Presuppositions, Subcategorization, Lexicalism, Coercion and the Mass-Count Distinction*

Alan Bale
Concordia University

Abstract This paper compares two potential explanations of sentential anomalies as it relates to the distribution of quantifiers and the mass-count distinction: one that is rooted in presupposition failure (Chierchia 1998), and another that employs syntactic features and constraints on subcategorization. The paper argues that the syntactic explanation fairs better than the presuppositional approach. However, the argument is not that Chierchia (1998) is wrong, but rather that the major insights of his theory should be recast in terms of a feature logic, feature interpretation and subcategorization. In essence, this paper is a defence of Chierchia 1998 with a slight modification which only partially separates the distributional characteristics of quantifiers from the semantic nature of nouns. This defence is particularly relevant given that even Chierchia himself (2010) rejects the underlying semantics that is essential to Chierchia 1998's explanation. Furthermore, it is argued that in order for subcategorization to adequately account for the facts, lexical items need to be stored with syntactic features and cannot be featureless roots.

Keywords: Mass-count Distinction, Quantifier distribution, Presupposition failure, Subcategorization, Lexicalism, Lexical features, Coercion

1 Introduction

In linguistics, a critical part of evaluating theories involves assessing how well they account for speaker intuitions. However, it cannot always be determined a priori how best to frame such an account. This especially holds for the empirical differences between catastrophic presupposition failure and subcategorization viola-

* I would like to thank Brendan Gillon, September Cowley, and two anonymous reviewers for commentary on earlier drafts of this paper. I would also like to thank the members of McGill's Semantics Reading Group. This research would not have been possible without a Social Sciences and Humanities Research Council of Canada (SSHRC) Insight Grant #2016-2021.

tions.[1] Consider some typical examples of odd-sounding combinations of quantifiers and nouns as shown in (1).[2]

(1) a. # Most **chair** can be easily destroyed with a good shovel.
(c.f., Most **chairs** can be easily destroyed with a good shovel.)

b. # Each **furniture** is heavy.
(c.f., Each **couch** is heavy.)

c. # Mary used a shovel to destroy every **couches**.
(c.f., Mary used a shovel to destroy every **couch**.)

Chierchia (1998) proposed that deviant sentences like in (1) are due to presupposition failure. According to him, quantifier-noun combinations are defined/coherent only if the nominal denotation satisfies certain semantic characteristics, such as being closed under group formation, being limited to atomic minimal parts, or being free of atomic minimal parts. These preconditions on well-formativeness track the semantic differences and similarities among singular count nouns, plural nouns and mass nouns.

For example, according to this explanation, a hearer can interpret a sentence like (1a) only if they assume that the denotation of the singular noun *chair* is closed under the group formation operator. In most contexts, this is an unaccommodatable presupposition. It violates common knowledge about the meaning of *chair*, namely that it denotes the set of singular chairs and prohibits groups.

In contrast to the presuppositional account, a syntactic account of the sentences in (1) would maintain that nouns can be divided into syntactic subcategories, such as mass, count and plural. The distribution of the quantifiers merely reflects the syntactic requirements that certain quantifiers select for certain types of subcategories. At first blush, the syntactic account might seem more appealing since the differences between quantifiers feels so categorical—at least in the examples in (1). (Coercion and category-flexible nouns make this distinction less "categorical" on the surface, although such flexibility is characteristic of all of the non-functional syntactic categories. I will address such data more thoroughly in section 4.)

Although I eventually defend a revised syntactic account, I think that this is faulty reasoning. The presuppositional account is equally based on categorical differences between the different types of nouns: they just happen to be abstract semantic differences rather than syntactic ones. Thus, the categorical nature of the

1 Indeed, the difficulty of teasing apart these types violations is one of the central topics in Chomsky 1965.

2 In this paper, the deviant status of sentences will be marked with the #-symbol. This should be interpreted in a theory neutral way that is underspecified for whether the oddness is due to presupposition failure or syntactic subcategorization.

deviation is an essential part of both theories. The question is whether the relevant categorical differences are syntactic or semantic in nature.

In this paper, I summarize Chierchia's (1998) arguments that the presuppositional account of the distribution of quantifiers is superior to existing syntactic theories. It not only accounts for the distribution of quantifiers but also accounts for two significant cross-linguistic generalizations: 1) there is no quantifier cross-linguistically that applies to both singular mass and count nouns to the exclusion of plural nouns; and 2) there is no quantifier that selects for both singular and plural count nouns to the exclusion of mass nouns. Traditional syntactic systems that use features like ±COUNT and ±PLURAL are unable to account for these types of cross-linguistic generalizations (see for example, Chomsky 1965; Gillon 1999). A similar criticism holds for more modern type-theoretical accounts of the mass-count distinction (as in Rothstein 2010).

Unfortunately, Chierchia's (1998) account also has some major weaknesses. First, it relies on the premise that the denotations of plural nouns do not contain any atomic minimal parts. Evidence from pluralia tantum nouns and from plural nouns in downward entailing contexts undermines this premise. Second, it predicts that nouns with identical denotations should have an identical distribution. In particular, nouns which denote the empty-set should theoretically be able to combine with any type of quantifier. Empirically speaking, this prediction is not borne out.

In contrast to the presuppositional account, this paper proposes a syntactic theory that builds off of Chierchia's (1998) insights and generalizations. It introduces three monovalent features, one that marks singular count nouns (SG), one that marks plural nouns (PL) and one that marks both plural and mass nouns (CL). Critically, plural and mass nouns share a feature to the exclusion of singular count nouns, and singular count nouns have no feature in common with either plural or mass nouns (other than the general nominal category feature). With fairly standard subcategorization rules, this syntactic system can account for the same cross-linguistic generalizations as Chierchia 1998 without committing to the premise that singular, plural and mass nouns must always have different types of denotations. However, this does not mean that the syntactic features have no semantic effect. Features are associated with semantic operations that determine the denotational characteristics of the nominal subcategories.

The syntactic theory is not without its problems. For example, the theory maintains that mass nouns and singular count nouns do not share any syntactic features: in particular they are not both singular. This hypothesis might seem controversial given that singular count nouns and mass nouns trigger similar types of verbal agreement. However, this paper provides evidence that such agreement can be explained by appealing to elsewhere conditions and/or impoverishment rules, similar to those discussed in Harley 2008 and Bobaljik 2001.

The main goal of this paper is to provide the best possible argument for a syntactic explanation of the distribution of quantifiers—one that is better than any present or future semantic explanation (i.e., not just the one presented in Chierchia 1998). By no means does this paper offer a knock-down argument, but it does present several significant road-blocks to a purely semantic account of the mass-count distinction. Furthermore, the syntactic account offered in this paper has two important corollaries: 1) the mass-count distinction (and the features involved in such a distinction) must be part of a noun's lexical entry (i.e., the lexicon is not made of featureless roots); 2) the same mechanisms that allow for the interpretation of novel words within a familiar syntactic context (a.k.a., nonce terms) also provides the best account of so-called "coercion" facts (i.e., the use of a prototypical mass noun is a count context and vice versa).

The outline of this paper is as follows. Section 2 reviews some of the literature on catastrophic presupposition failure before discussing Chierchia 1998. Section 3 outlines the basic strategy behind a syntactic account. Section 4 outlines some of the semantic details associated with the syntactic account. Section 5 summarizes the arguments and concludes this paper.

2 Presupposition Failure

Presupposition failure has often been invoked to account for the deviant nature of certain sentences, including odd combinations of quantifiers with nouns. For ease of exposition, I will define presuppositions along the lines of Stalnaker (1973): presuppositions are the propositions entailed by the common ground, where the *common ground* is what is known, and known to be known, (and known to be known to be known etc.) by both the speaker and hearer. In a slight abuse of terminology, I will sometimes characterize a sentence S as "presupposing" a proposition p if a precondition for interpreting S is that it be evaluated in a context where the common ground entails p. Often linguists and philosophers hypothesize that these *preconditions* are a result of partial functions or truth value gaps (see Frege 1892, Strawson 1952, van Fraassen 1966, 1968, among others). I will follow this tradition in my representation of truth conditions in this section, although nothing I say depends on this theoretical choice.

Technically speaking, sentences do not "presuppose" anything, only people do. However, it is well-known that certain phrases and words signal that the people engaged in a conversation have a certain type of common ground, as shown with the adverb *again* in (2a), the verb *know* in (2b) and the cleft construction in (2c).

(2) a. Hilary Clinton is campaigning to be president again.
b. My mother knows that Bertrand Russell co-wrote Principia Mathematica.

c. It was Oswald who shot John F. Kennedy.

In normal circumstances, even if one denies the statements in (2) or simply doubts their veracity, one still implicitly accepts that it is common ground that Hilary Clinton campaigned to be president sometime in the past, that Bertrand Russell co-wrote Principia Mathematica, and that somebody shot John F. Kennedy.

The rest of this section discusses the interaction between presupposition failure and accommodation. In particular, it reviews theories that account for sentential anomalies by appealing to a hearer's inability or unwillingness to accommodate the speaker. The section concludes by discussing and criticizing Chierchia's (1998) account of the distribution of quantifiers.

2.1 Oddity and Accommodation

A potential source of oddity arises when a sentence S with the presupposition p is uttered in a context where the common ground does not entail p. I say "potential" since, as discussed by Stalnaker (1973) among others (e.g., von Fintel 2008, Heim 1990, 1983, Karttunen 1974, etc.), the oddness of a sentence can often be mitigated through a process that Lewis (1979) called *accommodation*. Accommodation can be characterized as follows: If a statement is uttered and it has the presupposition p, but the current common ground does not entail p, then the hearer will often accommodate the speaker by updating the common ground so that it entails p before evaluating the statement.

As Stalnaker (1973) argued, this type of update is a natural extension of Gricean requirements on cooperation. According to the Cooperative Principle, the hearer assumes that the speaker's statement will be appropriately informative (MAXIM OF QUANTITY). However, if a sentence has a presupposition that is not entailed by the common ground, then that sentence would be completely uninformative (since it would be undefined and thus lack a meaning). To rescue the assumption that the speaker is being cooperative, the hearer needs to update the common ground. As Karttunen (1974) characterized the process...

> ...a sentence is always taken to be an increment to a context that satisfies its presuppositions. If the current conversational context does not suffice, the listener is entitled and expected to extend it as required...

The mental gymnastics that the hearer is expected to go through is not so different from the type of reasoning a hearer engages in when computing quantity implica-

tures.[3] Critically, speakers can plan their speech acts with the assumption that hearers will accommodate (within reason), much like they can assume that hearers will compute quantity implicatures. As Stalnaker (1973) characterized this process...

> ...as soon as there are established and mutually recognized rules relating what is said to the presumed common beliefs, it becomes possible to exploit those rules by acting as if the shared beliefs were different than they in fact are known to be...

In fact, blocking accommodation, theoretically speaking, would require either some extra grammatical principles or a careful redefinition of the MAXIM OF QUANTITY.

The pragmatic nature of accommodation might be helpful when trying to explain degrees of oddity. I will not spell out a detailed theory here but I can sketch out the basic outline of what such a theory should look like. The fundamental intuition is that the harder it is to accommodate a sentence, the more odd it sounds. For example, suppose I am engaged in a conversation with my friend John about a dinner he had on the previous night with our mutual acquaintance Mary. It would be completely acceptable for John to say the sentence in (3).

(3) I will have dinner with Mary again tomorrow night.

However, the exact same sentence would sound slightly odd if I had no idea whether or not John had dinner with Mary on the previous night and he knew I had no idea. Still, it would not be completely unacceptable since I could update my conception of the common ground to incorporate that he had dinner with Mary sometime in the past. In fact, the same sentence improves dramatically in the same type of context if the speaker makes explicit his expectation for me to accommodate, as in (4).

(4) a. **Me**: What did you do last night?
b. **John**: Let me put it this way. I will have dinner with Mary again tomorrow night.

This response is still not perfect, but it is much better than making the statement in (3) out of the blue.

Note also, that the sentence in (3) becomes quite a bit odder if John had previously uttered the sentence in (5).

3 The main differences between accommodation and quantity implicatures are two-fold. i) With accommodation, the ordering of the added information is critical—the common ground must be updated before the statement can be informative. This is not true with respect to quantity implicatures. The implied information can be added before or after the information contained in the literal meaning of the statement. ii) Quantity implicatures are derived through competition with alternative statements whereas accommodation is derived by calculating the truth conditions of a single sentence.

