Proceedings of the 21st Amsterdam Colloquium

21st Amsterdam Colloquium 2017

Edited by Alexandre Cremers,
Thom van Gessel & Floris Roelofsen

Foreword

This is a collection of papers presented at the 21^{st} Amsterdam Colloquium, organized by the Institute for Logic, Language, and Computation (ILLC) at the University of Amsterdam, December 20–22, 2017. The bi-annual Amsterdam Colloquia aim at bringing together linguists, philosophers, logicians, cognitive scientists and computer scientists who share an interest in the formal study of the semantics and pragmatics of natural and formal languages.

Besides the regular programme, the 2017 edition featured two workshops on *Causality and Semantics* and *Formal and Distributional Perspectives on Meaning*, respectively, and one evening lecture, jointly organized with the E.W. Beth Foundation. The programme included eight invited talks and 47 contributed talks.

We would like to thank the members of the programme committee and all the reviewers, listed below, for their efforts in selecting the contributed talks. We would also like to thank Patty den Enting, Luca Incurvati, and Peter van Ormondt for their help in organizing the colloquium.

Lastly, we would like to thank the ILLC, the E.W. Beth Foundation, the Netherlands Organization for Scientific Research (NWO), and the European Research Council (ERC) for financial support.

Alexandre Cremers
Thom van Gessel
Floris Roelofsen

Programme Committee

General programme

Robert van Rooij (chair)	ILLC, University of Amsterdam
Maria Aloni	ILLC, University of Amsterdam
Franz Berto	ILLC, University of Amsterdam
Paul Dekker	ILLC, University of Amsterdam

Workshop on Causality and Semantics

Franz Berto	ILLC, University of Amsterdam
Peter Hawke	ILLC, University of Amsterdam
Robert van Rooij	ILLC, University of Amsterdam
Katrin Schulz	ILLC, University of Amsterdam

Workshop on Formal and Distributional Perspectives on Meaning

Raquel Fernández Rovira	ILLC, University of Amsterdam

Reviewers

Luis Alonso-Ovalle	McGill University
Daniel Altshuler	Heinrich-Heine-Universitt Düsseldorf
Scott Anderbois	Brown
Nicholas Asher	IRIT, Université Paul Sabatier, Toulouse
Alan Bale	Concordia University
Moshe E. Bar-Lev	The Hebrew University of Jerusalem
Raffaella Bernardi	University of Trento
Bronwyn Bjorkman	Queen's University
Adrian Brasoveanu	UC Santa Cruz
Lisa Bylinina	Meertens Instituut
Ivano Caponigro	University of California San Diego
Lucas Champollion	New York University
Simon Charlow	Rutgers University
Emmanuel Chemla	LSCP, ENS, CNRS, Paris
Ivano Ciardelli	ILLC, University of Amsterdam
Ariel Cohen	Ben-Gurion University of the Negev
Cleo Condoravdi	Stanford University
Elizabeth Coppock	University of Gothenburg
Luka Crnič	The Hebrew University of Jerusalem
Chris Cummins	The University of Edinburgh
Kathryn Davidson	Harvard University
Christopher Davis	University of the Ryukyus
Henriette De Swart	Utrecht University
Marco Del Tredici	ILLC, University of Amsterdam
Jakub Dotlačil	University of Groningen

Regine Eckardt	University of Konstanz
Tim Fernando	Trinity College Dublin
Michael Franke	University of Tbingen
Alexander Goebel	University of Massachusetts at Amherst
Zsofia Gyarmathy	Heinrich Heine Universität
Valentine Hacquard	University of Maryland
Andreas Haida	The Hebrew University of Jerusalem
Robert Henderson	University of Arizona
Wesley Holliday	University of California, Berkeley
Julie Hunter	Universitat Pompeu Fabra, Barcelona and Université Paul Sabatier, Toulouse
Sabine Iatridou	Massachusetts Institute of Technology
Thomas Icard	Stanford University
Luca Incurvati	ILLC, University of Amsterdam
Gerhard Jaeger	University of Tuebingen
Stefan Kaufmann	Department of Linguistics, University of Connecticut
Magdalena Kaufmann	University of Connecticut
Chris Kennedy	University of Chicago
Hadas Kotek	New York University
Angelika Kratzer	University of Massachusetts at Amherst
Manuel Križ	Institut Jean Nicod, École Normale Supérieure
Jeremy Kuhn	Institut Jean Nicod, École Normale Supérieure
Fred Landman	Tel Aviv University
Daniel Lassiter	Stanford University
Sven Lauer	University of Konstanz
Emar Maier	University of Groningen
Edwin Mares	Victoria University of Wellington
Salvador Mascarenhas	École Normale Supérieure, Department of Cognitive Studies
Louise McNally	Universitat Pompeu Fabra
Sarah Murray	Cornell University
Reinhard Muskens	Tilburg Center for Logic and Philosophy of Science
Rick Nouwen	Utrecht Institute for Linguistics OTS
Edgar Onea	University of Graz
Paul Portner	Georgetown University
Jessica Rett	University of California, Los Angeles
Craige Roberts	The Ohio State University
Maribel Romero	University of Konstanz
Jacopo Romoli	university of ulster
Mats Rooth	Cornell University
Daniel Rothschild	Columbia University
Gillian Russell	The University of North Carolina at Chapel Hill
Kjell Johan Sæbø	University of Oslo
Julian Schlöder	ILLC, University of Amsterdam
Bernhard Schwarz	McGill University
Yael Sharvit	University of California, Los Angeles
Benjamin Spector	Institut Jean Nicod, École Normale Supérieure
William Starr	Cornell University
Shane Steinert-Threlkeld	ILLC, University of Amsterdam

