The Features and Structure
of the lower Noun Phrase

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy in the
Faculty of Humanities

2019

Christopher J Hicks

School of Arts, Languages and Cultures
## Contents

Abstract 6

Declaration 7

Copyright Statement 8

Acknowledgements 9

1 Introduction 11
   1.1 Features ......................................................... 12
      1.1.1 Feature complexity ........................................ 13
      1.1.2 Organisation of features ............................... 15
      1.1.3 Features vs structure ................................. 16
   1.2 Goals of the thesis ........................................ 17
   1.3 Organisation of the thesis ............................... 18

2 The Theoretical Approach 22

3 The Count/Mass Distinction 33
   3.1 Introduction ..................................................... 33
   3.2 Count and Mass in English ................................. 34
   3.3 Count and Mass in Mandarin Chinese ........................ 39
   3.4 What and where is the count/mass distinction? ........... 43
4.3.3 Returning to measure words ........................................ 116
4.3.4 The interaction of classifiers and measure words ............. 117
4.3.5 Adjectives .............................................................. 119
4.3.6 Comparison with English measure words ................. 121
4.4 Conclusion ............................................................... 122
4.5 A note on “kind” classifiers ........................................... 124

5 Gender - feature or functional head? ............................... 126

5.1 Introduction .............................................................. 126
5.2 Feature or functional head? .......................................... 127
  5.2.1 Gender as a feature ................................................. 127
  5.2.2 Gender as a functional head ...................................... 129
  5.2.3 Comparing the approaches ....................................... 133
5.3 Evidence for the acquisition of gender ......................... 134
  5.3.1 Gender and natural classes ..................................... 136
5.4 Acquiring gender as a feature ..................................... 136
  5.4.1 Internal structuring of gender ................................... 137
5.5 Gender and number in Spanish .................................... 138
  5.5.1 Number in Spanish ............................................... 139
5.6 Gender and number in German .................................... 142
5.7 Gender and functions ................................................. 144
  5.7.1 A mixed system? .................................................... 147
5.8 Conclusion ............................................................... 147

6 Gender in the absence of a noun .................................. 150

6.1 Introduction .............................................................. 150
6.2 Evidence from German and Dutch ................................. 150
6.3 The null noun in Dutch ............................................... 152
6.4 The null noun in German ............................................. 154
6.5 A potential incompatibility ............................................. 155
6.6 Reconciling *pro* with NumP and gender .......................... 156
  6.6.1 Problems with the plural ........................................... 157
  6.6.2 Defining “plural” *pro* ........................................... 159
  6.6.3 Problems with neuter *pro* ....................................... 161
  6.6.4 Defining neuter *pro* ............................................. 164
6.7 Gender or ±human ..................................................... 165
6.8 Conclusion ............................................................. 166

<table>
<thead>
<tr>
<th>7 Conclusion</th>
<th>168</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bibliography</td>
<td>172</td>
</tr>
<tr>
<td>Appendix</td>
<td>180</td>
</tr>
</tbody>
</table>

Word count: 47,645
Abstract

This thesis is an investigation into the behaviour of the features that exist, at least in some languages, in the lower reaches of the nominal projection and how they are structured. The features and structure of the lower noun phrase show a remarkable amount of crosslinguistic variation, despite the consistency in what types of information this level of the noun phrase must contribute to the full phrase and the wider grammar. Thus far, the complex and widespread variation in this domain has resulted in analyses which struggle to account for the variability in interpretation and in morphosyntactic marking under a single model, whilst also being something a language-learning child could acquire. This thesis aims to address some of the issues facing the features and structure of the lower noun phrase by taking a closer look at how the semantics, morphology, syntax and phonology interact, thus proposing an approach which considers, at its core, the nature of the acquisition of the grammar at the syntax-semantics interface.

Various features and phenomena are discussed, including the count/mass distinction in English and Mandarin Chinese, classifiers and measure words in Mandarin Chinese, gender in Spanish and German, and a construction in Dutch and German which suggests a null noun. For each of these case studies for my approach, I discuss the crosslinguistic variation at the levels of interpretation and marking, and consider the effect this has on the acquisition of any features within the syntax. In all cases, the variation can be attributed to differences in the feature specification of the noun, how the features of the noun shape its denotation, and how the denotation determines and interacts with the functions that are applied during the structure-building process. We find examples of features that are driven by a combination of meaning and marking, some that are primarily driven by meaning, and some that are primarily driven by marking, suggesting that the acquirer is sensitive to systematicity in the input on more than one level.

I conclude by arguing that the proposed approach provides the starting point for a model that has the potential to be powerful for a number of reasons. First, the model is functionally-driven. Nothing can exist within the grammar without a full functional motivation. This greatly constrains the system and aims to prevent any theory-motivated devices. Features within the grammar are not universal, they are created by the acquirer based on evidence within the input. This evidence can take various forms. A grammar based on the input will preserve those features that are transparently marked, but features with weaker marking may be candidates for loss or reanalysis. Finally, the model can account for the crosslinguistic variation we observed in the case studies and, in some cases, the differences in acquisition could be attributed to the strength of evidence for the acquirer and its effect on how a feature is created. There is a wide range of variation that still needs to be explored, but the potential of functionally-driven emergent features should be evident.
Declaration

I declare that that no portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.
Copyright

The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University’s policy on Presentation of Theses.
I could not have produced this work without the help and guidance of my supervisors: Professor Kersti Börjars and Professor John Payne. Kersti has unwaveringly been a supportive and motivating influence on my papers, my conference presentations, and with the production of this thesis. My development as a researcher during my PhD has been hugely helped by her belief and trust that I can get the job done (eventually), and her knowledge and experience of the workings of academia have been invaluable as I navigate the murky waters surrounding the bottom rungs of the ladder. Kersti has also been the voice of reason within supervisory meetings, when John and I find ourselves “down the rabbit hole” with the discussion of some peripheral phenomenon that was perhaps relevant to the conversation twenty minutes ago.

I must also thank John for his support and faith. His incisive theoretical comments have contributed to sharper analyses throughout my work, and he could always be relied on to find an exception to any generalisation – these were usually the entrance to the rabbit hole.

A special mention must go to the Arts and Humanities Research Council for their funding, as this research would not have been possible without their financial support.

I cannot forget my two fellow PhD-ers, Hannah Booth and George Bailey. We started our journeys together and somehow all made it through relatively unscathed. I do not think I could have survived the process without them and I thank them for the support they perhaps did not even know they were providing.

There are so many more fantastic members of the team at Manchester that I cannot mention them all, but I must thank Tine Breban, Martina Faller, Andrea Nini, Eva Schultze-Berndt and Julio Villa-García for their advice and support as colleagues and friends throughout my time at the university.

None of this would have been possible without my MPhil supervisor at the University of Cambridge, Dr Theresa Biberauer. I studied for a Masters without the full intention of continuing into a PhD, but her belief in my early view of grammar and linguistics was a driving force behind this work and I thank her for stoking the flames of my ambition.
My family have been full of support throughout my studies, and I thank them for always providing perspective and encouragement. Were it not for my mum, Ann, I never would have had the courage to return to university in the first place and would still be struggling along, unstimulated, at a telecoms company. I must thank my dad, Paul, for all the different ways he has helped me get through this, from providing distractions and relief to selflessly giving financial support. My sister, Rebecca, is the best and I could always rely on her to lift my spirits and to remind me that maybe everything is not as bad as it seems.

My life at the time of writing these acknowledgements is very different to how it was when I arrived in Manchester, bright-eyed and full of ideas. My family, and myself personally, have been through some difficult periods, but we have stuck together and the future looks bright. Not least because of my girlfriend Leanne, for whom I reserve these last and most important thanks. Your encouragement and support have been unmatched, and I am very lucky to have you by my side. Even luckier given your genuine interest in linguistics! I thank you for being a sounding board for my ideas, as well as an escape from them. I do not know what happens after this, but I am looking forward to discovering it with you.
Chapter 1

Introduction

The features and structure of the lower noun phrase show a remarkable amount of crosslinguistic variation, despite the consistency in what the noun phrase at this level must contribute to the full phrase and the wider grammar. Thus far, the complex and widespread variation in this domain has resulted in analyses which struggle to account for the variability in interpretation and in morphosyntactic marking under a single model, whilst also being something a language-learning child could acquire. This thesis aims to address some of the issues facing the features and structure of the lower noun phrase by taking a closer look at how the semantics, morphology, syntax and phonology interact, thus primarily proposing an analysis which considers, at its core, the nature of the acquisition of the grammar at the syntax-semantics interface.

This thesis is an investigation into the behaviour of the features that exist, at least in some languages, in the lower reaches of the nominal projection. The term feature is used throughout to refer, for the most part, to encode a minimal distinction between linguistic units in some linguistically relevant property. Putting any specific theory aside for now, this means that if two words are exactly the same except for one
property which cannot be further broken down, they should differ in terms of exactly one feature.

1.1 Features

Features evolved from early transformational rules that applied directly to morphemes, such as those proposed by Postal (1966) to capture determiner-noun agreement. According to these rules, an NP starts out as:

(1) \[
\text{[NP [Article un] [Noun Stem alumn] [Affix Gender M] [Number PL]]}
\]

An obligatory transformational rule applies to copy the nominal suffix -os to the determiner:

(2) \[
\text{[NP [Article un [Affix Gender M] [Number PL]]] [Noun Stem alumn] [Affix Gender M] [Number PL]]}
\]

Chomsky (1965) took these transformational rules and made them more abstract, replacing the morphemes with abstract featural content:

(3) \[
\text{Article → [α gender \ β number \ γ case] / ... [+N \ α gender \ β number \ γ case], where [Article ... N] is an NP}
\]

From this point, features have gained increasing importance within theories of syntax and morphology. This has happened in two different ways. As theories try to capture
ever more subtle distinctions, the features involved become more fine-grained. For example, the left periphery of the clause becomes ever more decomposed due to the different features and functional projections which evidence demonstrates can act independently of each other. Also, features have become the means by which operations are driven within the syntax. The process of movement in Government and Binding Theory, Move $\alpha$, was replaced by feature-driven movement, in that movement was triggered by the need for one feature to “check” another feature.

The theory of features has therefore needed to become more and more articulated as grammatical theory relies on them more and more. There is still, however, little consensus around both the nature and organisation of features.

### 1.1.1 Feature complexity

Feature complexity refers to how features are internally structured. Different theories, models and approaches take different views on how features should work for them in terms of whether features resemble labels or categories and what variation should be allowed within a single system of features. Features have therefore been applied within an approach in a number of ways, including the following (Sundaresan 2015a,b):

#### 1.1.1.1 Privative features

Privative features represent the least complex kind of feature: a privative feature is either present or absent. The template for such a feature can be represented by [F]. Categorial features, such as [D], [N], etc., are typically understood to be privative.
1.1.1.2 Binary features

Although the presence vs. absence of privative features may be considered binary, sometimes it is necessary to contrast binary values of a single feature. This is accomplished through one feature with \( \pm \) values. The template for a binary feature is \([\pm F]\). Definiteness is understood to be a binary feature. Given that in most analyses any element is either definite or not, this contrast is captured through a \([\pm \text{def}]\) feature. Definite elements are marked as \([+\text{def}]\), non-definite elements are marked as \([-\text{def}]\).

1.1.1.3 Attribute-value pairs (simplex values)

There may be a limited range of values for a feature, but more than the two allowed by a binary feature. In these cases, a feature can be structured as an attribute-value pair such that a single attribute can be valued by a range of defined values. The template for this is \([F:x]\). Gender may be thought of as an attribute in some languages, with values including masculine, feminine and neuter, for example. This would be represented as \([\text{gender}:\text{masc}]\).

1.1.1.4 Attribute-value pairs(feature values)

In some theories of grammar, it is possible for one feature to be valued by another feature. The template for this is \([F:[G:x]]\). The attribute \(G\) is valued as \(x\), and this attribute-value pair values the higher attribute \(F\). This type of feature is often associated with more functionally-driven frameworks, such as Lexical Functional Grammar. One driver behind this sort of feature is that making the feature more complex can take some of the complexity out of the syntactic structure.
1.1.2 Organisation of features

Moving on from the internal structure of features, there is also variation regarding how features may be structured relative to each other as objects in the morphosyntax.

1.1.2.1 Flat unordered features

Features may be taken to be unordered members of a set with no internal hierarchy, creating a “feature bundle”. The template is \([x, y, z]\). This type of feature is standardly assumed in Distributed Morphology and Minimalism.

1.1.2.2 Flat ordered features

Features are still gathered in flat structures, but they are now ordered within the set. This approach to features might be employed when a single head triggers multiple operations in a specific order. Müller (2010), for example, uses flat, ordered subcategorisation features to explain the fixed order of constituents in ditransitive sentences like “John gave the book to Sue”:

\[
\begin{align*}
\bullet P & \succ \bullet D \succ \bullet D
\end{align*}
\]

(4) Subcategorisation features:

In some theories, explaining the ditransitive through a hierarchical relation is not satisfactory, given the alternative structures available for ditransitives and the difficulty in promoting one of the constituents to dominate the other.
1.1.2.3 Ordered features - feature hierarchy/geometry

A final option is to structure features in a feature hierarchy or geometry. This can demonstrate the relationship between features and be used to explain any markedness effects observed between feature values. An example of a feature hierarchy is given in (5), taken from Harley & Ritter (2002):

(5) Referring Expression
   \[\text{PARTICIPANT} \quad \text{INDIVIDUATION}\]
   \[\text{Speaker} \quad \text{Addressee} \quad \text{Group} \quad \text{Minimal} \quad \text{CLASS}\]

There does, however, need to be a clear understanding of the link between feature hierarchies and syntactic structures and how the information on one relates to or combines with the information on the other.

1.1.3 Features vs structure

As we have seen, features are a flexible, abstract mechanism with the potential to be a very powerful device within any theory. Accordingly, any model of grammar should look to exploit the power of the tools at its disposal and aim to avoid proposing additional theoretical objects when the existing set could complete the task. This mirrors the Maximise Minimal Means (MMM) model proposed by Biberauer (2011) for acquisition, in which the acquirer checks whether any existing mechanisms in their grammar can capture a distinction before resorting to creating a new one.

However, how do we decide whether some distinction should be captured in the system of features or by the syntactic, structural arrangement of those features? For example, how can we determine whether a single functional head should bear different
combinations of features for different uses or whether multiple functional heads with their own features should account for the different interpretations/contexts?

The featural model and the structural model are two alternative ways of representing the same information, but they can make potentially very different predictions depending on our theories of feature structure and of structural hierarchy. Different frameworks can be distinguished according to where they fall in a continuum in the division of labour between features and structure:

(6) **More features, less structure**

- HPSG
- “anti-cartographic” Minimalism
- mainstream Minimalism
- cartography
- nanosyntax

**More structure, less features**

### 1.2 Goals of the thesis

The overarching goal of this thesis is to provide some clarity to the question of the division of labour between features and structure, considering whether they need to exist independently or whether one in fact drives the other. This will be achieved through a number of case studies into how features and structure interact in the lower projections of the nominal phrase.

More specifically, this thesis aims to address three research questions primarily:

1. Despite much having been written in the typological literature about the vari-
ation in the marking of features on nouns, there is no consensus as to what drives this marking. What motivates the creation of features within the grammar?

2. Furthermore, given the variation, what creates the different behaviours of features in different languages? For example, how might one language acquire a number distinction between singular and plural when another language also has dual number and another language might only have plural?

3. Additionally, how do these features find their place within the grammar? Do we find evidence for the placement of these features throughout the nominal projection, with some finding their home on the noun and some being hosted elsewhere? Is this variation completely free, such that there are no similarities between the placement of the same feature crosslinguistically?

1.3 Organisation of the thesis

The thesis is presented as follows. Chapter 2 provides an overview of my theoretical approach to the functionally-driven emergence of the grammar. I first consider the nature of features and how they might come to exist in the linguistic system, with some (simplified) discussion of the nature of acquisition. I make some key definitions with regard to the different types of features we might find in the grammar, and demonstrate how these types of features may correspond with each other to link the different dimensions of linguistic information. I provide an illustration of how the nominal phrase might be constructed according to this approach, using features that are acquired based on the link between semantic interpretation and overt realisation. I introduce the denotation of the noun as a join semilattice and discuss how features and functions operate on and restrict this denotation. This approach will be the
one employed in more detail in the discussions of specific phenomena in chapters 3, 4, 5 and 6.

Chapter 3 examines the count/mass distinction as a potential semantic divide that could be marked through a feature. I first discuss how the count/mass distinction as it exists in the grammar of English is not a perfect correlate of the perceptual divide between individuals and non-individuals. I review some existing analyses of the count/mass distinction in English and consider how the patterns observed are accounted for within these approaches. I then move on to look at the count/mass distinction in Mandarin Chinese, against the background claim of some linguists that all nouns in Mandarin Chinese are mass nouns. For the count/mass distinction to exist in the grammar, I argue that there must be some grammatically relevant impact of this distinction. This impact is found to exist in the observation that classifiers can only merge above nouns that can refer to individuals. There are, however, some key differences between the denotations of nouns in English and Mandarin Chinese – these are also discussed. I examine the options for the existence of the count/mass distinction in Mandarin Chinese, and judge that exists, but in a different way to English. I then move on to discuss the acquisition of these different systems, how (on a simplified level) this acquisition might take place, and how functions clearly link to denotations and features within the nominal projection. Morphological marking is used as a signpost for acquisition, and the lower complexity of morphological marking in one language compared to the other is a factor in the relative ease of acquisition, and the mismatches that are therefore created in the system.

Chapter 4 builds on chapter 3 by focusing in more detail on the classifier and how it compares with measure words. I review previous approaches to this type of element in Mandarin Chinese and discuss whether the difference between count and measure interpretations is a result of a difference in the feature specifications of classifier and measure words, or is down to the structural configurations they
appear in. I found there to be issues with these approaches which in some way stemmed from their treatment of classifiers and measure words as different types of the same element. To gain better insight, I took a closer look at the behaviour of these elements by conducting a small investigation to gather data from native Mandarin Chinese speakers. My investigation wanted to tease apart classifiers and measure words based on their modificational and combinatorial possibilities. The theoretical implications of my findings are then discussed and I incorporate these into my proposal for a treatment of classifiers and measure words in Mandarin Chinese. My approach takes into account the different status of these elements and can predict how they interact with other elements, and each other, within the nominal extended projection, based on their functional status and position.

Chapter 5 leaves individuation and Mandarin Chinese behind and moves on to examine the gender feature. The primary concern for this chapter is to determine whether a functional head can be motivated for the gender feature. This is based on the controversy amongst existing proposals around this issue, as both present good evidence in support of their position. After considering the models, I take a detailed look at gender in two languages which exhibit differences in their gender systems: Spanish and German. I then provide another application of my approach to features and functions which can capture the differences between Spanish and German through a combination of features and functions. Importantly, the mechanisms of my approach remain constrained by whether they are “acquirable”.

Chapter 6 moves on further to look at the presence of the gender feature in a nominal phrase that, based on our discussion so far, should not contain a gender feature at all. In previous discussion, we determined that the gender feature was attached to the noun in some way. In this chapter we look at constructions in Dutch and German that seem to lack a noun, but still have gender. More importantly, the gender specification has an impact on the form of determiners and the interpretation of
the nominal. I review existing proposals towards these constructions and find some
issues. Adopting some of the tenets of these approaches into my own approach, I
find that I can account for the behaviour observed without relying on additional
mechanisms whilst maintaining the possibility of acquisition.

Chapter 7 concludes with the strengths, and weaknesses, of my approach. I also sug-
gest some areas of future research for syntacticians, semanticists and those working
experimentally on child language acquisition.
Chapter 2

The Theoretical Approach

One of the primary aims of linguistic theory is to investigate and account for the structured variation observed in natural language. Cross-linguistic variation is structured in that, on the one hand, a large number of unrelated languages can exhibit the same patterns whilst, on the other, some hypothetically possible patterns are not exhibited in any language. The key question is whether these highly frequent cross-linguistic patterns and gaps occur due to genetic endowment, i.e. from Universal Grammar (UG), or due to functional factors of human language, or a combination of both.

At one end of this debate, all linguistic variation comes from an innate faculty for language – UG allows some structures but not others. At the other end, functional pressures on human capabilities shape the variation that is possible – UG does not exist. There is also the possibility that structured variation arises from a combination of formal and functional considerations. If this is the case, then how can we investigate the interaction between the two types of factors? Does one take precedence over the other? These questions place language acquisition firmly at the heart of linguistic variation and change, as it is during acquisition that the formal and
functional pressures interact to create individual grammars.

The syntactic approach I adopt in this thesis finds its roots firmly in the Minimalist Program as set out by Chomsky (2000), but within an emergentist framework like that of Chomsky (2005) and following in the spirit of Biberauer et al. (2014). The key idea of this approach is that grammatical parameters are not richly specified as part of a genetically-endowed UG, but emerge during acquisition as a result of interaction between the three factors of language design as proposed by Chomsky (2005: 6):

(7) a. Genetic endowment which recognises linguistic experience as part of the environment.

b. Trigger experience which leads to a narrow range of variation.

c. General strategies based on computational considerations not specific to the language faculty.

To go further than Biberauer et al. (2014), I argue that, in addition to emergent parameters, crosslinguistically variant formal features are acquired through the interaction of UG and experience. This hypothesis raises further questions: (i) what is a feature, and (ii) are there any features that are not crosslinguistically variant?

The question of what constitutes a formal feature, and therefore what constitutes a functional projection, is a key concern of research at the syntax-semantics interface. The cartographic approach (cf. Cinque (2002, 2006), Rizzi (2004)) assumes a rich functional structure, with a set of functional heads being specified by UG. The alternative view, often referred to as "what you see is what you get" (WYSIWYG) (Roberts & Rousson 2003), takes syntactic trees to be as small as possible. For a functional projection to be present in a particular clause, there should be empirical evidence, such as the overt realisation of a functional head. The fundamental
difference between the two approaches is that the WYSIWYG approach allows a particular functional projection to appear in only some sentences and in only some languages, whereas the cartographic approach insists that all existing functional projections are always present.

I align with the WYSIWYG approach and argue, along similar lines to Zeijlstra (2008), that formal features are created during acquisition and are not provided by UG. This means that commonalities amongst the sets of formal features found crosslinguistically are not a product of innate universals, but arise from third factor pressures on the mapping between cognitive/perceptual distinctions and the grammatical system. Crosslinguistic variation comes about as languages adopt different strategies to reflect perceptual distinctions in their grammar. Given that languages tend to express a similar range of semantic information, and as the third factor strategies employed in language design are taken to be universal, it is no surprise that similarities exist between the grammatical systems of different languages, but crucially these do not arise from genetically provided linguistic knowledge.

At this point, it is important to define the three distinct levels of the learning mechanism. First, we have purely cognitive/perceptual categories - those that are independent of language and are used to see and think. These play an important role in the acquisition process to form semantic categories through a probabilistic mapping. The child will map their perceptual categories onto semantic categories, which are more abstract or logical, such as “individual”. The semantic categories are formed through semantic features.

Lexical items are taken to consist of three kinds of features: phonological features, semantic features, and formal features. This discussion will for the most part put phonological features to one side and focus on the interaction of semantic features and formal features.
The third level is the syntax. The mapping to syntax from the semantic level involves recognising which of the semantic features should be "grammaticalised" as formal features.

Semantic features, and more specifically those that do not interact with formal features, are difficult to define in terms of a model of grammar. An example of a purely semantic feature might be *colour* or *opacity* which, as far as I am aware, do not have a syntactic reflex but nevertheless may be stored as information contributing to the meaning of a lexical item. I am not committed to whether these semantic features necessarily have an ontological status, nor am I concerned with how such semantic features are stored with items in the lexicon, or whether the set of semantic features is truly universal. A semantic feature stands for some element of semantic content that may or may not be relevant to the grammar.

Using examples from Pinker (2009), for the first mapping, the child thinks “If it’s a physical object, I’ll construe it as a kind of individual”. This takes a perceptual category and maps it to a semantic category “individual”, or marks that kind of object with the semantic feature [individual]. The child then thinks “If it’s the name of a kind of individual, then it’s definitely a noun”. This is the second mapping from semantic features to formal features. The acquisition of formal features relies on this second mapping, and I propose that all formal features, not just grammatical categories, are acquired in this way, as a conversion from a semantic feature to a formal feature, when a semantic feature has a systematic impact on the grammar.

The set of formal features present in a grammar is created during acquisition based on the transparency of the link between semantic information and its overt realisation. The variation in this transparency determines when and how a semantic feature is formalised in the grammar. This follows the WYSISYG approach under which a functional projection requires empirical evidence.
In this system, there are two methods of “systematic impact” which a semantic feature can have: (i) a semantic feature can be overtly realised on the item bearing that semantic feature, or (ii) a semantic feature can be overtly realised on another item or items, but this realisation is dependent on, or controlled by, the item bearing that semantic feature. This “overt realisation” can occur in a number of ways, through morphology, phrase structure, or speech sounds or intonation, as long as there is a recognisable difference in the overt production of the phrase. For the most part, I will focus on the presence or form of morphemes and functional items as my overt realisations.

It is these overt realisations that signal semantic features to the acquirer, and the link between the realisation and the semantic feature is encoded in the grammar as a formal feature. I propose that the more transparent, frequent and pervasive the link between a semantic feature and its realisation is, the earlier that semantic feature is grammaticalised as a formal feature.

This presents acquisition of a language as the development of a computational system which can find and decode the links between phonological form and our understanding of concepts and semantic information. This conception mirrors the “Y-model” architecture of the grammar as first described by Chomsky & Lasnik (1977: 431). Acquisition creates the grammar as the link between the Phonological Form perceived by the Articulatory-Perceptual system and the Logical Form understood by the Conceptual-Intentional system. The grammar breaks down expressions into their constituent parts, storing appropriate bundles of information as lexical items and encoding some semantic features as formal features. Once this computational system has begun to mature, the acquirer has the capability to reverse engineer these links to produce phonological material with the desired semantic content. Formal features in the grammar capture that semantic information which requires a phonological representation, whilst also allowing purely semantic and purely phono-
logical features to exist in the C-I and A-P systems entirely independent of the grammar.

The task of building functional projections is now presented as the process of converting, or "grammaticalising", semantic features into formal features in the grammar. This grammaticalisation process is a key ability in the acquisition of grammar and, coupled with the ability to recognise the systematic correlations between overt realisations and semantic content, could be the fundamental process for setting up the morphosyntactic elements of a grammar during a speaker’s early years.