(5) I have never had dinner with Mary.

Once the sentence in (5) has been integrated into the common ground, then I am no longer willing to accommodate the presuppositions introduced by (3). It is important to note, it is not that I could not see how to accommodate the sentence (I simply need to remove the proposition that John never had dinner with Mary before and then add the contrary proposition to my conception of the common ground), however this would require reanalyzing previous steps in our conversation. Although in principle it is not difficult for me to conceptualize how the sentence could be informative or to imagine a different conversation where the sentence would not sound odd, in practice my willingness to accommodate has limits.

If the inability or unwillingness to accommodate a sentence's presuppositions leads to a greater sense of oddity, then one could image that it might be possible for a sentence to have presuppositions that noone would be willing to accommodate into the common ground, no matter what context the sentence is uttered in. Following others, I will label this *catastrophic presupposition failure* and in the sections below I will assess whether this failure is a viable way to account for the oddity of certain quantifier-noun combinations.

2.2 Catastrophic presupposition failure and Chierchia 1998

Chierchia (1998) hypothesizes that the distribution of quantifiers among mass nouns, singular count nouns and plural nouns is determined by denotational preconditions imposed by the quantifiers themselves. This proposal runs directly contrary to a more syntactic approach, such as the one offered by Chomsky (1965), where lexical items are subcategorized to appear in certain types of phrase structure environments. I will talk about the syntactic alternative in more detail in section 3.

The basis of Chierchia's presuppositional account of the mass-count distinction relies on the denotational categories he associates with mass nouns, singular count nouns and plural count nouns. Chierchia (1998) adopts the fairly standard assumption that the denotation of singular count nouns is a set of atoms (i.e., singular entities that have no subparts). This assumption is supported by how speakers use demonstrative phrases and nominal predicates such as *that couch* and *is a couch*. Such demonstratives and predicates are only used to pick out singular entities and are only true of singular entities, respectively.

In contrast, demonstrative phrases with plural nouns (such as *those couches*) are generally used to pick out groups of two or more entities and similarly, nominal predicates with plural nouns (such as *are couches*) are generally only true of groups of two or more. Based on evidence like this, and following Link (1983), Chierchia (1998) analyses plural nouns as denoting a set of groups, where each group consists of more than one individual.

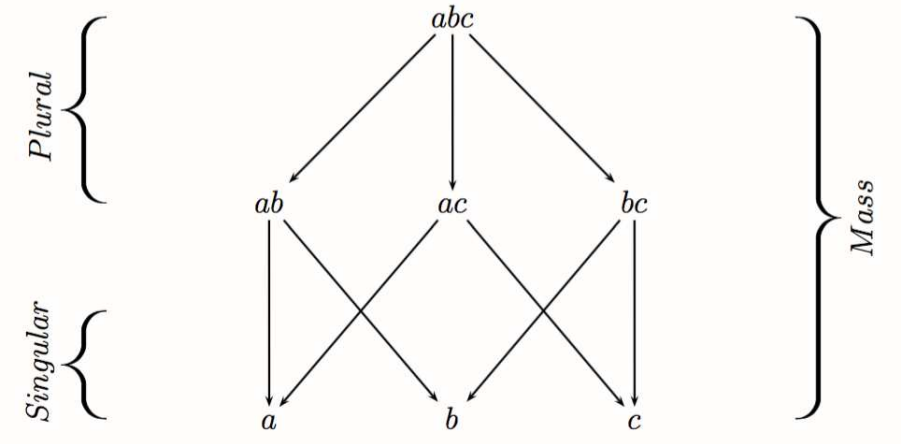


Figure 1 **A visual representation of the denotational restrictions.** The interpretation of singular nouns are restricted to sets of singular entities and the interpretation of plural nouns to sets of groups, while the interpretation of mass nouns is much broader. For example, if a, b and c are couches, and there are no other pieces of furniture in the domain other than these couches, then $[\![couch]\!] = \{a, b, c\}$, $[\![couches]\!] = \{ab, ac, bc, abc\}$, and $[\![furniture]\!] = \{a, b, c, ab, ac, bc, abc\}$.

Unlike plural and singular count nouns, demonstrative phrases with mass nouns (such as *that furniture*) can be used to refer to either singular entities (e.g., a single couch) or groups of entities (e.g., several couches). Similarly, predicates with mass nouns (such as *is/are furniture*) can be true of both groups and individuals. To account for these facts, Chierchia assigns mass nouns a denotation that is true of both groups and individuals. The division between singular count, plural count and mass nouns is represented in the graph in Figure 1.

To account for the distributional characteristics of quantifiers, Chierchia (1998) hypothesizes that most quantifiers specify certain preconditions that must be satisfied before the quantifier can coherently combine with its nominal complement. These preconditions critically rely on the denotational difference outlined in Figure 1. As Chierchia observes, in mass-count languages, there are four major categories of quantifiers:

i. **Singular Quantifiers**: Quantifiers like *every* require that the complement noun have a singular denotation;

ii. **Plural Quantifiers**: Quantifiers like *several* require that the complement noun have a plural denotation;

iii. **Non-Singular Quantifiers**: Quantifiers like *all* require that the complement noun have a non-singular denotation;

iv. **General Quantifiers**: Quantifiers like *no* impose no requirements.

To capture these facts, Chierchia (1998) proposes three different types of denotational checking functions, as outlined in (6).

(6) Let all potential domains D be closed under join, such that if * is the join closure operation, then $D = {}^*D$. Also, let D be generated by the set of atoms in D (i.e., if ATOM is the set of atoms in D, then $D = {}^*\text{ATOM}$).

a. **Singular Check**: Let S be a function from subsets of D to truth values such that for any set X, $S(X) = 1$ iff $X \subseteq \text{ATOM}$.

b. **Plural Check**: Let P be a function from subsets of D to truth values such that for any set X, $P(X) = 1$ iff there is a subset Z of ATOM such that $({}^*Z - Z) = X$ (i.e., elements in X are all the plural entities generated from a set of singulars, but X does not contain any of the singulars).

c. **Non-singular Check**: Let CL be a function from subsets of D to truth values such that for any set X, $CL(X) = 1$ iff ${}^*X = X$. (Note, CL stands for *closed* as in "X is closed under join.")

The function S is only true of singular count noun denotations. The function P is only true of plural count noun denotations. The function CL is only true of plural count noun and mass noun denotations.

With these functions in mind, the preconditions on quantifiers can be defined as follows.

(7) PRECONDITIONS ON QS

a. **Example of a Singular Q**: For any subset Z of the domain D, $[\![every]\!]^M(Z)$ is defined only if $S(Z) = 1$.

b. **Example of a Plural Q**: For any subset Z of the domain D, $[\![several]\!]^M(Z)$ is defined only if $P(Z) = 1$.

c. **Example of a Non-Singular Q**: For any subset Z of the domain D, $[\![all]\!]^M(Z)$ is defined only if $CL(Z) = 1$.

d. **Example of a General Q**: For any subset Z of the domain D, $[\![no]\!]^M(Z)$ is always defined.

Deliberately absent from these quantifier types is the contrast between *many* vs. *much*, *little* vs. *few*, and *less* vs. *fewer*. On the surface, *many*, *few* and *fewer* pattern like *several* in that they only appear with plural count nouns. However, *much*, *little*

and *less* do not pattern like any of the quantifiers mentioned above. Descriptively, all three quantifiers only seem to combine with mass noun complements.

Theoretically speaking, it is possible to specify a third denotational checking function that can account for this distribution, as specified in (8).

(8) A POSSIBLE MASS NOUN PRECONDITION: For any subset Z of the domain D, $[\![much]\!]^M(Z)$ is defined only if $CL(Z) = 1$ and $\neg(PL(Z))$.

However, there are two reasons why we might not want to have this fourth possible precondition. First, the existence of quantifiers that only appear next to mass nouns is rare cross-linguistically. Second, in languages that seem to have such quantifiers on the surface, there always exists a related quantifier that has the same meaning but exclusively applies to plural nouns.

Given these facts, it seems more reasonable to hypothesize that *much* and *many* are allomorphic variations of the same underlying quantifier. Likewise for *little* and *few*, and *less* and *fewer*. In other words, the root quantifier applies to both plural and mass nouns much like *most* however the variation in form is triggered by other factors in the syntactic environment. For example, the quantifier underlying *many/much* appears with both plural and mass nouns, but surfaces as *many* when appearing next to plurals (see Wellwood 2014 for a detailed analysis of *many* as an allomorph of *much*). For the rest of this paper, I will assume that this account of the quantifier distribution is on the right track and hence that there is no quantifier that has a mass-noun checking mechanism as a precondition. I will also spell out the allomorphic mechanism in more detail in section 3.

Chierchia's (1998) account of the distribution of quantifiers is a perfect example of catastrophic presupposition failure. For example, consider the distinctly odd sentences in (9).

(9) a. # All man will eventually be alone.
b. # Every men will eventually be alone.

The quantifier *all* signals to the hearer that they should implicitly assume that $[\![man]\!]^M$ is closed under join. Otherwise the sentence would not have an interpretation. Similarly, the quantifier *every* signals to the hearer that they should implicitly assume that $[\![men]\!]^M$ is a subset of the atoms in the domain.

However, in most contexts, these preconditions are not accommodatable. The noun *man* is singular and if there is more than one man in the context, it is impossible to interpret the noun as being closed under join. In contrast, the noun *men* is plural and if there is more than one man in the context, it is impossible to interpret the noun as being a subset of the set of atoms. Hence why, according to Chierchia, the sentences sound so odd.

There are several advantages to Chierchia's (1998) system. First, due to catastrophic presupposition failure, there is a fairly comprehensive explanation of both the acceptable and unacceptable judgments of all the sentences in (10).

(10) a. Every couch/*couches/*furniture will be stored in the warehouse.
b. Several couches/*couch/*furniture will be stored in the warehouse.
c. All couches/furniture/*couch will be stored in the warehouse.
d. Shakers used local American woods such as pine and cherry, but for most Shaker furniture/chairs/*chair, maple was the wood of choice.
e. No men/man/furniture will ever please Mrs. Donaldson.

For the sentences to be interpretable, *every* must combine with a singular denotation, *several* must combine with a plural denotation, and finally *all/most* must combine with denotations that are closed under join. In contrast, *no* can combine with any type of denotation. Presupposition accommodation would either require the hearer to hypothesize that the speaker has a non-conventional meaning assigned to the quantifier or that certain nouns are assigned to different semantic types than the grammatical rules dictate. Accommodations of this kind are not pragmatically reasonable in most contexts.

It is important to note that I have carefully chosen lexical items in (10) that are inflexible and resist coercion: *couch* resists a mass-noun interpretation and *furniture* resists a count-noun interpretation. The data becomes harder to analyze if we considered terms that sound acceptable in both mass and count contexts, such as *rock*, *water*, *judgment*, etc. However, Chierchia's general strategy for dealing with such categorial flexibility is to hypothesize categorical lexical distinctions and then let some type of coercion strategy or lexical ambiguity account for the full range of data. (This is the same strategy employed in Rothstein 2010, Link 1983, Gillon 1992, 1999, among others.) This seems reasonable enough and, in fact, I will adopt the same strategy when I present a more uptodate syntactic theory.[4] However, to highlight the underlying syntactic differences, I will continue to use nouns that resist coercion. Coercion and flexibility are discussed in more detail in section 4.3.[5]

4 For a thorough defence of this strategy, see Pelletier & Schubert 2002.

5 The nouns used here represent the extreme ends of the scale of countability developed by Allan (1980). Allan (1980) proposed an eight point scale of countability and hypothesized that quantifiers had different requirements with respect to this scale. Some quantifiers could combine with a 2 or higher whereas others could only combine with a 7. Thus, a certain amount of flexibility was built into the syntax itself. However, the data Allan (1980) discusses could also be accounted for by having two categories and coercion operators that work differently for different types of lexical items. Note also, other than completely flexible nouns (i.e., nouns that can be used in any nominal context), most of the so-called "in-between" examples in Allan's work consisted of collectives and pluralia tantum nouns. I will have more to say about such nouns in section **??**.

