginning of study of the topic – although often mentioned as asides and without awareness of the intricacies involved in them. Since this is where it all began, we start by reciting the very earliest discussions of the distinction in English, because that will give the reader a chance to see how the different contemporary positions on the CMD have evolved due to attempts by writers over the decades to deal with one or another of the many subtopics. And we can sometimes see how some modern positions can sometimes seem to be awkward due to some ill-chosen initial starting point. (Of course, most of these intermediate moves came in the decades prior to the present book, but one can see different modern emphases as reactions to one or another of the initial descriptions, sometimes at the expense of ignoring equally plausible alternative initial descriptions.)

Although Wikipedia\(^1\) tells us that English grammar books appeared as early as 1586, and that by the mid-1800s there had been more than 1200 books and booklets that had been published on English grammar, Henry Sweet’s two volume *A New English Grammar, Logical and Historical* (1892 and 1898)\(^2\) stands out to us for its clear (although perhaps naive, simplistic, incomplete and just plain wrong) account of a distinction among nouns that includes a place for what we would call mass and count nouns. (Sweet himself called them “material nouns” and “class nouns”, respectively, and he also had another group called “collective nouns”, which he distinguished from the class nouns.) All of these were what Sweet called “concrete nouns”, and in a later part of the book he discussed what he called “abstract nouns”. The fact that he has separated the abstract from the concrete nouns has the consequence that the features he assigns to the concrete material and concrete class nouns – those nouns most closely akin to our (concrete) mass and count nouns – do not carry over to the abstract nouns, except in those cases where Sweet assigns them a status of “half-abstract, or intermediate between abstract and concrete”. According to Sweet [§§150ff], a class noun is so-called “because it stands for a class or number of individual things having certain attributes in common”. On the other side are such material nouns are as


\(^2\) The material relevant to the CMD is in vol. 2, that is, in Sweet (1898).
iron, glass, bread, water. They do not express any definite thing, as the class-words tree etc. do, but each of them includes the whole mass of matter possessing the attributes implied by the word . . . Thus iron makes us think of hardness, weight, liability to rust, etc., associated together in a substance of indefinite form . . . When a material noun is used to express an individual object of definite shape, it is no longer a material noun, but a class noun. Thus iron in the sense of “implement to smooth cloth with”, or glass in the sense of “vessel to drink out of” are pure class-nouns.

We note that the examples provided are ones that are common in the literature of today. We further note Sweet’s suggestion that these two types of words bring to mind different conceptions: individual things versus a “mass of matter” – a common comment in current psychologically oriented literature. Furthermore, the examples mentioned at the end should remind us of current theories of the “coercion” of mass to count (and conversely) that occur in very many of the modern works on mass and count nouns.

One of the few things missing from Sweet’s account was provided by Otto Jespersen: the modern names for the distinction. Jespersen’s opus magnum was a seven-part work that was published volume-by-volume between 1909 and 1949. Jespersen discussed what we call the CMD in Part 2 of this work, published in 1914, and subtitled Syntax, First Volume (Jespersen 1914).3 Besides criticizing Sweet for not recognizing that concrete and abstract nouns both manifest a CMD, he gave a number of examples together with some explanations of the underlying semantic rationale for the distinction. His Section 5.2 (pp. 114ff ), called “Mass Words”, seems to be the first use of this term to describe these nouns4 (as also is his use of “Countable Nouns” in the section just before). Here he says

[A] form which implied neither singular nor plural would be even more called for when we left the world of countables (such as houses, horses; days, miles; sounds, words, crimes, plans, mistakes, etc.) and got to the world of uncountables. There are a great many words which do not call up the idea of some definite thing with a certain shape or precise limits. I call these “mass words”; they may be either material, in which case they denote some substance in itself independent of form, such as silver, quicksilver, water, butter, gas, air, etc., or else immaterial, such as leisure, music, traffic, success, tact, commonsense, and especially many “nexus-substantives” like satisfaction, admiration, refinement, from verbs, or like restlessness, justice, safety, constancy, from adjectives.5

Jespersen goes on (a) to emphasize the relevance of the singular-plural number distinction to the count–mass difference, (b) to cite various differences between languages as to what is a countable term vs. a mass-word, (c) to