Martin Stokhof	ILLC, University of Amsterdam
Yasutada Sudo	University College London
Eric Swanson	University of Michigan
Kristen Syrett	Rutgers University
Anna Szabolcsi	New York University
Jakub Szymanik	University of Amsterdam
Lyn Tieu	Macquarie University
Bob van Tiel	Bielefeld University
Kai von Fintel	Department of Linguistics & Philosophy, MIT
Klaus Von Heusinger	Universität zu Köln
Galit W. Sassoon	Bar Ilan University
Matthijs Westera	Universitat Pompeu Fabra, Barcelona
Aaron Steven White	University of Rochester
Yoad Winter	Utrecht University
Yimei Xiang	Harvard University
Seth Yalcin	Berkeley
Hedde Zeijlstra	University of Göttingen
Ede Zimmermann	Goethe Universität Frankfurt
Sarah Zobel	Eberhard-Karls Universität Tuebingen
Willem Zuidema	ILLC, University of Amsterdam

Contents

Invited

Workshop: Causality and Semantics

Workshop: Formal and Distributional Perspectives on Meaning

Contributed

Object Mass Nouns in Japanese

Kurt Erbach, Peter R. Sutton, Hana Filip, and Kathrin Byrdeck*

Heinrich-Heine University, Düsseldorf, Germany
erbach@hhu.de, peter.r.sutton@icloud.com, hana.filip@gmail.com, byrdeck@phil.hhu.de

Abstract

Classifier languages are commonly taken to have no grammaticized lexical mass/count distinction, but rather have this distinction encoded through the syntax and semantics of classifiers (e.g. [4], [5], [15], [17]). We contest this claim by drawing on data from Japanese. We provide novel empirical evidence showing that Japanese has quantifiers (e.g. *nan-byaku to iu* 'hundreds of') which directly select only for nouns denoting atomic entities (*onna no hito* 'woman') without requiring any classifier support. Moreover, the selectional restrictions of such quantifiers lead us to identify a class of object mass nouns in Japanese, i.e. nouns that have atomic entities in their denotation and yet are infelicitous in syntactic environments which are diagnostic of count nouns. This contradicts the prediction in [5] that object mass nouns should not exist in classifier languages. If Japanese has object mass nouns, then we should be ready to accept that Japanese nominal system is endowed with a grammatical mass/count distinction, and one which bears a certain resemblance to that which we find in number marking languages (e.g. English). We propose a novel semantic analysis of Japanese lexical nouns and classifiers, based on Sutton & Filip [21], a framework that unites notions of context in Rothstein [16] and Landman [12], and motivates the idea that counting contexts can remove overlap so that count nouns have disjoint counting bases while mass nouns do not.

1 Introduction

Japanese, a typical classifier language, is commonly taken to have no grammaticized lexical mass/count distinction, i.e. no lexical distinction between different kinds of nouns sensitive to countability that is reflected in the grammatical behavior of nouns. Instead, this sort of distinction is thought to be encoded through the syntax and semantics of classifiers (e.g. [4], [5], [15], [17]). However, we provide evidence that Japanese has quantifiers like (e.g. *nan-byaku to iu* 'hundreds of') that distinguish between mass and count nouns, whose denotation does not align with the semantic (ontological) non-atomic and atomic domains. This then motivates the existence of a group of nouns in Japanese with the two hallmark properties of object mass nouns: (i) they have atomic denotations, and (ii) are infelicitous in syntactic environments which are diagnostic of count nouns. Object mass nouns (alternatively *fake mass nouns*) are nouns such as *furniture* or *mail* in English, and are predicted to not exist in classifier languages [5]. Our results show that Japanese indeed has object mass nouns and *a forteriori* that the Japanese lexical nominal system has a mass/count distinction that is directly relevant to the grammar of Japanese. We do so by exploring the properties of Landman [12] and Sutton & Filip [21], we argue that the key factor underpinning the count/mass distinction is whether or not the entities that count as 'one' in the denotation of a noun (the counting base) overlap. Mass concepts have overlapping counting bases and count concepts have disjoint counting bases. Japanese quantifiers like *nan-byaku to iu* ('hundreds of'), we argue, can only compose with nouns that determine disjoint counting bases, without any classifier support. But this can be taken as evidence for the existence of bona fide count nouns in Japanese, and hence for countability having direct grammatical relevance for the Japanese grammar.

*This research is funded as part of DFG Collaborative Research Centre 991: The Structure of Representations in Language, Cognition, and Science. Our thanks to attendees of the 18th Szklarska Poreba Workshop, and especially to Yasutada Sudo and Eric McCready for valuable feedback. Thanks, too, to our consultants Kaori Fujita, Saki Kudo, Sebastian Steinfelder, Aiko Tendo, and Yuko Wagatsuma.

2 Background

Object mass nouns are of key importance in determining whether or not a language has a mass/count distinction, because they provide evidence for the mismatch between conceptual individuation, on the one hand, and grammatical mass behavior, on the other hand. We use the term *inherently individuable* to refer to entities that are *objects* as opposed to *substances* in the sense of Soja et al. [18]. Nouns with inherently individuable denotations can be count (e.g., *chair*, *cat*) or mass (e.g., *furniture*, *jewelry*). *Object mass nouns* are those nouns which have inherently individuable extensions, but that are nonetheless infelicitous in counting constructions (e.g. # *I bought two furnitures*). Chierchia's [5] explanation for object mass nouns is the *copy-cat effect*, according to which atomically stable nouns like *furniture* copy mass noun properties as a result of lexical choice. The theory of [5] predicts that object mass nouns are expected to be found in number marking languages like English, because their nouns are differentiated with respect to their denotations, and because lexical choice makes it simple to characterize a potential count noun as a mass noun. Object mass nouns cannot exist in classifier languages, according to [5], because all their nouns uniformly denote kinds, as they freely occur as bare nominal arguments and cannot directly compose with numericals (1):