The second advantage of Chierchia's (1998) theory is that it also accounts for certain distributional gaps. If the three denotational checking functions defined above are the only possible preconditions available to quantifiers, then it falls out as a consequence that there cannot be a quantifier that appears with both singular count and mass nouns to the exclusion of plural nouns. Similarly, it follows that there cannot be a quantifier that appears with both singular and plural count nouns to the exclusion of mass nouns.

(11) CROSS LINGUISTIC GENERALIZATIONS

a. There is no known quantifier that applies exclusively to singular count and mass nouns to the exclusion plural nouns.

b. There is no known quantifier that applies exclusively to singular count and plural nouns to the exclusion of mass nouns.

There are no clear examples of quantifiers that violate these predictions.[6]

In summary, Chierchia's (1998) system of preconditions accounts not only for the distribution of quantifiers but also for certain cross-linguistic distributional gaps. His system is able to do this without appealing to separate syntactic subcategories. Rather, the distributional characteristics of quantifiers and the cross-linguistic variations amongst quantifiers stem from catastrophic presupposition failures.

2.3 Summary and Criticisms

There are many aspects of presupposition failure that are promising with respect to the mass-count distinction. First, the distributional characteristics of Quantifiers rely on the semantics differences between mass nouns, singular count nouns and plural nouns: differences that are grounded in referential data independent of quantifier distributions. Second, Chierchia's (1998) theory can explain two significant cross-linguistic generalizations with respect to the distribution of quantifiers.

Despite its advantages, there is a significant potential flaw, especially with respect to nominal subcategories. The basic problem comes from the fact that sometimes nouns have an empty denotation. For example, suppose we are in a context where there is no furniture. In such a context, the nouns *chair*, *chairs* and *furniture* would have the exact same denotation, namely the empty set. (Perhaps, the following examples will be even easier to understand if we use an NP that is empty in pretty much any real world context, such as *martian chair*, *martian chairs* and

6 The closest violation is the fact that numerals in Western Armenian can modify either nouns with plural marking or nouns without (see Borer 2005, Bale et al. 2011a,b, Bale & Khanjian 2014, 2009). However, as discussed in Bale et al. 2011a,b, Bale & Khanjian 2014, 2009, non-plural nouns in Western Armenian do not have a singular interpretation.

martian furniture. I will include the modifier *martian* in brackets to help the reader remember that the denotation of the NP is completely empty.) The empty set is a perfectly suitable denotation for such nouns. It can be used to evaluate certain sentences as true, such as those in (12), and others as false, such as those in (13).

(12) a. There is no (martian) furniture here.
b. There are no (martian) chairs here.
c. There is no (martian) chair here.

(13) a. There is too much (martian) furniture here.
b. There are several (martian) chairs here.
c. There is a (martian) chair here.

Because this is a suitable denotation, the checking functions, *P*, *S* and *CL*, should be satisfied by the empty set. However, a problem arises when considering atypical combinations of quantifiers and nouns.

(14) Suppose we are in a context where there is no furniture and hence $[\![\textit{furniture}]\!]^M = [\![\textit{chair}]\!]^M = [\![\textit{chairs}]\!]^M = \emptyset$.
a. # There are several (martian) furniture here.
b. # There are several (martian) chair here.
c. # There is a (martian) chairs here.
d. # There is a (martian) furniture here.

If the anomalous nature of the phrases like *several (martian) furniture* were due to the fact that *(martian) furniture* has a different kind of denotation than *(martian) chairs*, then the anomaly should disappear when *(martian) furniture* and *(martian) chairs* have the same type of denotation, namely the empty set. The same holds for the distributional differences between *(martian) chair* and *(martian) chairs*, as well as those between *(martian) furniture* and *(martian) chair*. However, as demonstrated in (14) the distributional differences persist even when the denotations of the nouns are equivalent.

The empty denotation is not the only way that the denotation of different types of nouns could be equivalent. For example, another potential problem arises in a context where the only piece of furniture is a single chair. In this context, the denotation of *furniture* and *chair* would once again be equivalent, consisting of the singleton set containing the single chair. However, there is still a contrast between the sentences in (15).

(15) Suppose we are in a context where $[\![\textit{furniture}]\!]^M = [\![\textit{chair}]\!]^M = \{c\}$, where c is the only chair and the only piece of furniture.

a. There is a chair in the room.
b. There is less furniture in the room than I thought.
c. # There is a furniture in the room.
d. # There is less chair in the room than I thought.

Chierchia (1998) recognized this potential problem and proposed that the presuppositional functions P, S, and CL would have to hold across all possible contexts in order to account for the distribution of nouns in contexts where the extensional denotations are equivalent. In other words, the checking functions would need to hold intensionally even though the nominal denotation itself is only interpreted with respect to the present world/context. Let's reconsider the definitions of the presuppositional functions above with this kind of intensional definition spelled out in more detail. For the sake of simplicity, I will use world variables w to represent possible contexts and assume that models M consist of a domain of contexts and a lexical assignment function that is relative to a world/context.

(16) INTENSIONAL PRECONDITIONS ON QS: Let the denotation of any nominal predicate P be a function from worlds/contexts to sets that are a subset of the domain of w (i.e., functions of type $\langle s, \langle e,t\rangle\rangle$). Let generalized quantifiers be functions from nominal denotations to functions from worlds to truth values (i.e., functions of type $\langle\langle s, \langle e,t\rangle\rangle, \langle s,t\rangle\rangle$). For example, for all models M, $[\![boy]\!]^M = \lambda w.\{x : x \text{ is a boy in } w\}$ and $[\![no\ boy]\!]^M = \lambda P.\lambda w.(\{x : x \text{ is a boy in } w\} \cap (P(w)) = \emptyset)$.

a. **Example of a Singular Q**: For any nominal denotation Z, $[\![every]\!]^M(Z)$ is defined only if $\forall w.\ S(Z(w)) = 1$.
b. **Example of a Plural Q**: For any nominal denotation Z, $[\![several]\!]^M(Z)$ is defined only if $\forall w.\ P(Z(w)) = 1$.
c. **Example of a Non-Singular Q**: For any nominal denotation Z, $[\![all]\!]^M(Z)$ is defined only if $\forall w.\ CL(Z(w)) = 1$.
d. **Example of a General Q**: For any nominal denotation Z, $[\![no]\!]^M(Z)$ is always defined.

The intensional version of the checking functions avoid the problem of contexts where plural, singular and mass denotations happen to be the same. However, there is something vaguely uncomfortable with the idea that the presuppositional checking functions exploit the intensional nature of the nouns even when the entire sentence is only evaluated in the present context. Fortunately, vaguely uncomfortable feelings do not have much scientific weight when it comes to evaluating theories.

Perhaps less fortunately, there are two main arguments why these intensionalized versions of the checking functions are still off the mark. First, in this type of

explanation, there is an implicit assumption that *furniture* and *chair* will consistently map to the right kind of denotations across all worlds/context in any model. In other words, in all Models M and worlds w, $[\![chair]\!]^M(w)$ will always be a subset of the set of atoms in the domain associated with w, and similarly $[\![furniture]\!]^M(w)$ will always be a subset of the domain associated with w that is closed under group formation. The question becomes, how can/does our system implement the consistent behaviour of these lexical items? In other words, what prevents the possibility of a world w' where $[\![furniture]\!]^M(w')$ is mapped to an atomic set (for instance a set of chairs in w') and what prevents the possibility of a world w'' where $[\![chair]\!]^M(w'')$ is mapped to a set that not only contains all the chairs in w'' but also all groups that can be formed consisting of those chairs?

The standard answer to this question in model theoretic semantics is that *furniture* and *chair* belong to different types and that the lexical interpretation respects the restriction on types in all models and in all worlds within those models (see the discussion in Gillon 2012a, Carpenter 1997, Klein & Sag 1985). However, type restrictions require that the lexical items be marked in some way so that the interpretation function has the proper information about which type-category the lexical item belongs to. If the interpretation function is a map from syntactic objects (lexical items and phrases) to semantic denotations, then the standard treatment of typing requires that the syntactic object be marked in some way so that, for example, *furniture* is consistently mapped to closed sets and *chair* is consistently mapped to atomic sets. Hence, underpinning an intensional semantic category is a distinction among syntactic items: *furniture* and *chair* are marked as belonging to different types. However, this begs the question: If *furniture* and *chair* require some kind of syntactic distinction, why not provide a syntactic explanation for their different distributions?

The other problem with the intensional version of the checking functions has to do with contradictory nominal predicates: i.e., predicates that can never be satisfied by any referential subject. Although the intensional implementation of S, P and CL in (16) avoids problematic cases where mass nouns, singular count nouns and plural nouns happen to have the same empty denotation, they are still problematic when considering nominal predicates that denote the empty set in all possible worlds in any given model. Such predicates are listed in (17).

(17) a. (martian) furniture that is not (martian) furniture
b. (martian) chairs that are not (martian) chairs
c. (martian) chair that is not a (martian) chair

These phrases support a metaphorical meaning, but under a literal interpretation the predicates can never be satisfied. I am assuming here that metaphorical interpretations are not in the domain of the semantic interpretation function, however even

if one weakens this assumption, the phrases in (17) could be re-stated in ways that avoid metaphorical interpretations. For example, *furniture that is not furniture* can be re-stated as *actual furniture that is not actually furniture*. The points made below hold for all types of rephrasing that resist metaphorical interpretations.[7]

Perhaps unsurprisingly, these intensionally empty predicates pattern exactly like their non-intensionally counterparts. There is one minor difference: the intensionally empty predicates often yield contradictory or tautological statements. However, independent of this one difference, all such predicates are acceptable when they appear with certain quantifiers, as shown in (18) and (19), yet distinctly odd in other contexts.

(18) a. There is no (martian) furniture that is not (martian) furniture here (or anywhere else).
b. There are no (martian) chairs that are not (martian) chairs here (or anywhere else).
c. There is no (martian) chair that is not a (martian) chair here (or anywhere else).

(19) a. There is too much (martian) furniture that is not (martian) furniture here.
b. There are several (martian) chairs that are not (martian) chairs here.
c. There is a (martian) chair that is not a (martian) chair here.

(20) a. # There are several (martian) furniture that is not (martian) furniture here.
b. # There are several (martian) chair that is not a (martian) chair here.
c. # There is a (martian) chairs that are not (martian) chairs here.
d. # There is a (martian) furniture that is not (martian) furniture here.

As shown in (20), the nominal predicates [*(martian) furniture that is not (martian) furniture*] and [*(martian) chair that is not a (martian) chair*] sound much odder as complements of *several* than the nominal predicate [*(martian) chairs that are not (martian) chairs*]. Similarly, [*(martian) chairs that are not (martian) chairs*] and [*(martian) furniture that is not (martian) furniture*] sound much odder as complements of the indefinite determiner than [*(martian) chair that is not a (martian) chair*].

7 Other potential examples include *furniture that was made by a man who never has and never will make furniture*, *chairs that don't exist but have an existing man actually sitting on them*, and *chair that is absolutely wider and absolutely narrower than itself*, etc.

It is important to note that Chierchia 2010 employs a different semantic system than Chierchia 1998. This updated system does not explicitly address the distribution of quantifiers so it is difficult to compare the two theories in this respect, however Chierchia (2010) does imply that such a distribution will still follow from the presuppositional nature of the quantifiers. In the updated system, Chierchia adopts the same type of denotation for singular count nouns but changes the denotational types for plural count nouns and mass nouns: (i) plural count nouns denote sets that contain singular atoms and all the groups generated from those singulars and (ii) mass nouns denote a property that is only true of maximal elements in some join semi-lattice.[8] For example, suppose that a, b and c are all chairs and the only pieces of furniture in a given context. According to Chierchia (2010), in this context the singular noun *chair* would denote the set $\{a,b,c\}$, the plural noun *chairs* would denote the set $\{a,b,c,ab,ac,bc,abc\}$, while the mass noun *furniture* would denote the singleton set $\{abc\}$. Whatever the preconditions are on the quantifiers, the same problems arise in this new system. In contexts where there is no furniture, or in contexts where the only piece of furniture is a single chair, the denotations of the three types of nouns will be identical.

As discussed in the next section, the problem with these types of systems lies not with the semantic characterization of the different categories, nor with the nature of the presuppositional checking functions, but with the fact that quantifier distribution boils down to a syntactic and not semantic distinction.