To illustrate, imagine a language in which the lexical item used to denote an item that is blue in colour always ends in a morpheme with the phonological content of /i/. The semantic feature here is blue and it systematically links to an overt realisation in the suffix –/i/. During acquisition, the acquirer will notice this correlation and recognise that the semantic feature explains the overt realisation. This syntactically relevant semantic feature will therefore be grammaticalised as a formal feature.

(8) Lexical Item
      ↓
     Formal feature
       [+BLUE]
      Phonetic content Semantic feature
          /i/ ||blue||

In this illustration, the item that bears the semantic feature is itself overtly marked. This means that the formal feature created will be interpretable on this element.

Recall, however, that the overt marking may occur on other items, but be controlled by the item bearing the semantic feature. To illustrate this type of system, imagine a language in which the modifiers of a noun with a referent that is, for example,
opaque must end in a morpheme with the phonological content of /s/. The presence of the suffix –/s/ on a modifying adjective indicates the presence of the semantic feature opaque on the modified noun. The acquirer in this scenario has to recognise the link between the overt marking and the controlling semantic feature on another element. Formal features will be created on both the adjective and the noun, but the formal feature on the adjective will be uninterpretable whilst the formal feature on the noun will be interpretable. This is important for the acquisition process as, in this way, these “detached” agreement markers can be utilised by the acquirer as signposts for positing formal features on the controlling items in the grammar, and the acquisition of formal features will be aided by semantic features that are realised more frequently or pervasively in the system through agreement markers.

We now have three types of features, semantic features which, if they have a syntactic impact, are grammaticalised as interpretable formal features. One such syntactic impact might be that the semantic features trigger agreement on other elements. If so, these other elements will be marked with uninterpretable formal features. This separates uninterpretable formal features from semantic features, ensuring that agreement markers do not contribute semantic content to the marked element. Interpretable formal features are created on an element based on different evidence:

(9)  a. as a result of the semantic feature being overtly marked on the element itself
    b. as a result of dependent agreement marking being controlled by the element bearing the semantic feature
    c. as a result of both (9-a) and (9-b).

If an interpretable formal feature is acquired through (9-a) or (9-c), then the feature is grammaticalised as a functional head. Functional heads can be understood
as bundles of features. An interpretable formal feature acquired through (9-b) is grammaticalised, or rather lexicalised, as part of a lexical item. I maintain that functional heads are just that, functional. The formal feature acts as a function ensuring that the semantic feature it grammaticalises is applied as an operation on the interpretation of the phrase. The position of functional projections within the hierarchy reflects the successive application of functions to the basic semantic denotation provided by a root element.

To illustrate, I assume that the task of a nominal phrase is to identify an entity as its referent. The root of a nominal phrase is a noun, which enters into the derivation with a basic semantic denotation based on lexically-specified information. Functional heads within the extended projection of the noun apply their functions to this denotation to restrict the denotation to the intended referent. Both the lexical information on the root element and the functions that can apply vary cross-linguistically, based on which semantic features are grammaticalised into formal features during acquisition. Under this approach, it is also the case that only functional heads applying their function to the denotation of a phrase are projected in that phrase.

Take the imaginary noun taskle, which is lexically specified as denoting individuals. This means that, in this language, the semantic notion of individualisation is grammaticalised as an interpretable formal feature on the noun itself and the overt realisation is only visible on other elements in the phrase. A noun with a denotation of individuals is represented by a join semilattice of all individuals to which it can refer. Assume now that taskle can refer to only four individuals: a, b, c, and d. Introducing the word taskle into the syntax provides a denotation including these four individuals, and all the sets created by their union. The join semilattice is given in (10):
Note that if the semantic notion of individuation had not been stored as a formal feature of the noun, the denotation provided by the lexicon would be different.

At this point, the denotation of the noun is at its widest. The task of the nominal phrase is to signify the intended referent from those available by applying functions to this denotation. Functions are applied to the denotation through functional heads which arise from formal features grammaticalised with their own overt realisation. One function that can apply to a denotation of individuals is one that selects either individuals or sets of individuals. This is the functional head Number, which may be created based on the overt realisation of plural marking and how this correlates with the interpretation of having more than one individual. The functional head hosts plural morphology. In terms of the semilattice, depending on whether the intended referent is a singular individual or a set of individuals, the Number function will either select the atomic individuals, the bottom row of the semilattice, or the sets of individuals, the semilattice excluding the bottom row. Applying the Number function to taskle to select only sets of individuals leaves the denotation in (11):
Functions applying to the denotations need not only arise from functional heads. Modifiers and quantifiers also operate on the denotation through the construction of the nominal phrase. For example, to the output set above a further function can apply which selects sets of specific cardinalities. Numerals carry this function, which can only apply to denotations consisting of individuals. The cardinality the numeral selects for must also be compatible with the input denotation supplied to the function. For example, the numeral *one* will only return an output set if the denotation supplied contains only atoms, and numerals with a value higher than one will only return an output set if the denotation supplied contains only sets of individuals, and at least one set of the specific cardinality requested. Applying the function of the numeral *three* to the output set for *taskles* in (11) to select only sets of individuals with a cardinality of three leaves the denotation in (12):

(12) three taskles:

\[ abc \quad abd \quad acd \quad bcd \]

The nominal phrase *three taskles* must therefore have as its referent one of the four sets of individuals left in the denotation. Other functions could now apply to restrict this denotation further.

In this way, the nominal phrase is constructed based on an acquired set of formal features and functions. Each projection within the phrase is semantically motivated, but the separation between pure semantic features and those which are grammatic-
calised as formal features provides the basis for much of the crosslinguistic variation observed. The crosslinguistic variation in the overt realisation of semantic features creates differences in the information which is stored alongside nouns in the lexicon, and in the functions which can apply to these nouns throughout the construction of the phrase.

In summary:

(13) a. Formal features are created during acquisition, not innately given by UG.

b. Formal features are created in the grammar to capture the link between semantic features and their overt realisation via phonological content.

c. If the semantic feature is overtly realised independently or on the element which bears the semantic feature, then an interpretable formal feature is created and grammaticalised as a functional head.

d. If the semantic feature is only overtly realised on other elements within the phrase through agreement marking, then the interpretable formal feature will be created as part of the lexical information of the element which bears the semantic feature. The other elements which bear agreement marking receive an uninterpretable formal feature.

e. Formal features grammaticalised as functional heads apply the function determined by their semantic feature as an operation on the denotation of the phrase as constructed at the point of their projection.
Chapter 3

The Count/Mass Distinction

3.1 Introduction

It is likely to be the case that all languages can refer to both individuals and non-individuals. The world is classified along such lines and numerous attempts have been made in the fields of linguistics, specifically syntax and semantics, philosophy, and cognitive science to adequately describe how languages adopt various strategies to reflect this real-world distinction.

In many languages, this is reflected in a distinction between “count” nouns and “mass” nouns. Count nouns denote sets of individual, countable entities and can be modified directly by cardinal numerals. Mass nouns denote sets of non-countable entities and cannot be modified directly by cardinal numerals. An entity that inherently comes in countable units is an individual, as denoted by the English count nouns cow or chair. I will refer to an entity that comes in mass form and does not naturally divide into countable units as a non-individual, such as the referents of the English mass nouns beef or wind. Three cows and three chairs are grammatical phrases in English, whereas #three beefs and #three winds would usually be deemed
This suggests that the grammatical count/mass distinction follows directly from the real-world individual/non-individual contrast. There are, however, a number of problems with this correlation. Firstly, there exists a class of nouns known as “object mass nouns” which behave morphosyntactically as mass nouns but refer to sets of individuals. These are nouns such as cutlery, footwear, or the oft-cited furniture. Another problem is the great crosslinguistic variation in the grammatical count/mass distinction. The perceptual divide of entities into those that are individuals and those that are not is understood to be a pre-linguistic concept, so it is unclear why languages would vary in their classification of nouns as count or mass if the grammar is directly capturing the real-world contrast.

The difficult task for linguistic theory so far has been linking the grammatical count/mass distinction to the perceptual individual/non-individual divide. In this chapter, I argue that the count/mass distinction is an attempt to grammatically encode the individual/non-individual divide, but that the two are not always directly correlated. In more detail, I will be examining two languages, English and Mandarin Chinese, to investigate (i) whether the count/mass distinction exists in both languages, (ii) how the count/mass distinction is encoded in these languages, and (iii) the role of linguistic input during the acquisition of the count/mass distinction.

### 3.2 Count and Mass in English

I understand the count/mass distinction as a distinction made only in the grammar of a language – a linguistic reflection of the perceptual divide between individuals and non-individuals. However, as will be demonstrated, this is rarely a perfect
As suggested by Chierchia (2010), the “signature property” of a count noun is that it can be directly modified by a cardinal numeral, as in (14):

(14) one cow, two chairs, five trees, forty-seven cars

Common nouns without this property, such as those in (15), are taken to be mass nouns:

(15) #one beef, #two mud(s), #fifteen wind(s)

As we can see from (14) and (15), the count/mass distinction is linked to both countability and grammatical number in English. Count nouns can appear in plural form, whereas mass nouns cannot. If a mass noun does pluralise, it is understood as having shifted to a count interpretation in which it refers to a standard serving or a set of kinds. The standard serving interpretation can be read from the example in (16), in which three beers might denote ‘three pints of beer’ and one water might denote ‘one glass of water’ or ‘one bottle of water’.

(16) I ordered three beers for us and one water for our designated driver.

The sets of kinds interpretation can be read from (17), where three cheeses refers to ‘three different kinds of cheese’.

(17) After dessert, there is a selection of three cheeses and a glass of port for everyone.
The interpretation of these pluralised mass nouns can be understood from the context, which provides an appropriate countable measure of the non-individual entity.

In addition to modification by numerals and pluralisation, determiner selection is also sensitive to the count/mass distinction. There are some determiners which will combine only with count nouns, as in (18).

(18)  

a.   A cow is missing from the farm.
   
   b.   Every chair needs to be fixed.
   
   c.   Few trees grow taller than this.
   
   d.   Many cars are becoming more efficient.

Some determiners combine only with mass nouns, as in (19).

(19)  

a.   Much coffee is needed this morning.
   
   b.   Little effort is required for this task.

There are some determiners which combine with both mass nouns and plural count nouns, such as those in (20).

(20)  

a. He likes plenty of wine/biscuits.
   
   b. I want a lot of dancing/friends at the party.

Finally, some determiners will combine with either a count or a mass noun, as in (21).
(21)  
  a. I really enjoyed the film(s)/entertainment tonight.
  b. He has bought me some book(s)/jewellery for my birthday.

It therefore seems that, at least in languages like English, the count/mass distinction reflects a perceptual distinction. Nouns are divided into two major classes: count nouns, which refer to individuals, and mass nouns which refer to non-individuals. This means that the semantic feature individual appears to map to a formal feature [+COUNT] in English, as in (22).

(22)  

<table>
<thead>
<tr>
<th>Lexical Item</th>
<th>Formal feature</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[+COUNT]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phonetic content</th>
<th>Semantic feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Count syntax”</td>
<td></td>
</tr>
</tbody>
</table>

However, as we have already mentioned, this mapping is not so clear. One class of nouns highlighting that the count/mass distinction is not a direct reflection of the perceptual individual/non-individual divide is the group of nouns known as ‘object mass nouns’ (Barner & Snedeker 2005), ‘naturally atomic mass nouns’ (Rothstein 2010), or ‘fake mass nouns’ (Chierchia 2010). These nouns, examples of which include cattle, jewellery, and furniture, denote individuals but cannot occur with count syntax. Taking jewellery as an example, jewellery may refer to a collection of individual necklaces, bracelets, and earrings. Whilst these are all individuals, meaning jewellery refers to a set of individuals (like a plural count noun), it is ungrammatical to say I bought two jewellerys today. In fact, jewellery allows determiners like much, e.g. I simply have too much jewellery, suggesting that the grammar treats this noun as a mass noun despite it denoting a set of individuals.
Despite ‘object mass nouns’ being treated as mass nouns morphosyntactically, it is not the case that the grammar does not recognise that these nouns refer to sets of individuals. Some adjectives which usually cannot modify mass nouns, such as big and small, can combine with object mass nouns to modify the individuals in the denotation (Rothstein 2010; Schwarzschild 2011). (23) demonstrates how object mass nouns align with count nouns rather than mass nouns in this respect.

(23)  

a. I have lost all the small jewellery. [Object mass]  
b. I only like small earrings. [Count]  
c. #It is possible to drown even in small water. [Mass]

The grammar therefore shows sensitivity to object mass nouns denoting individuals, despite their overall treatment as mass nouns. Barner & Snedeker (2005) found that the perception of object mass nouns is also similar to that of count nouns. In their experiment, subjects consistently judged object mass nouns on quantity of individual entities, like count nouns, and not on overall size or volume, like mass nouns.

The existence of object mass nouns means that the grammatical count/mass distinction cannot only be based on the perceptual individual/non-individual divide, as a set of individuals is perceived as such despite a mass noun being used for reference.

Including object mass nouns (OMN), the following ‘perceptual to grammatical mapping’ is set up for English:

(24)  

<table>
<thead>
<tr>
<th>Grammatical: Count syntax</th>
<th>Mass syntax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceptual:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

OMN
Object mass nouns show that sets of individuals can be referred to with count nouns and mass nouns. There are also ‘flexible nouns’ (Barner & Snedeker 2005) which are nouns with both a count and mass form. In English, flexible nouns like *string*, *stone*, and *paper* can appear in both count and mass contexts, as seen in (25) and (26).

(25) a. This guitar only has five strings.
    b. How much string will I need?

(26) a. This wall is made of stones.
    b. This wall is made of stone.

(26) suggests that it must be possible for the denotation of a count noun to be the same as that of a mass noun. In this case, it is impossible to determine the count/mass distinction from the perceptual properties of the entity.

So far, we have shown that in English the count/mass distinction correlates closely with the perceptual individual/non-individual divide, but that there are notable differences in the existence of object mass nouns and flexible nouns. Next, we will examine a language which lacks the same indicators of a grammatical count/mass distinction: Mandarin Chinese.

### 3.3 Count and Mass in Mandarin Chinese

In Mandarin Chinese, a bare noun can refer to an individual, a set of individuals, or a non-individual. Nouns denote little more than “some amount of X”, as demonstrated in (27) from Zhang (2014):
The Features and Structure of the lower Noun Phrase

(27)  
\[
\begin{align*}
\text{zhou-shang} & \quad \text{you} \quad xigua \\
\text{table-on} & \quad \text{have} \quad \text{watermelon} \\
\text{‘There is a watermelon on the table.’} \\
\text{‘There are watermelons on the table.’} \\
\text{‘There is a slice of watermelon on the table.’} \\
\text{‘There are slices of watermelon on the table.’} \\
\text{‘There is a pile of watermelon on the table.’} \\
\text{‘There are piles of watermelon on the table.’}
\end{align*}
\]

This suggests that there is no clear count/mass distinction in Mandarin Chinese. The same noun can denote either an individual or a non-individual in the exact same context – taking the flexible nouns of Barner & Snedeker (2005) even further.

Recall Chierchia’s (2010) claim that the signature property of count nouns is their ability to be directly modified by a numeral. In Mandarin Chinese, no nouns have this property. To count an entity denoted by a noun, a “unit word” must appear between the numeral and the noun to identify a discrete countable unit in which the entity denoted by the noun can appear. This is true whether a noun refers to individuals or non-individuals, as in (28):

(28)  
\[
\begin{align*}
a. \quad san & \quad tou \quad niu \\
\text{three UW}_{\text{head}} \quad \text{cow} \\
\text{‘three cows’} \\
b. \quad san & \quad jin \quad mi \\
\text{three UW}_{\text{pound}} \quad \text{rice} \\
\text{‘three pounds of rice’}
\end{align*}
\]

The requirement for a unit word to make nouns countable in Mandarin Chinese aligns them very closely to English mass nouns, which require a measure word to combine with numerals, e.g. *three boxes of flour*. In fact, the parallels between nouns in Mandarin Chinese and mass nouns in English have led many to argue that all nouns are mass nouns in classifier languages like Mandarin Chinese (Allan 1980; Chierchia 1994, 1998a; Krifka 1995).
However, it is not the case that all nouns can take any unit word. Many researchers (e.g. Cheng & Sybesma (1998); Lyons (1977); Tai & Wang (1990)) have found that there are broadly two types of unit word: (i) those whose function is to pick out an individual, and (ii) those whose function is to denote a measure of either a non-individual or individuals. Henceforth, I will be referring to the former as classifiers and the latter as measure words.

Classifiers name the inherent individual unit being denoted by a noun, i.e. *tou* names an individual ‘cow’ from the denotation of *niu* which can then be counted. Classifiers form a closed class and each one relates only to a specific group of nouns – an association that must be learned. In this way, classifiers can only combine with nouns denoting individuals.

Measure words, on the other hand, identify a countable measure of the entity denoted by the noun, i.e. *jin* identifies that the denotation of *mi* ‘rice’ should be counted in measures of a ‘pound’. Measure words are an open class and can be used freely with any noun, regardless of whether they denote individuals or non-individuals, as in (29):

(29) a. *si he pingguo*
    four MW\textsubscript{box} apple
    ‘four boxes of apples’

b. *si he mianfen*
    four MW\textsubscript{box} flour
    ‘four boxes of flour’

The key difference is that classifiers can only combine with nouns denoting individuals, whereas measure words can combine with both individuals and non-individuals. This shows, importantly, that the grammar of Mandarin Chinese is sensitive to the individual/non-individual distinction. Despite the language having no nouns which
can directly be modified by a cardinal numeral, i.e. count nouns, Mandarin Chinese distinguishes between nouns which can denote individuals and those which cannot. Classifiers can only combine with nouns that can denote individuals – similar to the count syntax of English which will only permit count nouns.

It is important to note, however, that these cases are only similar and not identical. Classifiers combine with nouns that can denote individuals. In (27), we saw that bare nouns in Mandarin Chinese can refer to individuals, sets of individuals or non-individuals. The sense of some nouns includes the notion of “natural atomicity” (Rothstein 2010) and it is this to which classifiers are sensitive. If the entity denoted by a noun has the sense of an inherent discrete unit, then it is said to have natural atomicity. These include the referents of nouns such as apple, cow, or flower. Entities without an inherent discrete unit lack this property, such as the referents of the nouns mud, sand, and flour. A sense of natural atomicity is different to being grammatically count, as we have already seen from object mass nouns in English, e.g. cutlery and furniture. Rothstein’s (2010) natural atomicity mirrors what I have so far been referring to as individuals – a naturally atomic predicate has a denotation of individuals.

Returning to the ‘perceptual to grammatical mapping’, it is clear that the individual/non-individual divide is represented differently in Mandarin Chinese:

\[
\begin{array}{ccc}
\text{Grammatical:} & \text{Classifier} & \text{Measure word} \\
\text{Perceptual:} & \text{in} \text{dividual} & \text{n} \text{on-individual} \\
\end{array}
\]

Given that measure words can occur with nouns that denote both individuals and non-individuals, it seems that they are not sensitive to this distinction. Classifiers, on the other hand, are the key to investigating how this perceptual divide is present
The Features and Structure of the lower Noun Phrase in the grammar of Mandarin Chinese.

The mapping in (30) shows that Mandarin classifiers align somewhat with the count syntax of English in that they both require a noun with a denotation of individuals. In English, the count/mass distinction is made across nouns: count nouns denote individuals and mass nouns (for the most part) denote non-individuals. Nouns themselves do not make this distinction in Mandarin Chinese. All nouns in Mandarin Chinese can denote non-individuals, and some can also denote individuals or sets of individuals. Only with a classifier is the sense of individuals made explicit in Mandarin Chinese. In this way, count nouns and classifiers seem to be two different strategies addressing the same underlying perceptual notion (Borer 2005; Cheng & Sybesma 1998).

3.4 What and where is the count/mass distinction?

One of the most fundamental questions regarding the count/mass distinction has been whether it is (i) the grammatical encoding of the perceptual individual/non-individual divide, (ii) the grammatical encoding of the semantic property of being countable, or (iii) the result of a null syntactic classifier operating over flexible nouns. In this section, I will present and compare each of these three options.

3.4.1 Encoding the individual/non-individual divide

Count nouns and mass nouns can be defined in relation to the properties of individuals and non-individuals. Mass nouns are homogeneous, as proposed by Link (1983). Given that adding one quantity of milk to another quantity of milk results...
in a quantity of milk, this noun is upwardly homogeneous, or cumulative. Similarly, splitting a quantity of milk into two quantities results in two quantities of milk, so the noun is also downwardly homogeneous, or divisive. This contrasts with singular count nouns. Adding two individuals in the denotation of chair does not result in an entity in the denotation to chair (only the plural chairs), and dividing a chair into two parts will leave two parts which are not both a chair. Plural count nouns are, however, similar to mass nouns in terms of cumulativity and divisivity. Adding chairs to chairs will give chairs, and taking chairs from chairs will leave chairs (down to minimal parts).

In terms of singular nouns though, Link (1983) claims that mass nouns and count nouns denote different types of entities. Mass nouns, or homogeneous singular predicates, denote non-atomic Boolean semi-lattices. Singular count nouns, or non-homogeneous singular predicates, denote sets of atoms. The closure of these sets of atoms under sum gives the denotation of plural count nouns – an atomic Boolean semi-lattice. As mass nouns and plural count nouns both denote Boolean semi-lattices, they are both homogeneous, but the fact that the semi-lattice for plural count nouns is atomic means that their denotation is divisive down to minimal parts, rather than being fully divisive like mass nouns.  

Object mass nouns pose a problem for this approach, given that cutlery and footwear are clearly composed of individual parts. Object mass nouns cannot be divisive in the same way as mass nouns such as milk. Dividing cutlery might leave a fork, which is in the denotation of cutlery, but a piece of a fork is not. Object mass nouns seem to have minimal parts in the same way as plural count nouns.

A further problem with count nouns and mass nouns having different types of de-

---

2 Whether the denotation of mass nouns is truly fully divisive is debatable. For example, a molecule of water counts as water, but dividing it further does not give two quantities of water. I will not discuss here the matters of whether the denotation of mass nouns is truly built from minimal parts, or whether the semantic system should care about the sub-molecular level. For discussion, see Landman (2011).
notations, as highlighted by Chierchia (1998a), is the presence of near-synonyms like *carpets* and *carpeting*. *Carpets* will have the denotation of a plural count noun, whereas *carpeting* has the denotation of a mass noun. It is unlikely, or at least undesirable, that these two nouns, which can refer to exactly the same entity, denote different types of objects.

Chierchia (1998a) approaches the count/mass distinction in a different way. Count and mass nouns both have atomic denotations, but the “vagueness” of the atomic elements differs. As for Link (1983), a singular count noun denotes a set of atoms, and a plural count noun denotes the closure of this set under sum. A mass noun also denotes a set of atoms closed under sum, but given that mass nouns are inherently plural and there is no semantic operation triggering plurality, it is not clear what the atoms making this set are – the atoms of the plural set are “vague”. Under this approach, *carpets* and *carpeting* have the same denotation, an atomic Boolean semilattice, and whilst the denotation of *carpets* is built from the set of atomic *carpets*, the denotation of *carpeting* comes from the noun’s lexical meaning, which allows for vagueness around what an atomic piece of carpeting may be. In this way, object mass nouns can be neatly accounted for. Mass nouns proper have a denotation with total vagueness around their atomic parts. Object mass nouns have a denotation with less vaguely specified atoms and therefore fall somewhere in between pure mass nouns and count nouns.

Chierchia (2010) develops the notion of vagueness further. There is now vagueness in what counts as an instance of a noun and instability in what counts as an atom of that noun, and these two interact. In a given context, a count predicate will be true of some things, not true of some other things, and possibly true of the remaining things – there is vagueness in the denotation of the noun. To find the definition of the noun, the vagueness can be reduced. It may be the case that a mass predicate is only possibly true or not true, and there is much greater vagueness in the denotation.
Through the stages of reducing the vagueness, a count predicate will have stability in its atomic structure, meaning that what counts as an atom remains the same. A mass predicate is not stably atomic, meaning that what counts as an atom can change as the denotation of the noun becomes less vague and more defined. Singular count nouns denote sets of stable atoms, plural count nouns denote sums of stable atoms, and mass nouns denote sums of unstable atoms. It is the stability of the atoms in the denotation that determines whether a noun is countable or not, and therefore whether it is count or mass.

One problem with this approach is that not all count nouns are stably atomic. Take the noun *fence*. Rothstein (2010) demonstrates that it is not always clear what the atomic parts of a noun like *fence* are. Her example outlines a case in which the owners of two connected houses each decide to build a fence at the front of their properties. Each owner starts to build at the end farthest from the other and works inwards until they join in the middle. We can say that this is one fence built by two people or two fences each built by one person, but for either scenario we need to determine what counts as an atomic part of *fence*, meaning it is not stably atomic but context-dependent. There is also the opposite problem – some mass nouns, namely object mass nouns, are stably atomic.

In summary, both Link (1983) and Chierchia (2010) aim to link the count/mass distinction to differences in what counts as a minimal part of a count or a mass noun. Link (1983) proposed two separate domains, one atomic and one non-atomic. Chierchia (2010) proposed one atomic domain, with some atoms being stable and some being unstable. The problems for both approaches come from mismatches, such as object mass nouns like *cutlery*, homogeneous count nouns like *fence*, and flexible nouns or near-synonymous pairs like *carpets/carpeting*. Given the presence of these mismatches, some theories have decided to separate the count/mass distinction from the differences in the objects denoted by nouns to instead focus on how nouns
denote these objects.

### 3.4.2 A semantic basis for the count/mass distinction

A number of theories place the count/mass distinction as down to differences in the semantic properties of the nouns themselves, rather than the perceptual differences in the objects they denote.

Krifka (1989) starts with the observation that objects which are countable are denoted by count nouns, and therefore asks how countability is reflected in the grammar. Count nouns are derived from mass nouns through an operation which changes the denotation of the noun to make it countable.

Mass nouns are naturally of type \(<e,t>\). Count nouns incorporate the meaning of a measure expression through a two-place relation between numbers and entities: \(<n,\langle e,t \rangle>\). Take the measure expression *grain of rice*. *Rice* is a mass predicate and *grain* provides a unit for a measure relation between the predicate and a number. This process is demonstrated in (31).

\[
\begin{align*}
(31) \quad a. \quad & \text{rice} \quad \langle e,t \rangle \\
    b. \quad & \text{grain of rice} \quad \langle n,\langle e,t \rangle \rangle \\
    c. \quad & \text{three grains of rice} \quad \langle e,t \rangle \\
\end{align*}
\]

Count nouns have this measure operation built in, with the measure unit being a “natural unit” of the noun. They are derived from abstract mass nouns but emerge as type \(<n,\langle e,t \rangle>\). The process, as in (32), mirrors that in (31).

\[
(32) \quad a. \quad (\text{tree} \quad \langle e,t \rangle )
\]


The count noun has the meaning in (32-b), and can combine with a numeral, as in (32-c), to return to type $<e,t>$.  