(1) a. inu go-*(hiki)
dog five-CL$_{small.animal}$
'five dogs'[15, p. 73]

b. kagu itsu-*(tsu)
furniture five-CL$_{general}$
'five pieces of furniture'

c. yūbinbutsu go-*(bu)
mail five-CL$_{printed.material}$
'five pieces of mail'

d. mizu go-*(hon)
water five-CL$_{bottle}$
'five bottles of water'

The analysis of classifier languages in [5], and of the most influential to date, is couched in a compositional type-theoretic framework in which all nouns uniformly denote kinds ($\langle k \rangle$), and numericals are adjectival (of type $\langle \langle e,t \rangle, \langle e,t \rangle \rangle$); consequently, overt morphemes, namely classifiers of type $\langle k, \langle e,t \rangle \rangle$ must intervene between numericals and their nominal arguments.

There is, however, a growing body of work showing that a more nuanced view of the nominal system of classifier languages is warranted [1], [6], [9], [14], [19], [20]. For example, Inagaki & Barner [9] use comparison tasks in classifier-less 'more than' constructions, Japanese nouns like *kutsu* ('shoe') and *kagu* ('furniture') are compared according to cardinality of individuals, but substance nouns like *karashi* ('mustard') are judged according to volume. These 'more than' constructions were not only classifier-less but also lacked any other grammatical cues for individuation (i.e. the presence or absence of count syntax) that could have triggered a cardinality or volume comparison. Inagaki and Barner [9] take these results as evidence that some Japanese nouns encode the grammatical feature ±INDIVIDUATED even in the absence of classifiers or other count syntax.

In support of the stronger claim, that there are reflexes of the mass/count distinction in at least some classifier languages, Sudo [19], [20] argues that certain Japanese quantifiers differentially select for count nouns. For instance, *nan-byaku to iu* ('hundreds of') and *dono N mo* ('whichever' or 'every') are felicitous with count nouns (e.g. *hon* 'book') but infelicitous with mass nouns (e.g. *ase* 'sweat'). In [19], [20], this observation is taken to mean that there are nouns with count denotations in Japanese; i.e. the inherent individuation of extensions is directly encoded by Japanese nouns, rather than in count syntax via a classifier constructions.

This begs the question, however, why it is that count nouns can nonetheless not be directly modified by numerical expressions in Japanese. Sudo's [19] explanation of this (which mirrors one also found in Krifka [11]), is that the reason that numerical expressions in Japanese can only denote abstract objects of type $\langle n \rangle$. This differs from number marking languages, such as English, in which numerical expressions have a numerical determiner interpretation. On Sudo's analysis, classifiers denote functions which map entities of type $\langle n \rangle$ into expressions of

the adjectival modifier type $\langle s, \langle e, t\rangle\rangle$, which freely compose with common noun interpretations.

While Inagaki & Barner [9] show that Japanese nouns encode a feature ±INDIVIDUATED, Sudo [19], [20] makes the stronger claim that there are grammatical reflexes of a mass/count distinction in Japanese. However, if these reflexes were simply correlated with the atomic/non-atomic, or the individuated/non-individuated, distinction, then the analysis of Chierchia [3], [5] could be upheld by adding sensitivity to natural atomicity or individuation to the relevant parts of the grammar. In other words, a critic of Sudo could insist that classifier languages, such as Japanese, do not display a mass/count distinction in their nominal system, but merely mark the notional distinction between entities that are or are not inherently individuable.

One of the main contributions of this paper is to provide a means of resolving this dispute: *evidence for object mass nouns*. If the grammatical tests outlined by Sudo [19], [20] (such as felicitous combination with *nan-byaku to iu* 'hundreds of') can be shown to bisect the class of common nouns in a way that does not mirror the prelinguistic notional individuable/non-individuable divide, then we have evidence that the grammar encodes more than the mere notional distinction. In particular, if we find nouns with inherently individuable extensions that are infelicitous with e.g, *nan-byaku to iu* ('hundreds of'), we will have evidence that Japanese has grammatical reflexes a genuine lexical mass/count distinction. With this aim in mind we conducted an experiment designed to provide evidence for object mass nouns in Japanese.

3 Empirical Evidence

In English, object mass nouns, such as *furniture* have atomic denotations and yet are infelicitous with count quantifiers as for example *each* and *every* (2). For Japanese, it has been proposed that quantifiers such as *nan-byaku to iu* ('hundreds of') work similarly to *many*, in that it is felicitous with count nouns like *onna no hito* ('woman') but infelicitous with mass nouns like *yuki* ('snow') [19] as in (3).

(2) a. every dog b. *every furniture c. *every snow

(3) a. nan-byaku to iu onna.no.hito
what-hundred to say woman
'hundreds of women'

b. #nan-byaku to iu yuki
what-hundred to say snow
#'hundreds of snow'

3.1 Experimental Design

Building mainly on the observations about Japanese data in Sudo [20], we designed an acceptability judgment experiment in which we asked 49 native speakers (in an online survey on www.crowdworks.jp) to judge the acceptability of 120 sentences, including distractor sentences, on a five point Likert scale ranging from 1, *hen da* ('odd'), to 5, *yoi* ('good'). Each sentence contained a combination of the quantifier *nan-byaku to iu* ('hundreds of'), which does not require a classifier, with a noun. We tested 22 collective artifact nouns like *kagu* ('furniture') and *yūbinbutsu* ('mail') (6), alongside 11 nouns denoting discrete entities/individuals (e.g. *onna no hito* 'woman' in (5)) and 11 nouns denoting undifferentiated stuff like *yuki* ('snow') in (4). Sentences with an average acceptability rating higher than the neutral rating 3 were categorized as felicitous, whereas sentences with an average rating lower than 3 were categorized as infelicitous and marked accordingly in (4)-(6).