3 Syntactic Violations

Even in the earliest versions of phrase structure grammar (Harris 1946, Wells 1947 among others), the distributional characteristics of words was one of the central descriptive data-points that linguistic theory sought to account for. At the heart of these data-points was not only the observed distribution, but the feeling of oddness or ungrammaticality, and the lack thereof, when novel word combinations were presented to native speakers. Let's consider the difference between transitive and intransitive verbs as our prototypical example, as exhibited by the verbs *eat* and *devour* in (21).

(21) a. The furniture monster already ate our couch.
b. The furniture monster already ate.

8 For now, I will side-step the issue of vagueness, which, although is central to Chierchia's motivation for switching from one system to another, is orthogonal to the issue at hand in this section. Also, for simplicity, I will ignore kind denotations. Like Chierchia (1998, 2010), I will assume that the semantic system has operations that can freely coerce mass and plural denotations into kind denotations and vice versa.

c. The furniture monster already devoured our couch.

d. # The furniture monster already devoured.

Although *eat* is acceptable in both transitive and intransitive environments, *devour* is limited to transitive environments. As argued by Chomsky (1965), the fact that near synonyms pattern so differently make it unlikely that this kind of distribution can be directly attributed to the meaning of the verbs. Furthermore, the fact that the contrast between *devour* and *eat* persists even when novel (and unlikely) word combinations are used, as in (21), provides evidence against any account that appeals to familiarity or statistical probabilities. Chomsky (1965) argues that a much more likely explanation is that *eat* and *devour* have syntactic features which only allow the verbs to appear in certain syntactic environments: *devour* happens to be much more restrictive than *eat*.

There are many types of theoretical mechanisms that account for facts like these, but almost all versions involve marking lexical items with some kind of syntactic feature that either selects for or limits the number of syntactic environments it can appear in. For the sake of clarity, I will review two options: one based on the notation used in Chomsky 1965 and another based on the notation used in Chomsky 2000.

In Chomsky's (1965) system, lexical items contain a list of syntactic environments that they can be used in. The syntactic structure is grammatical as long as the environment the lexical item appears in satisfies at least one member of this list. Syntactic environments are symbolized using the underscore symbol plus the relevant syntactic category or feature that the lexical item is expected to be adjacent to at the point of insertion. For example, to symbolize that *eat* can appear in either transitive or intransitive environments, it would contain the list $\{\_DP, \_\#\}$, where # encodes the absence of a verbal complement. In contrast, *devour*'s list would only contain the transitive environment ($\{\_DP\}$). By assumption, the appearance of a lexical item in an improper environment would lead to a feeling of unacceptability.

In Chomsky 2000's system (see also Adger 2003 for a detailed introduction), lexical items contain uninterpretable features which can be deleted once the appropriate complement is selected. For example, a verb like *devour* would have an uninterpretable uD feature which can be "checked" by the introduction of a D complement (i.e., a determiner phrase). A key aspect of this type of theory is that so-called uninterpretable features must be deleted before the end of the derivation (i.e., before the syntactic structure is shipped to the cognitive-intensional interface). In essence, this system ends up being quite similar to the systems of subcategories developed in early phrase structure grammars (e.g., Wells 1947, Harris 1946) and in early versions of the Lambek Calculus (Lambek 1958) and other types of Categorial Grammars. It also inherits the problems that such systems have in trying to

account for the fact that certain lexical items (but not all) appear in more than one syntactic environment, as exhibited by *eat* in the examples above. In a selectional system, for *eat* to be transitive and intransitive, either *eat* must be ambiguous or the transitive and intransitive environments must be sub-environments of a broader distributional category. However, as noted by Gillon (2012a), a selectional system can be developed to handle this type of underspecification.

In the rest of this section, I discuss how traditional mass-count features fail to provide a better account of the distribution of quantifiers than Chierchia (1998). I also review a type-theoretical account of the differences between mass and count nouns (Rothstein 2010), demonstrating that it too suffers from some fundamental weaknesses. I then develop a better feature system for the mass-count distinction, one that is based on the empirical observations made in Chierchia 1998. This system is able to account for the distributional characteristics of quantifiers without encountering any problems with empty denotations. Overall, the new feature system provides a better account of the mass-count distinction than the purely presuppositional account.

3.1 The mass-count distinction and why traditional features fail

Given the weaknesses of the presuppositional account, it seems prudent to explore an account based on syntactic subcategorization rather than catastrophic presupposition failure. However, it is important to first note that a straightforward syntactic account does not fair better than the presuppositional account offered by Chierchia (1998). In this section, I discuss a naive syntactic theory of the mass-count distinction before presenting a revised theory in section 3.3.

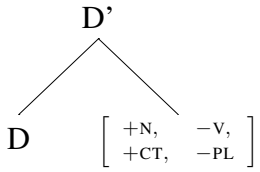
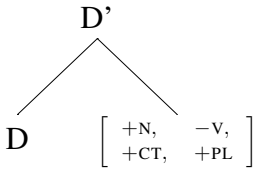
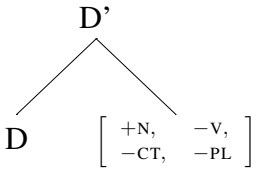
As hypothesized by Chomsky (1965), and discussed in detail by Gillon (1999), one way to represent the mass, count, and plural distinction is to use two binary features, [±CT] and [±PL]. In this type of theory, singular count nouns would contain the features [+CT, −PL], plural count nouns would contain the features [+CT, +PL], and prototypical mass nouns would contain the features [−CT, −PL]. As noted by Gillon (1999), there is even some evidence for a category of nouns with the features [−CT , +PL]. Nouns like *groceries*, *brains*, *remains* and *fireworks* (just to name a few) have overt plural morphology and trigger plural agreement but otherwise pattern like mass nouns.

With these features in place, the distribution of quantifiers can be characterized in terms of a set of subcategorization features. For example, quantifiers that only combine with singular count nouns like *every* could have the subcategorization set {_[+CT, −PL]} whereas quantifiers that only combine with plural count nouns like *several* could have the subcategorization set {_[+CT, +PL]}. In contrast, quantifiers that combine with either plural count nouns or mass nouns, like

most, could have the subcategorization set {_[+CT, +PL], _[−CT]}. Finally quantifiers that combine with any noun, such as *no*, could have the subcategorization set {_[+N,−V]}, or if we did not want to appeal to general nominal features, then equivalently {_[+PL], _[−PL]}.

These features give a list of possible quantificational complements that can appear to the right of the quantifier. Basically, there are three main environments as outlined in (22), (for now we will leave aside plural mass nouns, although our feature system here is compatible with such nouns).

(22)

Singular Count Environment	**Plural Count Environment**	**Mass Noun Environment**
		

A quantifier can appear in the slot marked by D only if one of members of the quantifier's subcategorization set is consistent with the features in the complement. Hence, a quantifier with the subcategorization set {_[+CT, −PL]} can only appear in the singular count environment whereas one with the subcategorization set {_[+CT, +PL]} can only appear in the plural count environment. In contrast, a quantifier with the subcategorization set {_[+CT, +PL], _[−CT]} can appear in either the plural count environment or the mass noun environment. As the reader can verify, these features (with the appropriate subcategorization frames) correctly account for the distribution of known quantifiers.

However there is a problem. Recall that Chierchia's system was able to account for two cross-linguistic generalizations, given in (11) and repeated below.

(11) CROSS-LINGUISTIC GENERALIZATIONS

a. There is no known quantifier that applies exclusively to singular count and mass nouns to the exclusion plural nouns.

b. There is no known quantifier that applies exclusively to singular count and plural nouns to the exclusion of mass nouns.

With the features ±CT and ±PL, this distributional gap would be coincidental. It would be easy to specify a subcategorization set that would characterize a quantifier that only attaches to singular and plural count nouns, namely {_[+CT]}. Similarly, it would be easy to specify a set that would characterize a quantifier that only attaches to singular nouns independent of whether they are mass or count, namely

{_[−PL]}. In other words, there is no principled explanation in this theory of why certain distributional gaps exist cross-linguistically.

The situation does not improve if we move to a representation of subcategorization using selectional features. In this type of theory, quantifiers have uninterpretable features that determine the type of complements they take. The range of uninterpretable features are as follows: {u[+CT], u[−CT], u[+PL], u[−PL]}. A lexical item would have a subset of these uninterpretable features. General determiners, like *no*, would not require any of these features but rather would just select for any noun (i.e., they would have a uN feature). In contrast, singular count-noun determiners, like *every*, would need the features u[+CT] and u[−PL] whereas the plural count noun determiners, like *several*, would need the features u[+CT] and u[+PL]. A problem arrises when it comes to determiners like *all* that select for either plural nouns or mass nouns. These two categories have no features in common and hence there is no single uninterpretable feature or set of such features that could yield the correct distribution.

Not only do selectional features have problems accounting for fact that plural nouns and mass nouns seem to form a natural class, they also over-generate much like the previous theory. For example, if a determiner only had the uninterpretable feature u[+CT], then it would select for singular and plural count nouns to the exclusion of mass nouns. Similarly, if a determiner only had the uninterpretable feature u[−PL], then it would select for singular count and mass nouns to the exclusion of plural count nouns. Neither type of quantifier is attested cross-linguistically.

3.2 Rothstein 2010 and Semantic Types

Generally, semantic textbooks, rather than syntactic ones, discuss the notion of a *semantic type*—i.e., the labelling of lexical items and other syntactic nodes in terms of the functional types associated with them, such as intransitive verbs with the type $\langle e,t\rangle$ and quantifier phrases with the type $\langle\langle e,t\rangle,t\rangle$ etc. However, such labelling has as much to do with syntax as it does semantics.

So-called *semantic types* serve two roles. First, they provide information about how lexical items and phrases are associated with a semantic category and hence they place limits on how interpretation functions map lexical items and phrases (e.g., a lexical item of type $\langle e,t\rangle$ will always be interpreted as a function that maps entities to truth values—or equivalently to a characterizing set of entities). Second, such types also provide information about which items can be combined together. Given that we are seeking an explanation about why certain quantifiers cannot combine with certain subcategories of noun, we will focus on this second role.

Rothstein (2010) introduces a type-theoretical distinction between mass nouns and count nouns. While mass nouns are of type $\langle e,t\rangle$—the standard type for pred-

icates that denote a set of individuals and/or groups[9]—count nouns are of type $\langle e \times c, t \rangle$, where c represents counting contexts and $e \times c$ represents a pairing between individuals/groups and counting contexts.

The motivations for this type difference stem both from semantic and distributional characteristics of mass and count nouns. Semantically, Rothstein notes that many count nouns are underspecified in terms of what counts as "being one". There are terms like *quantity* that vary from context to context. For example, the question "How many quantities of water do you have?" cannot be interpreted without first establishing what the contextual relevant quantities in question are. Similarly, the difference between whether there is one fence or two might depend on the importance of ownership and property lines: a single structure constructed by city planners that cuts across two property lines can be viewed as one long fence, whereas the same structure, if constructed by the property owners, might be considered two fences—the one that belongs to owner A and the other that belongs to owner B.

The pairing of entities with counting contexts explicitly represents this type of contextual sensitivity. In some pairings, an entity can count as one, whereas in others the very same entity can count as two or more. Note, however, the specification of a counting context need not be locally determined nor do such contexts need to be explicitly represented as being paired with entities—global contextual variables or implicit context set variables would do just as well in capturing facts about how context influences counting.[10]

Independent of the semantic facts, Rothstein also uses this type-difference between mass and count nouns to explain the distributional characteristics of certain quantifiers. Rothstein's idea is that numerals and other quantifiers like *each* and *every* select for nouns of type $\langle e \times c, t \rangle$—i.e., count nouns. If they are combined with a mass noun of type $\langle e, t \rangle$, there is a type-mismatch and the resulting structure is ungrammatical.

9 de Oliveira & Rothstein (2011) treat mass nouns as kinds of type $\langle e \rangle$ rather than predicates of type $\langle e, t \rangle$. However, they also employ the standard *lift* and *lower* coercion operators, $^{\cap}$ and $^{\cup}$ respectively, that map kinds to predicates and predicates to kinds. Given the coercion operators, the choice of type for mass nouns matters little in terms of the distribution of quantifiers, although it does have consequences for the kind versus existential interpretations of bare plurals and bare mass nouns in Portuguese. See de Oliveira & Rothstein (2011) for more details. For the sake of simplicity, I will not address semantic issues of kind interpretations in this paper. It is clear that kind interpretations are necessary to account for the semantic interpretations of mass nouns and plural nouns. However, it is also clear that both constructions sometimes yield predicate interpretations.