Krifka’s approach therefore sees the count/mass distinction as a lexical typal difference between count and mass nouns. Mass nouns are predicates of type $<e,t>$ and count nouns are relations of type $<n,<e,t>>$. There is no difference between singular and plural count nouns, this distinction is made based on the $n$ argument selected. Once the $n$ argument is satisfied, the typal difference between count nouns and mass nouns is lost. This poses a problem, given that determiners are sensitive to the count/mass distinction.

Another problem is that, through making counting analogous to measuring, Krifka has lost the ability to make the distinction between classifiers and measure words. Classifiers, as we saw in Mandarin Chinese, name an individual unit based on the inherent structure of the entity. Measure words, on the other hand, simply identify a quantity of the entity independent of this structure. Given the different semantic functions of classifiers versus measure words, it may be important to ensure that the operation built in to count nouns is the same as that carried out by a classifier rather than a measure word.

Rothstein (2010) builds on the approach of Krifka (1989), but emphasises the separation of counting and measuring as different operations. Rothstein (2010) maintains that counting is always specific to a noun of a certain kind, and the noun must define both the sort of entity being counted and the atomic minimal parts in its denotation. Count nouns provide this information, and are therefore able to be counted.

Like Krifka (1989), Rothstein (2010) derives count nouns from mass nouns, but
she does not set out counting as a special form of measuring. Given the context-dependent nature of determining what counts as an individual atom in the denotation of nouns like fence and wall, Rothstein (2010) shows that count nouns do not always have a denotation of inherently individual atoms – atomicity is always relevant to a context. An example of this issue, from Rothstein (2010), was given above. Two neighbours, each building one fence, might end up with one fence between them if the fences meet. The result is either one fence built by two people or two fences each built by one person. The context determines what counts as an atomic part of fence. The context-dependent atomicity must be encoded in the meaning of a count noun, and so the operation converting mass nouns into count nouns must return as the denotation of the noun the set of entities which count as atoms in a given context.

Rothstein (2010) claims that all nominals start from an abstract root noun, \( N_{\text{root}} \). The denotation of root nouns contains both natural atoms, i.e. stable individuals, and non-stable entities. Mass nouns are the same as root nouns, and therefore the operation MASS, when applied to \( N_{\text{root}} \), returns \( N_{\text{root}} \). The operation \( \text{COUNT}_k \) applies to root nouns and selects only the entities which count as semantic atoms in context \( k \). \( \text{COUNT}_k(N_{\text{root}}) \) returns the set of semantic atoms in context \( k \), which is the interpretation of a count noun in context \( k \). COUNT will always return a set of semantic atoms, irrespective of whether the input set was naturally atomic. This highlights the importance of context \( k \) – \( k \) has less impact in \( \text{COUNT}_k \) when applied to a naturally atomic set, as the choice of atom is constant, but when \( \text{COUNT}_k \) applies to a non-naturally atomic set, such as the denotation of fence, it is \( k \) that will determine what is returned as the set of atoms.

\( \text{COUNT}_k(N_{\text{root}}) \), or \( N_k \) for short, forms the input set for the plural operation, which returns the denotation of plural count nouns.

Rothstein’s approach to count nouns and mass nouns results in a typal difference.
A mass noun is of type $<e,t>$ and denotes a set of entities. A count noun is of type $<e \times k,t>$, denoting a set of entities indexed for context $k$. In contrast to Krifka’s (1989) approach, this typal difference is preserved throughout the building of the larger NP and DP and can explain the sensitivity of the wider grammar to the count/mass distinction.

The greatest advantage of this account, however, is the separation of natural atomicity and semantic atomicity. Natural atomicity can be represented by either mass or count nouns, but semantic atomicity, which is context-dependent, is grammatically encoded in count nouns.

### 3.4.3 A syntactic count/mass distinction

Similar to the accounts provided by Krifka (1989) and Rothstein (2010) above, these approaches also maintain that count nouns are derived from mass nouns. However, these theories argue that this “conversion” occurs in the syntax and that all bare nouns are, in principle, flexible.

From this, it follows that flexible nouns are the best support for these approaches. Consider the following sentences in (33).

(33) a. You could use rope / a rope to tie that together.
    b. Shall I order three beers and a water?
    c. By the end of the Christmas period there was tree all over the carpet.

Using a mass noun as a count noun, as in (33-a) and (33-b) is termed ‘packaging’. The count interpretation of the mass noun is said to come from it being the complement of a null classifier. *Three beers* and *a water* are likely to mean *three pints of beer* and *a glass of water*. With an explicit measure phrase, the plural marking
occurs on the measure word. In the case of a null classifier, as in (33-b), it is possible for plurality to be marked directly on the noun.

Using a count noun as a mass noun, as in (33-c), is known as ‘grinding’. There is, however, no operation applied here – the noun simply does not become the complement of a classifier. Without the null classifier to provide a count interpretation, (33-c) only says that there is some amount of ‘tree stuff’ on the carpet. All nouns are mass nouns until they are placed under a classifier, and even what we understand to be count nouns having undergone ‘grinding’ were never really count nouns at all.

Borer (2005) follows this approach, and places classifiers in the same syntactic position as grammatical number – under the syntactic head ‘Num’. Both apply a semantic function, DIV, to the denotation of the noun to divide the denotation into contextually relevant units. Once the denotation is divided, the set of units can be counted. When the syntactic ‘Num’ head is empty, i.e. not filled by a classifier or number marker, the noun is interpreted as a mass noun.

The benefit of Borer’s approach is that it can align number marking languages like English with classifier languages like Mandarin Chinese. Number marking and the null classifier occupy the same position as explicit classifiers. This also explains the lack of number marking in Mandarin Chinese.

### 3.4.4 Three approaches to consider

Three ways of approaching the count/mass distinction have been discussed. The first option of a grammatical encoding of the perceptual divide between individuals and non-individuals seems intuitive, but cannot fully account for the common mismatches of object mass nouns, homogeneous count nouns, and flexible nouns. The
second option, a grammatical encoding of semantic countability, separates semantic
atomicity from natural atomicity, allowing mismatches to be accounted for. The
third option of a syntactic operation which derives count interpretations predicts
that all nouns are mass unless a function, hosted in a syntactic head, applies to
create an atomic denotation. In this scenario, mismatches are not possible.

In the next section, I will examine these approaches against the proposed count/mass
distinction in Mandarin Chinese.

3.5 Examining the approaches: Mandarin Chinese

Recall that both classifiers in Mandarin Chinese and count syntax in English required
a noun with a denotation of individuals. In Mandarin Chinese, this is always true:
if the denotation of the noun contains atoms then it can take a classifier. In English
however, some nouns with denotations containing atoms, namely object mass nouns,
could not appear with count syntax.

Recall also that all bare nouns in Mandarin Chinese are understood to be flexible
nouns. The example sentences in (27) from Zhang (2014), repeated here as (34),
showed that the noun on its own holds no information in relation to its status as
count or mass.

(34)  Zhou-shang you xigua
table-on have watermelon
‘There is a watermelon on the table.’
‘There are watermelons on the table.’
‘There is a slice of watermelon on the table.’
‘There are slices of watermelon on the table.’
‘There is a pile of watermelon on the table.’
‘There are piles of watermelon on the table.’
Given that all nouns are flexible, it is impossible for mismatches to occur in Mandarin Chinese. Even when classifiers are used to create a count interpretation, these must be the correct classifier for the noun or the general classifier. Without mismatches, the individual merit of the three approaches outlined above is more evenly matched. It is plausible that, in Mandarin Chinese, classifiers can only be used with nouns that perceptually represent sets of individuals, following Link (1983) and Chierchia (1998a). It is also still possible that the semantic atomicity of the noun’s denotation determines whether a classifier can be used, following Krifka (1989) and Rothstein (2010).

The situation in Mandarin Chinese aligns somewhat with the conditions of the syntactic approach of Borer (2005) (given that this was one of the aims of the theory): all nouns are flexible and a syntactic operation creates count interpretations. However, it is not the case that the dividing operation of the classifier can apply to any noun. Only certain nouns can take classifiers, and it is precisely the lack of mismatches or flexibility with classifiers that causes a problem for this approach. In Mandarin Chinese, it seems that the perceptual divide is the stronger driving force for the count/mass distinction, as there is no motivation to separate semantic atomicity from natural atomicity.

Based on this observation, can it be the case that different approaches to the count/mass distinction might apply more or less successfully crosslinguistically? I argue that this is precisely the case, and that, despite the fact that each language is aiming to reflect in some way the same perceptual divide between individuals and non-individuals, the way in which they do this can vary depending on a number of factors, including the salience of this divide in the grammatical system and the competing pressures of encoding other properties in the grammatical system during acquisition.
3.6 Acquiring Count and Mass in English

3.6.1 Acquiring the count/mass distinction

English-acquiring children seem to be sensitive to the count/mass distinction by around the age of 24 months, as demonstrated by their use and comprehension of number marking and agreement (Barner et al. 2007; Clark & Nikitina 2009; Kouider et al. 2006). Bale & Barner (2012) outline an experiment in which 24-month-old children are shown to use verb agreement to guide their interpretation of plural nouns to sets of individuals. Fenson et al. (1994) show that acquirers start to use other quantifiers at around a similar time: more at 19 months, some and all at 26 months, the indefinite article a at 27 months, and another at 29 months. Numerals also start to emerge at around 2 years of age (Barner et al. 2009).

When learning new words, children are sensitive to the count/mass distinction by around 30 months of age. Soja (1992) found that children, when shown a novel object and taught a novel label for it, were more likely to infer that the word denotes an individual when used in count syntax than when used in mass syntax. Another experiment, from Barner & McKeown (2005), examined how acquirers interpreted novel words based on their representation in a syntactically neutral environment. Children were shown either an individual object or a set of ten objects of the same kind and then given a novel label with neutral syntax. They were then presented with a plural set and asked to label it. The children were more likely to use a plural form as the new label if the original word had labelled an individual than if it had labelled the ten objects. This suggests that the children took the original label as a mass noun if it labelled a plural set, given that it lacked plural morphology.
3.6.2 Clues for acquisition

The experiments outlined above suggest that, at least to begin with, the count/mass distinction in English is based on the perceptual divide between individuals and non-individuals. By around 24 months of age, English-acquiring children are noticing that number marking and number agreement highlight a perceptual difference between individuals and non-individuals. By around 30 months of age, children are using these clues to determine whether new words denote individuals or non-individuals.

Number marking is prevalent in the grammar of English, and is a relatively simple system: there is a plural suffix or there is no suffix. For the vast majority of nouns in English, the plural suffix is realised as /s/, /z/, or /ɪz/ and there is no singular suffix. Children quickly intuit that nouns that can appear with this plural suffix denote individuals. This correlation is salient, frequent and systematic. Similarly, nouns which cannot appear with a plural suffix denote non-individuals. This suggests that, for English, the acquisition of the count/mass distinction is closely related to the acquisition of grammatical number, and number agreement elsewhere will also contribute as overt evidence for the presence of the distinction.

It is important to note that the evidence provided by grammatical number is clearer for the acquisition of number than it is for the count/mass distinction. Nouns for which plural marking and count syntax are available denote individuals, and these nouns denote individuals regardless of whether this evidence appears in a given context, i.e. absence of plural marking and/or count syntax does not necessarily entail that the referent of the noun is a non-individual. This is different to the process of acquiring grammatical number for which, in general, the absence of plural marking on count nouns does entail lack of plurality, usually singular number. The acquirer works out the count/mass distinction through number marking – two interrelated
systems that rely on each other for their acquisition.

Bale & Barner (2012) claim that such a learning strategy requires inductive inference. The systematic use of particular forms of nouns to refer to a specific set of entities or, more exactly, the observation that singular individuals can only be referred to by bare nouns, leads a child to infer that those bare nouns with the ability to pluralise denote individuals. Bale & Barner (2012) point out that this type of inference is widespread in other areas of language acquisition, including the acquisition of category labels and their extension to novel referents. Citing Xu & Tenenbaum (2007), they also note that inductive processes like this, which draw on clear correlations, are easily adopted into inferential Bayesian models which mirror human learning.

3.6.3 Grammaticalising the distinction

At this point, the acquirer has linked the phonetic content of plural marking to the perceptual denotation of individuals. The salience and systematicity of this correlation is such that the acquirer determines that a formal feature should be posited within the grammar to capture this correlation, as the individual/non-individual divide is clearly relevant to the syntax. The acquisition facts suggest that the perceptual divide is grammaticalised relatively early in the acquisition sequence, and also that the acquisition cues are used from an early stage to determine the status of novel words in relation to the count/mass distinction. This means that the count/mass distinction in English is first acquired as an exact encoding of the perceptual divide.

The “signposts” for the count/mass distinction – number marking and other structural features – are not always present. The presence or absence of number marking in English must of course first represent a contrast in number. There is no mor-
The Features and Structure of the lower Noun Phrase

Phological content in English that uniquely represents the individual/non-individual divide; there are only structures that are compatible with nouns of the correct denotation in terms of individuals. This means, according to the approach I have followed so far, that the count/mass distinction cannot be formalised as its own function, and that the perceptual divide must be stored in the lexically-specified denotation of the noun itself. This follows the route outlined in (9-b) above: the feature is grammaticalised on an element bearing a semantic feature as a result of its controlling an overt realisation elsewhere on another element. This includes both the availability of number marking and count syntax. The count/mass feature, as will be a sufficient term for now, must be grammaticalised as a formal feature in the lexical specification of a noun.

This suggests that the count/mass distinction in English can best be accounted for by the approaches of Link (1983) and Chierchia (2010). The perceptual divide directly informs the count or mass status of a noun. Following the single domain approach of Chierchia (2010), a noun that refers to individuals has the denotation represented by an atomic Boolean semilattice, such as that in (35).

\[(35)\] Count noun:

\[
\begin{array}{cccc}
abcd & abc & abd & acd & bcd \\
ab & ac & ad & bc & bd \\
a & b & c & d \\
\end{array}
\]

The semilattice represents all possible singular individuals to which the noun could refer (the bottom row) and all possible sets of these individuals (other rows). That is not to say that in the lexical entry of, for example, person, all known persons are
represented and a new atom (and corresponding sets) is added every time a new person is met. To construct the relevant semilattice online for any given noun, the semantic content of that noun is used, i.e. the noun *person* will be understood as a kind of individual from the mapping between perceptual and semantic categories. The semantic content will be used for further identification in the construction of the nominal phrase.

A noun that refers to non-individuals is represented by an atomic semilattice with vague or unstable atoms, as in (36).

(36) Mass noun:

This shows how each vague atom, \( x \), can be divided into smaller quantities of \( x \) or accumulated into larger quantities of \( x \). Given that the atoms are vague, it is not possible to select any specific atom. These atoms cannot be counted and must be assigned a measure.

Based on the evidence, an acquirer will store nouns in the lexicon as one of these two types, using grammatical features such as [COUNT] and [MASS] to lexically encode the perceptual divide of individuals and non-individuals respectively.
3.6.4 Building the functional structure

At this point then, in English, acquirers have nouns listed in the lexicon as either count or mass, and these two types of nouns have different types of denotations. The noun forms the starting point for the referent-identifying task of the nominal phrase, which is carried out as the phrase is built, and the type of denotation of the noun will determine which functions can be applied.

As an illustration, take the noun *house*. This is a count noun which will have a lexically-specified denotation of stably atomic individuals. This denotation will have been attached to the noun based on the evidence provided by number marking and count syntax during the acquisition phase. *House* may not be one of the earliest nouns acquired, but it will match the distribution already understood from those earliest nouns (or perhaps even before the category noun is fully formed) in terms of the correlation of the evidence of the overt realisation and the interpretation of its status as an individual.

At this stage of constructing the lexicon, only four individuals in the world can be labelled with the noun *house*: *a*, *b*, *c*, and *d*. Introducing the word *house* into the syntax provides a denotation including these four individuals, and all the sets created by their union. The semilattice is given in (37):

\[
\text{(37) house:}
\]

![Diagram of semilattice](image)
The denotation here is at its widest. The task carried out by the construction of the nominal phrase is to signify the intended referent from those available by applying functions to this denotation.

Just as the acquirer posited a feature on nouns in the lexicon to reflect the perceptual individual/non-individual divide, the acquirer has linked the phonetic content of plural marking to the perceptual contrast between individuals and sets of individuals. The presence of one variant of a restricted set of sounds on the end of the noun links systematically with whether the interpretation of the noun is a set of individuals. This leads the acquirer to determine that the semantic feature of plurality should be grammaticalised as a formal feature – the atom/set of atoms divide is grammatically relevant. Given that, in this case, the noun itself is marked with the phonetic content, and that this phonetic content uniquely represents the perceptual divide, number is grammaticalised as its own function within the nominal phrase. In English, the Number function is the first referent-identifying function that can apply to the denotation during the syntactic construction of the nominal phrase.

Features and functions are therefore closely related. In the nominal phrase, formal features may be lexically-specified on the noun and reflected in the denotation that is presented to the structure-building process. These features may control agreement elsewhere in the structure. Features tied to specific interpretations with their own phonetic content create functions in the projection of the phrase, and are sequenced based on a combination of factors, including how they apply to the denotation and how salient they are to the acquirer.

As plural marking links uniquely to sets of individuals, the feature of number is grammaticalised as a function hosted in a functional head which, for theoretical consistency, we will call Num (following Ritter (1991)). This functional head hosts plural morphology. The Number function selects from the denotation of the noun
either individuals (singularities), or sets of individuals (pluralities). The denotation must include stable atoms for the Number function to apply. The Number function selecting individuals will select only the bottom row of the semilattice, whilst the Number function selecting sets of individuals will select all but the bottom row of the semilattice. To return to our illustration using the noun house, The Number_{sg} function will select the bottom row of the semilattice, leaving only a, b, c, or d as the possible referents for the nominal phrase being constructed. Applying Number_{pl} to the denotation of house to select only sets of individuals leaves the denotation in (38). This now represents the denotation of the bare noun houses:

(38) houses:

Functions will only apply to appropriate denotations. This means that both Number_{sg} and Number_{pl} cannot apply to the same denotations, as one function will rule out the application of the other. Both can apply to the denotation supplied by a noun that is lexically-specified as count (i.e. with a denotation of individuals), but neither can apply to the denotation supplied by a noun that is lexically-specified as mass (i.e. with a denotation of unstable or vague atoms). However, this rules out the availability of such phrases as three beers and one water. In cases such as these, I argue that the context can “stabilise” the atoms in a denotation. Once the atoms become more stable (or less vague), it becomes possible for the Number function to apply to the denotation. By linking this to the context, the observation that certain mass nouns are more susceptible to ‘packaging’ than others can be accounted for.
The denotations that are the outputs of the Number function may form the input for further functions in the nominal phrase. One such function could be called the Cardinality function. This function applies through a functional head we will call Numeral, which hosts numerals. The Cardinality function can only apply to a denotation consisting of individuals, just like the Number function. The Cardinality function selects from the denotation those entities which contain the number of individuals represented by the numeral hosted in the functional head. For example, the Cardinality function applied by the numeral *two* will select from the denotation those entities which are a set of two individuals.

In English, the Cardinality function and the Number function are closely linked, and the Number function must always apply to the denotation first. If the Number$_{pl}$ function applies to a count denotation, it selects all sets of individuals. A Cardinality$_{2+}$ function can apply to this set and will select sets of the designated cardinality. If the Number$_{sg}$ function applies to a count denotation, it selects all individuals. A Cardinality$_{1}$ function can apply to this set and return the same denotation as its input – only the individuals. Applying Cardinality$_{2+}$ to the output of Number$_{sg}$ will return the empty set, and similarly applying Cardinality$_{1}$ to the output of Number$_{pl}$ will return the empty set. This captures the requirement in English for sets of individuals to be marked as plural if they are being modified by a numeral of two or above, and for individuals to be marked as singular if they are being modified by the numeral one.

Returning to the illustration, applying the Cardinality function of the numeral *three* to the denotation of *houses* in (38) selects only the sets of individuals with a cardinality of three. This leaves the denotation in (39), representing the nominal phrase *three houses*:
The referent of the nominal phrase *three houses* must therefore be one of the four sets of individuals left in the denotation. Other functions could now apply to restrict this denotation further, such as those applied by a demonstrative (*those...*), a possessive (*Leanne’s...*), or a relative clause (*...that we saw yesterday*). The uppermost functions of the nominal phrase will be those most concerned with identifying a single referent for the noun phrase in the real world. This links with the ideas of Longobardi (1994) who associated the D head, standardly assumed to be the highest functional projection in nominals, with referentiality, argumenthood and “looking outside” the nominal phrase.

A mass noun with its lexically-specified denotation of vague atoms enters into the nominal phrase in the same way. The Number function cannot apply to this denotation, as there are no stable individuals or sets of individuals to select. This accounts for the inability for mass nouns in English to be marked for number (except in those “stabilised-by-context” cases already discussed). There are functions which can apply to a denotation of vague atoms, and some of these will be discussed in more detail in Chapter 4.

Based on the grammaticalisation of these functions through functional heads, the following structures of the lower noun phrase are constructed. (40) shows how the nominal phrase for a counted plural count noun is built. Each functional head is
split into the content it hosts and the function it applies.

(40) Counted plural count noun: *three houses*

\[
\begin{array}{c}
\text{NumP} \\
\text{NumeralP} \\
\text{Numeral} \\
\text{three} \\
\text{CARDINALITY}_3 \\
\text{–s} \\
\text{NUMBER}_{pl} \\
\text{N} \\
\text{house}_{COUNT}
\end{array}
\]

This clearly shows how the feature `count` is lexically specified in English, but that the Number and Cardinality functions are grammaticalised as functional heads hosting any associated phonetic content.

The –s suffix may be realised on N through N to Num movement. Given that numerals can be modified, e.g. *almost three houses*, I am assuming that Numeral projects to the phrasal level before optionally merging with the extended projection of the noun, which at this stage in this example is NumP. In English, there is a separation between Num and Numeral, reinforced by Greenberg’s (1963) “Universal 20”: in prenominal position, the order of demonstrative, numeral, and adjective (or any subset thereof) conforms to the order Dem > Numeral > A > N. This demonstrates that Num and Numeral do not form one unit in English, but, as we will see, similar functions to NUMBER may work in different ways.

(41) shows how the nominal phrase for a counted singular count noun is con-
(41) Counted singular count noun: *one house*

The trees represent the application of functions, in order, and mirror whether functions can legitimately apply as the structure is built. (42) represents how neither the Number function or the Cardinality function can apply to a mass noun. Again, the feature mass is lexically specified in English. The incompatibility of Number and Cardinality with a denotation containing unstable or vague atoms results in a structure with no functional projections, as neither of these functions can apply, although this is not to say that none could apply later in the derivation. Further functions may be compatible with such a denotation.

(42) Mass noun: *sand*
3.6.5 Creating mismatches

The count/mass distinction in English is acquired relatively early, based on the strong correlation of number marking and count syntax with the perceptual divide between individuals and non-individuals. From early on in acquisition, these signposts are being used to judge whether novel words are count or mass, and therefore which denotations should be given to these nouns in the lexicon. This means that judgements regarding the count/mass status of nouns are quickly taken away from the perceptual divide, and start to be made based on the noun’s environment in the input. Based on this, nouns that appear in mass contexts will be acquired as mass nouns, and nouns that appear in count contexts will be acquired as count nouns, regardless of their perceptual status.

3.6.5.1 Object mass nouns

This creates a problem for object mass nouns. On the surface, the use of object mass nouns in mass contexts means that they are stored in the lexicon as mass nouns, with a denotation consisting of unstable atoms. Although object mass nouns refer to sets of individuals, the morphosyntax does not initially provide the clues to the acquirer that the denotation should consist of stable atoms. This mismatch between the lexical denotation and the perceptually atomic nature of object mass nouns is managed by moving towards the approach of Rothstein (2010). Barner & Snedeker (2005) propose that a good plan for the acquisition of the count/mass distinction in English is to first assume that the syntax is providing the right information and, when this creates inconsistency, change the denotation of the noun itself as an exception. This process is similar to the way in which children learn irregular plural forms.

To alter the denotation of object-mass nouns to reflect their atomic status, whilst re-
maining mass nouns, the atoms of the denotation need to be “stabilised” to some extent. Contrast (43) with the denotation of a pure mass noun, as given in (36).

(43) Object mass noun:

By indexing the unstable atoms, they become proper minimal parts of the denotation without fully stabilising. This maintains the cumulative and divisive (to an extent) nature of the mass noun. It also ensures that the Number function cannot apply to this denotation, as there are no fully stable individuals, i.e. \{a\}, \{b\}, which can create sets of individuals, i.e. \{ab\}.

This is a different solution to that proposed for instances of “packaging” mass nouns, such as *three beers* and *four waters*. In those situations, I argue that the context provides stabilised atoms in the denotation based on some contextually-defined unit of measure. With object mass nouns, the denotation of the noun is altered permanently, as an exception, in such a way that unstable atoms may become full minimal parts, but they do not become truly stable in such a way that functions can operate on them as individuals. This has the effect of limiting the divisive nature of the denotation, such that minimal parts can be reached and downwards homogeneity is not limitless.
3.6.5.2 Flexible nouns

Flexible nouns could also present a problem for this approach, especially given that nouns are said to be lexically-specified as either count or mass from an early stage of acquisition. However, this specification only takes place based on the presence of strong evidence one way or the other. If a noun is appearing in the input in both count and mass contexts, how will the acquirer determine the type of denotation to give the noun?

For Dutch, in which the count/mass distinction manifests in a similar way, Van Witteloostuijn (2013) found that acquirers attained adult-like use of flexible nouns in count contexts from around 8 years of age (as opposed to the age of 4 for pure count nouns) and that none of the ages studied, up to the age of 12:6, had adult-like performance for flexible nouns in mass contexts (as opposed to the age of 8 for pure mass nouns). This suggests that, with the lack of consistent morphosyntactic evidence for the count/mass distinction, the denotation of flexible nouns is very difficult to specify.