(4) kinō yuki ga fu-tta. #nan-byaku to iu yuki wa mō toke-te
yesterday snow NOM fall-PST; #what-hundred to say snow NOM already melt-TE
shima-tta
finish-PST
'It snowed yesterday. #Hundreds of snow melted already.'

(5) toranpu-shi ga daitoryō ni na-tta ato, nan-byaku to iu
Trump-president NOM president ACC become-PST after; what-hundred to say
onna.no.hito ga washinton de neriarui-ta
woman NOM Washington LOC march-PST
'After Trump became president, hundreds of women marched in Washington DC.'

(6) kono yūmei-na aidorugurūpu wa fanretā ga aoku-te pinku no fūtō dake
this famous-ADV band TOP fanletter NOM blue-TE pink GEN envelope only
de mora-tte iru. #senshū mo nan-byaku to iu yūbinbutsu o mora-tte
with become-TE PRG; #lastweek too what-hundred to say mail ACC get-TE
i-ta
PRG-PST
'This famous band gets fan letters exclusively in pink and blue envelopes. Last week they got #hundreds of mail.'

3.2 Results

The main results are summarized in Figure 1. Across participants, judgments were found to be consistent using the Friedman test [8], meaning there was little variance in judgment per test item. Nouns denoting discrete entities (e.g. *onna no hito* 'woman') were judged to be felicitous with *nan-byaku to iu* ('hundreds of'), with the average judgment of 3.92. Nouns like *yuki* ('snow') denoting undifferentiated stuff had an average judgment of 2.08, and were infelicitous with *nan-byaku to iu* ('hundreds of'). The collective artifact denoting noun *yūbinbutsu* ('mail') is also infelicitous with *nan-byaku to iu* ('hundreds of'), receiving an average judgment of 2.25.

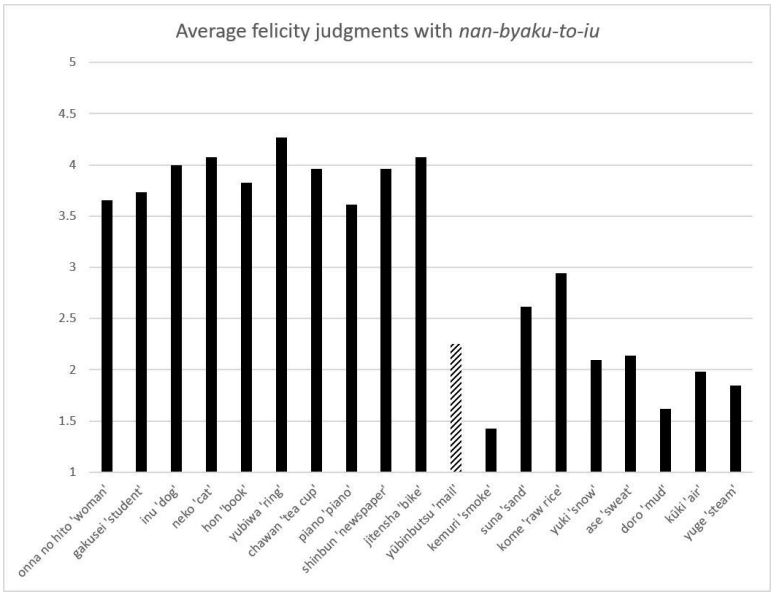

Figure 1: Bi-partite split of Japanese nouns based on compatibility with *nan-byaku to iu* ('hundreds of')

3.3 Discussion

The two competing hypotheses regarding the selectional restrictions of *nan-byaku to iu* ('hundreds of') are: (i) *nan-byaku to iu* ('hundreds of') is a suitable test of whether the extension of its argument noun has inherently individuable structure; (ii) *nan-byaku to iu* ('hundreds of') is a suitable test of whether its argument noun is count. If hypothesis (i) were correct, we would expect to see low judgement scores for all nouns that lack inherently individuable extensions and high scores for all nouns that have inherently individuable extensions. Evidence against hypothesis (i) and in favor of hypothesis (ii) would be for felicity scores with *nan-byaku to iu* ('hundreds of') to form a partition that does not mirror the individuable/non-individuable divide.

The results for *yūbinbutsu* ('mail') provide exactly the evidence we were looking for in support of hypothesis (ii). Although *nan-byaku to iu* ('hundreds of') is infelicitous with all nouns that denote substances (which lack an inherently individuable structure), *nan-byaku to iu* ('hundreds of') is not felicitous with all nouns that denote objects (which have an inherently individuable structure), namely *yūbinbutsu* ('mail'). In the absence of an alternative explanation for this pattern, we have good reason to conclude that Japanese has grammatical reflexes of the lexicalized mass/count distinction, and what is more, it also has object mass nouns. Both of these conclusions conflict with the common view of the nominal system in classifier languages, as, for instance, implemented in Chierchia's [5] recent analyses of the nominal semantics for classifier languages.

One possible counterargument to our conclusions, however, would be that *yūbinbutsu* ('mail') actually does not denote entities with an inherently individuable structure (at least in the way that Japanese speakers perceive of them). To reject this counterargument, we have begun to test native speaker judgements using the 'more than' test [9]. If a noun denotes entities with an inherently individuable structure, then there should be a felicitous *cardinality comparison* reading available for questions with 'more than'. If a noun denotes entities which lack an inherently individuable structure, then there should only be a felicitous *measure comparison* reading available for questions with 'more than'. To determine which of these options applies to *yūbinbutsu* ('mail'), we presented native speakers with sentences in which a measure or cardinality comparison is possible between two groups of items. Each sentence used one of our test nouns, and each had a group larger in volume but smaller in cardinality—e.g.s (7)-(9).