3 Jespersen also repeated much of this discussion in his 1924 popular work, The Philosophy of Grammar (Jespersen 1924).
4 Other than his earlier forward reference in Section 4.17 (pp. 72–3) to this section.
5 Jespersen’s nexus-substantives apparently means what is now more commonly called nominalizations from verbs and adjectives.
remark on cases “where the same word has to do duty now as a mass-word and now as a thing-word”, (d) to remark on cases where a mass-word can “denote a mass” of something that is often designated by a thing-word (Jespersen cites Oak and beech began to take the place of willow and elm), and (e) of the tendency to use the same word for animals (countable) as for the food of that animal (mass) (Jespersen cites fish). Jespersen also remarks on (f) how “immaterial mass-words” can become countables when they come to stand for a single act or instance of the quality (a stupidity and a beauty are cited). This usage, he says, “is not so universal in English as in many other languages, and the best rendering of eine unerhörte Unverschämtheit is a piece of monstrous impudence”. He also points to “the use of a mass word to denote one kind of the mass” (offering us This tea is better than the one we had last week; Various sauces; The best Italian wines come from Tuscany).

Finally, Jespersen is sensitive to the fact that (g) the CMD is drawn differently in different languages, although this is not made into a major topic. He remarks (p. 200) that “in English, but not in Danish, tin is used for a receptacle made of tin (for sardines, etc.). In English, bread is only a mass-word, but the corresponding word in many languages is used for what in English is called a loaf: un peu de pain, un petit pain = a little bread, a small loaf.”

All of the topics in (a)–(g) form current strands of research within studies of the CMD, and each of them form a part of one or another of the papers in this volume. Although it is not explicitly mentioned by Jespersen, perhaps it is implicit in (b) and (g) that there is a possibility of languages not manifesting a CMD directly in the lexical nouns but instead by means of classifiers, or perhaps not even manifesting a CMD directly at all (examples of which are discussed in many of the papers of this volume).

2. Distinguishing Features of the CMD

There are different aspects of the CMD that are emphasized by different modern authors. All of them take a set of morpho-syntactic, also called grammatical, features as driving the division between count and mass nouns. The following list is usually assumed for English count and mass nouns, many of which are also present in other languages.

- singular/plural contrast
- direct combination with numerals
- constructions with classifier phrases
- restricted selection of determiners/quantity modifiers

Some of these characteristics are considered more relevant than others or even distinctive for countability. Chierchia (1998a, 2010; this volume, Chapter 2) thinks of the ability of some nouns to occur directly with cardinal
numerals to be “the signature property” of count nouns, whereas non-count nouns (normally) require a “classifier phrase” in order to occur with cardinal numerals, as in *While John examined two blades of grass, Mary extracted two shovelfuls of mud.*

Besides these overt differences in the CMD, many theories rely heavily on certain referential properties of count and mass nouns, such as mass nouns having a homogeneous reference, i.e. being divisive and cumulative. These properties lead many authors to decide on an ultimate determinant which divides nouns into count and mass, such as (contextual) atomicity for count nouns (Rothstein 2010), overlap in the generator sets of mass nouns (Landman 2016) and the vagueness of mass atoms (Chierchia 2010) to name a few.

With regard to the morpho-syntactic features of the CMD and the referential properties of these nouns, we can observe many inconsistencies in English as well as cross-linguistically which have been discussed thoroughly. These inconsistencies can be divided in two groups: one consists of mismatches between the actual properties of the references, and the other is a matter of what we might call “variation” – nouns that exhibit properties of both mass and count nouns.