Given that we have seen examples of gradable stability in atoms, perhaps it is possible to give flexible nouns a mixed denotation. Both count and mass contexts are available, and both count and mass uses of a noun can apply the same referent. The denotation of flexible nouns could therefore be represented as in (44).
This denotation would be the one stored for a flexible noun like *stone* or *chocolate*. Based on both count and mass contexts being available for these nouns, the functions available for both count and mass contexts need to be able to apply to the denotation. Similar to the denotation of object mass nouns, there is a link between unstable and stable atoms. Between the fourth and fifth rows of the denotation in this example, stable atoms can combine or divide to become unstable atoms and unstable atoms can combine or divide to form stable atoms. This accounts for examples such as those in (45).

(45) a. I've just dropped a glass, so there will be broken glass all over the floor.

b. When you are finished with your papers, make sure they are collected in the blue bin, which is for paper only.
A count use of *glass* dividing into a mass interpretation is shown in (45-a), whilst (45-b) shows a count use of *paper* combining into a mass interpretation.

From this flexible denotation, the Number function can only apply to the stable individuals (the bottom section of the diagram) so that mass nouns do not appear with number marking. The same goes for the Cardinality function. The full denotation, which represents the bare noun, includes both mass and count interpretations – it is not possible to tell which interpretation the bare noun has on its own – and the context will apply the appropriate functions to this denotation to determine the correct interpretation.

The argument for the combined denotation for both uses in English comes from the lack of evidence for a lexical specification either way, and the unattractive alternative of treating mass and count uses of the same word as different words entirely. This is not strong evidence for this treatment, as the approach here could be extended further into pairs such as *carpet* and *carpeting* which, although having different word forms, could be treated as the same entity. The upcoming discussion of flexible nouns in Mandarin Chinese will shed some further light on this issue.

### 3.7 Acquiring Count/Mass in Mandarin Chinese

#### 3.7.1 Acquiring the count/mass distinction

The acquisition of the count/mass distinction is later in children acquiring Mandarin Chinese than in those acquiring English. Whereas in English, children linked individuals to count syntax as early as the age of 2;6, and showed adult-like use by the age of 4;6 (Soja et al. 1992; Barner & Snedeker 2005), children acquiring Mandarin Chinese do not achieve adult-like performance of the count/mass distinction until 6
or 7 years of age (Huang et al. 2005).

Acquirers of Mandarin Chinese understand that classifiers include information such as the shape of the noun’s referent by the age of 4. Li et al. (2008) found that four-year-olds could match classifiers to nouns based on shape, but that, by the age of 6, children were still pairing classifiers with nouns with a non-solid referent. Cheung et al. (2010) further found that children under the age of 6 do not infer that a novel noun refers to an individual from the presence of a default classifier. The observation that classifiers can only be used with nouns with denotations that include individuals is made later for acquirers of Mandarin Chinese than the link between number marking and denotations of individuals is made by acquirers of English.

This prompts an interesting question: why is the same semantic feature grammaticalised earlier in English than in Mandarin Chinese? In English, the acquisition of the count/mass distinction was proposed to be the result of how systematic and transparent the number marking system is – the number system provided strong evidence for the grammaticalisation of the individual/non-individual divide. Do we see the same strength of evidence in Mandarin Chinese?

### 3.7.2 Clues for acquisition

Firstly, there is no morphological marking on the noun in Mandarin Chinese. As we can see in this example from Zhang (2014), repeated again here as (46), a bare noun can refer to both an individual or a non-individual.

\[(46) \quad \text{zhou-shang you xigua} \]
\[\text{table-on have watermelon} \]
\[\text{‘There is a watermelon on the table.’} \]
\[\text{‘There are watermelons on the table.’} \]
\[\text{‘There is a slice of watermelon on the table.’} \]
We have discussed that the realisation of the count/mass distinction in Mandarin Chinese is at the classifier level – only those nouns which can denote individuals can combine with a classifier. Recall that the evidence for the grammaticalisation of a semantic feature is a systematic link with some overt phonetic content. In English, number marking on the noun was a signpost for the count/mass distinction. Compared to English number marking, the classifier system of Mandarin Chinese is much more complex. McEnery & Xiao (2010) found 159 different types of unit classifiers in their corpus study of Mandarin Chinese. In addition to their variety, each classifier can also contain semantic information, such as shape, animacy, function, etc., that is largely irrelevant to the count/mass distinction. There are also a number of similarities between classifiers and measure words, which at least superficially share many aspects of their distribution. In sum then, there is a wide range of classifiers, which contain additional semantic information, and share similarities with another element which does not provide evidence of the individual/non-individual divide.

Therefore, the evidence for a count/mass distinction in Mandarin Chinese is much weaker than that in English. The correlation between a denotation of individuals and the presence of a classifier is not immediately salient and, given the variety of forms, does not present with any great frequency or systematicity. During early acquisition, there is not enough evidence for the perceptual individual/non-individual divide being syntactically relevant information that should be stored in the lexicon – a key difference with English.
3.7.3 The denotation of nouns

The lack of clear acquisition cues regarding the count/mass distinction means that nouns are not stored in the lexicon as count or mass. It is not the case, however, that all nouns can refer to individuals or non-individuals. In Mandarin Chinese, there is a binary distinction between nouns which can refer to both, and nouns that can only refer to non-individuals. For example, *pingguo* ‘apple’ can receive a count interpretation and a mass interpretation, whilst *tang* ‘soup’ can only receive a mass interpretation. This presents a different divide amongst nouns as is present in English. Rather than a count/mass distinction, there is a flexible/mass distinction.

This suggests that, in Mandarin Chinese, nouns are listed in the lexicon with one of two types of denotations: one potentially mirroring the denotation given to English flexible nouns or one mirroring the denotation given to English mass nouns. The same arguments presented for the late acquisition of flexible nouns in English and Dutch can be presented here in support of the late acquisition of the flexible/mass distinction in Mandarin Chinese – albeit this is the only distinction made on nouns in this language. The lack of consistent morphosyntactic evidence for the distinction makes the denotation of nouns difficult to specify at an early stage of acquisition.

In Mandarin Chinese then, there are two types of nouns with different lexical denotations. We have the mass denotation, as represented in (47), and some type of flexible denotation.
Nouns that can only refer to non-individuals have the denotation represented in (47). As all the atoms are unstable, it is not possible to select any specific atom. These atoms cannot be counted and must be assigned a measure.

One proposal for the denotation of flexible nouns in Mandarin Chinese is the same as that put forward for flexible nouns in English, as in (48).
Our treatment of flexible nouns in English needs now to be made more secure. There are two options available: (i) given the lack of evidence for the count/mass status of a noun, the noun retains its inherent flexible denotation. The count or mass feature, once grammaticalised, selects a section of the inherent denotation and discards the rest (as further functions do), or (ii) given that there is evidence for both a count and a mass status of a particular noun, individual count and mass denotations are both created and combined to give the flexible denotation.

The same dilemma is facing the denotation of all nouns in Mandarin Chinese. Are all nouns inherently flexible, and the mass interpretation selected, or are all nouns inherently mass, and a stabilising process takes place once the evidence provided by classifiers is decoded?

These options make different assumptions about the types of errors that might be
expected during the acquisition process, meaning we can consider the experimental evidence to make a judgment. If all nouns are inherently mass and a stabilising process takes place, then we can expect that early acquirers of Mandarin Chinese cannot refer to individuals. A difficulty arises here, however, given the late acquisition of the classifier system, the only reflex of the count/mass distinction, when bare nouns can refer to individuals or non-individuals. Li et al. (2010) demonstrate that it is not until 4 years of age that children start to link classifiers to individuation. Until that point, the picture is confused, as children are working out the other functions of the classifier. An experiment on two- and three-year-old children found that knowledge of classifiers as signalling a denotation of individuals relies on knowing some number words and perhaps also having acquired the general classifier ge, which has no additional semantic content other than functioning as an “individuator”. These other elements might also contribute as early evidence for individuation. What can be taken from this is, though, that some children of 2 or 3 years of age have acquired the general classifier. The general classifier relies on a denotation which included individuals, in some way. There is no other evidence to suggest that acquirers of Mandarin Chinese cannot refer to individuals early in acquisition, but the finding that the general classifier is acquired at an early stage points to a noun’s denotation including individuals from an early stage also.

Based on this, the preferred option for the denotation of flexible nouns is that all nouns are inherently flexible and some nouns are at a later stage lexically-specified as mass. This means that they can no longer be used with classifiers or, more technically, their denotations no longer allow the function applied by a classifier to apply.

Furthermore, I would tentatively suggest that the default denotation for a noun at the start of acquisition is that of a flexible noun, as represented in (48). If there is no systematic evidence for a divide, be it a full count/mass distinction as in English
or a mass specification as in Mandarin Chinese, a language will, by default, provide
a flexible denotation to nouns. The mixture of stable and unstable atoms means
that the noun can refer to either individuals or non-individuals.

This analysis is preferable to an entirely syntactic approach, whereby all nouns of all
languages are flexible and the features [COUNT] and [MASS] are read directly from the
syntax for a number of reasons. Firstly, I have discussed the available “mismatches”
and the various operations of packaging and grinding, e.g. *Could we get two wines
over here please?*, in which a mass noun is “packaged”, and *There’s tree all over the
road*, in which a count noun is “ground”. Whilst these are permissible, the intuition
that these uses are “exceptional” relies on the fact that they go against the standard
use of the noun. I argue that this understanding is due to the lexical denotation
of the noun being probabilistically determined as count or mass. If all nouns were
flexible, there would be nothing exceptional about these uses and there would be no
sense of “coercion”.

Furthermore, in a syntactic analysis where count/mass features operate over nouns
with flexible denotations, how do we ensure that the system observed in Mandarin
Chinese is permitted? Packaging through the use of classifiers is not found in Man-
darin Chinese – it is impossible for a noun with a mass denotation to have a syntactic
feature like [COUNT] applied by a classifier to become atomic. There needs to be
a lexical specification of count/mass to account for the crosslinguistic variation in
both coercion and lack of coercion.

### 3.7.4 Building the functional structure

As we have discussed, a key difference between English and Mandarin Chinese is
that the count/mass distinction is acquired later in the latter. There are no clear
clues in Mandarin Chinese that can shape the acquirer’s understanding of novel
words in terms of their count/mass status.

It is also the case that the functional structure of the nominal phrase begins to be built around the noun before the acquirer fully understands how classifiers relate to the denotation of the noun in terms of individuation. The denotations of the nouns are set up as full flexible nouns, including individuals and non-individuals. Links between classifiers and their nouns are set up, however, based on properties such as shape, animacy, function, etc. and classifiers first find their place in the functional structure based on these connections. Learning how classifiers categorise nouns based on these properties initially develops separately to any notion that classifiers provide individuation information, and, at 3 years of age, children perform better at identifying referents based on shape properties than on whether they are individuals (Li et al. 2010).

At some stage during the acquisition of classifiers, the child will find a set of nouns that cannot refer to individuals. This may be made apparent by the incompatibility of these nouns with the general classifier which, according to Erbaugh (1986), is used more prevalently in early production as children tend to be more conservative whilst they learn more specific classifiers. This is not a particularly strong correlation for the acquirer to pick up on, hence why it is acquired late, but it is systematic enough for the denotation of these nouns to be altered in the lexicon via the grammaticalisation of a mass feature. This sets up the flexible/mass distinction that exists in Mandarin Chinese.

To illustrate, take the noun pingguo ‘apple’. This is a flexible noun which will have a lexically-specified denotation of individuals and non-individuals. When the bare noun is used in isolation, it is not apparent whether an individual apple or some quantity of an apple substance is being referred to. Imagine that, at this stage of constructing your lexicon, you only know of four individuals which can be labelled with the noun pingguo: a, b, c, and d. Introducing the noun pingguo to the syntax
provides a denotation including these four individuals, the sets of individuals created by their union, and the unstable atoms created by the combining or dividing of these individuals. The denotation is given in (49):

(49) pingguo:

Functions can now apply to this denotation to identify the intended referent. The acquirer has, at this point in acquisition, registered the correlation between classifiers and restricting the flexible denotation to include only individuals. Given that classifiers are required for the identification and counting of individuals, the link is deemed syntactically relevant. It is likely that the classifier has already been granted its own syntactic head, based on its other categorising function, but it now gains a new function to grammaticalise the connection between classifiers and individuals, which we will call Individuation.
The Individuation function requires as its input a denotation that contains individuals and returns a denotation that consists of only individuals and sets of individuals. The Individuation function therefore selects only the stable atoms and the sets of stable atoms from the denotation. Applying the Individuation function to the denotation of *pingguo*, via the addition of the semantically appropriate classifier *ke*, returns the denotation in (50).

(50) ke pingguo:

The denotation now mirrors the lexically-specified denotation of a count noun in English, in that it contains only stable atoms and sets of stable atoms. The next function that would apply in English would be the Number function, but there is no evidence for this function – or at least not exactly the same function – in Mandarin Chinese. There is no overt evidence for a function that will select only the atoms or only the sets of atoms. The lack of number marking on nouns in Mandarin Chinese also means that there are no interdependencies between counting with numerals and number marking.

Therefore, the output of the Individuation function can directly form the input for the Cardinality function, which also exists in English. The Cardinality function can only apply to a denotation consisting of individuals, and is applied by the Numeral head. It selects those entities which contain the number of individuals represented by the numeral hosted in the head. The function works in exactly the same way as
Applying the Cardinality function of the numeral liang ‘two’ to the denotation of ke pingguo ‘apple’ in (50) selects only the sets of individuals with a cardinality of two. This leaves the denotation in (51), representing the nominal phrase liang ke pingguo ‘two apples’.

(51) liang ke pingguo: \( ab \quad ac \quad ad \quad bc \quad bd \quad cd \)

The referent of the nominal phrase liang ke pingguo must therefore be one of the six sets of individuals remaining in the denotation.

A mass noun in Mandarin Chinese works in much the same way as a mass noun in English, with the same denotation. The denotation for the noun tang ‘soup’ is given in (52)

(52) tang:

Given that the denotation consists only of vague atoms, the Individuation function cannot apply. The Individuation function relies on the ability to select stable atoms individually. This accounts for the inability of classifiers to combine with non-
individual-denoting nouns.

The following structures of the lower noun phrase in Mandarin Chinese can now be constructed. (53) shows how the nominal phrase for a counted individuated noun is built. The noun is not marked for a feature, as it has provided the full inherent denotation to the structure. Each functional head is split into the content it hosts and the function it applies.

(53) Counted individuated noun: liang ke pingguo

Similar to the structure for English, ClassP is taken as a level in the extended projection of the noun (when projected). In contrast to English, however, NumeralP is integrated with ClassP. In English, recall that NumeralP was separated from NumP to allow for the position of adjectives below numerals. In Mandarin Chinese, as we will discuss in further detail later, the standard position for adjectives is below classifiers, i.e. Numeral > Classifier > Adjective > Noun. The evidence is, in fact, that classifiers are almost always attached to numerals (Aikhenvald 2000) and that there are no languages with Num > N > Class or Class > N > Num (Hall 2015).

In (54), the structure of a mass noun is given. The mass feature is marked on the
noun, as this creates the lexically-specified denotation that the noun provides to the structure. As neither the Individuation function or the Cardinality function can directly apply to this denotation, the structure of such a nominal phrase is limited.

(54) Mass noun: *tang*

\[
\text{NP} \\
\text{N} \\
\text{tang}_{\text{mass}}
\]

### 3.7.5 Mismatches in Mandarin Chinese

Mismatches in terms of the count/mass distinction are not observed in Mandarin Chinese. This can be explained by a number of reasons.

Primarily, the morphosyntactic signposts for the count/mass distinction are much more difficult for an acquirer to notice. Unlike in English, where the speed of acquisition means that the syntax rather than the semantics informs whether a novel noun should be listed as count or mass in the lexicon, it takes Mandarin Chinese acquirers longer to link classifiers to the count/mass distinction. At this later point, the perceptual divide between individuals and non-individuals is better understood and is the primary driver of the grammatical distinction – the system is built slowly enough as to not create mismatches between the syntactic context and the perceptual divide.

Also, given that there is morphological marking on nouns at all, the inherent denotation of a bare noun survives much longer. This, in effect, reduces the opportunity for mismatches. All nouns either denote individuals and non-individuals, or non-individuals only. As classifiers only combine with noun denotations which
include individuals, the only possible mismatch that could arise is between a classifier and a non-individual denoting noun. When we recall that classifiers, as well as individuating, also provide information regarding the nouns they combine with in terms of shape, animacy, function, etc., it would be very difficult for an acquirer to create a systematic mismatch between a classifier and a noun that only denotes non-individuals. Children must learn the appropriate classifier for each noun (or type of noun), so there would be no classifier available for such a noun other than the general classifier.

3.8 Conclusion

Returning to the initial questions, I set out to examine two languages, English and Mandarin Chinese, to investigate (i) whether the count/mass distinction exists in both languages, (ii) how the count/mass distinction is encoded in these languages, and (iii) the role of linguistic input during the acquisition of the count/mass distinction.

We found that both English and Mandarin Chinese make distinctions that in different ways reflect the individual/non-individual divide. Nouns in English are divided into count nouns and mass nouns, with this difference being encoded in the lexical specification of nouns. The count/mass distinction is acquired early, based on the prominence of count syntax within the grammar. Once the count/mass distinction is grammaticalised, it is the signposts of count syntax that determine whether a novel noun is listed as count or mass. This moves the count/mass distinction away from the perceptual divide between individuals and non-individuals, and creates mismatches within the system.

In Mandarin Chinese, the flexible/mass distinction is acquired much later, as there
are no prominent systematic morphosyntactic signposts for the distinction available to the acquirer. As a result of this, and a more general lack of morphological marking, nouns maintain an inherent flexible denotation much later, until there is evidence for a lexically-specified denotation of a mass noun. Based on the use of flexible nouns, and that classifier-noun associations must be learned, mismatches between the grammatical and perceptual distinctions are not possible.

In terms of the parallels between count syntax and classifiers, both require a denotation consisting of individuals. In English, the majority of nouns have denotations already consisting of either individuals or non-individuals, so the function applied by number marking, Number, ensures that mass nouns are not marked for number. In Mandarin Chinese, no noun has a denotation of only individuals, so the function created by the count/mass distinction on classifiers, Individuation, actively selects individuals from the inherent flexible denotation. In this way, the Individuation function can be said to be a true reflex of the count/mass distinction, as it uniquely returns a denotation of a count noun that would otherwise be unavailable in Mandarin Chinese. The Number function in English, however, is only a near parallel of the individual/non-individual divide, which signposts the acquirer to early acquisition of the count/mass distinction but does not provide a one-to-one mapping. Earlier acquisition prompted by prominent morphosyntactic clues provided by a different, but related, perceptual divide results in a less accurate mapping between the lexical information of the noun and its semantic interpretation.

There are two sets of true parallels between the languages, in terms of the denotations that the features/functions return, although these come about in different ways. The COUNT feature that is lexically-specified on nouns in English applies the same operation as the Classifier function in Mandarin Chinese. Both take a flexible denotation and return only the stable atoms and sets of stable atoms. The MASS feature that is lexically-specified on nouns in both languages applies the same op-
eration, selecting the unstable atoms from the flexible denotation. In English, this is set up early as part of a count/mass distinction. In Mandarin Chinese, this is grammaticalised later once classifiers have been acquired. It is clear how the same perceptual divide is encoded in the two grammars in different ways, based on a range of factors. I argue that the acquisition of the features shapes and maintains how the grammars individually treat this perceptual divide.

In terms of cross-compatibility, there is the question of whether it would be possible to have a language with an English-style count/mass distinction and Mandarin Chinese-style classifiers, or a language with a Mandarin Chinese-style flexible/mass distinction and no classifiers. Theoretically both would be possible, but it would depend on how these distinctions are marked. For the former to be possible, there would need to be independent marking of individuals vs. non-individuals – this could be linked to number or not – and then classifiers could operate over the marked count nouns. The presence of the individuation function would, however, be called into question, as the denotation would already contain only stable atoms. Classifiers may exist only for their other semantic classification functions. For the latter language to exist, there would have to be some marking on nouns whose denotation consists of atoms which cannot be stabilised. This is not outside the realms of possibility. Similarly, for such a language to have no classifiers of the type in Mandarin Chinese is not impossible, but there would likely need to be some other operator/function/process by which the atoms in the denotation of flexible nouns can be individuated, at least if quantification is to be possible.
Chapter 4

Unit Words in Mandarin

Chinese

4.1 Introduction

Across languages, there is a strong link between the count/mass distinction, or individuation, and the ways in which it is possible to count and/or measure entities. In English, count nouns, which denote individuals, combine directly with a numeral and receive plural marking, as in (55). A mass noun cannot combine directly with a numeral without forcing a count reading\(^1\). The noun must appear with a measure word which allows counting, as in (56).

(55) Count:
\[ \text{two books} \]

\(^1\)Combining a mass noun directly with a numeral in English coerces a count reading of the noun, e.g. \textit{two waters} or \textit{two beers}. This was discussed in Chapter 3
Based on Carlson’s (1977) division of nouns into predicates, which denote a set of entities, and arguments, which denote kinds, Chierchia (1998b) suggested a typology of languages with regard to the count/mass distinction. An N that is a predicate must combine with a determiner to become an argument, whereas an N that is an argument is available to fill an argument position immediately. Assuming that all languages must have nouns that are at least one of these two types (predicate or argument), Chierchia proposed a three-way typology of languages: [+predicate, +argument], [–predicate, +argument], and [+predicate, –argument]. In English, bare singular count nouns are taken to be ungrammatical and therefore must be of type [+predicate], whereas bare mass nouns can occur in argument position and therefore must be of type [+argument]. This makes English a [+predicate, +argument] language.

In Mandarin Chinese, there is no such simple division between count and mass nouns. Consider (57), adapted from Zhang (2014:2):

(57)  

The watermelon here could be a count noun, in either the singular or the plural, or
it could be a mass noun with a partitive or collective measure, also in the singular or the plural. The sentence in (57) suggests that bare nouns in Chinese are closer in interpretation to English mass nouns than English count nouns, as the sentence says no more than “there is some amount of watermelon on the table”. As nouns in Mandarin act like mass nouns in English, Chierchia (1998b) claims Mandarin as a [-predicate, +argument] language – all nouns can occur in argument position.

When counting entities in Chinese, nouns behave more like English mass nouns in that they cannot combine directly with a numeral. A classifier is required, and a numeral takes a classifier+N as its complement rather than a bare noun, as in (58) and (59):

(58) \(\text{liang ben shu}\)
    two    \text{CL}    book
    ‘two books’

(59) \(\text{liang bei shui}\)
    two    \text{CL}    water
    ‘two glasses of water’

Given that all nouns are [+argument], they must denote a mass or a kind. This explains the lack of morphological number marking in languages like Chinese. Bare nouns in Mandarin have a number-neutral, or general number, interpretation which includes the plural – the flexible denotation referred to in Chapter 3.²

The main stumbling block for the proposals of Chierchia (1998b) is that there are a large number of languages that do not neatly fit into one of the three types. In Brazilian Portuguese, for example, bare singular count nouns have been shown to

²For completeness, an example of Chierchia’s (1998b) third type of language, [+predicate, –argument], is French. Every noun is a predicate, meaning none can appear by itself as a bare argument. However, the count/mass distinction may still exist in the same way as for type 1 languages, like English. There are no [-predicate, –argument] languages, as all nouns need to be interpreted.
exist by Schmitt & Munn (2002) and then claimed to actually be mass nouns by Pires de Oliveira & Rothstein (2011). This suggests that a single noun could be both a predicate and an argument.

Another problem is the variation in number marking crosslinguistically. Turkish, for example, has a morphological marker of plural, the suffix –lar, but bare nouns already express general number, meaning that their denotation includes both atomic entities and pluralities:

(60) \( \text{çocuk} \)
\( \text{boy} \)
\( \text{‘boy’ or ‘boys’} \)

(61) \( \text{çocuk-lar} \)
\( \text{boy-PL} \)
\( \text{‘boys’} \)

Interestingly, a noun being counted through combination with a numeral is not marked as plural:

(62) \( \text{iki çocuk} \)
\( \text{two boy} \)
\( \text{‘two boys’} \)

The grammatical number system in Turkish has a binary distinction between general number and plural number. Corbett (2000) notes several languages that, like Turkish, have a general vs. plural number contrast, in which the absence of plural marking does not result in a singular interpretation. Bare nouns typically represent a categorial meaning rather than representing a member of that category. For example, \( \text{insan} \) refers to the category \( \text{human} \), \( \text{bir insan} \) refers to a single member of the category \( \text{human} \), and \( \text{insanlar} \) indicates an indefinite number of members of the
category *human*. The issue for the analysis of Chierchia (1998b) here is that a term supposedly giving a kind denotation *can* be marked as plural.

Chierchia (2010) puts forward an alternative analysis for three types of languages divided instead according to where and how the count/mass distinction is made. Chierchia proposes classifier languages as one type, in which the count/mass distinction shows up at the classifier level, number marking languages, in which the distinction is made at the noun level and morphologically marked, and a third type, in which the distinction is made at the noun level without morphological marking. Mandarin Chinese is placed in the first type and English in the second. This less rigid classification of languages is an improvement, but still does not account for the number of languages in which there is both classifiers and number marking, such as Hungarian. Csirmaz & Dékány (2014) show that, in Hungarian, numerals can be followed either by a classifier+N or by a bare, naturally atomic noun, e.g.:

(63) \[ \textit{két (fej) hagyma} \]
\[ \text{two (Cl\textsubscript{head}) onion.ATOMIC} \]
\[ \text{‘two onions’} \]

(64) \[ \text{a. \#három kosz} \]
\[ \text{three dirt} \]
\[ \text{b. három darab szemét} \]
\[ \text{three Cl\textsubscript{general} dirt} \]
\[ \text{‘three pieces of dirt’} \]

There is also a morphological singular/plural distinction as expected in count/mass languages, e.g.:

(65) \[ \text{a. könyv} \]
\[ \text{book.sg} \]
\[ \text{‘book’} \]
This suggests that there needs to be an allowance for languages which have a count/mass distinction but also allow for flexible nouns.

### 4.2 Mandarin Chinese

The classifier element that appears between numerals and nouns in Chinese has been categorised in a number of different ways. Chao (1968) proposes five types: individual measure, group measure, partitive measure, container measure, and standard measure. Tai & Wang (1990) suggest two major groups: classifiers and measure words. In this paper, I will follow Tai & Wang (1990) and use the terms “classifier” and “measure word”. Classifiers identify the unit in which an entity denoted by a noun naturally appears. Measure words, on the other hand, provide a unit of measure for the individual or substance denoted by the noun.