10 For example, the interpretation function could be parameterized for a counting context such that every interpretation needs to be superscripted with a counting context in the same way the interpretation function needs to be superscripted with a model. Alternatively, as argued in Westerstahl (1985), there is independent evidence demonstrating that each noun is restricted by a contextually determined context set. This context set could also be responsible for determining what counts as one rather than having a pairing between entities and counting contexts.

The advantage of a type-theoretical explanation is that the type differences between mass and count nouns remains even if their denotations are empty—thus, Rothstein's account avoids the weaknesses of an explanation based on the characteristics of denotations alone. However, there are still two problems with Rothstein's account. First, Rothstein (2010) does not make any type distinctions between singular and plural count nouns—both are of type $\langle e \times c, t \rangle$. Thus, there is no clear explanation of why *two boys* is well formed but *two boy* is not. She could appeal to Chierchia's (1998) presuppositional theory, but this would run into the same problem of empty denotations that was discussed in the previous sections. She could also appeal to the idea that plural marking with respect to certain quantifiers is due to agreement (as suggested in Krifka 1995) and does not represent any type of semantic pluralization, although it should be noted that this is not the view that Rothstein (2010) adopts. It is also possible to adopt a third semantic type for plural count nouns, but as discussed below, adding more types compounds some problems with compositionality.

A second problem with Rothstein's theory, a problem that is thoroughly discussed in Rothstein 2010 and that will persist under any type-theoretic account of the mass-count distinction, is that if mass and count nouns have different types, then this means that nominal modifiers that apply to both categories need to be (systematically) ambiguous. (And furthermore, if plural count nouns are also of a different type, then modifiers would need to be three-way ambiguous.)

This class of modifiers not only includes every type of adjective—whether gradable (*big furniture*, *big chairs*), non-gradable (*dead insects*, *dead wildlife*), intensional (*alleged secretary*, *alleged equipment*), or privative (*fake gun*, *fake weaponry*)—it also includes relative clauses and prepositional phrases (*the man in the bedroom*, *the woman that I like*, *the furniture in the bedroom*, *the cutlery that I like*). Each one of these modifiers must have two version, one that modifies nouns of type $\langle e \times c, t \rangle$ and another that modifies nouns of type $\langle e, t \rangle$ (and perhaps a third that modifies the type assigned to plural count nouns). Furthermore, it would include all of the quantifiers that apply to both count nouns and mass nouns (e.g., *all*, *most*, *a lot of*, *no*, *the*, *some*, etc.).

There are two ways to account for this type of ambiguity. One (not-so-appealing) possibility is that there are at least two separate lexical entries for each term (perhaps three if plural count-nouns have their own separate type). This has the unwelcome prediction that a phrase like *[all the big chairs]* has eight possible parses, seven of which create a type mismatch and hence an ungrammatical expression, and one of which forms a coherent determiner phrase. Such an explanation also implies that it is merely coincidental that all relative clauses, prepositional phrases and adjectives modify all three types of nouns—without exception. The type-theoretical account technically leaves it open that there could be adjectival modifiers that only

modify mass nouns, or for that matter, only count nouns. Despite the fact that some modifiers seem to be sensitive to whether a denotation has atomic minimal parts or not—such as *big*—there are no modifiers that apply exclusively to one category but not the other. (Recall that *big* can modify mass nouns like *furniture* and *equipment*). Furthermore, this fact, at least on the surface, is consistent cross-linguistically; adjectives, prepositional phrases and relative clauses are never limited in their distribution to certain subcategories of nouns.

The other (and more appealing) possibility to account for the systematic ambiguity of certain quantifiers and modifiers is that there are type-shifting rules that are employed when needed to resolve a type mismatch. For example, an adjective like *expensive* is underlyingly of type $\langle\langle e,t\rangle,\langle e,t\rangle\rangle$ but is converted to type $\langle\langle e\times c,t\rangle,\langle e\times c,t\rangle\rangle$ when it combines with a count noun. Similarly, a quantifier like *no* is underlyingly of type $\langle\langle e,t\rangle,\langle\langle e,t\rangle,t\rangle\rangle$ but is type-shifted to $\langle\langle e\times c,t\rangle,\langle\langle e,t\rangle,t\rangle\rangle$ when it combines with a count noun. Note, this shift cannot be in the other direction—from $\langle\langle e\times c,t\rangle,\langle\langle e,t\rangle,t\rangle\rangle$ to $\langle\langle e,t\rangle,\langle\langle e,t\rangle,t\rangle\rangle$. If it were, then phrases like *every furniture* would be acceptable, contrary to fact.

There are two weaknesses with respect to the coercion account. First, it seems to be missing a major generalization. In almost all known languages, prepositional phrases, relative clauses and adjectives are nominal modifiers independent of the existence of any nominal subcategories. The best account of this would be that these modifiers modify a single category, namely nouns. Assigning different compositional types to the nominal subcategories does not capture this generalization.

Second, the coercion account starts to become even more complicated once plural count-nouns are assigned a different semantic type. There are quantifiers that only apply to plural count nouns (e.g., *several*). Thus, there cannot be a general coercion mechanism that converts a function that applies to plural count-nouns to one that applies to mass nouns (or singular count-nouns for that matter). There are also quantifiers that only apply to singular count nouns (e.g., *every*). Thus, there cannot be a general coercion mechanism that converts a function that applies to singular count nouns to one that applies to mass nouns (or plural count-nouns for that matter).

Keeping these two observations in mind, note also that there are quantifiers that apply to all nouns (e.g., *no*) and others that only apply to plural count-nouns and mass nouns (e.g., *all*). Neither of these two types of quantifiers could be stored as underlyingly applying to count nouns, whether singular or plural, since there cannot be any coercion mechanisms that would be able to covert them into a form that would apply to count nouns. However, if the coercion is from a function that underlyingly applies to mass nouns, then it would be very difficult to explain why *no* and *all* have such different distributions: how could a coercion mechanism prohibit

one function from applying to singular count nouns (namely $[\![all]\!]^M$) and yet enable another to apply to the same type of nouns (namely $[\![no]\!]^M$)?

In summary, although semantic types avoid the problems of empty denotations, they complicate the grammar in other ways. One either needs to hypothesize massive ambiguity (which seems undesirable) or the existence of coercion rules. However, both options seems to be missing a fairly well established generalization, namely the three subcategories of nouns are treated as being the same compositionally when it comes to modification. Furthermore, the coercion rules would have to limit type-shifting in fairly arbitrary ways.

3.3 An updated version of the mass-count contrast

An alternative account of the mass-count distribution takes advantage of Chierchia's distributional observations by employing four monovalent features: *n*, SG, PL and CL which represent the intuitive idea of *noun*, *singular*, *plural* and *closed* respectively. There are two main advantages of this syntactic system. First, unlike the system proposed in section 3.1, this theory can explain why there are certain distributional gaps cross-linguistically. Second, by accounting for the mass-count distinction syntactically, the system avoids problems where nouns share the same denotation but yet have a different distribution. This is beneficial not only with respect to empty denotations, but also with respect to evidence that plural denotations have atomic minimal parts.

3.3.1 Monovalent features and the distribution of quantifiers

Before discussing the interaction between monovalent syntactic features and semantic interpretations, it is important to fully spell out how the features *n*, SG, PL and CL can characterize the mass-count distinction and also how these features interact with subcategorization processes. This section explores the syntactic details of using these four monovalent features. It first outlines how singular count nouns, mass nouns and plural count nouns are featurally encoded before discussing the range of possible quantifiers in terms of subcategorization.

Building off of Chierchia's (1998) observations, a distinction between singular count nouns, plural count nouns and mass nouns can be represented as follows:

- singular count nouns have the features [*n*, SG],
- plural count nouns have the features [*n*, PL, CL] and
- mass nouns have the features [*n*, CL].

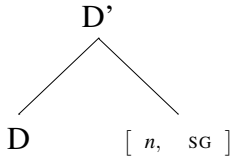
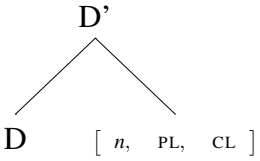
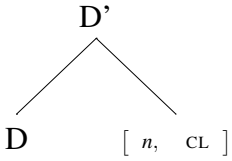
An important aspect of this feature system is that the plural feature cannot appear independently of the CL feature. Semantically speaking, this simply states that plural denotations are closed under join. However, to capture the syntactic relationship between these two features, I will assume that the following redundancy rule holds:

(23) PL → CL

In other words, the presence of PL implies the presence of CL.

This type of feature system yields three environments for determiners, as outlined in (24).

(24)

Singular Count Environment	**Plural Count Environment**	**Mass Noun Environment**
		

Given these environments, there is only a limited range of possibilities with respect to subcategorization. In what follows, I discuss these limitations both in terms of feature selection and constraints on grammatical environments.

Let's first consider an account of subcategorization that appeals to constraints, as outlined in Chomsky 1965. Recall that within this theory, each lexical item has a set of environments, a subcategorization set, that delineates all possible syntactic complements. The simplest type of subcategorization set is a singleton set. For example, {_*n*}, {_SG}, {_PL}, or {_CL} are all potential subcategorization sets, where the *blank* represents the insertion site (i.e., the D-head) and the elements to the right of the blank represent a required feature of the complement. These sets characterize the following distributions:

- Quantifiers with the singleton set {_*n*} are compatible with any noun (i.e., the distributional characteristics of *no*).
- Quantifiers with the singleton set {_SG} are only compatible with singular count nouns (i.e., the distributional characteristics of *every*).
- Quantifiers with the singleton set {_PL} are only compatible with plural nouns (i.e., the distributional characteristics of *several*).
- Quantifiers with the singleton set {_CL} are compatible with both plural and mass nouns, but not singular count nouns (i.e., the distributional characteristics of *all*).

Hence, the singleton sets alone account for the distribution of English quantifiers.

More interestingly, there is no subset of features that would limit the distribution of a quantifier to singular count and mass nouns. The closest possible subcategorization set would be {_SG, _CL} which states that a quantifier is compatible with a complement that has the feature SG **or** the feature CL. However, both mass and plural nouns have the feature CL. Hence, the subcategorization set {_SG, _CL} is equivalent to the singleton set {_*n*}. This limitation of the feature system could help explain why there are no languages with quantifiers that take singular count noun and mass noun complements to the exclusion of plural count nouns.

Although there are some advantages to this feature system, it still does not account for both of the cross-linguistic generalizations mentioned in Chierchia 1998. For example, the subcategorization set {_SG, _PL} would yield a quantifier that, in principle, would be compatible with singular and plural count nouns. As mentioned earlier, such a quantifier is unattested. However, if we move to a subcategorization system that specifies featural selections rather than sets of compatible environments, we can capture this cross-linguistic generalization as well.

With the features *n*, SG, PL and CL, there are four basic selectional features, namely u*n*, uSG, uPL, and uCL. Determiners that have one or more of these selectional features would have the following distributions:

- Quantifiers that only have the selectional feature u*n* would select for any noun, whether it be mass, count, singular or plural (i.e., the distributional characteristics of *no*).
- Quantifiers that only have the selectional feature uSG would select for any singular count nouns (i.e., the distributional characteristics of *every*). Note, having the selectional feature u*n* in addition to uSG would not alter the distribution. In contrast, having either the feature uPL or uCL in addition to uSG would yield an unusable quantifier.
- Quantifiers that only have the selectional feature uPL would select for any plural noun (i.e., the distributional characteristics of *several*). Note, having the selectional feature u*n* and/or uCL in addition to uPL would not alter the distribution.
- Quantifiers that only have the selectional feature uCL would select for mass or plural nouns (i.e., the distributional characteristics of *all*). Note, having the selectional feature u*n* in addition to uCL would not alter the distribution.

Not only does this feature system capture all of the possible distributions of English quantifiers, it also explains why there can never be a quantifier that either (i) selects for singular mass and count nouns to the exclusion of plural nouns, or (ii) selects

for singular and plural count nouns to the exclusion of mass nouns. For example, the only feature shared by singular mass and count nouns is the general nominal feature, namely *n*. The only feature shared by singular and plural count nouns is this same general nominal feature. Thus, any quantifier that selects for singular and plural count nouns will automatically select for mass nouns too. Similarly, any quantifier that selects for mass and singular count nouns will automatically select for plural count nouns.