The first group comprises so-called fake mass nouns and homogeneous object nouns. Fake mass nouns (also called furniture-nouns, superordinates, object mass nouns, or aggregate nouns) are mass nouns according to the grammatical features, but have a reference similar to count nouns being individuated in units. *Five furnitures/silverwares is ungrammatical since furniture and silverware require a classifier to enable counting pieces: five pieces of furniture/silverware.* The difference between fake mass nouns and ordinary mass nouns (or object mass vs. substance mass) is the mode of measurement which in substance mass nouns is volume and in object mass nouns the number of individual entities. This difference was further supported by experiments conducted by Barner and Snedeker (2005) which employed quantity judgement tasks for identifying these classes of mass nouns. Since then, quantity judgement tasks have been used targeting the mode of counting in many languages besides English. Likewise, homogeneous objects nouns (also known as fence-nouns) are count nouns with a homogeneous reference, such as wall, fence or sequence. A wall can be divided vertically into several parts, each of which is still a wall. These two subclasses of nouns pose a challenge for formal theories that consider atomicity to be a distinctive property of count nouns and require adjustments to account for such nouns as well. Besides fake mass nouns and homogeneous object nouns, there is also an issue with nouns that have the same type of reference but different countability properties, such as *rice vs. lentils or onions vs. garlic.*

The other group of inconsistencies can be subsumed under the term *variation.* It regards nouns that cannot be classified straightforwardly as either
count or mass, because they have features of both count and mass nouns. Such nouns have often been called dual-life nouns, flexible or elastic nouns. An example for such a noun is *cake*, which can be used as count and as mass, as illustrated below.

(1) a. John ate five cakes.
   b. John ate five pieces of cake.
   c. John ate much cake over the last weekend.
   d. John ate many cakes over the last weekend.

Another phenomenon related to variation regards nouns that are classified as count nouns but permit also mass uses albeit with an implied difference in interpretation or meaning. Two such interpretations are commonly used in natural language and have been acknowledged in the literature on the CMD as a shift or coercion from mass into a count use: packaging (Bach 1986) as in (2) or sorting (Bunt 1985) as in (3).

(2) a. We would like three waters.
   b. He put four milks in the fridge.

(3) a. The restaurant that Kim suggested offers three wines.
   b. The hotel we visited last year provided pools with different waters.

In a similar way some nouns which are normally considered to be countable can – under certain circumstances – occur as mass although this shift is not as productive as the mass-to-count shifts in (2) and (3). Count-to-mass shifts can appear in advertising such as (4).

(4) More car for less money.

Some researchers assume that almost every count noun is capable of turning into a mass noun by means of applying Pelletier’s thought machine, the Universal Grinder (Pelletier 1975), which aims to grind the reference of a count noun, a hat or carrot for instance. The outcome of it should be referred to by a mass expression, as the following sentence with the bare use of *carrot* suggests.

(5) There is carrot all over the floor.

When it comes to the semantics of count and mass nouns, we need to bear in mind that besides linguists, cognitive psychologists and philosophers also show interest in the CMD. However, the different fields that study the CMD have different conceptions of what the term “semantic” encompasses. On the whole, philosophers of language think of the term “semantics” as describing a relationship between language and the world. Cognitive psychologists on the other hand think of it as describing a relationship between language and mental concepts. And formal semanticists aim to offer model-theoretic
interpretations of the nouns’ references as a functioning input for a compositional analysis.

So, not surprisingly we get different viewpoints when it comes to explaining the semantics of count nouns and mass nouns. However, all our theorists seem to agree that there is a notion of individuation that plays a role such that only count nouns presuppose a principle of individuation while mass nouns do not. (Count nouns may inherit their principle from features of the context of utterance; mass nouns may in fact have such a principle but it is not activated in the context.) Depending on the area that a theorist calls home, the difference between *three chairs* and *three bloods* is:

a. chairs are individual objects; blood is a stuff and not an individual [philosopher];

b. one’s conception of a chair is of some independent object, whereas blood is conceptualized as an undifferentiated puddle [psychologist];

c. chairs are a part of an atomic lattice whereas blood designates a non-atomic mereology [formal semanticist].

The articles of this book tackle many of the above-mentioned issues that regard the semantics of count and mass nouns across languages theoretically and/or empirically.

3. The Articles in this Volume

This volume follows the organizational structure of the conference. It contains four longer, invited papers, followed by twelve shorter papers that touch upon or expand on or take issue with certain aspects of the invited papers and thus provide many independent observations and conclusions concerning the CMD. These twelve papers are grouped into three general areas of four papers each: Implications from Individual Languages; Compositional Analyses and Theoretical Issues; and New Empirical Approaches to the Semantics of the CMD.