(7) Yuma wa futa-tsu no fūtō o uketo-tta. Hito-tsu wa atarashī shigoto
Yuma TOP 2-CL GEN envelopes ACC receive-PST. 1-CL TOP new work
no keiyaku de, mō-hito-tu wa apāto no keiyaku da. Satomi wa
GEN contract and, another-1-CL TOP apartment GEN contract COP. Satomi TOP
itsu-tsu no chīsai fūtō o uketo-tta. Doremo tomodachi kara no chīsai
5-CL GEN small envelope ACC receive-PST. Both friend from GEN small
tegami o fukun-de iru.
letter ACC contain-TE IRU.
'Yuma received two large envelopes, one with her new work contract and one with her apartment contract. Satomi got five small envelopes, each containing a short letter from a friend.'

(8) Mai wa yot-tsu no ōkī koshikake o ka-tta. Hiroaki wa itsu-tsu no
Mai TOP 4-CL GEN big armchair ACC buy-PST. Hiroaki TOP 5-CL GEN
kodomo-yō no chīsai isu o ka-tta.
child-use GEN small chair ACC buy-PST.
'Mei bought four large arm chairs. Hiroaki bought five small chairs for children.'

(9) Toma wa hito-tsu no ōkī yukidaruma o tsuku-tta. Mizuki wa itsu-tsu no
Toma TOP 1-CL GEN big snowman ACC make-PST. Mizuki TOP 5-CL GEN
yuki no tama o tsuku-tta.
snow GEN ball ACC make-PST.
'Toma made a big snowman. Isuki made five small snowballs.'

Following the presentation of each scenario, we asked the speakers to judge who has more *yūbinbutsu* ('mail')*isu* ('chair') and *yuki* ('snow'), respectively. In our pretest, *yūbinbutsu* ('mail') and *isu* ('chair') were judged according to cardinality comparison, while *yuki* ('snow') was judged according to volume. This is evidence that the extension of *yūbinbutsu* ('mail') has an inherently individuable structure.

In sum, the above data leads us to the conclusion that the Japanese nominal system does not only distinguish the notional individuable/non-individuable divide, but, in fact, has grammatical reflexes of the mass/count distinction, as attested by the presence of nouns which denote entities with an individuable structure, but nonetheless pattern, when combined with *nan-byaku to iu* ('hundreds of'), with substance denoting nouns. In other words, for Japanese we found evidence for the existence of object mass nouns, namely, *yūbinbutsu* ('mail').

4 Analysis

Our quantification and quantity judgment data respectively show clear grammatical and notional differences between Japanese nouns. To account for these grammaticized lexical differences in Japanese, we build on Sutton & Filip [21], who argue that the grammaticized lexical mass/count distinction is grounded in the (non-)resolution of overlap (also see [12]). To their model of lexical entries, we add a field for presuppositions (or, more neutrally, preconditions) for composition. We use presuppositions in two main ways: (i) in the entries for sortal classifiers, they capture selectional restrictions on the nouns with which they combine (e.g., that the argument noun must denote printed items); (ii) in the entries for sortal classifiers and count quantifiers, they require counting bases of argument nouns to be disjoint. In Section 4.1, outline an account of the semantics of the mass/count distinction in English (based on [21] and [13]). In Section 4.2, we extend this account to cover the Japanese data by providing an analysis of Japanese numerical expressions, classifiers, and *nan-byaku-to-iu* ('hundreds of').

4.1 Counting in context

Sutton & Filip [21] provide a cross-linguistic analysis of collective artifact nouns, such as *furniture* and *kitchenware*, in English. The puzzle they address is why collective artifact nouns stubbornly resist count-to-mass coercion when directly modified with a numerical expression (*# three furnitures/kitchenwares)*. Their solution is based on exploiting two types of counting contexts: *specific counting contexts*, which remove overlap in counting bases (the set of entities for counting); and *null counting contexts*, which allow overlap in counting bases.

Recent semantic analyses of the count/mass distinction [12, 13, 21] advocate representing the lexical entries of concrete nouns using ordered pairs. For example Landman [13] represents CN entries as $\langle body, base \rangle$, a pair of *base*, the counting base set, and *body*, a subset of the upward closure of *base* under sum. Following Krifka [10], Sutton & Filip analyze the lexical entries of nouns as including qualitative and quantitative criteria of application in the lexical entries of nouns. They are presented as ordered pairs, $\langle P, counting\_base(P) \rangle$. P is a property for the qualitative criteria of applying the noun concept. $counting\_base(P)$ specifies the quantitative criteria for applying the noun concept, which, crucially, includes information regarding: (a) whether or not the extension is inherently individuable; and (b) whether or not potentially countable entities are conceptualised in terms of a disjoint individuation schema (formalised in terms of counting contexts). Counting goes wrong when the counting base is an overlapping

set. Grammatical counting is possible when the counting base is a disjoint set.

Here we combine some elements of Landman's account (distinguishing our *body* from *base*) and some of Sutton & Filip's (inclusion of an individuation function interpreted at a counting context). Furthermore, following Filip & Sutton [7], we introduce a third projection to record preconditions and/or presuppositions relating to e.g., selectional restrictions, so CN entries have the form $\langle extension,\ c\_base,\ presup\rangle$.

The components *extension* and *c_base* are formed from up to three ingredients: A predicate, $P_{\langle e,t\rangle}$, an $\mathbf{IND}_{\langle\langle e,t\rangle,\langle e,t\rangle\rangle}$ function, and a counting context $c_{\langle\langle e,t\rangle,\langle e,t\rangle\rangle}$. For example $\mathbf{IND}(CAT)$ denotes the disjoint set of single cats. However, the **IND**-set is not always disjoint for other predicates. For example, the set of things that count as one for collective artifact nouns overlaps [12, 13, 21] e.g., a nest of tables ($a_1 \sqcup a_2 \sqcup a_3$), and the individual tables in the nest (a_1, a_2, a_3) each count as one with respect to *furniture*): $\{a_1, a_2, a_3, a_1 \sqcup a_2 \sqcup a_3\} \subset \mathbf{IND}(FURN)$.