In summary, the revised monovalent feature system is able to account for the distribution of quantifiers while also restricting the range of possible quantifiers. As we will see, the advantage of this system is that it is purely syntactic and thus will avoid many of the problems encountered by Chierchia's (1998) theory.

3.3.2 Two potential objections

Before discussing the semantic advantages of this system, I would like to address two potential objections. One potential objection is that the system outlined above does not have a feature associated with count nouns. Some might consider this problematic since, intuitively, the morphological distinction between *much* and *many* correlates with the mass-count distinction. If it is assumed that *much* and *many* are allomorphs of a single morpheme (which seems likely), then one might suspect that a count feature would trigger this variation. However, as more thoroughly discussed in Wellwood 2014, since *many* is restricted to plural count nouns, this allomorphic distinction can be characterized using a plural feature rather than a count feature. In other words, the quantifier surfaces as *many* when appearing next to a plural and surfaces as *much* elsewhere, as outlined by the vocabulary insertion rules in (25).

(25) Where MUCH is the underlying morpheme for both *much* and *many*:
MUCH ⇔ /mɛni/ / _PL
MUCH ⇔ /mʌtʃ/ elsewhere

Another potential objection is that, in this system, singular mass and count nouns do not share any features independent of plural nouns. However, in terms of verbal agreement, singular mass and count nouns pattern together to the exclusion of plural nouns. For example, both require third person singular copulas and both trigger third person singular agreement in the present tense.

(26) a. Every chair is/*are in the back room.
b. All the furniture is/*are in the back room.
c. The furniture stays/*stay in the back room.
d. Every chair stays/*stay in the back room.

One might object that this evidence demonstrates that singular mass and count nouns share a feature. However, such data does not necessitate any type of feature sharing between the two categories.

First of all, in most languages other than English (and German), such an agreement pattern would not be problematic due to the fact that a contrast between singular and plural agreement only requires the plural feature. In other words, plural morphology could be due to agreement with a plural feature, whereas so-called singular morphology could just be the result of a default/elsewhere form of the morpheme.

However, the English example is particularly problematic since the plural form appears to be the same as the default/elsewhere realization of the morpheme. For example, the plural form *are* is the same as the form used with second person singular subjects (e.g., *you are in the back room*). Similarly, the plural form *stay* is the same as the form used with first and second person singular subjects (e.g., *you stay in the back room*). Thus, the 3rd person singular realization of the morphemes, *is* and *stays*, must somehow be distinguished featurally from the plural forms without appealing to the plural feature.

Fortunately for the theory advanced in this section, it is more likely that the 3rd person feature is at play than a feature associated with being singular. Although the 3rd person feature is shared by all nouns (mass, count, singular and plural), there is evidence in English that this feature might be deleted from the verbal agreement morpheme when there is a plural subject. As argued by Harley (2008) and Bobaljik (2001), the lack of certain distinctions across paradigms indicates that an impoverishment rule might be active in the grammar, i.e., a rule that deletes certain features after the point of interpretation but before vocabulary insertion, essentially making the morphological system blind to certain features in certain environments. In English, plural agreement markers on verbs and plural forms of irregular verbs systematically lack any person distinctions across all paradigms, this despite the fact that person distinctions exist in the so-called singular forms. Such a systematic lack of a distinction suggests that person features in the agreement morpheme are deleted before the point of vocabulary insertion when co-occurring with plural features (see also the discussion in Bale et al. 2011b).

(27) **Impoverishment Rule**: In verbal (agreement) heads, person features are deleted in the environment of PL. (Bale et al. 2011b)

Given the evidence for this type of impoverishment rule, it is likely that plural agreement morphemes lack any person features and hence the so-called third-person singular forms could more accurately be described as third-person forms (absent the singular distinction). In other words, what *third-person* mass nouns and singu-

lar count nouns share is that they are both *third person*.[11] Although plural nouns also share this feature, there is independent evidence that a morphological impoverishment deletes this feature before the agreement morpheme is phonologically realized.

4 Lexical vs. Non-Lexical Feature interpretations

With the features SG, PL and CL, we do not need to rely on differences in the semantic types associated with singular, plural and mass nouns to characterize the distribution of quantifiers. For example, *several* cannot combine with *furniture* or *chair* because it requires its complement to have the PL feature. Similarly, *all*, *most* and *more* cannot combine with *chair* because they require their complements to have the CL feature. Finally, *every*, *each* and *a* cannot combine with either *chairs* or *furniture* because they require their complements to have the SG feature. These subcategorization restrictions apply even when *furniture*, *chair* and *chairs* have equivalent denotations (e.g., when their denotations are empty or singleton sets).

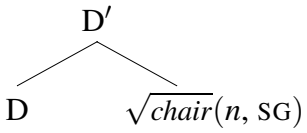
However, even though the distribution of quantifiers is determined syntactically, this does not mean that the syntactic features cannot affect the semantic nature of nouns. This section discusses how these nominal features might influence the denotational characteristics of nominal expressions and argues that the features SG and CL must be lexical in the sense that (at least some) roots must be stored with these features. This thesis contradicts the claims in Bale & Barner (2009) and references therein. The problem with non-lexical theories is that they are unable to avoid the distributional deficiencies that led to the adoption of syntactic features in the first place. Below, I discuss the differences between lexical and non-lexical theories and review how non-lexical theories cannot explain the distribution of quantifiers before outlining the benefits of a lexicalist approach.

4.1 Lexical Features

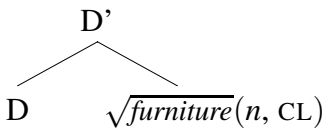
The difference between a lexicalist versus a non-lexicalist approach can be characterized by the syntactic trees in (28) and (29). In (28), the features SG and CL are part of the lexical entries for *chair* and *furniture* respectively, whereas in (29) these features are introduced by the morpho-syntactic system. (For the sake of clarity, I represent stored lexical items using the $\sqrt{\ }$ symbol.)

11 Note that the claim advanced here is not that all mass and count constructions are third person, but rather that the agreement marker on verbs for third-person mass and count nouns is due to their shared third-person feature rather than a shared singular feature. This leaves open whether first person and second person pronouns can be classified as being mass or count.

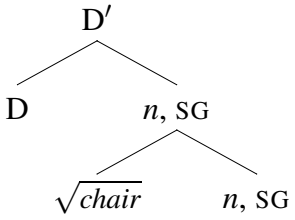
(28) a.



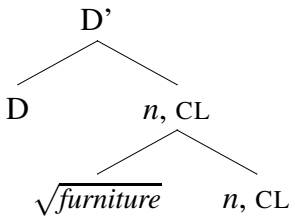
b.



(29) a.

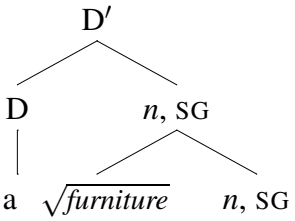


b.

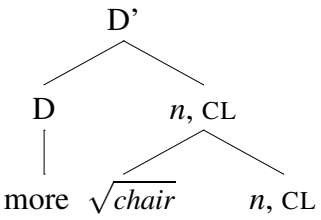


Recall that the purpose of introducing selectional restrictions was to explain why, for example, quantifiers like *all*, *most*, and *more* coherently combine with *furniture* but not *chair*, while quantifiers like *every*, *each* and *a* coherently combine with *chair* but not *furniture*. However, non-lexical theories can only explain these facts if there is a restriction on the distribution of the roots. In other words, the trees in (30) would have to be unusable or grammatically dispreferred.

(30) a.



b.



Most attempts by non-lexical theories to explain the unacceptable nature of these phrases relies on placing semantic restrictions on the distribution of the roots. For example, the feature SG could be mapped to a partial identity function that checks if its arguments are sets that contain only singular atoms.

(31) For all models M, $[\![n, \text{SG}]\!]^M = f_{sg}$ where for all sets X, $f_{sg}(X)$ is defined only if $X \subseteq \text{ATOM}$. When defined, $f_{sg}(X) = X$.

If $[\![\sqrt{\text{furniture}}]\!]^M$ were inherently closed under join (i.e., for all models M, $[\![\sqrt{\text{furniture}}]\!]^M = {}^*([\![\sqrt{\text{furniture}}]\!]^M)$), then $\sqrt{furniture}$ would not always be usable as a sister to [n, SG]. For instance, it would not be usable whenever there are two or more pieces of furniture in the domain of discourse.

Similarly, the feature CL could be mapped to a partial identity function that checks whether its arguments are closed under join.

(32) For all models M, $[\![n,\ \text{CL}]\!]^M = f_{cl}$ where for all sets X, $f_{cl}(X)$ is defined only if ${}^*X = X$. When defined, $f_{cl}(X) = X$.

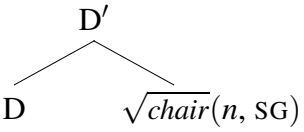
If $[\![\sqrt{\text{chair}}]\!]^M$ were inherently singular (i.e., for all models M, $[\![\sqrt{\text{chair}}]\!]^M \subseteq \text{ATOM}$), then $\sqrt{chair}$ would not always be usable as a sister to $[n, \text{CL}]$. For example, it would not be usable if there are two or more chairs in the domain of discourse.

However, just as with the quantifiers, problems arise for these kinds of semantic restrictions when the denotations of $\sqrt{furniture}$ and $\sqrt{chair}$ are either singular (i.e., containing only one piece of furniture/one chair) or empty. In these contexts, there is no distinction between the denotations associated with $\sqrt{furniture}$ and $\sqrt{chair}$.

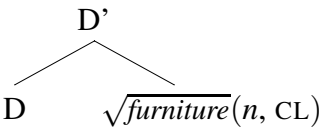
It is important to note that this problem runs just as deep as the one with quantifiers. There is no hope of there being a better interpretation of SG and CL that could rescue the non-lexical account. We need $\sqrt{furniture}$ and $\sqrt{chair}$ to sometimes be empty in order to account for the truth of sentences like *There are no (martian) chairs in the room* or *There is no (martian) furniture in the room* and the falsity of sentences like *I have a (martian) chair* and *I have more (martian) furniture than you*. If there are contexts where the denotations of $\sqrt{furniture}$ and $\sqrt{chair}$ are identical, then there are no possible semantic restrictions that can be introduced that distinguish these roots.

A better account of the distributional data involves marking the roots associated with *furniture* and *chair* with the CL and SG features respectively. In other words, the syntactic trees might be more accurately represented in (28), repeated below.

(28) a.



b.



If the roots associated with *furniture* and *chair* are stored with these formal features as part of their lexical entry, then the unacceptability of *a furniture* and *more chair* would follow as a consequence of the selectional restrictions of the quantifiers *a* and *more* respectively (e.g., *a* selects for SG and *more* selects for CL).

However, just because these features are part of the stored root does not mean that they cannot have an effect on the semantic form of the lexical items. In particular, it is possible that these features determine the shape of a noun's denotation through restrictions on the lexical interpretation function associated with each model.

(33) For all models M, let i symbolize the lexical interpretation function in M, let ATOM be the set of all atoms in the domain of M, and let * be the join-closure operation on sets.

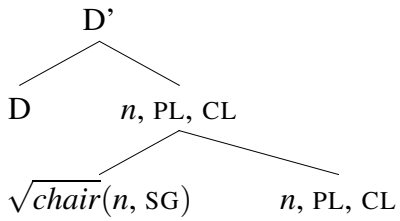
a. If SG is a lexical feature, then for all lexical items x with the feature SG, $i(x)$ maps to a subset of ATOM.
b. If CL is a lexical feature, then for all lexical items x with the feature CL, $i(x)$ maps to a set Z such that Z is subset of D and ${}^{*}Z = Z$.

Thus, each lexical item with the SG feature is necessarily mapped to a set of atoms in the model. It is easy to imagine a simple method for determining such a restriction, namely intersecting the conceptual denotation with the set of atoms in the model. For example, if CHAIR were the underlying concept associated with the root for *chair*, then $i(\sqrt{chair}(n, \text{SG})) = \{x : \text{CHAIR}(x)\} \cap \text{ATOM}$. Similarly, each lexical item with the CL feature is necessarily mapped to a set that is closed under join. Once again, it is easy to imagine a simple method to ensure this restriction, namely applying the join closure operator to the conceptual denotation associated with the lexical item. For example, if FURNITURE were the underlying concept associated with the root for *furniture*, then $i(\sqrt{furniture}(n, \text{CL})) = {}^{*}\{x : \text{FURNITURE}(x)\}$.