We also include, at the end of the volume, language and subject indices, as well as a common-to-all-papers reference list. The language list is intended to provide locations where the CMD is discussed in languages other than English. English constructions are listed in this index only when they are being described as unique concerning a certain type of feature.

3.1. Invited Papers

The invited speakers at the conference offered us broad-scale accounts of what they took to be central issues in the study of the CMD – as, for instance, what is responsible for the distinction? Or how does it come to be differently
realized in different languages? Or what is the role of plurality and its relation to the CMD? Or is there any relation or dependence between different inconsistencies which underlie the CMD?

Gennaro Chierchia’s approach classifies all human languages into three groupings based on the way these languages manifest countability. This, of course, requires a criterion for what makes a noun be count, and Chierchia defends his choice by assuming a cognitive contrast between count and mass nouns motivated by the findings of Carey and Spelke (1996), according to which children in pre-linguistic age distinguish objects (as well defined units which retain their identity upon moving through space and entering in contact with each other) from substances (concepts like “water” or “sand” which do not have readily accessible minimal parts, and whose samples don’t retain an identity when moving or congregating). Chierchia uses the term *cognitively count* for nouns which refer to “Spelke objects” and *cognitively mass* for nouns that denote “Spelke substances”. He assumes that the CMD – as one dividing nouns into cognitively count and cognitively mass – is universal, but languages have different means to express this division. While Type I languages, such as English, have strong morpho-syntactic differences, Type II languages, such as Korean (also called classifier languages), do not present an overt difference in syntax or morphology since both (cognitively) count and mass nouns require classifier phrases. Count and mass nouns can, however, be identified by the choice of the classifier which is sensitive to (cognitively) count and mass nouns. Type III languages, such as Yudja, do not provide an overt difference between count and mass nouns either, since they all can combine directly with numerals, but the interpretation of numeral-noun combinations depends on the type of the noun. Cognitively count nouns have an interpretation that quantifies over individual items, while cognitively mass nouns imply a (hidden) classifier, or container for counting or measuring. Chierchia proposes a semantic mechanism that accounts for count vs. mass nouns and offers a compositional analysis for numeral-noun constructions in Type I–III languages both for inclusive and exclusive plural readings. His logic can also account for the lexical variation and elasticity in the CMD which is present in every language in that certain modifications mirror coercion between count and mass references.

Susan Rothstein questions the relation between countability and pluralization. These features have been widely claimed to be necessarily interrelated: pluralization implies countability and countability is expressed by the ability to pluralize (cf. Chierchia 1998a, 2010; this volume, Chapter 2). She finds motivation for her concern in mass nouns which pluralize but do not refer to a multiplicity of kinds, but instead have an abundance interpretation, such as mass plurals in Greek, or the abundance reading of nouns such as *rain* or *fog* in English as well as the pluralization of events and abstract mass nouns. In addition to Greek and English, Rothstein provides further evidence for the
non-relatedness of plurality and countability by means of a cross-linguistic study of several languages: Yudja, Dëne Sųliné, Modern Hebrew, Wayaro, Panará, Sakurabiat, Ye’kwana and Taurepang, many of which are peculiar for their relation between countability and pluralization. For instance, Dëne Sųliné provides a contrast between count and mass nouns but no plurality marking, and Taurepang is a language with a plural marker which attaches to every noun except for animate nouns that require a special plural -damök. Based on observation stemming from different languages, Rothstein concludes that the relation between plurality and countability allows for much more variation, and hence this relationship has to be inspected with more caution.

Jenny Doetjes re-evaluates the notions count and mass from different perspectives by elaborating on different meanings of nouns and different types of quantity expressions. She presents asymmetries that are evident throughout the CMD in the English language but also across languages. In particular, she discusses asymmetries in Mandarin, Dëne Sųliné, French, Halkomelem, Blackfoot, Nez Perce, Hungarian, Cantonese, Yudja, Brazilian Portuguese and Indonesian. The asymmetries are expressed either in grammar or in meaning. With regard to these asymmetries, and motivated by the fact that count is (and mass is not) marked in the grammar of language, Doetjes argues that mass should be interpreted as the absence of count. Such a relational explanation of count and mass is further supported by a blocking principle in quantity expressions. She introduces a typology of quantity expressions according to which quantity expressions are divided in three groups: count quantity expressions; non-count quantity expressions; and anti-count quantity expressions. Crucial to this typology is the blocking principle according to which the existence of non-count expression is explained through the absence of count counterparts. Unlike this, anti-count quantity expressions, such as much, have a count counterpart, namely, many. Non-count quantity expressions, however, are those which combine with both mass and plural count nouns, such as all, and are not strictly reserved for mass nouns due to the lack of anti-count expressions.