Further 'perspectives' on **I**ND-sets are represented with counting contexts (of type $\langle\langle e,t\rangle,\langle e,t\rangle\rangle$) which come in two varieties:

Specific counting contexts $c_{i\geq 1}$: map sets onto maximally disjoint subsets. Intuitively, the specific counting contexts represent the imposition of a disjoint individuation schema.

The null counting context $c_{i\geq 1}$: is the identity function. It does not remove overlap if present: $\forall P \forall x[\mathrm{IND}(P)(x) \leftrightarrow c_0(\mathrm{IND}(P))(x)]$

Overlapping counting bases give rise to mass predicates, since grammatical counting requires disjointness. Therefore, evaluated at a specific counting context, the set denoted by $c_{i\geq 1}(\mathbf{IND}(FURN))$ is disjoint and disjoint counting bases mean grammatical countability. Evaluated at the null counting context c_0, the set denoted by $c_0(\mathbf{IND}(FURN))$ is overlapping and overlapping counting bases mean grammatical non-countability. Notice, however, that if an **IND**-set is anyway disjoint, there is no difference whether it is evaluated at a specific counting context or at the null counting context. Sutton & Filip [21] argue that this accounts for cross-linguistic variation in mass/count lexicalization patterns for collective artifact nouns. Whether or not a lexical entry indexes the **IND**-set to the null counting context or to a specific counting context is essentially a matter of lexical 'choice' (a parameter set language by language and noun by noun). This explains why nouns such as *cat*, and its cross-linguistic counterparts are all lexicalized as count ($\forall c_i[c_0(\mathbf{IND}(CAT)) = c_i(\mathbf{IND}(CAT))]$). It also explains why nouns which denote inherently individuable entities, but for which the **IND**-set of entities that count as one overlap can be lexicalized as either count or mass cross- and intra-linguistically. For example, we have an account for why we find the count noun *meubel* ('(piece of) furniture', Dutch) as well as the mass nouns *furniture* and *meubilair* ('furniture', Dutch).

Sutton & Filip also argue that predicates for substances and objects are semantically distinguished, which is supported by the ability of pre-linguistic infants to distinguish substances from objects [18]. Formally, this translates as there being no **IND** function in the lexical entries for substance denoting nouns (nouns which denote stuff that lacks an inherently individuable structure). Importantly however, the distinction between substances and objects does not perfectly mirror the mass/count distinction, as seen in the behavior of nouns like *furniture* which have objects in their denotation, yet grammatically pattern with nouns that denote substances, liquids, and gases. (For an explanation of why substance denoting nouns are almost always, but not universally lexicalized as mass, see Sutton & Filip [22].) Examples of a range of lexical entries are given in (1a–1f):

$$[\![cat]\!]^{c_i} = \lambda x.\langle c_i(\mathbf{IND}(\mathrm{CAT}))(x),\ \lambda y.c_i(\mathbf{IND}(\mathrm{CAT}))(y),\ \varnothing\rangle \tag{1a}$$

$$[\![cats]\!]^{c_i} = \lambda x.\langle {}^*c_i(\mathbf{IND}(\mathrm{CAT}))(x),\ \lambda y.c_i(\mathbf{IND}(\mathrm{CAT}))(y),\ \varnothing\rangle \tag{1b}$$

$$[\![meubel]\!]^{c_i} = \lambda x.\langle c_i(\mathbf{IND}(\mathrm{FURN}))(x),\ \lambda y.c_i(\mathbf{IND}(\mathrm{FURN}))(y),\ \varnothing\rangle \tag{1c}$$

$$[\![meubels]\!]^{c_i} = \lambda x.\langle {}^*c_i(\mathbf{IND}(\text{FURN}))(x),\ \lambda y.c_i(\mathbf{IND}(\text{FURN}))(y),\ \varnothing\rangle \quad (1d)$$

$$[\![furniture]\!]^{c_i} = \lambda x.\langle {}^*c_0(\mathbf{IND}(\text{FURN}))(x),\ \lambda y.c_0(\mathbf{IND}(\text{FURN}))(y),\ \varnothing\rangle \quad (1e)$$

$$[\![mud]\!]^{c_i} = \lambda x.\langle {}^*c_0(\text{MUD})(x),\ \lambda y.c_0(\text{MUD})(y),\ \varnothing\rangle \quad (1f)$$

Each entry contains *extension* (the truth conditions for applying the noun), *c_base* (the individuation schema for the noun concept), and *presup* (a slot which can specify extra lexical or compositional information or restrictions). The singular nouns *cat* and *meubel* ('(item of) furniture', Dutch) in (1a) and (1c) are interpreted at the specific counting context of utterance c_i. This removes any overlap in the counting bases. Their application conditions and individuation schemas express the same properties (the sets of single cats/items of furniture) and we get the grammatical count nouns *cat* and *meubel* ('(item of) furniture', Dutch). The plural forms ((1b) and (1d)) require the extensions to be single cats/items of furniture or sums thereof. The mass nouns *furniture* and *mud* in (1e) and (1f) are interpreted relative to the null counting context c_0. This allows overlap in the counting bases (i.e. different overlapping partitions of mud-stuff or different overlapping partitions of furniture into items), and so we get the grammatically mass nouns *mud* and *furniture*.

In short, the only difference in the entries for the plural count noun *meubels* and the mass noun *furniture* is whether the counting base is interpreted at c_0 or c_i. Interpretation at the null or at a specific counting context is essentially a matter of lexical choice. Hence, we expect both count and mass terms, cross-linguistically, to express this concept.