4.2 The Plural Morpheme

To account for the distribution of plural nouns with respect to quantificational subcategorization, the plural morpheme must satisfy at least two conditions. First, the plural morpheme must contain the features PL and CL and second these features must project. An example of a syntactic structure that satisfies both of these conditions is shown in (34), where the plural morpheme is represented by the feature bundle [n, PL, CL].[12]

(34)



With this kind of structure, the plural morpheme itself can be interpreted as a join-closure operator. This type of interpretation can be directly associated with the CL feature—the same feature that is critically involved in the lexical interpretation of mass nouns. In particular, just like $i(\sqrt{furniture}(n, \text{CL}))$ is mapped to ${}^{*}\{x : \text{FURNITURE}(x)\}$, so too can $[\![[\sqrt{\text{chair}}(\text{n, SG})\ [n, \text{PL, CL}]]]\!]^{M}$ be mapped to

12 Whether the n feature is part of the plural morpheme or exceptionally projects from the singular noun is inconsequential to the discussion in this section. For ease of exposition, I have assumed that n is part of the plural morpheme.

$^{*}(i(\sqrt{chair}(n,\ \text{SG}))) = {}^{*}\{x :\text{CHAIR}(x)\}$. In fact, this equivalency can be more directly represented if CL is interpreted as the join closure operator whereas *n* and PL are semantically vacuous. For example, if we interpret feature bundles through functional composition and treat the interpretations of *n* and PL as identity functions, then $[\![[n,\ \text{PL},\ \text{CL}]]\!]^{M}$ would be equivalent to (ID ∘ ID ∘ *) which in turn would be equivalent to *.

A consequence of this semantic system is that the denotation of plural nouns do not exclude atomic parts. For example, in (34), if we assume that $[\![\sqrt{chair}(n,\ \text{SG})]\!]^{M}$ is the set $\{a, b, c\}$, then the interpretation of the plural noun *chairs* would be equivalent to the join closure operation applied to $\{a, b, c\}$, i.e. $\{a, b, c, ab, ac, bc, abc\}$. This denotation contains the atomic parts *a*, *b* and *c*. Both Link (1983) and Chierchia (1998) assumed that plural denotations only contained plural entities, e.g. $[\![chairs]\!]^{M}$ = $\{ab, ac, bc, abc\}$. This assumption is based on the fact that demonstrative phrases like *those chairs* can only pick out plural objects and plural predicates such as *are chairs* are only true of plural objects.

However, as thoroughly discussed in a variety of different papers (see Croft 1990, 2nd ed. 2003, Krifka 1989, Sauerland 2003, 2008, Sauerland et al. 2005, Spector 2003, 2007, Bale et al. 2011b, among others), the "strict plural" nature of plurals seems to be due to competition with singular counterparts rather than true denotational restrictions. There are at least two sources of evidence supporting this analysis. First, plural nouns that do not have singular counterparts, e.g. pluralia tantum nouns such as *scissors*, are not restricted to plural entities in either their demonstrative or predicative instantiations. For example, *those scissors* can be used to pick out a singular entity (i.e., one pair) and *Those are scissors* can be a statement about the existence of a singular object. (Note, The fact that one pair of scissors counts as a singular object can be demonstrated by the fact that the phrase *two scissors* is used to pick out two pairs.)

Second, restrictions on plural denotations are weakened in certain types of downward entailing environments. For example, although the statement *John has children* implies that John has more than one child, the antecedent of the conditional *If John has children, then he will get a tax break* is satisfied even when John only has one child. Weakening of restrictions in downward entailing contexts is often a sign that such restrictions are due to competition rather than semantic limitations.

An important aspect of the current theory is that whether or not plurals are restricted to non-singular denotations is completely independent of the distribution of quantifiers. It could very well be that in a given context, the denotations of *furniture* and *chairs* are identical, especially if the only pieces of furniture in the context are chairs. However, this coincidental identity does not affect what types of quantifiers a speaker can use with either *furniture* or *chairs*. The distribution of the quantifiers is determined by the syntactic features alone. In Chierchia's (1998) system, it is

critical that mass nouns and plural nouns are distinct in terms of their denotations since quantifier distribution is based on denotational differences.

In summary, although the distributional characteristics of quantifiers is completely determined by syntactic features and subcategorization constraints, the interpretation of these features still account for the denotational characteristics of mass, count and plural nouns. Not only can it capture many of the semantic generalizations observed in the mass-count literature, it does so while permitting denotational overlap in contexts where this overlap is empirically justified.

4.3 Coercion, flexible Roots and nonce terms

Once a lexical theory of the mass count distinction is adopted, there are certain consequences for the treatment of coercion and flexible roots. *Coercion* can be defined as the acceptable (although slightly marginal) use of a noun in one type of subcategorial context when it is more typically used in another (such as a typical mass root being used with a count quantifier or vice versa). In contrast, *Flexible roots* are lexical items that sound natural in either subcategorial context (e.g., *string*, *rock*, *tile*, *stone*, *paper* etc.).

(35) Examples of Coercion

a. Every *water* on the table is contaminated.

b. The groomer thought there was just too much *dog* to deal with.

(36) Examples of Flexible Roots

a. Each *string* had a gem tied to it.

b. My aunt often tells me that you can never have too much *string*.

In either Chierchia's (1998) presuppositional theory, or in a theory with category-neutral roots, coercion is purely semantic: the mapping of one denotation to another. For example, one could imagine a *universal grinding* function (to borrow a term from Pelletier 1975) that maps sets of atoms to a non-atomic set that is closed under join. Similarly, one could imagine an atomizing function (several different types of functions for that matter) that map non-atomic semi-lattices to a set of singular atoms. The details of how these grinding and atomizing functions work are fascinating but not important for the discussion at hand (see Chierchia 1998, Pelletier 1975, Bunt 1985, Bale & Barner 2009, Gillon 1999, 2012b, Rothstein 2010 and references therein for further details and discussion). What is more important is how such functions relate to the empirical generalizations about the distribution of quantifiers.

If the difference between mass and count nouns is lexical, then coercion functions must be a mapping between the syntactic features SG and CL as well as a

mapping from one kind of denotation to another. In other words, the coercion must be both semantic and syntactic in nature. The details of such a function would have to be complex and nuanced to say the least. To give just one example of a potential complication, a key component of the quantifier data is that some nouns cannot appear as mass nouns and others cannot appear as count nouns (hence why it is difficult for certain quantifiers to appear with certain nouns). If a coercion mechanism were fully productive, both semantically and syntactically, then one would expect all nouns to be acceptable in all nominal environments—contrary to fact. An adequate coercion mechanism would have to be constrained, either semantically or syntactically. I have attempted to outline some constraints on coercion in previous work (see Bale & Barner 2009), but this is not the tactic I want to pursue here.

Instead, I would like to suggest another possibility to account for coercions—namely that the grammatical mechanisms that underly coercion are not much different from the interpretation of nonce terms used in psychological experiments. Nonce terms are novel words that speakers of a language have never seen or heard before. As a result, they have no established conventional meaning. It is well known that speakers, even children as young as 3 years old, will use syntactic cues to guess the meaning of such terms and that this "guess" is partly influenced by whether the novel term appears in a count or mass context (Gordon 1985, Bloom 1999). For example, upon hearing a sentence like "there is a dax on the table," both children and adults will assume that the term *dax* denotes a set of individuals of some kind rather than a substance. Similarly, upon hearing a sentence like "there is too much blicket on the table," both children and adults will generally assume that the term *blicket* denotes some kind of substance or "mass of stuff" rather than a set of individuals.

This inference from syntactic contexts to meaning is not absolute: it represents general trends in how people guess—trends that can be violated. For example, as discussed in Barner & McKeown (2005) and Bale & Barner (2012), given certain contextual cues of how objects are arranged in a scene, it is possible to get children to interpret novel mass nouns as denoting a set of individuals. However, even though speakers can be influenced by non-grammatical contexts, they can never violate the semantic constraints of the syntactic category. For example, a count noun can never be interpreted as a substance denoting term.

Given that people have the capacity to interpret nonce terms in a way that is sensitive to both grammatical and non-grammatical factors, it is possible that these same capacities could account for coercions. For instance, suppose that the term *spider* appeared in a sentence like the one in (37).

(37) The evil jerk squashed the bug and now there is spider all over his shoe.

The term *spider* is normally used in count syntax. In (37), it appears as a singular, bare noun—a syntactic construction typically reserved for mass nouns. The sen-

tence is slightly awkward but perfectly interpretable: the term *spider* in this sentence denotes spider-stuff—the squashed guts and exoskeleton smeared on the shoe.

This interpretation could have been derived from a coercion rule (with some grammatical or extra-grammatical constraint to explain why it sounds awkward), but it also could have been derived from the strategies used to interpret nonce terms. To explain how, imagine a grammar without any coercion rules but with lexical items that are marked as count or mass. In such a grammar, the English word *spider* would be mark as count and should not be used as a mass noun. As a result of these assumptions, the term *spider* in (37) could not be the same as the English word *spider* since it is being used as a mass noun. Thus, a hearer could and should treat it as a novel word—a term that has no conventionalized meaning. Because it is being used as a mass noun, hearers will usually assume that it is a substance-denoting term. Also, given its similarity in sound to the count-noun *spider*, it is reasonable to expect that people might favour a meaning that bears some connection to the conventionally established meaning of the count noun *spider*. The *subtance* that spiders are composed of seems like a natural fit. No coercion operators are necessary to derive the meaning and furthermore there is a natural explanation of the *awkwardness* of the sentence: the sentence contains a nonce term that does not have a conventionalized meaning.

There are two main advantages of adopting a nonce-term approach to coercion. First, it simplifies the grammar. Instead of having one strategy for nonce terms and another for instances of coercion, there would be only one mechanism for both. Second, it provides a natural explanation of why coercions sound slightly awkward. Under this type of explanation, coercion involves interpreting a novel term without a conventionalized meaning.[13] However, it should be noted, that this type of explanation should not be extended to flexible terms that can be freely used as both mass and count nouns. These terms could be treated as being associated with two homophonous lexical items—one mass and the other count. The similarities

13 There is a potential third advantage of the nonce term approach to coercion. It is well known that people, when interpreting nonce terms, avoid interpreting those terms as denoting or referring to objects that already have a conventionalized name or that were previous named with a different term in the discourse (see Markman 1989, Clark 1987, 1988). This strategy might help explain why object-mass nouns such as *furniture* and *equipment*—mass terms that are used to refer to objects and denote sets of objects—sound particularly bad when used as count nouns in an attempt to refer to items of furniture and equipment (e.g., **three furnitures*, **each equipment*).

Note that although the avoidance of synonymy is a general learning strategy for the acquisition of new words and for the interpretation of a nonce term, it is not a principle of grammar. Different constructions can be used in language to talk about and refer to the exact same things. For example, *items of furniture* versus *furniture*. Hence, appealing to an "avoidance of synonymy" to account for the awkwardness of *three furnitures* makes the most sense if the mechanisms that underly coercion are the same as those that underly the interpretation of nonce terms.

in meaning could perhaps have a historical explanation—a coerced meaning of a nonce term becoming conventionalized over time.

In summary, there are some reasons to favour a nonce-term approach to coercion. However, such an approach requires lexical items to be marked as mass or count (CL or SG). For example, hearers would not be able to assign *spider* in (37) a non-conventional meaning without first being able to recognize that it cannot be the same term as the count noun *spider*.

4.4 DP modification

As pointed out by Rothstein (2010), there seems to be a mass-count distinction at the DP level as well as the NP level. For example, consider the sentences in (38).

(38) a. Much (*many) of the furniture was purchased in San Diego.
b. Many (*much) of the chairs were purchased in San Diego.

When the embedded definite DP has a mass noun as a head noun, it can be modified by *much* and not *many*. When it has a plural count noun as a head noun, it can be modified by *many* and not *much*. The theory presented in this paper offers no explanation of these facts, unlike Rothstein 2010.

However, the facts are not as clear as they appear to be in (38). The link between the acceptability of many/much with the embedded NP versus the whole DP is not easy to characterize. As noted by Allan (1980), there are certain plural nouns that are not compatible with *many*, such as *poor* or *rich*. However, the DPs created from these nouns are compatible with *many*.