It verges on the trivial to state that the CMD has semantic impact, and hence is addressed within semantics. But it should also be clear that any compositional semantic analysis must find its roots in syntax; this holds in particular if the CMD is severed from lexical categories, i.e., if the CMD itself becomes the result of the compositional combination of syntactic items. Following the general syntactic architecture developed in Borer (2005), Hagit Borer and Sarah Ouwayda deal with the syntactic structure of the nominal spine in Arabic, and in particular with the somewhat surprising co-occurrence of a classifier and plural realizations. Following T’sou’s generalization, according to which classifiers and plural realization should not co-occur in the same syntactic structure, Borer and Ouwayda discuss the (Lebanese) Arabic
feminine ending -\textit{ah}, which can be added to nouns that are either mass or of indeterminate quantity, and act as a classifier in these constructions. Yet, plural marking is possible after -\textit{ah} has been added as a supposed classifier. The construction, however, is not fully identical to a plural realization of a count noun. To solve this apparent paradox, Borer and Ouwayda propose a more fine-grained syntactic structure of the nominal projection, where cardinals are categorically distinguished from quantifiers. In a broader context, the paper suggests that the distinction between inclusive and exclusive plural readings should be addressed by assuming different syntactic structures for the different plural types (and hence, different compositional analyses). The apparent contradiction to T’sou’s generalization is eventually resolved by assuming that the realization of the -\textit{ah}-marker cannot be analysed as a classifier (divider), but instead as an agreement marker with a cardinality specifier. The obligatory presence of the specifier yields the pertinent interpretations.

3.2. \textit{Implications from Individual Languages}

The invited papers have made claims that are to be applied to a wide swath of human languages – or even all – and which culminate in empirical remarks about how the languages of the world treat matters of count and mass – or alternatively, how the count–mass features would differentially appear in these languages. Papers in the next subsection of the book describe classes of apparent problems for some of these claims based on facts about certain languages, which in turn might thereby cast doubt on the universal applicability of some sweeping claims that were made.

The paper by Christine Hnout, Lior Laks and Susan Rothstein focuses attention on a mass-to-count shift that occurs in the Galilee dialect of Palestinian Arabic (PA). As in other Arabic dialects, a morphologically marked shift in gender (cf. also Borer and Ouwayda, this volume, Chapter 5) can result in contrasts between mass and count interpretations, so that masculine nouns are mass, while derived feminines are count. After discussing the peculiarities of the CMD in PA, the paper addresses constraints on the input to the mass to count operation, observing that nouns denoting solids and granular substances can be affected, while liquid- and powder-denoting nouns cannot. The input denotation of the operation is not only constrained, but it also determines the output (which sometimes can be idiosyncratic). A particularly relevant output condition governs the shift for \textit{s’enf} nouns, which are also analysed as mass nouns. The present analysis supports a distinction between natural and semantic atomicity in general. The analysis suggests in particular that count–mass shifts should not be thought of as the result of a contextually determined partition on the set denoted by a mass noun (as e.g. proposed in Chierchia 2010). Finally, the lexical coercion operation is contrasted with a
highly constrained coercion, which involves the addition of a covert classifier in the syntax.