Bare nouns freely occur in Mandarin Chinese, and, as is shown for the bare noun *xigua* in (57) above, this can provide a denotation that is count or mass, singular or plural. Relating this back to Chierchia’s (2010) three language groups, there seems to be no count/mass distinction made on nouns in Mandarin. This has led many linguists to propose that there is no count/mass distinction at all in Mandarin Chinese.

In the previous chapter, I argued that there is a grammatical reflex of the individual/non-individual divide and agreed with many before that the distinction is made by the classifier system rather than the nouns themselves. Approaches of this type tend to follow the two-way distinction I outline above between classifiers and measure...
words (although a variety of terms are used). Cheng & Sybesma (1999) proposed that classifiers select for count nouns and “massifiers” select for mass nouns. This would mean that classifiers and massifiers are taking advantage of some lexically-encoded distinction between count nouns and mass nouns that had previously been argued not to exist in Mandarin.

4.2.1 The feature specification of classifiers/measure words

The distinction between classifiers and measure words is an important one to define. Li (2013) makes this division based on the function of the classifier/measure word within the interpretation of the larger nominal, that of either counting or measuring entities. Two binary features, [+Count] and [+Measure], are proposed, which create the following four types:

(66) a. [+Count, –Measure]
counting operators by default, e.g. classifiers

b. [–Count, +Measure]
measuring operators by default, e.g. measure words
c. [+Count, +Measure]
count or measure, e.g. measure words for containers, groups, and partitions
d. [–Count, –Measure]
refer to kinds

By having two features rather than one, it is possible to have four logical possibilities, recognising that specific words are capable of fulfilling the role of a classifier and a measure word. These words are specified as [+Count, +Measure]. Li (2013) notes that ping ‘bottle’ can refer to either a concrete countable bottle or an ab-
stractic measure unit equal to the quantity a bottle holds, supported by the following sentences (Li 2013:135):

\[(67)\]

\begin{enumerate}
  \item \textit{wo ling le liang ping jiu, zuo shou yi ping, you shou yi ping}\
  \hspace{1cm} I lift PFV two CL wine left hand one CL right hand one CL\
  \hspace{1cm} ‘I carried two bottles of wine, one in the left hand, one in the right.’
  \item \textit{ta-de jiuliang shi liang ping hongjiu}\
  \hspace{1cm} his drinking-capacity be two CL red-wine\
  \hspace{1cm} ‘His drinking-capacity is two bottles of red wine.’
\end{enumerate}

In the first sentence, the bottles must be concrete and countable, while in the second, there is an abstract measure reading with which there is no requirement for real bottles to be present – assuming a bottle of wine is 750ml, we can understand that his drinking capacity is 1500ml of wine regardless of the container.

Li & Rothstein (2012) provide examples to show that the [+Count, +Measure] group fully align with pure classifiers ([+Count, –Measure]) when a count interpretation is intended, and with pure measure words ([–Count, +Measure]) when a measure interpretation is intended. They make the observation that numerals can be dropped before classifiers, but not before measure words. When a [+Count, +Measure] element is used, dropping the numeral forces a count reading, or the numeral cannot be dropped with a measure reading.

There are also subtle differences when considering the additive modifier \textit{duo} ‘more’, which can appear in two contexts (where \(x\) and \(y\) are the positions for the classifier/measure word):

\[(68)\]

\begin{enumerate}
  \item \textit{Num + duo + x + N}
  \item \textit{Num + y + duo + N}
\end{enumerate}
Either a classifier or a measure word can appear in position $x$, allowing either a count or a measure reading. Only a measure word can appear in position $y$, allowing only a measure reading. This also holds for the [+Count, +Measure] elements – appearing in position $x$ will result in the phrase being ambiguous between count and measure, but appearing in position $y$ will force a measure reading.

Count interpretations are available with [+Count, –Measure] and [+Count, +Measure] elements, and measure interpretations are available with [+Count, +Measure] and [–Count, +Measure] elements. While these two binary features are useful for capturing the observation that some words can be either count or measure elements, they cannot capture the distinction between interpretations as observed when dropping the numeral or with the placement of duo. There is a class of classifiers that can only give a count reading, there is a class of measure words that can only give a measure reading, and there is a class of words that can give either, potentially dependent on their position or other surrounding elements. This would suggest that it is not only the classifier/measure word on its own that determines the interpretation, but that there are other factors at work. After all, in (68), either the position of duo or the position of the classifier/measure word changes between the two examples. The position change of either element is responsible for the loss of the count reading, not the featural specification of the classifier/measure word.

Zhang (2012) provides an analysis of the count-mass distinction such that it depends on two binary features of the noun: [±Numerable] and [±Delimitable]. The first feature determines whether the noun can combine directly with a numeral. A noun that is [+Numerable] can combine with a numeral and thus represents a count noun in English. A noun that is [–Numerable] is non-count. The second feature relates to whether the noun can combine with a shape, size or boundary modifier, i.e. a delimitive adjective. A noun that is [+Delimitable] can combine with such a modifier and thus is a non-mass noun.
Expanding the count-mass distinction to two binary features allows an extra dimension of variation – it is not the case that a noun that is not count is necessarily mass, and vice versa. This is exemplified in the following table:

(69)

<table>
<thead>
<tr>
<th>[±Numerable]</th>
<th>[±Delimitable]</th>
<th>Countability status</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
<td>Count with a delimitable feature</td>
<td>apple</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>Count without a delimitable feature</td>
<td>belief</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>Non-count, non-mass</td>
<td>furniture</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>Mass</td>
<td>oil</td>
</tr>
</tbody>
</table>

We can see here that there are mass nouns and two types of count nouns, but also a class of non-count, non-mass nouns, words such as furniture in English, that can combine directly with a delimitive adjective but not a numeral.

Zhang (2012) claims that no nouns in Chinese are count nouns, but, under the two feature system proposed, this does not entail that all nouns are mass nouns. All nouns are [–Numerable] – they cannot combine directly with a numeral – but there is a distinction between [+Delimitable] non-mass nouns and [–Delimitable] mass nouns.

In Mandarin Chinese, all nouns are [–Numerable], explaining why they cannot combine directly with a numeral. Classifiers and measure words provide the [±Numerable] feature. The features are defined based on the combinatorial capability of the element, i.e. [+Numerable] nouns can directly combine with a numeral, [+Delimitable] nouns can directly combine with a delimitive adjective. This moves away from a preference within Minimalism to semantically motivate the postulation of
features. Also, in terms of feature acquisition, it seems unlikely that nouns in Chinese are marked \([-\text{Numerable}]\) given that there is no instance within the language when an acquirer would have associated “numerability” with a noun. In a feature system based on the availability of combinations, privative features might be preferable.

What is clear from the analyses is that the count-mass distinction is more complex than a single binary classification. The key difference between them is determining where the count-mass distinction exists: on the noun or at the level of the classifier. Regardless, it seems that the distinction is not made in the lexical specification of classifiers themselves. Some linguists have also argued for a difference in the structural configurations in which classifiers and measure words occur.

### 4.2.2 Structural configurations of classifiers/measure words

Zhang (2012) argues that expressions with measure words have a different syntactic structure from those with a classifier. The former have a left-branching structure \([[\text{Numeral-Cl}] \ N]\) as in (70), whilst the latter have a right-branching structure \([\text{Numeral} \ [\text{Cl-N}]\) as in (71).
Numerability, whether a noun can combine directly with a numeral, and delimitability, whether a noun can combine directly with a shape, size, or boundary modifier, are represented by two projections: UnitP and DelP respectively. Classifiers and measure words are both Unit heads. Numerals are base-generated in the specifier position of UnitP and the occurrence of either a classifier, a measure word, or a
numeral requires the projection of UnitP – this closely links the Unit head to the numeral, such that a numeral cannot occur without a Unit head, but also allows structures in which there is no Unit projection. Delitive adjectives appear in the specifier of DelP.

Plural markers, as far as they exist in forms such as reduplicate unit words (RUWs), are hosted in NumP. UnitP and NumP are different functional projections, so there is no correlation between the projection of UnitP and the occurrence of a plural marker.

Li (2013) also proposes different syntactic structures depending on whether the reading is to be count or measure. This is most important for the [+Count, +Measure] words, such as those denoting a container, group, or partition, for which there is no apparent featural difference to identify the intended reading. For the counting reading, the structure in (72) is claimed:

(72) NumeralP
    /   
   Numeral CIP
     /   
    Cl NP
   Classifier

Here, the projection CIP is headed by the classifier, and the CIP can itself be a complement of a higher functional projection, NumeralP. For the measure reading, the structure in (73) is proposed:
Here, the numeral and the measure word form necessary parts of a complex modifier, “Numeral-Cl”, which is adjoined to NP – neither internal element can be dropped.

Proposing a system in which the same word can be either a classifier or a measure word makes it necessary for Li (2013) to claim that there is a difference in syntactic structure between a count reading and a measure reading. For the count reading, the classifier and the numeral are included in the extended projection of the noun as functional heads, but for the measure reading, the measure word and numeral combine as a complex modifier that adjoins to the noun phrase. The system proposed, therefore, has (i) no distinction on the noun, (ii) four possible types of classifiers, and (iii) two different structures. There seems to be no explanation of how the system ensures that the [+Count] elements appear in only one structure and the [+Measure] elements appear in only the other. Also, if for the measure reading, (73), the Numeral+Cl complex modifier adjoins to NP, then it must be explained how it is then impossible to merge a classifier and numeral above that to give the count reading – effectively building both structures in one. Whilst this would never be motivated semantically, having one structure work through adjunction and the other through functional heads in the extended projection does not rule this out. It seems likely that there are in fact structural differences between the two readings, but also likely is that the elements involved are of the same “status”, i.e. both adjuncts or head-based, such that one somehow rules out the other. That the numeral
carries out exactly the same function in both structures further points to more uniformity in its structural position – forming a complex modifier in one and being a functional head in the other does not seem to accurately reflect the clear parallels in the function of the numeral across these two interpretations.

Zhang (2012) argues for a subtler difference in structure which seems preferable, but there is no exploration of why classifiers or measure words appear in certain structures. How would the differences between classifiers and measure words link to the [±Numerable] and [±Delimitable] features on the noun?

To find greater clarity on the structural configuration of classifiers and measure words, it is therefore important to further investigate the details of the different configurations available to each.

4.2.3 Pre-classifier and pre-measure word adjectives

Tang (1990) shows that certain adjectives can occur between the numeral and the classifier or measure word, as in (74) (Tang 1990:418):

(74) a. yi da zhang zhi
   one big CL paper
   ‘one large sheet of paper’

   b. na yi xiao xiang shu
   that one small MW book
   ‘that one small box of books’

Very few adjectives are allowed in this context; Cheng & Sybesma (1998) claim that the only productive ones are da ‘big’ and xiao ‘small’.

This stands in opposition to Li (2013), who claims that pre-classifier adjectives occur only in counting contexts. This is because this class of adjectives must adjoin in the
specifier position of Cl, as in (75):

\[
\text{(75)} \quad \text{NumeralP}
\quad \text{Numeral} \quad \text{ClP}
\quad \text{adjective} \quad \text{Cl'}
\quad \text{Cl} \quad \text{NP}
\]

In the measure structure, the complex modifier Numeral+Cl is not a full projection of Cl and thus the adjectives cannot occupy the specifier position.

Zhang (2012) shows that adjectives can occur before both classifiers and measure words, and rules out a movement analysis based on the following examples:

\[
\text{(76) a. san xiao di shui}
\text{three small MW water}
\text{‘three small drops of water’}
\]

\[
\text{b. *san di xiao shui}
\text{‘three MW small water}
\]

\[
\text{(77) a. san da ka fang xigua}
\text{three big CL square watermelon}
\text{‘three big square watermelons’}
\]

\[
\text{b. san xiao pian yuan shuye}
\text{three small CL round leaf}
\text{‘three small round leaves’}
\]

(76) shows that there are certain adjectives that can only occur in the higher position, since mass nouns cannot be directly modified by size delimitive adjectives. This goes against a movement analysis which would suggest the same adjective could
occur in either position. (77) shows how it is possible to have both a preceding and following adjective.

Her (2012) observes that the scope of the adjective depends on whether it occurs with a classifier or a measure word. A pre-classifier adjective modifies both the classifier and the noun, while a pre-measure word adjective modifies the measure word only:

(78)  a.  yi da xiang pingguo
       one big MW apple
            ‘one big box of apples’

       b.  yi xiang da pingguo
            one MW big apple
                ‘one box of big apples’

(79)  a.  yi da ke pingguo
       one big CL apple
            ‘one big apple’

       b.  yi ke da pingguo
            one CL big apple
                ‘one big apple’

Also, given that a pre-classifier adjective modifies “through” the classifier, it follows that a pre-classifier adjective and a pre-noun adjective in the same phrase cannot contradict each other. On the other hand, a pre-measure word adjective can occur with what would be a contradictory pre-noun adjective, as the first will modify the measure word only and the second will modify the noun only:

(80)   yi da xiang xiao pingguo
       one big MW small apple
            ‘one big box of small apples’
Adjectives can also be “stacked” before a measure word, suggesting either a recursion of projections or multiple specifier positions (Li 2000:57):

(82)  si  xiao  bo  pian  mianbao
four small thin MW bread
‘four small thin slices of bread’

(83)  si  da  chang  chuan  tang-hulu
four big long MW sugar-fruit
‘four big long strings of sugared-fruit’

4.2.4 Summary and implications of these approaches

Zhang (2012) proposes that nouns are not simply divided into those that are count and those that are mass, but that the noun is specified in terms of two binary features: [±Numerable], which determines whether the noun can combine directly with a numeral, and [±Delimitable], which determines whether the noun can combine directly with a delimitive adjective. Nouns in Chinese are [–Numerable] as they cannot combine directly with a numeral. Count nouns in English, however, are [+Numerable].

Li (2013), on the other hand, places the distinction on the classifier, which is specified with two binary features [±Count] and [±Measure]. The feature specification is based on the function of the classifier as a “counter” or a “measurer”.

In terms of structure, both Zhang (2012) and Li (2013) assume that classifiers and measure words form a single grammatical category: Unit for Zhang and Cl for Li.
This forces the proposal of a difference in syntactic structure between a count and a measure reading. Classifiers appear in a right-branching structure and measure words appear in a left-branching structure.

This leaves a number of interesting open questions:

(84) a. Where is the count-mass distinction made? Does the interpretation of the noun in terms of count or mass depend on the information provided by the lexicon or is it based on the elements it combines with in the syntax?

b. What is the category of the classifier and/or measure word? If all classifiers are of the same category, how are differences amongst them explained?

c. If different structural configurations lead to these differences, how do we ensure that the right types of classifier/measure words appear in the right structures?

These questions clearly relate directly to the grammar of Mandarin Chinese, but they also impact on the wider topics of number, countability, and the count-mass distinction cross-linguistically and within the theory.

4.3 Investigating numeral classifiers

With these questions in mind, I designed and carried out an investigation into the grammaticality of nominal phrases involving classifiers in Mandarin Chinese. I distributed an online questionnaire (provided in the appendix) to native speakers of Mandarin Chinese which included constructions taken from the following configuration options:
The Features and Structure of the lower Noun Phrase


Above, $Num$ stands for numeral, $Adj$ for adjective, $CL$ for classifier, and $MW$ for measure word. Elements in brackets are optional.

I focused on several binary parameters of variation:

(86)

<table>
<thead>
<tr>
<th>Either...</th>
<th>or ...</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Classifier</td>
<td>Measure word</td>
</tr>
<tr>
<td>b. Count noun</td>
<td>Mass noun</td>
</tr>
<tr>
<td>c. Adjective before CL/MW</td>
<td>Adjective before noun</td>
</tr>
<tr>
<td>d. Delimitive adjective</td>
<td>Other adjective</td>
</tr>
<tr>
<td>e. Adjective before or after CL/MW</td>
<td>Adjective before and after CL/MW</td>
</tr>
<tr>
<td>f. Unmodified adjective</td>
<td>Adjective modified by $hen$ “very”</td>
</tr>
</tbody>
</table>

Of the configurations presented to the participants, some were assumed to be grammatical for all, some were predicted to be ungrammatical for all, and the remainder were predicted to be variable in grammaticality. Participants were asked to rate each construction on a three-point scale: fully grammatical, fully ungrammatical, or somewhere in between. Participants were also reminded that they should make their judgements based on their personal intuition and understanding of the language as native speakers, to avoid any prescriptive bias. The phrases were either taken from the Lancaster Corpus of Mandarin Chinese or newly constructed.
4.3.1 Findings

A prevalent view of Mandarin Chinese is that there is a single grammatical category combining both classifiers and measure words. However, my data supported the view that there are, in fact, differences in the behaviour of classifiers and measure words.

Classifiers were found to combine only with nouns that can refer to entities that are discrete countable units, i.e. with a denotation that, at least in some part, is the same as the denotation of count nouns in English. Classifiers could not combine with mass nouns. Measure words, on the other hand, could combine freely with either mass or count nouns. It was equally possible to have two boxes of books, a discrete countable unit, as two boxes of sand, a substance with no natural divisions.

There were two positions for adjectives, before the CL/MW and before the noun. All adjectives were allowed before a measure word and before a count noun. Only delimitive adjectives were found to be grammatical before a classifier, whereas delimitive adjectives were found to be less acceptable before a mass noun than a count noun. Similarly, adjectives could be modified by hen, “very”, more readily in the contexts in which there were no restrictions on them, i.e. before a measure word and before a count noun.

It is ungrammatical for a pre-classifier adjective and a pre-noun adjective to have contradictory meanings when appearing in the same phrase, whereas this is acceptable with a pre-measure word adjective and a pre-noun adjective.

Interestingly, some participants noted that, in colloquial speech, it is not always necessary to have a classifier when counting entities. This only seemed possible with the general classifier ge in phrases like that in (87):
4.3.2 Theoretical implications

To start, and as outlined in the previous chapter, I am maintaining the flexible/mass distinction amongst nouns in Mandarin Chinese. Flexible nouns are those that may denote entities with natural divisions, as well as entities without, and mass nouns only denote substances without natural divisions. Classifiers combine only with flexible nouns – nouns that are lexically provided with some idea of a natural unit division. The classifier serves to “realise” this unit in the syntax, allowing it to be counted. This also explains why classifiers cannot readily appear with mass nouns.

As discussed in the previous chapter, there is the possibility of a classifier+mass noun combination to coerce a count reading, similar to adding plural morphology to mass nouns in English, i.e. *two beers*. This sets up two levels of the count/mass distinction, such that the lexically-supplied semantic notion of natural division can be overridden in the syntax, through the "stabilising" of unstable atoms. This coercion of a count reading from a lexical mass noun is quite common in English, but I found no evidence of a classifier being used with a mass noun in Mandarin Chinese. It may be the case that Mandarin Chinese disallows this stabilisation effect introduced by the functional structure and relies more insistently on the lexical count/mass distinction than English does. What this does put forward, quite strongly, is that there is clear evidence against the previous claims that there is no type of count/mass distinction in Mandarin Chinese. The count/mass distinction on nouns in Mandarin

(87)  
\[ \text{a. } liang \text{ ge ren} \]  
\[ \text{two } \text{CL people} \]  
\[ \text{‘two people’} \]  
\[ \text{b. } liang \text{ ren} \]  
\[ \text{two } \text{people} \]  
\[ \text{‘two people’} \]
Chinese is the strongest justification for the lack of this type of coercion.

Classifiers exist, primarily, for functional reasons – they do not contribute independent *semantic* content of their own to the phrase and stand to identify some relevant part of the meaning of the noun, be it for reasons of categorisation or individuation. For this reason, it is to be expected that they cannot be modified and that any pre-classifier adjectives in fact modify either the noun “through” the classifier, or the classifier + noun combination. The restricted class of adjectives that are found before the classifier makes it difficult to choose between these two options, but it is clear that a pre-classifier adjective cannot contradict a pre-noun adjective:

(88) *yi da ke xiao pingguo
     one big CL small apple

On the other hand, measure words make a clear contribution to the interpretation of the phrase, defining the amount of the entity referred to by the noun that will be counted. Measure words can combine directly with either mass or count nouns. The fact that the noun has natural divisions makes no different to the interpretation in Chinese, for example *liang xiang pingguo* “two boxes of apple(s)” can be interpreted as “two boxes of whole apples”, or “two boxes of pieces of apple”, or “two boxes of crushed apple”; without naming the natural division of the noun with a classifier, it is not realised in the syntax.

This strengthens the proposal of a flexible/mass distinction. This lexical flexible/mass distinction must exist in Chinese, as there must be a natural unit in the denotation of a flexible noun for the classifier to realise, and there must be no stable atoms in the denotation of the mass noun, else classifiers would be able to combine with them also. The stable atoms in the denotation of a flexible noun can, however, be ignored, as we see in the phrases containing a measure word with a count noun. As in the example above, *liang xiang pingguo* “two boxes of apple(s)”, measure
words will operate over the full denotation of the flexible noun if one is selected, and both count and mass interpretations remain.

The full contribution made by a measure word to the semantics means that it can be modified by any adjective, similar to any noun. It is possible to have contradictory adjectives before the measure word and before the noun, as one is modifying the interpretation of the measure word and the other the interpretation of the main noun.

On the possibility of dropping classifiers, this would make sense as a case for grammatical change in light of our definition of classifiers above – that classifiers name the natural division of the entity denoted by the noun. If the noun lexically provides a sense of the unit in which the entity it denotes naturally comes – if the denotation of the noun inherently includes stable atoms – then can the semantic motivation for the classifier truly be maintained, over and above the function of “realising” the atom for the syntax?

Her (2012) relates the optionality of classifiers to their “transparency” in terms of the scope of preceding elements. As discussed, pre-classifier adjectives can modify the noun “through” the classifier, whereas a pre-measure word adjective cannot modify any elements after the measure word. The same can be argued for the scope of numerals, classifiers can be omitted and the numeral can quantify the noun with no change in meaning, whereas a missing measure word will lead to a different interpretation of the phrase. For example (Her 2012:4):

(89)  a.  \textit{wu} (\textit{zhang}) \textit{bing er} (\textit{tiao}) \textit{yu weibao wuqian} (\textit{ge}) \textit{ren}
      five (CL) loaf two (CL) fish feed-full five-thousand (CL) person
      ‘5 loaves and 2 fish fed 5000 people.’

      b.  \textit{wu} *(\textit{lan}) \textit{bing er} *(\textit{xiang}) \textit{yu weibao wuqian} *(\textit{zu})
      five *(MW) loaf two *(MW) fish feed-full five-thousand *(MW)
The lack of semantic contribution from the classifier makes it a prime candidate for omission, and explains differences in the scope of adjectives. Measure words have greater value for the interpretation of the phrase and therefore cannot be dropped.

Based on these differences between classifiers and measure words, I believe there is motivation for different syntactic structures. The structure of a phrase with a classifier is provided in (90) and the structure of a phrase with a measure word is provided in (91).

This is the structure that I am proposing for phrases with classifiers. Starting from the bottom, the noun merges, then an optional adjective might merge – all elements in brackets are optional. The lower position is the standard position for adjectives.

Then comes the classifier, which, as previously proposed, has its own functional head, Class, with its own functional projection. The functional head applies the
Individuation function which stands to “name” the natural, discrete countable unit denoted by the noun or, more technically, selects from the flexible denotation of the noun only the stable atoms and sets of stable atoms.

Within the ClassP is the higher position for adjectives that hosts the restricted set of delimitive adjectives. In this structure, the higher adjectival position is adjoined to a projection of the functional Class head, accounting for the limitation on the range of adjectives allowed in this position.

The numeral merges in the specifier position of ClassP, mirroring the previous analyses in their structural marking of the dependence between classifiers and numerals.

Perhaps most crucially, the classifier is categorised as a functional head in the extended projection of the noun – its purpose is to fulfil a functional role as part of the referent-identification task of the nominal phrase. Its functional status explains why the classifier is “invisible” in terms of the scope of the adjective, even though this adjective is merged as part of the ClassP.\footnote{This formulation links to the idea of semi-lexical heads as proposed by van Riemsdijk (1998) and Vos (1999). In direct partitive constructions, measure nouns taking “main” nouns show the behaviour of single projections rather than dual projections. In their proposal, quantifier nouns are closed class items analysed as functional heads, whereas other types of nouns observed in the first position in these constructions are analysed as semi-lexical heads. I argue that classifiers in Mandarin Chinese are functional heads, and are therefore in the same extended projection of the noun. Measure words could be acting as semi-lexical heads in some constructions, but I would argue for a consistent dual projection approach given that quantification is possible between the higher measure word and a lower classifier. As van Riemsdijk (1998: p.15) claims: “if the lower noun were the head of an independent projection, we would expect it to be accompanied by all the elements that characterise a DP” – “functional heads of the D/Q type”. Numerals are possible in this lower position, suggesting two independent projections.}
This is the full structure I propose for phrases with measure words, including all the elements related to individuation. Based on the characteristics of measure words that I have observed, I am taking measure words to be more closely linked to lexical nouns – sharing many but not all behavioural properties and perhaps having started down a path to grammaticalisation. The key difference between measure words and lexical nouns, which aligns them more with classifiers, is that they can directly combine with a numeral to be counted. This places the measure word in a grey area between lexical and functional justification. A measure word can still act as a noun, and provide its own denotation to the structure-building process – it is still possible to talk about “boxes”, for example, without needing to talk about what
they contain. This is an important distinction to make, especially given that a large number of nouns can be employed as measure words.

The “measure” function therefore needs to be discussed in more detail. This function, as well as being able to apply to the denotations of both flexible and mass nouns, also needs to legitimise the availability of phrases such as that in (92), in which the measure function applies to a fully-formed numeral+classifier+noun combination.

(92) san xiang si ge pingguo
three MW four CL apple
‘three boxes of/for four apples’

What the Measure function must accomplish is a different task to that of the Individuation function, which simply selects the stable atoms from the denotation of a flexible noun. It must be able to apply to the denotation of flexible nouns, mass nouns, and what are the equivalent of count(ed) nouns, whilst also being able to operate as a full lexical noun itself. My proposal now needs to account for elements that are somewhere between lexical and functional.

### 4.3.2.1 Lexical projections in the nominal phrase

Before considering how to treat elements that are both functional and lexical, it would be sensible to consider in more detail how purely lexical items operate in this semantically-driven system. We have discussed how the base noun in any nominal projection contributes its lexically-specified denotation to the structure-building process as its starting point, but what about higher lexical projections such as adjective modifiers? The key difference between purely functional and purely lexical projections is that functions apply to the existing denotation only, whereas lexi-
cal projections contribute their own denotation with which the existing denotation intersects.