(39) a. The poor are often ignored.
b. * Many poor are unable to get health insurance.
c. Many of the poor are unable to get health insurance.

Also, as discussed in Pinker 2014, these partitive constructions are syntactically anomalous. For example, consider the following sentences paying particular attention to the underlined verb.

(40) a. * The older, male scholars that often likes to cite his own work wrote a scathing review.
b. One of the older, male scholars that often likes to cite his own work wrote a scathing review.

As shown in (40a), when the subject is a non-partitive DP construction that contains a relative clause with an elided subject, the agreement marker on the embedded verb has to match the number associated with the head noun of the DP (i.e., since the

head noun is *scholars*, the embedded verb should be *like* and not *likes*). The lack of plural agreement on *likes* in (40a) leads to an unacceptable sentence.

However, judgments change when the subject DP is a partitive as in (40b). The verb embedded in the relative clause must have singular agreement even though *scholars*, which is the only noun in the entire DP, is plural. These facts demonstrate that there is something quirky about the status of the noun in partitive constructions. Since the facts surrounding DP modification are somewhat unclear and the constructions are somewhat quirky, I don't think they should be used as a deciding factor between two theories without a more in-depth investigation.

5 Conclusion

This paper discussed the differences between subcategorization violations and catastrophic presupposition failure with respect to the distribution of quantifiers. It argued that the basic strategy outlined in Chierchia 1998 is on the right track, namely hypothesizing 1) a unique property that is only true of singular count nouns, 2) a unique property that is only true of plural nouns, and 3) a shared property between mass and plural nouns that is not shared by singular count nouns. However, for Chierchia (1998), these properties were semantic in nature.

Connecting the distributional characteristics of quantifiers to the denotational properties of nouns faces some challenges. For example, there are contexts where nouns from different subcategories seem to have identical denotations (e.g., contexts where the nominal denotations are empty). In such contexts, quantifiers are still restricted to appear next to only certain nouns, despite the fact that the nominal denotations are identical. Also, since there are quantifiers that are only felicitous with plural nouns, Chierchia (1998) needed to maintain that mass nouns and plural nouns never have the same denotation, even in contexts where the nouns are used to talk about the same things (e.g., *furniture* and *chairs* in a context where chairs are the only pieces of furniture, or *furniture* and *pieces of furniture* in almost any context). Chierchia (1998) adopts Link's (1983) hypothesis that plural nouns only denote groups, whereas mass nouns denote both groups and atomic individuals. However, evidence from pluralia tantum nouns and downward entailing contexts demonstrates that plural nouns, like mass nouns, denote groups and individuals.

The syntactic theory does not face similar challenges since it partially separates the properties relevant to quantifier distributions from the denotational characteristics of nouns. Beyond the general nominal feature (e.g., *n*), the syntactic theory involves three monovalent features: SG, PL and CL. As with Chierchia's (1998) theory, there is a unique property that is only true of singular count nouns (the syntactic feature SG), a unique property that is only true of plural nouns (the syntactic feature PL), and a property that is shared by both plural and mass nouns to the exclusion of

singular count nouns (the syntactic feature CL). Because the properties are syntactic, it no longer matters if two nouns from separate subcategories have identical denotations. Furthermore, since the basic strategy for the distribution of features mirrors Chierchia's strategy with respect to the denotational characteristics, the syntactic theory can account for the same cross-linguistic generalizations.

However, a critical aspect of the syntactic theory is that lexical items must be marked as either being count or mass (SG or CL). This type of lexicalism has interesting consequences for coercion and the flexible use of nouns in both count and mass contexts. It suggests that completely flexible nouns are ambiguous. It also opens up the possibility that the same underlying grammatical and non-grammatical mechanisms underly coercion and the interpretation of novel nouns (i.e., nonce terms).

References

Adger, David. 2003. *Core syntax*. Oxford: Oxford University Press.

Allan, Keith. 1980. Nouns and countability. *Language* 56(3). 541–567.

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

Bale, Alan & David Barner. 2012. Semantic triggers, linguistic variation and the mass-count distinction. In Diane Massam (ed.), *Count and mass across languages*, 238–260. Oxford University Press.

Bale, Alan, Michaël Gagnon & Hrayr Khanjian. 2011a. Cross-linguistic representations of numerals and number marking. In *Proceedings of semantics and linguistic theory SALT XX*, Cornell University, Ithaca, NY: CLC Publications.

Bale, Alan, Michaël Gagnon & Hrayr Khanjian. 2011b. On the relationship between morphological and semantic markedness: the case of plural morphology. *Morphology* 21(2). 197–221.

Bale, Alan & Hrayr Khanjian. 2009. Classifiers and number marking. In Effi Georgala & Jonathan Howell (eds.), *Proceedings of semantics and linguistic theory SALT XVIII*, Cornell University, Ithaca, NY: CLC Publications.

Bale, Alan & Hrayr Khanjian. 2014. Syntactic complexity and competition: The singular/plural distinction in Western Armenian. *Linguistic Inquiry* 45(1). 1–26.

Barner, D. & R. McKeown. 2005. The syntactic encoding of individuation in language and language acquisition. In A. Brugos, M. R. Clark-Cotton & S. Ha (eds.), *BUCLD 28: Proceedings of the 28th annual boston university conference on language development.*, 61–72. Somerville, Mass: Cascadilla Press.

Bloom, Paul. 1999. The Role of Semantics in Solving the Bootstrapping Problem. In Ray Jackendoff, Paul Bloom & Karen Wynn (eds.), *Language, logic and concepts*, 285–309. MIT Press.

Bobaljik, Jonathan. 2001. Syncretism without paradigms: Remarks on Williams 1981, 1994. In G. Booij & J. van Marle (eds.), *Yearbook of morphology 2001*, 53–86. Dordrecht: Kluwer.

Borer, Hagit. 2005. *Structuring sense*. Oxford: Oxford University Press.

Bunt, Harry C. 1985. *Mass terms and model-theoretic semantics*. New York: Cambridge University Press.

Carpenter, Bob. 1997. *Type-logical semantics*. Cambridge, MA: MIT Press.

Chierchia, Gennaro. 1998. Plurality of mass nouns and the notion of "semantic parameter". In Susan Rothstein (ed.), *Events and grammar*, 53–104. Kluwer Academic Publishers.

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

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Massachusetts: M.I.T. Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–156. MIT Press.

Clark, Eve. 1987. The principle of contrast: A constraint on language acquisition. In Brian MacWhinney (ed.), *Mechanisms of language acquisition*, 1–33. Hillsdale, New Jersey London: Lawrence Erlbaum Associates.

Clark, Eve V. 1988. On the logic of contrast. *Journal of Child Language* 15(2). 317–335.

Croft, William. 1990, 2nd ed. 2003. *Typology and universals*. Cambridge UK: Cambridge University Press.

von Fintel, Kai. 2008. What is presupposition accommodation, again? *Philosophical Perspectives* 22(1). 137–170.

van Fraassen, Bas. 1966. Singular terms, truth value gaps and free logic. *Journal of Philosophy* 63. 481–495.

van Fraassen, Bas. 1968. Presupposition, implication and self-reference. *Journal of Philosophy* 65. 136–152.

Frege, Gottlob. 1892. Über sinn und bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100. 25–50.

Gillon, Brendan. 1992. Towards a common semantics for english count and mass nouns. *Linguistics and Philosophy* 15. 597–640.

Gillon, Brendan. 1999. The lexical semantics of English count and mass nouns. In E. Viegas (ed.), *The breadth and depth of semantic lexicons*, 19–37. Dordrecht: Kluwer.

Gillon, Brendan. 2012a. Implicit complements: a dilemma for model theoretic semantics. *Linguistics and Philosophy* 35(4). 313–359.

Gillon, Brendan. 2012b. Mass terms. *Philosophy Compass* 7(10). 712–730.

Gordon, Peter. 1985. Evaluating the semantic categories hypothesis: the case of the mass/count distinction. *Cognition* 20. 209–242.

Harley, Heidi. 2008. When is a syncretism more than a syncretism? Impoverishment, metasyncretism, and underspecification. In Daniel Harbour, David Adger & Susana Béjar (eds.), *Phi theory: Phi features across modules and interfaces*, 251–294. Oxford: Oxford University Press.

Harris, Zellig. 1946. From morpheme to utterance. *Language* 22. 161–183.

Heim, Irene. 1983. On the projection problem for presuppositions. In Daniel P. Flickinger (ed.), *Proceedings of WCCFL 2*, 114–125. Stanford University, Stanford, California: CSLI Publications.

Heim, Irene. 1990. Presupposition projection. In R. van der Sandt (ed.), *Presupposition, lexical meaning and discourse processes: Workshop reader*, University of Nijmegen.

Karttunen, Lauri. 1974. Presupposition and linguistic context. *Theoretical Linguistics* 1. 181–193.

Klein, Ewan & Ivan A. Sag. 1985. Type-driven translation. *Linguistics and Philosophy* 8(2). 163–201.

Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In Renate Bartsch, Johan van Benthem & Peter van Emde Boas (eds.), *Semantics and contextual expressions*, 75–116. Dordrecht: Foris Publications.

Krifka, Manfred. 1995. Common nouns: A constrastive analysis of English and Chinese. In G. N. Carlson & F. J. Pelletier (eds.), *The generic book*, 398–411. Chicago: Chicago University Press.

Lambek, Joachim. 1958. The mathematics of sentence structure. *American Mathematical Monthly* 65(3). 154–170.

Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8(1). 339–359.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Baeuerle, C. Schwarze & Arnim von Stechow (eds.), *Meaning, use and interpretation of language*, 302–323. DeGruyter.

Markman, Ellen M. 1989. *Categorization and naming in children : problems of induction*. Cambridge, Mass: MIT Press.

de Oliveira, Roberta Pires & Susan Rothstein. 2011. Bare singular noun phrases are mass in brazilian portuguese. *Lingua* 121(15). 2153 – 2175.

Pelletier, F. Jeffry. 1975. Non-singular reference: some preliminaries. *Philosophia* 5(4). 451–465.

Pelletier, Francis Jeffry & Leonard K. Schubert. 2002. Mass expressions. In D. Gabbay & F. Guenther (eds.), *Handbook of philosophical logic*, vol. 10, 1–87. Kluwer Academic Publishers.

Pinker, Steven. 2014. *Sense of style: The thinking person's guide to writing in the 21st century*. New York, New York: Viking.

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

Sauerland, U., J. Andersen & K. Yatsushiro. 2005. The plural is semantically unmarked. In S. Kepser & M. Reis (eds.), *Linguistic evidence*, Berlin: Mouton de Gruyter.

Sauerland, Uli. 2003. A new semantics for number. In *Proceedings of SALT 13*, 258–275. Cornell University, Ithaca, NY: CLC Publications.

Sauerland, Uli. 2008. On the semantic markedness of ϕ-features. In Daniel Harbour, David Adger & Susana Béjar (eds.), *Phi theory: Phi features across modules and interfaces*, 57–82. Oxford: Oxford University Press.

Spector, Benjamin. 2003. Plural indefinite DPs as PLURAL-polarity items. In J. Quer, J. Schroten, M. Scorretti, P. Sleeman & E. Verheugd (eds.), *Romance languages and linguistic theory 2001*, 295–313. Amsterdam: John Benjamins.

Spector, Benjamin. 2007. Aspects of the pragmatics of plural morphology: On higher-order implicatures. In U. Sauerland & P. Statev (eds.), *Presuppositions and implicatures in compositional semantics*, 243–281. Palgrave-Macmillan.

Stalnaker, Robert. 1973. Presuppositions. *Journal of Philosophical Logic* 2(4). 447–457.

Strawson, P. F. 1952. *Introduction to logical thoery*. London, UK: Methuen & Co. Ltd.

Wells, Rulon S. 1947. Immediate constituents. *Language* 23(2). 81–117.

Wellwood, Alexis. 2014. *Measuring predicates*: University of Maryland dissertation.

Westerstahl, D. 1985. Determiners and context sets. In J. van Benthem & A. ter Meulen (eds.), *Generalized quantifiers in natural language*, 45–71. Dordrecht: Foris Publications.

Alan Bale
Linguistics Program, Concordia University
1455 de Maisonneuve O., Montréal, Québec
Canada H3G 1M8
alan.bale@concordia.ca