The contribution by Kurt Erbach, Peter Sutton, Hana Filip and Kathrin Byrdeck reminds us that object (or fake) mass nouns do not align with ordinary intuitions concerning the relationship between stuff and things. It is for this reason that they say that such nouns provide a good testing ground for theories of the CMD. As they remark, there are a number of theories in the literature to account for the peculiarities of object mass nouns. But, they also note, these theories have all been for those languages that overtly manifest a CMD: the type of languages that Chierchia (1998a, 2010; this volume, Chapter 2) calls Type I languages. They wonder whether other language types – particularly classifier languages such as Mandarin and Japanese – might also have object mass nouns. Here they argue first that it is at least in theory possible that such languages have object mass nouns, and to this end they proceed to test the determiner *nan-byaku to iu* (‘hundreds of’) with a number of nouns, using native speakers as subjects. The idea is that any noun that refers to collections of discrete objects, but which fails the countability test with *nan-byaku to iu* is a candidate for being an object mass noun. The authors provide a partial semantic analysis of how such nouns might fit into a grammar of Japanese (and presumably of other classifier languages, if they also had a determiner of the same nature as *nan-byaku to iu*). Although they are careful to say that their analysis is preliminary, nonetheless, if it does stand up it would raise serious objections to most theories of the cross-linguistic description of the CMD in all languages.

Kayron Beviláqua and Roberta Pires de Oliveira propose a novel methodology for testing how different languages differentially treat the various types of nouns that are in play when considering issues about the CMD. Using five types of nouns (Bare Singulars, Bare Plurals, Singular Flexible Nouns, Plural Flexible Nouns, and Mass Nouns), they propose testing how speakers of four different languages (English, Brazilian Portuguese, Rioplatense Spanish, and Cape Verdean) will judge the different types of pictorial representations of situations that compare amounts (size, number, and so on) of the five different types of nouns. (The tests are extensions of the ones made popular by Barner and Snedeker 2005.) Their pilot studies give some intriguing results, categorizing – in the terminology of Chierchia (1998a, 2010; and this volume, Chapter 2) – some of the investigated languages as number-marking, others as classifying, and, perhaps most surprisingly, another as neither. Although as they say, further more-detailed studies need to be performed, this methodology could open up a better way to test which of Chierchia’s categories a language falls into than those methods currently employed.

Following upon Chierchia’s (1998a, 2010; and this volume, Chapter 2) distinction in different count–mass language types, Helen Koulidobrova addresses
the CMD in American Sign Language (ASL) and asks to which type ASL belongs. The structure of ASL initially suggests that ASL differs strongly from English and other Type I languages in Chierchia’s proposal in that a CMD is not visible; thus it should be grouped together with Yudja and other Type III languages. Her investigations, however, show that such conclusions are not warranted and that ASL bear more resemblance to Type I than Type III languages, while clearly rejecting classification as a (Type II) generalized classifier language. Crucial empirical evidence for this conclusion is provided by the grammar of the quantifier-split in ASL. Koulidobrova shows that the quantifier-split in ASL is sensitive to the CMD after all.

3.3. Compositional Analyses and Theoretical Issues

While the CMD classifies nouns into (at least) two groups, one has to bear in mind that it has much wider implications beyond the individual representation of a word (a “lemma”). That count and mass nouns have different semantic representations is the initial point of a compositional structure. The fact that certain nouns can combine with numerals directly while others require a classifier phrase, tells us that numerals and classifier phrases need to have such a semantic and syntactic structure which is sensitive to the noun they are modifying. Likewise, the indefinite article and other quantity modifiers require an analysis which is compatible with the nouns they are accompanying. Compositionality is central in the papers of this section, which investigate the syntax and semantics of particular phenomena related to the CMD.