4.2 Nominal semantics in Japanese

Lexical entries for common nouns. On our analysis, lexically simple Japanese nouns have lexical entries that closely match those in number marking languages. Count nouns like *isu* ('chair') are interpreted at a specific counting context that specifies disjoint counting base 2b. Object mass nouns like *yūbinbutsu* ('mail', 2c) and *yuki* ('snow', 2a) have entries saturated with the null counting context c_0, but *yuki* ('snow', 2a), as a substance denoting noun is interpreted without the **IND**-function. The one difference between e.g., Japanese and English is that, since Japanese has a highly restricted (and even then, optional) use of plural morphology, lexically simple Japanese nouns have number neutral extensions (that include entities and sums thereof).

$$[\![yuki]\!]^{c_i} = \lambda x.\langle {}^*c_0(\text{SNOW})(x),\ \lambda y.c_0(\text{SNOW})(y),\ \varnothing\rangle \quad (2a)$$

$$[\![isu]\!]^{c_i} = \lambda x.\langle {}^*c_i(\mathbf{IND}(\text{CHAIR}))(x),\ \lambda y.c_i(\mathbf{IND}(\text{CHAIR}))(y),\ \varnothing\rangle \quad (2b)$$

$$[\![yūbinbutsu]\!]^{c_i} = \lambda x.\langle {}^*c_0(\mathbf{IND}(\text{MAIL}))(x),\ \lambda y.c_0(\mathbf{IND}(\text{MAIL}))(y),\ \varnothing\rangle \quad (2c)$$

Counting with classifiers in context. Both count nouns and object mass nouns can be combined with a numerical expression when there is an intervening sortal classifier. Following Krifka [11], we assume that numericals (e.g. 3a) denote numbers of type $\langle n\rangle$. Key to our analysis are four functions for Japanese, sortal classifiers, e.g. *bu* ('printed item') in 3b: (1) they map type n expressions to expressions of the type for numerical determiners; (2) they also ensure that the counting base predicate provided by the argument noun is evaluated at the counting context of utterance. For example, if the argument noun is saturated with the null counting context, then the equivalence in 3c ensures that overlap is removed in the resulting counting base predicate; (3) they add a presupposition that the counting base is disjoint (so as to be fit for counting); (4) they add a presupposition that the argument predicate is of some restricted sort. For example, for *bu* ('printed item'), it is the presupposition that the argument predicate denotes a subset of PRINTED.ITEM (this also acts to filter out combination

with substance denoting nouns).[1]

$$[\![san]\!]^{c_i} = 3 \tag{3a}$$

$$[\![bu]\!] = \lambda n.\lambda P.\lambda c.\lambda x.\langle \pi_1(P(x)),\ \mu_{card}(x, \lambda y.c(\pi_2(P(x)))(y)) = n, \tag{3b}$$
$$\text{DISJ}(\lambda y.c(\pi_2(P))(x)(y))\ \wedge\ \lambda x.\pi_1(P(x)) \subseteq \text{PRINTED.ITEM}\rangle$$

$$\forall P.\forall c.[c(c_0(P)) \longleftrightarrow c(P)] \tag{3c}$$

$$[\![y\bar{u}binbutsu\ san\text{-}bu]\!]^{c_i} = \lambda x.\langle {}^{*}c_0\mathbf{IND}(\text{MAIL}(x),\ \mu_{card}(x, \lambda y.c_i(\mathbf{IND}(\text{MAIL})(y)) = 3, \tag{3d}$$
$$\text{DISJ}(\lambda y.c_i(\mathbf{IND}(\text{MAIL})(y))\ \wedge$$
$${}^{*}c_0\mathbf{IND}(\text{MAIL}) \subseteq \text{PRINTED.ITEM}\rangle$$

The result of combination, expressed in 3d, is the set of items of mail that have cardinality 3 at the counting context of utterance under the presupposition that the set of single items is disjoint and that mail is a type of printed material.

The quantifier *nan-byaku-to-iu* ('hundreds of') has, on our analysis, a semantics that closely resembles that of a numerical combined with a sortal classifier. The key difference is that it does not introduce a new context variable (intuitively, it does not provide information for selecting a disjoint individuation schema). Other differences are that the cardinality it specifies is underspecified (which we represent with the context-determined type n variable n_c), and *nan-byaku-to-iu* ('hundreds of') does not place extra restrictions (e.g., printed item) on the argument noun.

$$[\![nan\text{-}byaku\text{-}to\text{-}iu]\!] = \lambda P.\lambda x.\langle \pi_1(P(x)),\ \mu_{card}(x,\ \pi_2(P(x))) \geq n_c,\ \text{DISJ}(\pi_2(P(x)))\rangle \tag{4a}$$

$$[\![nan\text{-}byaku\text{-}to\text{-}iu\ isu]\!] = \lambda x.\langle \text{CHAIR}(x),\ \mu_{card}(x,\ \lambda y.\ c_i(\mathbf{IND}(\text{CHAIR})(y)) \geq n_c, \tag{4b}$$
$$\text{DISJ}(\lambda y.\ c_i(\mathbf{IND}(\text{CHAIR})(y))\rangle$$

$$[\![nan\text{-}byaku\text{-}to\text{-}iu\ y\bar{u}binbutsu]\!] = \lambda x.\langle \text{MAIL}(x),\ \mu_{card}(x,\ \lambda y.\ c_i(\mathbf{IND}(\text{MAIL})(y)) \geq n_c, \tag{4c}$$
$$\text{DISJ}(\lambda y.\ c_0(\mathbf{IND}(\text{MAIL})(y))\rangle$$
$$\Rightarrow \text{FALSE PRECONDITION!}$$

This simple difference is enough to capture the selectional restrictions of *nan-byaku-to-iu* ('hundreds of') since it predicts that *nan-byaku-to-iu* ('hundreds of') will only straightforwardly felicitously combine with count nouns. In 4b, $\lambda y.\ c_i(\mathbf{IND}(\text{CHAIR})(y)$ is a disjoint set, so *isu* ('chair') is count and *nan-byaku-to-iu isu* is felicitous. In 4c, $\lambda y.\ c_0(\mathbf{IND}(\text{MAIL})(y)$ is not disjoint, so *yūbinbutsu* is mass and *nan-byaku-to-iu yūbinbutsu* is infelicitous.