To illustrate, we will briefly return to the example of number marking in English. A noun is supplied from the lexicon with its lexically-specified denotation. In this example, the noun is *house*. This has been acquired as a count noun, and so has a denotation that includes only stable atoms and sets of stable atoms. The function Number\textsubscript{PL} applies through the merge of the functional head Num. This restricts the denotation to sets of individuals and contributes the suffix -s to the phonological form. This function clearly only operates on the existing denotation.

Next, adjectival modification takes place – the adjective *blue* merges. Note that this modification takes place through adjunction. The adjective, as a lexical item, contributes its own denotation. Put simply for this example, the set of things which are blue. This contribution presents a different process for the denotation of the noun and acts as a type of filter. Only those atoms in the denotation of the noun that are in the intersection with the denotation of the adjective will be selected. This is not a function selecting from the existing denotation, this is the intersection of two denotations being selected. This can apply recursively, such that the output denotation that is the intersection of the denotation of “houses” and the denotation of “blue” may now form the input for modification by another adjective, for example *big*. The structure will now undertake the intersection of “blue houses” and “big” to return a further restricted denotation.\footnote{I will not be discussing the relative ordering of adjective modifiers, but nothing here is incompatible with the notion that some subtle semantic distinctions can determine which denotation intersections must take place first.}

This approach to lexical projections extends to apply to further types of nominal modification, such as relative clauses and prepositional phrases. The contribution of its own lexical content, its own denotation, is a clear difference with functional projections.
4.3.3 Returning to measure words

Measure words remain trapped in between the lexical and the functional, perhaps demonstrating the grammaticalisation process at work. In the synchronic grammar though, their application still needs to be explained.

Measure words can act as full nouns. In fact, most are full nouns that can be employed for use as measure words. When used as nouns, they act as full nouns. They are introduced into the structure with full flexible denotations and need classifiers to be individuated and counted. The crucial point is that, when acting as full nouns, they contribute their denotation entirely unhindered – when entering the structure as the base noun, they exist on their own with their full denotation and there is no result from their intersection with another denotation.

This is different to their use as a measure word. In the simplest form of this use, there is only Numeral–N_{MW}–N. One of the most complex uses of a measure word is shown in (91). In these types of uses of the N_{MW}, there is an intersection of denotations. The denotation of the measure word intersects with the full flexible denotation of a lower noun, the full mass denotation of a lower noun, or the denotation of a lower classifier+noun combination. What we have here is an example of nominal modification which is not technically an example of pure intersection. After all, in the example boxes of sand, the output is not the set of things that are both boxes and sand.

When a noun is being employed as a measure word, and therefore cannot provide its own full denotation to the structure, it may apply its new-found functional power. The Measure function takes the denotation of the lower noun and applies its denotation to create new sets of atoms in the lower denotation, as set out by the lexical content of the measure word. For example, with the lower noun sand, the Measure function applied by the measure word box will rearrange the denotation
The Features and Structure of the lower Noun Phrase

of sand, which consists of unstable atoms, into sets of unstable atoms bound by the denotation of the measure word box. This Measure function, in some way, is superimposed over the existing denotation, such that both still exist. This reflects the “in-between” nature of the function – it is not a case of lexical intersection and it is not a case of a function applying only to an existing denotation. What is crucial in the case of measure words is that the application of the Measure function creates stable atoms in the resultant denotation, meaning that the Cardinality function can immediately apply. The superimposition of the measure word denotation necessarily realises individual groups of the atoms of the lower noun, whether stable or unstable.

4.3.4 The interaction of classifiers and measure words

Finally, we need to consider the interaction of classifiers and measure words and ensure that the theory predicts the observed behaviour. We have examples of a measure word and a classifier in the same phrase with two different numerals. However, according to the data, there can only ever be one classifier per full lexical noun, and it must be lower than any measure words interacting with the same lexical noun. Measure words are always higher and, in contrast, can be stacked, i.e. three boxes of four bottles of water.

What we must ensure is that classifiers cannot operate over nouns that are acting as measure words. This can be accounted for through the function Measure. Recall that the difference between these elements acting as nouns and as measure words regards whether there is an existing denotation to which they can apply the Measure function. If there is no existing denotation, the element is acting as a full noun and the Measure function does not apply. This means classifiers can apply their function to the flexible denotation supplied. If there is a denotation, i.e. the element is
acting as a measure word, then the Measure function applies. This application of the Measure function must create an output that is an incompatible input for the Individuation function of a classifier.

To account for this, I will return to degrees of stability and the evidence that some nouns may now be counted without a classifier. The denotation of a flexible noun, as previously discussed, contains both stable and unstable atoms. The Individuation function has been proposed as the method of selecting only the stable atoms of the denotation, making counting of these atoms possible. However, my experiment participants remarked that some nouns can now be counted in the absence of a classifier. It is usually understood that the general classifier ge is being omitted as it provides no additional categorising content of its own. Where a classifier is not providing this content, it can be omitted in certain contexts. This suggests that the stable atoms in the denotation of these nouns, in these contexts, are stable enough to have the Cardinality function apply without the need for the Individuation function. This mirrors the behaviour of measure words applying the Measure function, whose denotation would otherwise require Individuation.

I propose that this is due to the relative stability of these atoms in the denotation of the noun. These “innovative” nouns are heading in the direction of count nouns in English, in the way that the individual atoms in their denotations are becoming fully stable and part of the lexical specification of the noun. This demonstrates that once atoms are lexically specified as stable, there is no need for the Individuation function and suggests that the denotations of flexible nouns contain stable atoms that are slightly less stable than those in the denotation of a true count noun. The Individuation function, therefore, takes these “nearly” stable atoms and, in effect, tops them up to 100% stability – this is what “realising individuals for the syntax” equates to.

The question is: why can’t classifier apply to measure words acting as measure
words? More technically, why can’t the Individuation function apply directly to the output of the Measure function? I put this down to the nature of the output of the Measure function, which is a denotation including stable atoms which are themselves sets of atoms, either stable or unstable. This two-tiered denotation, which reflects the crossover of a lexical and functional projection, is not compatible with the Individuation function – the “outer” tier is already stable and the “inner” tier, whether stable or unstable, is inaccessible. The Individuation function, as a simple functional projection, requires a simple input.

4.3.5 Adjectives

The structure in (91) also captures the observation that measure words can be modified by an unrestricted set of adjectives by aligning them with full lexical nouns – there is no restriction on the lower, pre-noun position for adjectives, so why should there be on the higher, pre-measure word position if they are both noun-modifying positions? Also, recall that measure words blocked the modification of the lower noun by the higher adjective. This was supported by the possibility of contradictory adjectives in the two adjective positions. The structure accounts for this observation by linking the higher adjective position to the higher noun, i.e. the measure word, and the lower position to the lower main noun. In this configuration, there is no reason to expect a clash of adjectives as they modify different entities.

---

5 This notion of an inner, inaccessible tier suggests to me that perhaps the lower noun is in fact closed off to the higher measure word, and, in reference to the structure in (91), perhaps there is an additional functional layer between the lower ClassP (if projected) and the higher N_{MW} which restricts true functions applying further to the denotation of the lower structure. The question remains as to how much of the extended projection of the lower, main noun is required prior to it forming the complement of the N_{MW}. 
4.3.5.1 Delitive adjectives

The range of adjectives allowed in the position before a classifier was highly restricted to the so-called delitive adjectives. There is some consensus amongst current research that these include size and shape adjectives, but there is no widely agreed set or reasoned definition for the grouping.

Interestingly, size and shape adjectives are exactly those that are prohibited from modifying nouns that have mass denotations, providing further evidence that the adjunction of these adjectives relies on the denotation of the noun – it must contain stable atoms.

The question remains as to why the full range of adjectival modification is not allowed in the higher pre-classifier position. Assuming that classifiers have in some way developed from the grammaticalisation of nouns, it may be the case that the classifiers lost their ability to be modified by adjectives at some point along this journey, whilst much of their own semantic content was also lost. We have seen how pre-classifier adjectives modify the noun “through” the classifier. The contribution of the classifier is, however, to apply the Individuation function and create a countable unit in the syntax. Muromatsu (2001) and Truswell (2004) argue that these “dimension” adjectives such as big and tiny must merge above the head which creates countable entities. There is good evidence for the explanation as to why these adjectives cannot modify mass nouns, but they can appear in the lower pre-noun position for flexible nouns according to my data.

What is clear is that the classifier is not truly being modified, given its lack of semantic content and “invisibility” to modification. The ability for delitive adjectives to appear in the higher pre-classifier position could be a result of their modification being dependent on a denotation of only stable atoms – a status that is only granted by the Individuation function of the classifier. Given that the result of the classi-
fier+noun combination is a restricted set of the noun’s original denotation, this is a logical approach. The presence of these delimitive adjectives in the lower pre-noun position is therefore a result of the same “degrees of stability” argument, and any variation could be down to interspeaker variation or context dependence. These adjectives are licensed in both positions whilst being dependent on a denotation of stable atoms. Their presence in the higher pre-classifier position demonstrates that this modification may only take place in this restricted position, for some speakers/contexts, as it is the earliest opportunity for such modification to take place, i.e. the first time that the denotation of the noun is compatible for the intersection with the denotation of the adjective.

4.3.6 Comparison with English measure words

English measure words are similar to those in Mandarin Chinese, but there are some differences. Measure words in English can be modified independently of the noun they measure, i.e. *three big boxes of small apples* is not contradictory, and they can appear above full DPs, i.e. *three boxes of two apples*. However, English measure words can also be transparent for adjectival modification, i.e. *a delicious box of chocolates, a hot cup of coffee*. This is a clear departure from measure words in Mandarin Chinese. Comparatively, therefore, it seems that (at least some) measure words in English are slightly further along the journey from lexical to functional than those in Mandarin Chinese, and are candidates to become semi-lexical heads, of the type discussed by van Riemsdijk (1998) and Vos (1999), sooner.
4.4 Conclusion

In this chapter, the focus has swung towards the differences between classifiers and measure words, and how they interact with the flexible/mass distinction that we previously argued to exist in Mandarin Chinese.

Many previous analyses had chosen to bundle together classifiers and measure words as the same type of element, providing a featural or structural distinction for the difference in their behaviours. I have proposed that the difference between classifiers and measure words has much deeper roots, arising directly from a difference in terms of both category and syntactic configuration.

A true classifier is a functional element that merges in the extended projection of a noun. It applies the Individuation function which selects from the flexible denotation of a noun only the stable atoms or sets of stable atoms. For this reason, classifiers cannot combine with mass nouns.

As a functional element, the classifier is merged with little, or more likely no, grammatically-relevant semantic content. This means that direct modification by adjectives is not licensed, and any pre-classifier adjective modifies the lower main noun. This is the behaviour expected from a functional head in the extended projection.

Measure words present an interesting case on the continuum between the lexical and the functional. They can combine with flexible nouns, mass nouns and classifier+noun combinations. On the one hand, measure words allow full adjectival modification and provide lexical information for the interpretation of the phrase – behaviour aligning them with nouns. On the other hand, measure words can be directly modified by numerals, aligning them with classifiers. This suggests that measure words are somewhere along the grammaticalisation path from nouns to
classifiers.

Given their characteristics, and the observation that all measure words can also operate as full nouns, I chose to align measure words with nouns, but provided them with the Measure function that applies dependent on the point at which they enter the structure-building process. As part of an element’s progress through grammaticalisation towards becoming functional, it must at some point acquire a function. This is, I propose, the Measure function. When a measure word enters a derivation which already includes a denotation from another noun, the Measure function applies. This converts the lexical denotation of the measure word into a function and superimposes it on the existing denotation, in a different process to the intersection which takes place with standard modification.

Measure words are, therefore, full nouns which, when placed in a non-standard structural position for full nouns, must apply their acquired function. The result of the Measure function is an output which is an incompatible input for the Individuation function, but is a compatible input for the Cardinality function. This relates to the stability and complexity of the atoms. I propose that the Individuation function requires simple stable atoms, whereas the Cardinality function can apply to any set of stable atoms.

I also discussed the behaviour of adjectives in their interaction with classifiers and measure words, and found that the restrictions on adjective modification may also stem from the types of denotation available as the point of modification. It was clear that delimitive adjectives relied on a denotation with stable atoms, and this explains why in some contexts they cannot merge until the Individuation function has applied.

In sum, measure words and classifiers are entirely different elements that appear in entirely different structural positions for entirely different reasons. I aligned
measure words with nouns, but it is likely that grammaticalisation is in progress here. As measure words move in a more functional direction, it will be interesting to observe how their nominal properties are lost and whether there is a structural reconfiguration that adopts them into the extended projection of the noun. Given that they must contribute some semantic content, there may well be a limit to how functional these elements can become.

4.5 A note on “kind” classifiers

As well as classifiers and measure words, there are also elements that denote kinds or types, as in (93):

(93) liang zhong pingguo
    two CL apple
    ‘two types of apples’

These “type words” function somewhat like classifiers and measure words, but they neither select stable atoms from the denotation through the Individuation function, like classifiers, or bundle atoms into lexically-defined measures through the Measure function, like measure words.

From my data, type words cannot be dropped, as they have content relevant to the interpretation that would be lost if omitted. They cannot be directly modified by an adjective in the pre-type word position, perhaps given that their semantic content is not strongly “contentful”, i.e. in the phrase “two A types of N”, it is impossible for A to modify “types” and not N. The function applied by a type word is unlikely to result from a lexical denotation, like measure words, and is more likely to operate over the existing denotation, much like classifiers and other purely
functional projections. What we have is a type of element that has some content, like a measure word, but cannot be modified, like a classifier.

The semantic content provided by a type word is on the same level as the majority of classifiers, although it is of a different quality. Recall that, with the possibility of omitting classifiers, it was only ever possible on the basis that the dropped classifier is the general classifier $ge$. Dropping any other classifier loses the interpretation provided by the additional categorising information inherent to that classifier. The same situation is found with type words. They do not identify some salient property of the noun for categorisation as classifiers do, but rather identify a grouping within the denotation of the noun based on some understanding of what a “type” or a “kind” is.

For this reason, I align type words with classifiers as functional elements. The function of a type word must be able to apply to the denotations of flexible and mass nouns alike – a type of apple is as likely as a type of sand. It must appeal to the interpretation of the noun as to what constitutes a type, and create individuated groups based on this information. The Type function is different to the Individuation function then, in that it applies to both mass and flexible denotations and groups atoms, both stable and unstable, into conceptually-driven types. There is no selection of a subset of the denotation, the Type function primarily arranges the existing denotation into countable sets of atoms without restricting the denotation in any way.
Chapter 5

Gender - feature or functional head?

5.1 Introduction

In this chapter, we move away from issues of the count/mass distinction and the Individuation function, and consider the nature of grammatical gender. This is standardly assumed to be the lowest of all the nominal features, given that the feature values are almost always inherently linked with the lexical specification of a given noun.

Unlike the proposals for an independent functional projection for Number, as discussed above and going back to Ritter (1991), which have generally been accepted, there are a number of proposals for the syntactic treatment of gender within the nominal phrase that will be considered through the course of this chapter.

Finally, I discuss gender as a feature and as a functional head, based on the possibilities of its acquisition.
5.2 Feature or functional head?

5.2.1 Gender as a feature

Ritter (1993) proposes that gender is a feature realised on one of the existing syntactic heads of the noun phrase, and that the choice of which syntactic head is subject to cross-linguistic variation. Comparing Romance languages and Hebrew, she assumes that the basic noun phrase structure is the same:

\[(94)\]
\[\text{DP} \ D \ [\text{NumP} \ Num \ [\text{NP} \ N ] ]] \]

The intermediate functional category of Num hosts plural suffixes. In Romance languages, the gender feature is base-generated as a feature on Num and appears on N through N-to-Num raising. In Hebrew, the gender feature is lexically inherent to the noun and thus appears on the noun stem “at all levels of syntactic representation”. In support of this hypothesis, Ritter shows how gender switching is a productive strategy for deriving new nouns. For example, adding a feminine suffix to an inanimate masculine noun stem derives a related feminine noun:

\[(95)\]
\[\begin{align*}
\text{a.} & & (M) \text{ magav} & \text{‘wiper’} \\
& & (F) \text{ magev-et} & \text{‘towel’} \\
\text{b.} & & (M) \text{ maxsan} & \text{‘warehouse’} \\
& & (F) \text{ maxsan-it} & \text{‘magazine’} \\
\text{c.} & & (M) \text{ amud} & \text{‘page’} \\
& & (F) \text{ amud-a} & \text{‘column’}
\end{align*}\]

Both the masculine noun stems and the derived feminine nouns can be inflected for number, showing clearly that gender is both anchored to the noun and separate from number. This contrasts with the Romance languages in which such gender
manipulation typically does not form new words, except in a restricted number of cases involving animate or human reference that produce compositional interpretations. Gender is a purely inflectional feature on nouns, like number, and cannot be employed as a derivational strategy.

To support her hypothesis that gender is on Num in Romance languages, Ritter (1993) turns to Walloon, a Romance language spoken in Belgium, in which there is no N-to-Num movement (based on Bernstein (1991)). This means that plural marking is not realised on nouns. Walloon has a feminine plural marker ´es which Bernstein analyses as the overt realisation of Num, stating that ´es is best analysed as a number marker with inherent gender features. This, Ritter claims, is sufficient evidence to suggest that in the majority of Romance languages, in which N-to-Num movement is commonplace, the gender feature is found together with the the number specification on the functional head.¹

Fuchs et al. (2015) refer to this approach as the *bundling* model, given the bundling of features on the existing Num head. Accordingly, all gender morphology is either:

(96)  a. hosted on Num  \[\text{NumP} \ldots \lbrack \text{Num} (\text{Number}, \text{Gender}) \rbrack\]

b. expressed on Spec,NumP \[\text{NumP} \text{XP}_{(\text{Gender})} \lbrack \text{Num} (\text{Number}, \text{Gender}) \rbrack\]

In either option, gender features do not project independently of number, and the valuation of gender presupposes a valuation of number. This is inspired by the observation that many languages combine their gender and number information. Empirically, few language systems have both gender and number take part in agreement entirely independently.

¹It is perhaps important to note that Bernstein (1993) later suggests that the prenominal plural marker can be decomposed into gender, i.e. feminine ´e, and number, i.e. plural s).
Bundling gender onto an existing syntactic head also mediates how it is a feature that does not have a consistent syntactic reality or consistent semantic content. Gender on almost all nouns is lexically-specified – a noun has a specification for gender regardless of its position within the syntactic structure. As will be discussed, the impact of gender within the syntax is minimal – in terms of gender, the syntax goes along with what the lexicon provides and not much more. This is a clear difference to number which, as discussed in previous chapters, applies a function which in some way alters the denotation of the noun. Number is a prime candidate for a functional head, given that it applies a function. It is difficult on the current evidence to argue for gender operating in the same way.

Furthermore, the projection of functional heads has so far resulted from the application of functions on the denotation. The gender feature on inanimate nouns is uninterpretable. If features are primarily motivated by the semantics and the task of identifying reference, how can an uninterpretable feature make a contribution? For inanimate nouns, there is no consistent semantic content that may lead to the acquisition of the Gender feature. This suggests that the gender feature is not motivated by semantics, providing good motivation for gender to be bundled together with a semantically-motivated feature such as number. If this is the case, what are the mechanisms by which gender is acquired and how does the systematic evidence lead gender to be bundled with number?

5.2.2 Gender as a functional head

Picallo (1991) claimed that gender projects its own functional phrase within the DP, between NP and NumP, labelled Gen(der)P. The relative position of GenP reflects the fact that gender is expressed directly on the noun stem and that number is expressed outside gender, as in Spanish:
The Features and Structure of the lower Noun Phrase

(97) a. mes-a-s
    table-FEM-PL
    “tables”

Following the approach outlined, the affixes for gender and number are checked (cf. Chomsky (1995)) as the noun stem raises from N to Num, the highest functional projection below D.

This idea is further developed by Picallo (2008), who puts forward the proposal that gender inflection in a common noun reflects an interpretable feature that is hosted in a functional projection in the nominal structure. The interpretation of common nouns is generally not affected by belonging to a particular gender and grammatical gender is not interpretable on N. For gender inflection to reflect an interpretable feature, Picallo proposes the interpretable functional feature [CLASS], which is hosted in a functional projection c immediately dominating N:

(98) \[ c \[ \text{CLASS} \] [N N] \]

The key claim is that [CLASS] translates the non-linguistic processes of entity categorisation to the grammatical system. Commonly in Indo-European language, the presence of [CLASS] is manifested in a formal gender system for nouns, but in other language families the feature may be realised in other ways, such as in a system of noun classes where semi-lexical items or morphemes categorise nouns according to some, possibly non-linguistic, factor. Picallo (2008) maintains that grammatical gender, noun classifiers and noun classes are all the same kind of functional element.

In Romance languages, markers of gender usually follow the noun. Previously, Picallo (1991) suggested that the post-nominal position of gender inflection obtains
from head raising of N to the head of a functional category. Picallo (2008) moves forward from this to consider the alternative that the pre- or post-nominal position of realisation of the [CLASS] feature follows from a parametric difference: the locus of its valuation.

In Romance languages, [CLASS] in c is interpreted but is selected as unvalued. The gender inflection that appears as a suffix on the noun is the result of agreement between [CLASS] in c and the formal feature on N, which is lexically inherent but not interpretable on N. [CLASS] in Romance has to relate to a fully inflected N in order to be valued. The crucial difference here is that the locus of interpretation is the unvalued probe [CLASS] in the functional c, with its matching valued feature in the goal N that is not interpretable.

This suggests that gender inflection involves two features: a lexically specified gender feature on the noun that is valued but uninterpretable and the [CLASS] feature on the functional head c, immediately dominating NP, that is unvalued but interpretable. Agreement between these two features ensures that gender is valued and interpretable in c. It’s possible that the parametric variation leading to noun classifiers and noun classes, as opposed to a gender system, is that these systems have an interpretable and valued [CLASS] in c.

Fuchs et al. (2015) refer to this approach as the split model, given that gender and number are separate and gender is given its own functional projection, GenP, under NumP:

\[(99) \quad \text{[NumP [GenP … ] ]}\]

The move for syntactic separation is supported by languages in which gender and number morphology can be separated and there is a clear and consistent ordering of
the number and gender morphemes. Spanish is an example of a language in which
the order is consistently Stem-Gender-Number, as in (100):

(100) a. [[[libr] - [GenP 0] - [NumP s]]
    libros
    ‘books’

b. [[[libr] - [GenP 0] - [NumP ø]]
    libro
    ‘book’

The advantage of this approach over the bundling model is that it can predict the
consistent ordering of number and gender. Crosslinguistically, there is a tendency
for gender or class information to take a structural position closer to the noun, and
a bundle of features on the same syntactic head cannot account for this. The split
model predicts a strict ordering based on the order of projection of the functional
heads.

The argument for the split model is strengthened further by instances of nominal
ellipsis. In both Greek and Spanish, the gender features of a nominal phrase being
elided must match those of its antecedent, whereas no match in number features is
required (Fuchs et al. 2015). This is demonstrated in (101):

(101) a. Juan visitó a su tío, y Pedro visitó a
    Juan visit.3SG.PST to 3SG.POSS uncle, and Pedro visit.3SG.PST to
    los suyos.
    DET 3.POSS.M.PL
    ‘Juan visited his uncle, and Pedro visited his (uncles).’

b. *Juan visitó a su tío, y Pedro visitó a
    Juan visit.3SG.PST to 3SG.POSS uncle, and Pedro visit.3SG.PST to
    la suya.
    DET 3.POSS.F.SG
    (‘Juan visited his uncle, and Pedro visited his (aunt).’)

132
The sentence in (101-a), with the number mismatch between the elided nominal in the plural and the antecedent in the singular, is grammatical, whilst the sentence in (101-b), with a gender mismatch between the elided nominal with feminine gender and the antecedent with masculine gender, is ungrammatical. This points to an ordering restriction and a closer relationship between gender and the noun that can be neatly accounted for by the split model. The bundling model cannot accurately explain this data, given the lack of reasoned justification for why one of the bundled features need not match while the other one must.

There is, therefore, good evidence for the separation of gender and number. One of the main drivers of the bundling model is that gender and number tend to take part in agreement together, but this is not ruled out by the split model. This approach allows opportunity for gender and number to participate in agreement together and separately.

5.2.3 Comparing the approaches

What is consistent between the two models is that the gender feature exists in the grammar, and that it is linked to the lexical specification of the noun in some way.

Ritter (1993), in the bundling model, states that gender must be recorded in the lexical representation of the noun, but in Romance languages it is not lexically anchored to the noun stem and thus can be base-generated as a feature on Num, attaching to N through N-to-Num raising. This could be achieved functionally, whereby a higher functional projection “realises” the lexical gender specification in the syntax.
Picallo (2008), in the split model, turns to agreement and a parametrically-defined interaction between $[\text{CLASS}]$ on $c$ and gender on $N$ for Romance languages. This may operate via a similar process, but under the assumption that there is a functional head specific to gender that applies a function specialised for “realising” gender, or indeed any of the non-linguistic categories similar to grammatical gender, in the grammar.

The key difference therefore comes down to whether the function that gives non-linguistic gender a syntactic reality exists solely for that purpose with its own functional projection, or is attached to another functional head.

Before making a judgement as to which of these options may be preferable, it is first important to understand how the gender feature itself may be acquired as part of the lexical specification of the noun.

### 5.3 Evidence for the acquisition of gender

As discussed in previous chapters, the acquisition of a formal feature must be semantically motivated, and the correlation of a difference in interpretation with a difference in overt realisation is the strongest motivator for the creation of a formal feature. However, what is the semantic basis for grammatical gender, or noun classes, as they apply to most nouns in a language – inanimate nouns?

The gender on inanimate nouns is uninterpretable. During acquisition, there is no evidence in the interpretation of inanimate nouns that should match with the differences in overt realisation, wherever and however they are marked. In terms of the clues made available to the acquirer, they seem to only be getting half the story.
Nevertheless, the evidence in the overt realisation does result in the acquisition of gender. Many European languages exhibit gender through agreement with determiners, meaning that an acquirer’s selection of a determiner for a given noun can be tested.

Dutch has two grammatical genders: common and neuter. Children acquiring Dutch first go through a stage in which articles are omitted (Blom et al. 2008), after which they overgeneralise the appropriate article for common nouns, *de*, for use with both genders. This may be a result of the input – common nouns are around twice as frequent as neuter nouns (Van Berkum 1996). *De* is also used for both genders in the plural. Van der Velde (2003, 2004) finds that this overgeneralisation continues until acquirers are at least six years old, at which point 79.2% of neuter nouns are produced with the correct article *het*. There is no evidence for the overgeneralisation of *het* in the other direction.