In his contribution, Alan Bale discusses the implications of interaction between three different theories of the semantics of the CMD with two theories of number marking. Link (1983) has proposed that count and mass expressions must be mapped to two different domains, while Rothstein (2010, this volume, Chapter 3) assumes a difference in type theory. Finally, proposals like Chierchia (1998a, this volume, Chapter 2) assume that the semantic difference is captured neither by a distinction of domain nor of type theory. Bale combines the predictions of these models with two theories of (morpho-)syntactic number marking, where number is either assigned low, i.e. within the NP, or – following Sauerland (2003) – high, i.e. as a sister node of a Determiner Phrase (DP). Various issues concerning coercion emerge in the different theories, including the topic of how single-domain-for-both-count-and-mass denotations can employ coercion, and as well, what it would mean for coercion to effect a change in domain for two-domain theories. Given the current state of affairs, Bale does not profess a favourite for any of the theories he describes, since advantages are counter-balanced by drawback. In conclusion, a “best overall theory” will involve a very intricate and detailed syntactic–semantic account that interweaves many different levels of explanation.
Following previous work (Acquaviva 2008; Alexiadou 2011; Butler 2012; Wiltschko 2012; Mathieu 2014) which proposed that morphemes corresponding to number can occupy several functional positions depending on the individual semantics of these morphemes, Myriam Dali and Éric Mathieu investigate some plural forms which require a yet-different syntactic position. They analyse bare nouns in Turkish and Western Armenian that possess a general (or transnumeral) number (cf. Corbett 2000) and argue that pluralization of such bare nouns is a two-step process which requires a renominalization effect as familiar from collective nouns in Tunesian Arabic (Dali and Mathieu 2016, 2021). The pluralization process requires two NumP projections, a lower one with the null exponent providing a singular form and a higher one which operates morpho-syntactically on the singular and returns a set of atoms. Such a procedure is known as a case of morphological compositionality where one number is built on another.

Starting with a discussion of different interpretations of pseudo-partitive constructions (PPC), Peter Sutton and Hana Filip propose analyses with respect to the individual interpretation of such PPCs. The main difference among these interpretations is the one between container+contents readings and ad hoc measure interpretations. The former makes both the container as well as the content anaphorically available while the latter blocks anaphoric reference to a container (+content) but licenses reference to the measured stuff. Sutton and Filip analyse PPCs in a dynamic, mereological framework which incorporates several ingredients that have previously been suggested to be relevant in a theory of the CMD, such as, for instance, Rothstein’s contextual influence onto the counting units (Rothstein 2010) and Landman’s counting bases (Landman 2016). Counting is permitted in case the context makes an individuation schema available which, applied to the counting base of singular count nouns, yields a quantized predicate as defined by Krifka (1989).

Assuming a pre-linguistic notion of countability inspired by Soja, Carey, and Spelke (1991), Gordon (1985) and Bloom (1990), Hanna de Vries and George Tsoulas discuss a set of asymmetries ranging over flexible nouns and various types of coercion. Comparing the processing costs of packaging, sorting and grinding with flexible nouns and elaborating on different proposals to this type of variation, both in terms of lexical reanalysis and covert classifier constructions, they consider a combination of two structurally different theories as the most suitable solution. Their paper presents an intriguing attempt to combine a constructionist approach with a lexicalist framework in order to account for asymmetries relating to countability across languages. This hybrid theory embodies Borer’s exoskeletal syntax with Landman’s Iceberg semantics in order to account for various relations of countability and number neutrality. The resulting theory correctly does not allow non-number-neutral noun phrases with uncountable nouns – a combination considered non-existent.
across languages. In this way, they offer a framework in which only [+countable] nouns can be [-number neutral] and “stuff reference” is not reduced to uncountable number neutrality.

3.4. New Empirical Approaches to the Semantics of the CMD

Empirical approaches to the study of language can come in many different guises. Corpus studies have been used to offer an objective state of the natural language use of many phenomena relevant in linguistic theories. The CMD has also been studied in this way. We furthermore observe a rise in experimental methods in the last twenty years in many different languages. In this subsection we include papers that exhibit some of the different strands of empirical approaches: psycholinguistic investigations by means of different kinds of experiments and significant corpus studies enriched with investigations of relevant lexical resources. Of course, other papers already introduced also can be seen as fitting under this description, but we see these four papers as especially relevant in this regard.