5 Conclusion

Our novel empirical evidence confirms that the Japanese quantifier *nan-byaku to iu* 'hundreds of' is a suitable diagnostic test for the count status of Japanese nouns. Moreover, and more importantly, we show that Japanese has object mass nouns, contrary to the prediction in [5] that they should not exist in classifier languages. This has not been shown in any previous work on classifier languages, to the best of our knowledge. Based on our findings for Japanese, we reject the common view that the mass/count distinction in *all* classifier languages is solely reflected in the syntax and semantics of their classifier systems, advocated for in [4], [5] or [15] among others. In Japanese, we find direct grammatical reflexes of the grammaticized

[1] Lexical entries for classifiers make use of product types (e.g.[2]). For example, an expression $\langle X_a, Y_b, Z_c\rangle$ is of type $\langle a \times b \times c\rangle$. We also use projection functions π_1 and π_2 such that $\pi_1(\langle X_a, Y_b, Z_c\rangle)=X_a$ and $\pi_2(\langle X_a, Y_b, Z_c\rangle)=Y_b$.

lexical mass/count distinction, as we argue. If there are classifier languages like Japanese that have a grammatical mass/count distinction in the lexicon, then the nominal system of such classifier languages are typologically closer to the nominal systems in languages with a bona fide lexical mass/count distinction, like English, than has previously been assumed. This conclusion requires a novel formal analysis of Japanese nouns, numericals, classifiers, and quantifiers, which we have provided based on [21].

References

[1] Alan Bale and David Barner. Semantic triggers, linguistic variation and the mass-count distinction. In *Count and mass across languages*, pages 238–260. Oxford University Press, 2012.

[2] Bob Carpenter. *Type-logical semantics*. MIT press, 1997.

[3] Gennaro Chierchia. Plurality of mass nouns and the notion of "semantic parameter". In *Events and Grammar: Studies in Linguistics and Philosophy Vol. 7*, pages 53–103. Kluwer, 1998.

[4] Gennaro Chierchia. Reference to kinds across language. *Natural language semantics*, 6(4):339–405, 1998.

[5] Gennaro Chierchia. Mass nouns, vagueness and semantic variation. *Synthese*, 174(1):99–149, 2010.

[6] Jenny Doetjes. Count/mass distinctions across languages. In Claudia Maienborn, Klaus von Heusinger, and Paul Portner, editors, *Semantics: An International Handbook of Meaning*, volume 3, pages 2559–2580. Walter de Gruyter, 2012.

[7] Hana Filip and Peter Sutton. Singular count nps in measure constructions. *Semantics and Linguistic Theory*, 27(0), forthcoming.

[8] Milton Friedman. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the american statistical association*, 32(200):675–701, 1937.

[9] Shunji Inagaki and David Barner. Countability in absence of count nouns: Evidence from japanese quantity judgments. *Studies in language sciences*, 8:111–125, 2009.

[10] Manfred Krifka. Nominal reference, temporal constitution and quantification in event semantics. In Renate Bartsch, J. F. A. K. van Benthem, and P. van Emde Boas, editors, *Semantics and Contextual Expression*, pages 75–115. Foris Publications, 1989.

[11] Manfred Krifka. Common nouns: A contrastive analysis of English and Chinese. In G.N. Carlson and F. J. Pelletier, editors, *The Generic Book*, pages 398–411. Chicago University Press, 1995.

[12] Fred Landman. Count nouns-mass nouns, neat nouns-mess nouns. *Baltic International Yearbook of Cognition, Logic and Communication*, 6(1):12, 2011.

[13] Fred Landman. Iceberg semantics for count nouns and mass nouns: The evidence from portions. *The Baltic International Yearbook of Cognition Logic and Communication*, 11, 2016.

[14] Peggy Li, Yarrow Dunham, and Susan Carey. Of substance: The nature of language effects on entity construal. *Cognitive psychology*, 58(4):487–524, 2009.

[15] Keiko Muromatsu. Classifiers and the count/mass distinction. *Functional structure (s), form and interpretation: Perspectives from East Asian languages*, pages 65–96, 2003.

[16] Susan Rothstein. Counting and the mass/count distinction. *Journal of semantics*, 27(3):343–397, 2010.

[17] Susan Rothstein. *Semantics for Counting and Measuring*. Cambridge University Press, 2017.

[18] Nancy N Soja, Susan Carey, and Elizabeth S Spelke. Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition*, 38(2):179–211, 1991.

[19] Yasutada Sudo. The semantic role of classifiers in japanese. *Baltic International Yearbook of Cognition, Logic and Communication*, 11(1):10, 2016.

[20] Yasutada Sudo. Countable nouns in japanese. *Proceedings of WAFL*, 11(1):11, 2017.

[21] Peter Sutton and Hana Filip. Counting in context: count/mass variation and restrictions on coercion in collective artifact nouns. *Semantics and Linguistic Theory*, 26(0):350–370, 2016.

[22] Peter Sutton and Hana Filip. Individuation, reliability, and the mass/count distinction. *Journal of Language Modelling*, 5(2):303–356, 2017.