Karmiloff-Smith’s (1979) study looked to assess the predictive power of different types of cues for gender in French, comparing two alternative propositions: the “semantic” hypothesis, under which grammatical gender is acquired as an extension of natural gender, and the “linguistic” hypothesis, according to which grammatical gender emerges based on the links between certain morphological endings on nouns and other parts of speech. One of the experiments in the study found that, from three years of age, children were capable of choosing the correct form of the article in terms of gender based on the ending of the noun – of the four pseudowords tested, 100% was reached for three. Based on this and other observations in the study, Karmiloff-Smith deduced that the linguistic hypothesis was stronger, with the predominant driver for the creation of gender in the linguistic system being linguistic information.

This is good, and perhaps expected, news, given that the gender categorisation of inanimate nouns in many languages has no semantic basis. The results from
Karmiloff-Smith (1979) demonstrate that the gender system in French can be acquired based on phonological cues alone, and that this system is acquired from an early age.

5.3.1 Gender and natural classes

In Italian and other Romance languages, a specific gender can be associated with a natural class, i.e. masculine for the names of rivers and feminine for the names of cities. The data from Karmiloff-Smith (1979) suggests that phonological cues could be sufficient for gender assignment from an early age, but Thornton (2009) finds that, at least for gender assignment to loanwords and neologisms, semantic gender assignment rules prevail over phonological/morphological ones. Thornton (2009) proposes a constraint on semantic gender assignment such that a noun’s gender is inherited from its basic level hypernym. This would explain why names of rivers are often masculine in Italian, given the masculine *il fiume*, and why names of cities are often feminine, given the feminine *la città*. This semantic gender assignment relies on the acquisition of the gender of the hypernym, however, and the correct establishment of the hypernymy relation. Further investigation of the role of semantic classes in the acquisition of gender is required.

5.4 Acquiring gender as a feature

The acquisition of gender, therefore, relies on a different type of evidence to the count/mass distinction and the Number feature that have previously been discussed. For the majority of nouns in an acquirer’s input, there is no semantic information linked to the gender morphology observed – there is no systematic impact on the interpretation that correlates with the overt realisation.
Despite this lack of impact on the interpretation, the acquirer clearly recognises that there is a system in operation in the grammar. A wealth of research into how children acquire grammatical gender crosslinguistically has shown that the system is constructed based largely on distributional evidence (Popova 1973; Karmiloff-Smith 1979; Smoczyńska 1985; Mills 1986; Cain et al. 1987; Levy 1996), and that this acquisition takes place relatively quickly (Smoczyńska 1985). This at least seems to be the case in languages with simple, transparent systems of gender marking. Smoczyńska (1985) finds that, with a more opaque system of marking, the acquisition of gender can be delayed and, in some cases, semantic information may be used at first to unlock the system (Mulford 1985).

The acquisition of gender therefore takes place quickly and easily when the system of morphological marking is simple and transparent, even in the absence of semantic cues. The question of where and how this feature is created depends on the specific language being acquired – we must continue to reject the assumption that any feature is structured in the same way crosslinguistically, even if there is strong evidence for similarities. In the same way, the potential for the bundling and the split models to each be applicable to different languages must also be considered.

5.4.1 Internal structuring of gender

For gender-marking languages, all nouns will be allocated to a particular gender. This means that, during acquisition, the overt clues will make it clear to the acquirer which gender to assign. However, it is not always the case that all the genders will be morphologically specified. There may be one gender or genders that require morphological marking while one does not.

Furthermore, as we saw for individuation in Mandarin Chinese, it is not always the case that all the values of the feature are semantically specified. Recall that the
count/mass distinction in English was an even semantic specification between the two groups, but, in Mandarin Chinese, we found a flexible/mass distinction such that the denotation of flexible nouns included the denotation of mass nouns. A different type of distinction was being made across the two languages.

Gender may be structured in the same way for some languages, with one gender almost operating as a subset of another. There is also the opportunity for sets of individuals with different genders. Does one gender get selected over the other for mixed sets? Is the selected gender the one that is morphologically marked, or the one that is not? These questions need to be considered as we try to deduce whether gender deserves a functional head, in addition to its status as a feature, especially against the lack of semantic content.

First, we will consider the gender system of Spanish.

5.5 Gender and number in Spanish

In Spanish, there is agreement in both gender and number between a head noun and any associated determiners, adjectives, and participles. This agreement in number and gender is also maintained in anaphors.

The gender feature in Spanish has two values: masculine and feminine. In terms of morphological marking, there is a general correlation between the ending -o and masculine gender, and the ending -a and feminine gender. This is not a perfect one-to-one mapping though, Harris (1991) points out that neither of these endings is exclusively used with a specific gender. The masculine ending -o can be found on feminine nouns, such as la mano “the hand”, on nouns that can be masculine or feminine, such as el/la testigo “the witness”, and on words that are not even nouns at all, such as dentro “inside”. The feminine ending -a is no better behaved,
appearing on masculine nouns, such as \textit{el dia} “the day”, nouns that can be masculine or feminine, such as \textit{el/la turista} “the tourist”, and other words that are not nouns either, such as \textit{afuera} “outside”.

Although gender marking on the noun may not be a perfect system, gender is accurately marked elsewhere. Determiners demonstrate a reliable system of gender marking on all forms: \textit{un} vs. \textit{una}, \textit{el} vs. \textit{la}, \textit{los} vs. \textit{las} and \textit{unos} vs. \textit{unas}.

With regard to the semantic specification of gender in Spanish, masculine gender is the one used when referring to a group of nouns of mixed gender. For example, \textit{los padres} is ambiguous between ‘parents’ and ‘fathers’ whilst \textit{las madres} can only refer to ‘mothers’. In the same way, if two or more nouns of both genders are coordinated, the agreement marker on the adjective has to be masculine:

\begin{quote}
\textit{El libro y la pintura son caros} / *caras} \\
\text{The book.M and the painting.F are expensive.M.PL / expensive.F.PL} \\
\text{‘The book and the painting are expensive.’}
\end{quote}

\subsection{5.5.1 Number in Spanish}

There are two options for number in Spanish, the singular and the plural. The plural is marked morphologically with the suffix \textit{-s}, as in English. There is no morphological marking of the singular. Morphological marking of number in Spanish takes place after, or outside, gender marking, as in (103):

\begin{quote}
\begin{enumerate}
\item a. \textit{El libro} ‘the book’ \\
\quad \textit{Los libros} ‘the books’ \\
\item b. \textit{La pintura} ‘the painting’ \\
\quad \textit{Las pinturas} ‘the paintings’
\end{enumerate}
\end{quote}
In terms of semantic specification, plural number can be said to include singular reference, in much the same way as English again. Asking “Do you have children?” in Spanish will receive an affirmative response, even if the respondent only has one child. In this way, singular reference is included in the use of the plural.

5.5.1.1 Neuter gender marks mass nouns

In Asturian, a Romance language spoken in Asturias in Spain, adjectival endings stemming from Latin neuter morphology seem to mark a count/mass distinction. Traditionally known as “neutro asturiano”, there is a three-way distinction on adjectives and pronouns reflecting masculine, feminine and mass. This exists alongside the binary masculine/feminine contrast on articles. For example (data from Neira Martínez (1955: p.70-72) and Neira Martínez (1978: p.260), taken from Loporcaro (2016: p.934)):

(104) a. Feminine count:  
la casa fria  
‘the cold house’

b. Feminine mass:  
la tsiche frio  
‘the cold milk’

(105) a. Masculine count:  
el pie friu  
‘the cold foot’

b. Masculine mass:  
el cafè frio  
‘the cold coffee’

In such a system as this, it seems that the neuter morphology has been retained and repurposed for the marking of mass nouns, perhaps given that the Latin neuter hosted a large number of mass nouns. The binary gender contrast remains the
same and is marked through determiner agreement, but, for mass nouns, any gender agreement on adjectives and pronouns is replaced with the neuter morphology marking mass. This is a clear difference to Spanish, and demonstrates how different agreement targets may mark different features of the noun. In Asturian, there is good evidence for the count/mass distinction being grammaticalised and split from gender.

5.5.1.2 Number changes gender

The neuter plural has also survived into some Romance languages, but now only marks plural and has ultimately been reanalysed as a marker of feminine gender. Where this marker survived, it expanded beyond neuter nouns to a large part of the lexicon, although remaining restricted to inanimate nouns. This has resulted in an effect known as genus alternans, whereby nouns that are masculine in the singular take feminine plural forms.

Consider the following examples from Italian:

(106) a. Masculine singular:
    *braccio*
    ‘arm’
 b. Feminine plural:
    *braccia*
    ‘arms’

(107) a. Masculine singular:
    *uovo*
    ‘egg’
 b. Feminine plural:
    *uova*
    ‘eggs’
These forms slightly cloud the acquisition picture for the child, but do not derail the entire acquisition task. In a number of Romance languages, such as Vegliote, Ligurian and Tuscan, there are only a few lexemes demonstrating the feminine plural in this way, with many referring to individuals found in sets or to a mass, as might be expected from the neuter origins (Maiden 2016). In Italian, morphologically regular masculine plurals usually exist alongside the feminine plural form, with differing semantic differentiation from lexeme to lexeme, speaker to speaker and region to region (Thornton 2011).

5.6 Gender and number in German

In Spanish, gender and number are constructed and accessed independently in the morphology, as is consistent with the split model. However, this may not be true for all languages. A language with bundled gender and number might have its gender features or a subset of those features accessed via number. German may be an example of such a language. In the singular, German distinguishes three genders: masculine, feminine and neuter. In the plural, there are not always clear distinctions between grammatical genders.

(108) a. Der Student ‘the student’
    Die Studenten ‘the students’

    b. Die Nation ‘the nation’
    Die Nationen ‘the nations’

    c. Das Thema ‘the topic’
    Die Themen ‘the topics’

From the examples in (108), the number morphology is clear, but there are few clues to unlocking the gender system – even in the singular. Concentrating on the determiners, gender is only ever marked on determiners in the singular. There
is a masculine singular definite article, a feminine singular definite article, and a
neuter singular definite article, whereas there is one plural definite article (which
has the same form as the feminine singular). This suggests that gender is only
realised in the singular, and forms a bundled set of features with the plural, i.e.
masculine, feminine, neuter, plural, with only one feature being expressed at any
one time.\footnote{Assuming the bundling model for German and the split model for
Spanish requires that we accept that gender originates from different categories in
different languages. In Spanish, the evidence suggested that gender is realised by
its own functional projection, GenP, which is dominated by NumP. Gender and
number are entirely independent features. In German, gender and number are
bundled together so, following the bundling model, the lexical specification of
gender on nouns in German can only be realised in the syntax at the level of Num.
Consider the following contrast:}

(109) a. Spanish:
\[
[\text{DP} \ldots [\text{Num} \ldots \text{Num}_\text{(Number)}] [\text{GenP} \ldots \text{Gen}_\text{(Gender)}] [\text{NP} \ldots \text{N}_{\text{gender}}] ]]
\]

b. German:
\[
[\text{DP} \ldots [\text{Num} \ldots \text{Num}_\text{(Number+Gender)}] [\text{NP} \ldots \text{N}_{\text{gender}}] ]]
\]

In both German and Spanish, nouns are specified for gender in the lexicon. In
Spanish, there is a functional projection, GenP, dominating NP that allows gender
and number to behave independently. In German, there is no GenP and thus
gender as a feature cannot operate independently of number.

\footnote{Some animate nouns allow feminine morphology which is carried over to the
plural, i.e. \textit{der Student} (masc.), \textit{die Studentin} (fem.) and \textit{die Studentinnen} (plural). I propose that these
dem feminine forms are not inflected forms of the masculine and are stored separately
in the lexicon. The plural form therefore does not contain a true "feminine" inflection
as would be expected in a Spanish-like split system.}
5.7 Gender and functions

As we have seen, gender on inanimate nouns may be acquired based primarily on phonological cues where a language has a transparent system of marking. This allows the gender feature to be acquired and stored as a lexical feature on the noun. Under all the approaches considered, gender was agreed to be lexically-specified on the noun.

This prompts the question: what type of feature is gender? The previous features we have discussed were semantically motivated, and therefore have as their basis some element of meaning which determines or restricts an interpretation. In this simple system, features of the semantics are “grammaticalised” to create formal features. Gender, for the most part, does not have this semantic basis, yet the acquirer creates the system in their grammar.

We have a formal feature being created without semantic motivation. There is a systematic impact on the grammar of a distinction that is not semantically marked. Whilst perhaps this is not desirable for a functionally-driven syntax, the observations are undeniable. The acquirer creates a formal feature for gender in their grammar, but not one that is “grammaticalised” from a semantic feature on the noun. The formal feature may, nonetheless, be attached to the lexical specification of the noun, based on the evidence provided by a particular language.

Whether there is a functional projection for gender, therefore, depends on how gender is morphologically marked. This is not ideal for the system, as we have a functional projection with no function – it exists only to host the gender morphology. If gender is marked individually in a language, as it is in Spanish, the acquirer will create a functional head Gen, dominating the noun, which hosts the gender marker and “realises” the lexically-specified gender of the noun within the syntax. This realisation may, in fact, be the function, but it is difficult to describe how
the denotation of the noun is affected by this non-semantic information. The key difference, though, is that there is a functional head hosting the morphology.

We can assume that nouns in German have a similar lexical specification for gender as nouns in Spanish, given that the system will be acquired in the same way. What is different for German is that the morphological marking is suggesting a close interaction with number, such that number dominates of gender – the marking of number determines the marking of gender. Therefore there is no gender marking independent of number and all this marking must take place via the Num head. The evidence for the acquirer will create such a system, with the bundles of features that generate the morphological marking being hosted in the functional head Num.

Does the Number function therefore work in the same way? In German, there is a range of plural endings, but the singular is unmarked. A count/mass distinction ensures that nouns are lexically-specified to have denotations of individuals and non-individuals. The Number function operates as it does in English, selecting either atoms or sets of atoms depending on its value. If sets of atoms are selected, no gender marking takes place – there is one morphological marker for the Number$_{pl}$ function. The Number$_{sg}$ function, which selects only individual atoms from the denotation, does not mark number, but allows gender to be marked based on the noun that has been selected. The individual atoms within a German noun’s denotation carry the gender information, not the sets of atoms.

In this way, we can see how the result of the Number function in German results in the correct morphological marking, and why the features appear bundled. The denotation of nouns in German includes a specification for count/mass and a specification for gender. Only individual atoms within the denotation carry the gender information though, whether stable or unstable. For mass nouns, or those with denotations of unstable atoms, each atom within the denotation carries a gender feature which will be realised morphologically by a higher projection, likely the de-
The Features and Structure of the lower Noun Phrase

terminer or a modifying adjective. Mass nouns cannot be counted, so their gender should always be realised. For count nouns, only the individual stable atoms, not the sets of these atoms, carry gender. In the denotation, gender is not transmitted upwards through what we have represented as a join semilattice. If the Number$_{pl}$ function applies, only sets of atoms are selected and the gender information is lost. This explains why the grammar, specifically determiners and modifying adjectives, can no longer access the gender specification. If the Number$_{sg}$ applies, only the individual atoms are selected, which are those that carry the gender specification. This means that gender is realised in the singular and is available for agreement within the nominal phrase.

There is, therefore, a fundamental difference between how the marking of gender in German and Spanish affects the lexical denotation of nouns. In Spanish, the marking of gender in both specifications of number allows gender to be stored throughout the denotation of the noun, on atoms and sets of atoms. The marking on the noun that is specific to gender allows its acquisition, its proliferation throughout the denotation, and the creation of a functional head to realise the feature within the syntax. The Number function does not interact with gender at all. In German, the marking of gender in only the singular means that, in the denotation, it is only stored on individual atoms. Any function operating on individual atoms will allow gender to be preserved, but the Number$_{pl}$, which in effect removes the individual atoms, also removes any gender specification at the same time. The marking of gender in a language clearly affects how the gender feature is distributed in the denotation of a noun.
5.7.1 A mixed system?

Some languages, like French, have an overt manifestation of gender on a class of nouns, but not on other nouns. Several studies have identified very strong phonological cues to gender on singular nouns for certain noun endings. In French, (Tucker et al. 1977: p.37) found that nouns ending in /z/ and /e/ are, respectively, 94.2% and 90% masculine. This could be argued as strong enough evidence for a split system, as in Spanish. However, other noun endings provide a less reliable cue for a noun’s gender – around one in four French nouns ending in /i/ are masculine, for example. The lack of strong, Spanish-like systematicity in the gender marking of the singular, coupled with the lack of gender marking in the plural in French – there is one plural definite article in French, les, aligning more closely with German die than with Spanish los/las – provides strong evidence for the bundling model overall.

5.8 Conclusion

The aim of this chapter was to determine whether, in those languages that have grammatical gender, it exists in the system as a feature or a functional head. From a theoretical standpoint, all systems with grammatical gender have it operate as a feature, so the dichotomy we are debating is more accurately whether this feature projects through a functional head or not.

Gender presented a challenge for the semantically- or functionally-motivated syntax that has been proposed so far, given that grammatical gender is marked on nouns with no apparent semantic basis, but is still easily and quickly acquired in transparent systems. Acquisition studies found that this acquisition takes place based on phonological clues alone.
The Features and Structure of the lower Noun Phrase

We discussed existing proposals on both sides of the argument, contrasting the bundling model based on ideas from Ritter (1993) with the split model based on ideas from Picallo (1991), and found that both find good evidence for their individual approaches. Given that gender is treated as a feature in both, this gave rise to the possibility that gender appears as a functional head in some languages and not others.

As we found, there was no semantic motivation for gender, so providing the feature with a function was difficult. However, if the feature has no semantic motivation yet has clear systematic impact, then maybe a function can be motivated through the same types of reasoning. With Number, we acquire a functional head that applies a function and hosts number morphology. Even in the absence of a function, it seems that a functional head can still host gender morphology, but only where the marking uniquely reflects gender, as in Spanish. Specific markers for gender in Spanish lead the acquirer to create a functional head for gender, even though there is no true function applying to the denotation. Spanish therefore provides evidence for the split model, and gender as a functional head.

German does not have specific markers for gender, given a close interaction with number which results in the suppression of gender morphology in the plural. Without this independence, there is no motivation for a functional head for gender. Gender is only realised when a higher function applies, for example Number. This sheds light on how gender marking impacts the specification of gender within the denotation. Given that German only has gender marked in the singular, gender is only specified on the atoms in the denotation of the noun, such that if a set of atoms is selected, by the Numberₚₑ for example, gender is not carried forward as part of the output denotation. Spanish has gender marked regardless of number, so all atoms within the denotation are marked for gender.

This treatment of gender does not broach the question of gender agreement external
to the nominal phrase, i.e. agreement with verbs or coreferent pronouns, but it should not be incompatible with existing analyses. The proposal investigates the origin of gender within the nominal phrase and how this feature is realised in the syntax. Regardless of origin, the “DP-external” presence of gender can operate in the same ways (and encounter the same issues).

To conclude, the crosslinguistic variation observed in the gender system can be accounted for by how and where these features are created and exist within the grammar. There are undoubtedly more options to investigate, but, through varying both how the gender feature is stored in the lexicon and how it is realised in the syntax, there is flexibility to account for a wide range of realisations. The syntactic reality of gender must, however, be supported by overt morphology, whether through a higher functional head or its very own GenP.3

---

3Various phenomena discussed here have alternative morphological analyses in terms of affixation/derivation, extended exponence, impoverishment, etc. My approach relies on an acquisition-based generalisation such that the grammatical system allows “no expletive morphology”. This is based on work led by Ian Roberts as part of the University of Cambridge’s Rethinking Comparative Syntax project. Inflectional distinctions consistently encode formal features and morphology must encode something syntactically relevant. This dependence on rich agreement is an issue of learnability. If there is morphology, acquirers posit relevant features. If there are no features, there is no morphology. If there is no morphology, acquirers do not know which features are present and sometimes need to rely on other evidence.
Chapter 6

Gender in the absence of a noun

6.1 Introduction

In the previous chapter, we proposed gender as a feature that is included as part of the lexical specification of the noun. This gender feature is acquired as part of the acquisition of a noun and is realised in the syntax through some functional head, be it a specialised Gender head or another functional projection in the nominal phrase.

However, there are contexts where gender seems to exist in the absence of noun. In these cases, where can the gender specification be said to come from?

6.2 Evidence from German and Dutch

In German, all adjectives can be used as nouns (Durrell 2011):
In these examples, we can see the adjective *alte* ‘old’ being used as a noun. The genders shown on the definite determiners in (110-a), (110-b) and (110-c) are the only indication of how the phrase should be interpreted in terms of gender. Masculine and feminine “adjectival nouns” tend to have human referents, whereas neuter adjectival nouns usually refer to abstract or collective entities. The adjective can also be marked as plural, as in (110-d). As previously discussed, gender is only specified in the singular, but the plural has an interpretation of a set of human referents, where it could as easily have, in this example, had the interpretation of ‘the old things’.

A similar phenomenon is found in Dutch. Here, there is only a two-way distinction in gender with no distinction between the masculine and the feminine (Donaldson 1981):

(111) a. de rijke
    ‘the rich-person’

b. de blinde
    ‘the blind-person’

c. de rijken
    ‘the rich(-people)’

d. de blinden
    ‘the blind(-people)’
e. het mooie  
   ‘the nice-thing’

f. het stomme  
   ‘the stupid-thing’

Again, the adjectives preceded by a neuter article refer to abstract entities and the adjective can be inflected as plural.

In each of these examples, the definite determiner is indicating the gender of the referent – nothing is revealed by the adjective. As discussed earlier, gender is understood to be part of the lexical specification of the noun, so where is this gender specification coming from? There is seemingly no noun to provide this feature. One possible explanation is that there is, in fact, a null noun in the structure that provides the gender feature.

6.3 The null noun in Dutch

Kester (1996), building on work by Lobeck (1993, 1995), extends the distribution of the non-lexical pronoun *pro* to also involve the empty categories found in elliptical constructions, as well as the standard instances of null subjects in languages like Italian. This means that *pro* can correspond both to a DP with referential content, as in Italian null subjects, and to an N which is non-referential, as in elliptical constructions. Kester (1996) employs this approach in an account for null nouns in adjectival contexts in Germanic and Romance languages.

In her analysis, Kester (1996) relies on the presence of inflectional morphology in Dutch. Dutch adjectives are inflected by the morpheme *-e* in both the adjectival noun contexts, as can be seen in (111), and ellipsis. This proposal is supported by the many languages in which inflection and morphologically ‘rich’ agreement seem
to be important for the licensing and identification of null nouns.

The morpheme -e is realised with a null noun where it would not be with a lexical noun:

(112) a. de besproken person
    the talked-about person
    ‘the person talked about’

   b. de besprokene pro
    the talked-about
    ‘the person talked about’

The morpheme -e, therefore, is not an example of standard adjectival agreement and must surface for some other reason.

To give the interpretations available with adjectival nouns, Kester (1996) claims that pro has inherent features. There are three possible feature specifications for pro:

(113) a. [ +human, +plural ]
   b. [ +animate, +count ]
   c. [ –animate, –count ]

A pro with the features in (113-a) appears in constructions like blinden pro ‘blind (people)’ and gives rise to the special adjectival morphology exhibited. The ending -n is said to license pro and make its features visible for interpretation.

The two types of pro in (113-b) and (113-c) are used in singular cases. Kester (1996) follows two lexical redundancy rules of grammatical gender in Dutch, such that the feature specification in (113-b) reduces to [–neuter] and that in (113-c) to [+neuter]. This means that the pro in (113-b) occurs with the non-neuter definite article de.
The features and structure of the lower noun phrase

The pro in (113-c) occurs with the neuter definite article het. It is the grammatical gender features expressed by the definite articles that make the different features of pro visible for interpretation.

This analysis follows the proposal that the gender specification comes from the noun, but it depends on the feature specification of pro in the lower nominal position. The suggested feature specifications on pro correctly predict the resultant overt realisations, but do not rely on a gender feature at all. Can it truly be the case that gender is realised on the basis of other features? Might this stand as evidence for a lack of gender, or a gender driven through the semantic content of other features, in Dutch?

6.4 The null noun in German

Can the proposals of Kester (1996) be applied to adjectival nouns in German? German differs from Dutch in the fact that there is a three-way distinction in the gender system in the singular, whilst plural adjectival nouns remain possible:

(114) a. der Alte
   ‘the old-man’

b. die Alte
   ‘the old-woman’

c. das Alte
   ‘the old-thing’

The pro in (114-d) is the same as that in (113-a), with features [+human, +plural]. According to the proposal, the adjectival morphology – the ending -n – may license
pro and make it visible for interpretation. There is therefore a similarity between Dutch and German in terms of the realisation of gender with plural number.

The pro in (114-c) is the same as that in (113-c) and is similarly used to denote an abstract entity with neuter gender.

For (114-a) and (114-b), we have a clear reflection of gender in the interpretation – this is not strictly grammatical gender given that there is a semantic interpretation. It must be the case, however, that German has an extra pro with a feature contrast not available in the Dutch system that splits (113-b) into masculine and feminine, rather than non-neuter.

This difference between German and Dutch raises questions around the relation between grammatical gender and interpretable gender, and whether they have different realisations in the grammar. As demonstrated by the determiners, the grammatical gender follows from the interpretable gender, or perhaps the interpretable gender is only interpretable from the marking of grammatical gender. We will return to the presence of semantically-motivated gender in German, and discuss whether gender in Dutch can be reduced to a combination of other features as suggested.

### 6.5 A potential incompatibility

In the previous chapter, we found that the structure of the nominal extended projection for German included a NumP dominating N, but no GenP. This is represented in (115):

\[
(115) \quad [\text{DP} \ldots [\text{NumP} \ldots \text{Num} [\text{NP} \ldots \text{N}]]]
\]

Although potentially motivated by different features, the gender marking in Dutch
works in the same way.

Kester (1996) claims that the feature specification of *pro* in constructions like the Dutch *blinden pro* ‘blind (people)’ is [+human, +plural]. The *pro* should correspond to N in the above structure, but its featural specification inherently includes plural. This should be provided by the higher functional head Num. The proposal of *pro* as having a bundle of features therefore needs closer examination.