The paper by Francesca Franzon, Giorgio Arcara and Chiara Zanini remarks that one source of information that is weak or even missing from the discussion of countability in language is specific data on the actual occurrence of count vs. mass nouns in naturally produced cases of language, and they set out to discover this data as it occurs in Italian. Their review of the psychological literature on the question shows that there are unresolved differences in people’s representation, processing and learning of count vs. mass nouns. And this paper suggests that a part of this inconsistency might be due to each study relying only on one speaker’s intuitions as to what are relevant count and mass nouns. To help resolve this, they investigated the actual occurrence of mass and count nouns, by having subjects rate the subjective frequency of occurrence of 225 words (concrete nouns) in the singular and also in the plural. The results seem to show that there is a gradation here, not a binary distribution of “always singular” vs. “sometimes plural” – it seems there are “different levels of countability”. To determine how many of the “always singular” turn out to be sometimes plural, the authors turned to the Italian itWaC corpus, to investigate the existence of a correlation between occurrences of a same noun in mass syntax and in count syntax. The results of this showed no significant inverse correlation between (allegedly) count and (allegedly) mass nouns. One of their overall conclusions is that their data is most easily explainable in a framework that does not make countability into a lexical feature.

The contribution by Peter Lauwers addresses the role of lexical plurals, a topic that has often been neglected in investigations of the CMD. Lauwers points out that although lexical plurals show plural morphology, the inherent number specification must be separated from counting. As lexical plurals do
not pass ordinary countability tests, such as the combination with cardinals, they provide evidence for a dissociation between countability and plurality. Here, we note one of the common themes across the individual contributions, in this case relating Lauwers’ paper to both Rothstein’s and Borer and Ouwayda’s papers. Lauwers not only argues that lexical plurals show a discrepancy between number marking and countability, but proposes that lexical plurals actually show more properties of mass than of count nouns. According to the ensuing analysis, lexical plurals are less count than other expressions, but neither are they entirely mass either. Lauwers thus takes up recent proposals (briefly mentioned in the contribution by Kiss et al. in Chapter 17 as well) that the CMD, although the name suggests otherwise, should not be seen as a binary opposition, but more like a continuum. Lauwers places his methodology in the perspective of the “empirical turn” on the CMD, and uses both corpus studies (which are focused in the present paper) and sentence rating tasks on thirty lexical plurals from French, which represent a broad variety of lexical fields.

Keith Allan (1980) was possibly the first to develop a categorization of nouns according to how many (and which ones) of the various tests for countability a given noun obeys. In turn this leads to a (sub-)classification of types of noun-countability. Scott Grimm and Aeshaan Wahlang largely extend Allan’s database by using all nouns occurring in four of the five subcorpora of COCA (Corpus of Contemporary American English), filtered through their occurrence in CELEX, and determine whether Allan’s classification can be maintained after such an extension. Using machine learning algorithms such as gradient boosting (random forest classification) and non-parametric clustering methods such as DBSCAN, they predict the actual occurrence patterns of the nouns. It turns out that the ability of a noun to occur bare (without determiners) in the singular or in the plural makes for the most accurate prediction. They conclude that there is a much more varied distribution of “types of countability” than any of the theoretical models would predict, which seems to be a common conclusion of all papers in this subsection. Of course, this conclusion provides severe challenges for many (small-scale) studies on the CMD.

Leaving syntactic proposals aside, many analyses assume that the CMD is a property of words. Tibor Kiss, Jeff Pelletier and Halima Husić question this conclusion from two directions: first they note that lexical items typically have more than one sense, yielding all kinds of ambiguities; secondly, they point out that the evidence presented for count or mass properties often rests on a very small set of nouns, which can lead to researchers missing the micro-variation within lexical classes. Using both BECL, a database of 11,000 noun–sense pairs with annotated countability status and also further corpus analyses, Kiss et al. distinguish four types of ambiguity, two of which are of particular relevance. They use these ambiguities as touchstones for proposals such as
Chierchia (1998a, 2010; and this volume, Chapter 2) and Rothstein (2010; this volume, Chapter 3). BECL allows a quantitative specification regarding count-to-mass and mass-to-count shifts, and a way to differentiate “dual life” nouns from cases of grinding. Following their typology of the CMD in particular, Kiss et al. argue that their findings are compliant with an analysis of the CMD where count expressions show more denotational structure than mass expressions.

The papers in this volume have “moved the yardstick” in our understanding of how the CMD should be understood. They have done this both by looking at very large-scale views of all natural languages and at detailed descriptions of specific topics within one language or comparisons between different languages. If some readers find some of the directions outlined by our authors to be less than completely convincing, these directions nevertheless will certainly challenge such readers to investigate alternative possible explanations.