### 6.6 Reconciling *pro* with NumP and gender

According to Kester (1996), the structure of *blinden pro* ‘blind (people)’ is:

\[(116) \quad \begin{array}{c} \text{DP} \\
\text{D'} \\
\text{D} \\
\text{FP} \\
\text{AP} \\
\text{blinden} \\
\text{F'} \\
\text{F} \\
\text{NP} \\
\text{N} \\
\text{pro} \end{array} \]

This follows Cinque (1994) by placing prenominal adjectives in the specifier positions of different functional projections.

Our structure of functional projections has been based on overt morphology and the order in which functions must apply to a denotation. NP is dominated by NumP, which applies the Number function. After this, the Cardinality function can
be applied by numerals. At the top, we have assumed the DP which situates the identified referent of the phrase within a wider context. Adjectives are usually found in a position lower than a numeral.

This ordering mirrors a standardly assumed structure of the nominal phrase as proposed by Julien (2002):

(117)

\[
\begin{array}{c}
\text{DP} \\
\text{D} \quad \text{CardP} \\
\text{Card} \quad \text{AgrP} \\
\text{Agr} \quad \text{NumP} \\
\text{Num} \quad \text{NP} \\
\text{N}
\end{array}
\]

Adjectives in this structure form adjective phrases that are hosted in the specifier position of AgrP. AgrP can be recurring, with each instance hosting one adjective in its specifier. This proposal also follows Cinque (1994) by placing adjectives in the specifiers of functional projections.

Whilst these two structures are not incompatible, some intricacies of how \textit{pro} interacts with each projection should be discussed.

6.6.1 Problems with the plural

Consider the following contrasting constructions:
The Features and Structure of the lower Noun Phrase

(118) a. Ik heb aardige zieken pro bezocht
    I have nice sick.PL (pro) visited
    ‘I have visited nice sick people’

b. Ik heb aardige zieke mannen bezocht
    I have nice sick man.PL visited
    ‘I have visited nice sick men’

Isolating the nominal phrases, we have:

(119) a. aardige zieken pro
    nice sick.PL (pro)
    ‘nice sick people’

b. aardige zieke mannen
    nice sick man.PL
    ‘nice sick men’

With a lexical noun, as in (119-b), the plural morpheme -n is realised as a suffix on the noun. This is as we expect. The noun enters the lexicon with a count denotation and a specification of common gender.\(^1\) The functional head Num merges next, which applies the Number\(_{\text{pl}}\) function to select only sets of atoms. The plural marker is also hosted by this functional head.

Above this, the adjective zieke is merged. As previously discussed, lexical projections provide a denotation with which the existing denotation intersects. The adjective therefore selects from the denotation of mannen, i.e. the denotation of man restricted to pluralities, those sets that are within the intersection with the denotation of sick. The process is repeated for the adjective aardige.

Assuming that gender exists only on individual atoms in the denotation of full nouns, as it does in German, then this structure is built according to the mechanisms we have already outlined.

\(^1\)For the purposes of this discussion, I refer to the genders of Dutch as “neuter” and “common”. Common gender includes masculine and feminine when the distinction is not being made.
We can also consider Kester’s (1996) proposal of pro along the same lines. The key difference between the realisations that needs to be explained, given the lack of gender in the plural, is the appearance of number marking on the adjective. However, we also need to define how pro might work in our denotation-/function-driven approach, if at all.

6.6.2 Defining “plural” pro

Recall that Kester (1996) claims three types of pro with different feature specifications.

(120) a. [ +human, +plural ]
    b. [ +animate, +count ]
    c. [ –animate, –count ]

The feature specification of pro in (120-a) is used for referring to plural sets of humans, such as in (119-a): zieken. This feature specification is not compatible with our functional structure though, as plural seems to be specified on pro itself. To correct this, could we reduce the feature specification of this pro to only include a [+human] feature? This, as a feature, would exist on the lexical specification of a noun. For example, consider the gender system as far as it exists in English – on pronouns. This could be said to rely on the existence of the [+human] feature in the specification of a noun (or perhaps [+animate]). A similar system may exist for this type of Dutch pro.

The denotation of this pro, then, is the set of all things that are [+human]. This pro is the base for the nominal projections in the following examples:
This does not yet answer the question of how the adjective gets marked for number, as in (121-a) and (121-b). In fact, no additional mechanisms need to be proposed to explain the number marking on the adjective. Given that this pro has a denotation, it behaves within the structure as a full noun, despite its lack of phonetic content.

This pro enters the lexicon with its specification [+human]. The functional head Num merges next, which applies the Number_{pl} function to select only sets of atoms. The plural marker is also hosted by this functional head. At this point, there is no overt element to which this plural marker can attach.

Above this, the adjective rijke is merged and its denotation intersects with the existing denotation. The adjective therefore selects from the denotation of sets of [+human] atoms those sets that are within the intersection with the denotation of rijke. The adjective also provides overt phonetic content to which the plural marker can attach, resulting in the form rijken with the correct denotation.

The parallel can be demonstrated as in (122):

(122) a. *rijke+mann+en*
    b. *rijke+∅+en*

As we allowed the acquisition of gender without semantic content, it seems that we have evidence for a category pro in the absence of phonetic content. The denotation of [+human] must originate from some element, and this element has no overt
realisation. [+human] is a feature on nouns, so we have no choice but to propose a phonetically-null element to provide this semantic information.

### 6.6.3 Problems with neuter pro

We have specified one type of pro within our approach, the one used to refer to groups of people, but what about the other two types of pro proposed by Kester (1996)?

In the singular, the gender distinction is making a distinction between humans and things:

\[(123)\]

a. de besprokene pro
   the talked-about (pro)
   ‘the talked-about person’

b. het besprokene pro
   the talked-about (pro)
   ‘the talked-about thing’

c. de besprokenen pro
   the talked-about (pro)
   ‘the talked-about (people)’

d. *het besprokenen pro
   the talked-about (pro)
   ‘the talked-about things’

As can be seen from (123-c) and (123-d), only the human interpretation is available in the plural. This relies on the pro defined in the previous section. Given that pro has been separated from number, contra Kester (1996), there is no justification for this same type of pro not also being the noun in structures like that in (123-a). The denotation of this pro is appropriate, and the Number\textsubscript{sg} function restricts this denotation to only individual atoms.
The pro that refers to things, that exhibits neuter gender, is more restricted in its usage. Firstly, a neuter adjectival noun must be preceded by the definite article het:

(124)  
\begin{align*}
\text{a. } & \text{het besprokene } pro \\
& \text{the talked-about (pro)} \\
& \text{‘the talked-about thing’}
\end{align*}

\begin{align*}
\text{b. } & \text{*een besprokene } pro \\
& \text{a talked-about (pro)} \\
& \text{‘a talked-about thing’}
\end{align*}

\begin{align*}
\text{c. } & \text{*dit besprokene } pro \\
& \text{this talked-about (pro)} \\
& \text{‘this talked-about thing’}
\end{align*}

\begin{align*}
\text{d. } & \text{*besprokenen } pro \\
& \text{talked-about (pro)} \\
& \text{‘talked-about things’}
\end{align*}

Another difference with neuter pro is that it doesn’t allow multiple adjectives unless they are coordinated, unlike the common gender example in (119-a):

(125)  
\begin{align*}
\text{a. } & \text{*het mooie goed pro} \\
& \text{the beautiful good (pro)} \\
& \text{‘the beautiful good thing’}
\end{align*}

\begin{align*}
\text{b. } & \text{het mooie en goede } pro \\
& \text{the beautiful and good (pro)} \\
& \text{‘the beautiful and good thing’}
\end{align*}

This pro is behaving much less like a full noun. It does not allow a range of determiners, only the definite article, and it will permit only one adjectival modifier. It cannot appear in the plural, as in (124-d). Returning to our structure in (117). The only functional projection above N that we have not yet ruled out or restricted is “CardP”.

162
Returning to Kester’s (1996) feature specifications for pro, the neuter pro had the specification [–animate, –count]. As –count, the Cardinality function would be incompatible, given that there would be no stable atoms individuated in the denotation to be counted.

With common gender (human) pro, numerals are possible:

(126) twee blinden pro  
    two  blind  (pro)  
    ‘two blind people’

Much of the restriction on the nominal phrase with neuter pro can be explained by considering the denotation of this null noun.

However, the neuter singular adjectival nouns also appear in a “quality” construction:

(127) a. het onvoorspelbare pro  in z’n karakter  
    the unpredictable (pro) in his character  
    ‘the unpredictable thing in his character’

b. dit onvoorspelbare pro  in z’n karakter  
    this unpredictable (pro) in his character  
    ‘this unpredictable thing in his character’

c. het vaak geobserveerde onvoorspelbare pro  in z’n karakter  
    the often observed unpredictable (pro) in his character  
    ‘the often observed unpredictable thing in his character’

d. *een onvoorspelbaar pro  in z’n karakter  
    a  unpredictable  (pro)  in  his  character

Here, we have the well-established example with the definite article het. (127-b) provides a grammatical example with the demonstrative dit, although the demonstrative tends to have a discourse function rather than a deictic function here (as brought to my attention by Ad Neeleman, p.c.). (127-c) demonstrates that fur-
ther attributive modification is permitted, and that the additional modifier can be phrasal.

The crucial example is (127-d). An indefinite would be semantically felicitous in this context; “an unpredictable thing in his character” makes sense. However, (127-d) is ungrammatical. The adjective in the indefinite neuter form lacks the final “–e”, as expected from the inflectional paradigm for attributive adjectives. This suggests that the –e on the other licensed nominalised adjectives is the regular attributive ending, but also that the presence of this ending could somehow license pro.

### 6.6.4 Defining neuter pro

The behaviour of neuter pro is as follows: it cannot appear in the plural, it can be modified by more than one adjective in the quality construction, it cannot be counted by a numeral, and it can only appear with definite determiners like het and dit.

Some of this behaviour can be predicted based on the denotation of neuter pro that is introduced to the structure when it is selected. Firstly, the lexical specification of this pro needs to include the [-human] feature. When using this pro, an interpretation of a human referent is never available. Also, the interpretation of this pro does not include any information with regard to the count/mass status of its referent.\(^2\)

With no count/mass information, we must assume that the neuter pro contributes a flexible denotation to the structure, including both stable and unstable atoms. For this reason, the Number function cannot apply. Recall that the Number function requires a denotation of only stable atoms from which it picks either individuals or sets. This is not the denotation provided by neuter pro. If the Number function

---

\(^2\)In my glosses, I have translated this pro as “thing”, but this is not entirely accurate. Het besprokene more accurately translates simply to “the talked-about stuff” with no commitment to a count or mass interpretation.
cannot apply, then the Cardinality function cannot apply, also ruling out the appearance of numerals. Furthermore, if the Number function cannot apply, then certain determiners are also ruled out. This explains the ungrammaticality of (124-b) with *een “a” which requires a denotation of individual atoms. Given that *het and *dit are compatible with both count and mass neuter nouns, both are expected to be licensed in these constructions.

The only part of the behaviour of neuter pro yet to be accounted for is the restriction to one adjective. To explain this, we return to Cinque’s (1994) proposal that prenominal adjectives are merged in the specifier positions of different functional projections. With neuter pro, there is a complete lack of functional projections, given that the denotation of the null noun prevents any functions from applying. The one permitted adjective is likely to merge at some lower functional level dominating N that I have not discussed, such as n (Dobrovie-Sorin 2012).

In terms of defining pro in these adjectival noun constructions, we have one type that is specified as [+human] and has a count denotation and another type that is specified as [–human] and has a flexible denotation. All other behaviour in these null noun constructions should be predictable from the denotation.

6.7 Gender or ±human

One final question with regard to pro is that we have reduced gender to a ±human distinction. This still does not explain how the articles are realised in specific gender forms. The ±human distinction is the dividing line for these null noun constructions, but there must also be a gender feature somewhere that is providing the source of the agreement for the articles.

This, of course, only happens in the singular, as the plural article does not show
gender (and the neuter pro does not occur in the plural anyway). Given that there is not a one-to-one mapping of gender on pro to ±human on nouns, i.e. there are non-human nouns of common gender that would have to be referred to with neuter pro, it is difficult to see the correlation is formed. One option may be that acquirers reverse-engineer the gender of pro based on the determiners it appears with in the input. In this way, the links between +human and common gender and –human and neuter gender may form, the genders can be added to the specification of each pro and the determiners find a source of gender specification with which they can agree.

6.8 Conclusion

In this chapter, we considered an interesting case of gender in the absence of an overt noun. In both German and Dutch, there is a construction in which it seems adjectives can act like nouns in the absence of a noun. The form of the determiner and the interpretation of the nominal demonstrates that there is a gender specification for the phrase, even though there is no noun. This contradicts the previous chapter in which we claimed gender to be lexically-specified on the noun.

We considered the proposals of Kester (1996) which claimed that these types of constructions do include a noun – the null noun pro. There were three types of pro with different feature specifications which gave rise to the morphology exhibited and the gender found on the articles.

However, the feature specifications on pro were incompatible with the functional structure I propose. Namely, it was impossible for pro to be specified for number when this specification takes place in the higher functional Num projection.

This led to an examination of pro as a bundle of features, and we looked more
closely at the behaviour of pro in these “adjectival noun” constructions. What became apparent is that pro was still contributing some semantic content to the interpretation of the phrase, meaning that it must have a denotation like any other noun. The differences in behaviour of the types of pro in these constructions could then be reduced to a difference in the denotation of these null nouns.

The pro used to refer to humans has a count denotation and is specified as +human. The pro used to refer to things has a flexible denotation and is specified as –human. The differences in the denotations predicted all the differences in behaviour observed. To link back to gender, the acquirer creates the link between the determiner used and the ±human specification.

In terms of number marking, the Number function could apply to the denotation of pro in the same way it would for any noun. The difference is that the number morphology needs an element to attach to. This was found in the next functional projection up, the modifying adjective.

These constructions can therefore be accounted for through the standard mechanisms of my approach. Whether pro operates in this way, with its own denotation, more widely in cases where ellipsis is not taking place is an interesting question for future research.
Chapter 7

Conclusion

The features and structure of the lower noun phrase show a remarkable amount of crosslinguistic variation, as we have seen from a small number of languages. The lower noun phrase has been seen to contribute the same types of semantic information in different languages, but this is achieved in different ways and along different lines. This thesis demonstrates that, whilst the variation is widespread, it can be accounted for using a single model that allows for different grammars to be created during acquisition based on evidence in the input. In the comparison of different phenomena, we reviewed previous approaches which were found to have fallen short in their coverage or were theoretically incompatible with the analyses provided for the same phenomenon in another language. Through the case studies, an account for a range of phenomena and the variation within them can be given that should demonstrate the strength of the model.

The primary research goal of this thesis was to ascertain the motivation for the creation of features on nouns and throughout the nominal phrase. I have proposed that formal features in the grammar are not universal, provided by some inherent linguistic knowledge, but are created as part of the language acquisition process.
Language-learners are sensitive to systematicity in the input and can recognise regular patterns in interpretation, in marking, or in the correlation between both. I proposed that the transparency of the systematicity, relative to the complexity of the system overall, determine the ease and pace of acquisition of these features. This was supported by the experimental findings comparing the acquisition of the count/mass distinction and number in English with the accurate production of classifiers in Mandarin Chinese. The more complex system of Mandarin Chinese led to later acquisition of the distinction than in English, which was deemed to have a more transparent system. By comparing how two different languages choose to reflect the same perceptual distinction in their grammar, it was possible to examine how the differences in the input may contribute to how a feature is acquired.

Recall the count/mass distinction, English and Mandarin Chinese reflect the individual/non-individual divide differently in their grammar. Based on the transparency of the English number marking system, the count/mass distinction is acquired early, before the inventory of nouns is very large. The early evidence suggests that count/mass information is always stored on nouns and correlates with the individual/non-individual divide. This means that the denotations of nouns, as they are acquired, are structured as either count or mass. As new nouns are acquired, they fall into one of these two camps based on the evidence of their marking, suggesting that transparent marking is a preferred type of evidence to interpretation. This means that mismatches are possible in the system, as English does not have a perfect one-to-one mapping between the count/mass distinction and the individual/non-individual divide. Based on the complexity of the classifier system, Mandarin Chinese acquirers display later understanding of a grammatical reflex of the perceptual individual/non-individual divide. This means that nouns can be acquired based on their interpretation rather than their marking for this information. Bare nouns are not marked for this information, but a difference in their denotation is created based on the interpretation alone. This means that mismatches are not possible in Mandarin Chinese. The
differences in the denotations of the nouns are a result of the transparency of the marking and the pace of acquisition. Although English has a more transparent system which speeds up acquisition of the feature, the feature is immediately put to work in further acquisition which allows “mistakes” to be made.

Features were not reserved exclusively for nouns though. As demonstrated with English number, features with their own system of marking were proposed to receive their own functional head to host the morphology and to ensure that functions apply to the denotation of the noun in the right order. The nominal phrase is motivated by the need to identify a referent, and functions apply to the denotation of the noun to narrow the range of possible referents. Features stored on the noun, like count/mass information, become part of the lexical specification of that noun, but features can be built into a nominal during construction of a phrase. There is variation in how this takes place also. Recall the comparison of gender in Spanish and German. Nouns in Spanish are marked for gender by a marker that only signals gender – gender is marked as an independent feature, so receives an independent functional head. In German, the marking of gender relies on the number specification of the noun. As a feature then, gender in German can only be marked through some interaction with the application of the Number function. This was accounted for by returning to the denotation of the noun and specifying that gender information is only retained in the denotation if the Number function selects singular. The interaction of lexically-specified features and the functions above them explains the difference in marking here.

The proposed approach provides the starting point for a model that has the potential to be powerful for a number of reasons.

1. The model is functionally-driven – nothing can exist within the grammar without a full functional motivation. This greatly constrains the system and aims to prevent any theory-motivated devices.
2. Features within the grammar are not universal, they are created by the acquirer based on evidence within the input. A grammar based on the input will preserve those features that are transparently marked, but features with weaker marking may be candidates for loss or reanalysis.

3. Evidence for feature acquisition takes various forms. Findings suggest that the strongest evidence for a feature is morphological marking, but formal features can be also acquired through systematic variation in interpretation.

4. The model accounts for crosslinguistic variation. The differences in the acquisition of features, and therefore the grammar, could be attributed to the strength of evidence for the acquirer.

There is a wide range of variation that still needs to be explored, but the potential of functionally-driven emergent features should be evident.

There are three obvious avenues for future research. One is to focus on the semantic side of the discussion and investigate the nature of denotations and how functions might apply in more technical detail. I proposed a distinction between lexical and functional projections that needs to be verified against more than prenominal adjectives. The second avenue is to develop the weight of evidence supplied by acquisition experiments and to compare the nature of “transparency of systematicity” in the input. It would be valuable to look at both the relative ordering of the acquisition of different features in one language and the relative ordering of the acquisition of the same feature in different languages, compared with the input. Finally, it would be interesting to investigate whether there is a comparative model available to the clausal domain. I propose that there is, given the well-researched similarities between the two extended projections, but exploring if the denotation of a verb works in the same way and how any acquired functions might apply to it would certainly be of great value in uniting the domains and providing support for the approach.
Bibliography


Biberauer, Theresa. 2011. In defence of lexico-centric parametric variation: two


Chierchia, Gennaro. 1998a. Plurality of mass nouns and the notion of “semantic parameter”. Events and Grammar 70. 53–103.


Clark, Eve V. & Tatiana V. Nikitina. 2009. One vs. more than one: Antecedents to plurality in early language acquisition. *Linguistics* 47. 103–139.


The Features and Structure of the lower Noun Phrase


Krifka, Manfred. 1995. Common nouns: A Contrastive Analysis of Chinese and
The Features and Structure of the lower Noun Phrase


Maiden, Martin. 2016. Number. In Adam Ledgeway & Martin Maiden (eds.), The
The Features and Structure of the lower Noun Phrase


Pires de Oliveira, Roberta & Susan Rothstein. 2011. Bare singular noun phrases are mass in Brazilian Portuguese. *Lingua* 121. 2153–75.


Sundaresan, Sandhya. 2015a. Lecture 1: Features - an overview. IGRA 08: Agree(ment) and the nature of syntactic features, Universit¨ at Leipzig, 9 April.

Sundaresan, Sandhya. 2015b. Lecture 2: Agreement - an overview. IGRA 08: Agree(ment) and the nature of syntactic features, Universit¨ at Leipzig, 9 April.


Appendix – Questionnaire

Respondents were asked to judge the acceptability of the following phrases. Whenever possible, these phrases were taken or adapted from the Pinyin portion of the Lancaster Corpus of Modern Chinese\(^1\) and transcribed into Chinese characters by a native speaker, along with questionnaire instructions. The examples were chosen to investigate the interaction of several binary parameters of variation:

\[(128)\]

\[
\begin{array}{|c|c|}
\hline
\text{Either...} & \text{or ...} \\
\hline
\text{a. Classifier} & \text{Measure word} \\
\hline
\text{b. Count noun} & \text{Mass noun} \\
\hline
\text{c. Adjective before CL/MW} & \text{Adjective before noun} \\
\hline
\text{d. Delimitive adjective} & \text{Other adjective} \\
\hline
\text{e. Adjective before or after CL/MW} & \text{Adjective before and after CL/MW} \\
\hline
\text{f. Unmodified adjective} & \text{Adjective modified by } \text{hen} \text{ “very”} \\
\hline
\end{array}
\]

The examples provided were as follows.

\(^1\)McEnery, Tony and Xiao, Richard. (2004). The Lancaster Corpus of Modern Chinese. Lancaster University
<table>
<thead>
<tr>
<th>Feature Structure</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCM</td>
<td>three waters</td>
</tr>
<tr>
<td>NCC</td>
<td>three words</td>
</tr>
<tr>
<td>NMM</td>
<td>two cups of tea</td>
</tr>
<tr>
<td>NMC</td>
<td>two rows of teeth</td>
</tr>
<tr>
<td>NDCM</td>
<td>eight big snows</td>
</tr>
<tr>
<td>NDCC</td>
<td>two small books</td>
</tr>
<tr>
<td>NDMM</td>
<td>two small pieces of charcoal</td>
</tr>
<tr>
<td>NDMC</td>
<td>two big stacks of books</td>
</tr>
<tr>
<td>NOCM</td>
<td>two sour milks</td>
</tr>
<tr>
<td>NOCC</td>
<td>three stupid people</td>
</tr>
<tr>
<td>NOMM</td>
<td>two tasty cups of coffee</td>
</tr>
<tr>
<td>NOMC</td>
<td>two talented groups of players</td>
</tr>
<tr>
<td>NhDCM</td>
<td>four very small milks</td>
</tr>
<tr>
<td>NhDCC</td>
<td>two very big people</td>
</tr>
<tr>
<td>NhDMM</td>
<td>eight very small cups of coffee</td>
</tr>
<tr>
<td>NhDMC</td>
<td>four very big groups of players</td>
</tr>
<tr>
<td>NhOCM</td>
<td>three very dangerous snows</td>
</tr>
<tr>
<td>NhOCC</td>
<td>eight very interesting books</td>
</tr>
<tr>
<td>NhOMM</td>
<td>four very dry pieces of charcoal</td>
</tr>
<tr>
<td>NhOMC</td>
<td>two very cheap stacks of books</td>
</tr>
<tr>
<td>Feature and Structure</td>
<td>Example 1</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>NCDM</td>
<td>four small milks</td>
</tr>
<tr>
<td>NCDC</td>
<td>eight little eyes</td>
</tr>
<tr>
<td>NMDM</td>
<td>three cups of big coffee</td>
</tr>
<tr>
<td>NMDC</td>
<td>two pairs of big hands</td>
</tr>
<tr>
<td>NCOM</td>
<td>three dangerous snows</td>
</tr>
<tr>
<td>NCOC</td>
<td>two new passwords</td>
</tr>
<tr>
<td>NMOM</td>
<td>200ml of dilute hydrochloric acid</td>
</tr>
<tr>
<td>NMOC</td>
<td>two pairs of new shoes</td>
</tr>
<tr>
<td>NChDM</td>
<td>very big snows</td>
</tr>
<tr>
<td>NChDC</td>
<td>eight very small books</td>
</tr>
<tr>
<td>NMhDM</td>
<td>two pieces of very small charcoal</td>
</tr>
<tr>
<td>NMhDC</td>
<td>four stacks of very big books</td>
</tr>
<tr>
<td>NChOM</td>
<td>four very sour milks</td>
</tr>
<tr>
<td>NChOC</td>
<td>two very interesting people</td>
</tr>
<tr>
<td>NMhOM</td>
<td>three cups of very tasty coffee</td>
</tr>
<tr>
<td>NMhOC</td>
<td>two groups of very talented players</td>
</tr>
<tr>
<td>NDCOM</td>
<td>two big sour milks</td>
</tr>
<tr>
<td>NDCOC</td>
<td>one small red cherry</td>
</tr>
<tr>
<td>NDMDM</td>
<td>eight big pieces of small charcoal</td>
</tr>
<tr>
<td>NDMDC</td>
<td>four small stacks of big books</td>
</tr>
</tbody>
</table>
The Features and Structure of the lower Noun Phrase

<table>
<thead>
<tr>
<th>NDMOM</th>
<th>four big cups of tasty coffee</th>
<th>四大杯可口咖啡</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDMOC</td>
<td>three small groups of talented players</td>
<td>三小组天才球员</td>
</tr>
<tr>
<td>NOCDM</td>
<td>three dangerous big snows</td>
<td>三危险个大雪</td>
</tr>
<tr>
<td>NOCDC</td>
<td>eight interesting little books</td>
<td>八有趣本小书</td>
</tr>
<tr>
<td>NOCOM</td>
<td>four dangerous sour milks</td>
<td>四危险个酸牛奶</td>
</tr>
<tr>
<td>NOCOC</td>
<td>two talented interesting people</td>
<td>两天才各有去人</td>
</tr>
<tr>
<td>NOMDM</td>
<td>two tasty cups of small coffee</td>
<td>两可口杯小咖啡</td>
</tr>
<tr>
<td>NOMDC</td>
<td>two talented groups of big players</td>
<td>两天才组大球员</td>
</tr>
<tr>
<td>NOMOM</td>
<td>three cheap cups of tasty coffee</td>
<td>三低廉杯可口咖啡</td>
</tr>
<tr>
<td>NOMOC</td>
<td>two new pairs of cheap shoes</td>
<td>两心双低廉鞋</td>
</tr>
</tbody>
</table>

The codes in the first column indicate the interacting categories:

- Presence of (N)umeral
- Presence of “(h)en” – adjective modifier (very)
- (D)elimitive or (O)ther adjective
- (C)lassifer or (M)easure word
- Presence of “(h)en” – adjective modifier (very)
- (D)elimitive or (O)ther adjective
- (C)ount or (M)ass noun