Heinrich Heine Universität Düsseldorf

How to count stuff

Psycholinguistic Investigation of Noun Countability

Inaugural-Dissertation
zur Erlangung des Doktorgrades der Philosophie (Dr. phil.)
durch die Philosophische Fakultät der
Heinrich-Heine-Universität Düsseldorf
vorgelegt von

Natalja Selina Peiseler
geborene Beckmann

aus Sprockhövel

Hauptprüfer: Prof. Dr. Dr. Peter Indefrey
Zweitgutachter: Prof. Dr. Ruben Van de Vijfer

SFB 991
Integriertes Graduiertenkolleg
„Structures of Representation"
an der Philosophischen Fakultät
SToRE SFB 991

Abstract

Noun countability is a linguistic phenomenon that allows speakers to grammatically mark the conceptual distinction between entities that are countable and entities that are not. A given noun can be combined with a certain set of determiners and quantifiers in order to form a grammatical noun phrase. This set of determiners is different for countable nouns (1) than for uncountable nouns (2). So-called dual nouns (3) occur with both sets of determiners and slightly change interpretation depending on which syntactic context they occur in.

(1) Count Nouns
   a. ein Bus, zwei Busse, viele Busse
      a bus, two busses, many busses
   b. *viel Bus, *etwas Bus, *kaum Bus
      much bus, some bus, barely any bus

(2) Mass Nouns
   a. viel Sand, etwas Sand, kaum Sand
      much sand, some sand, barely any sand
   b. *ein Sand, *zwei Sande, *viele Sande
      a sand, two sands, many sands

(3) Dual Nouns
   a. ein Tuch, zwei Tücher, viele Tücher
      a cloth, two cloths, many cloths
   b. viel Tuch, etwas Tuch, kaum Tuch
      much fabric, some fabric, barely any fabric

This noun property has sparked the interest of linguists and cognitive researcher in recent years, not in the least because of the fascinating interplay between grammar and conceptual (world-) knowledge it represents, as the grammatical distinction between countable and uncountable nouns is often reflective of a conceptual distinction between entities that represent objects (in a sense) and substances (in a sense). This property has been termed boundedness, individuatedness, objecthood or atomicity and is said to be applicable not only to physical objects and substances, where such a difference is more or less easy to establish, but also to abstract entities.
Interestingly, it is precisely this property that can be purposefully changed when inserting a noun into an incongruent noun phrase, which provokes a shift in interpretation:

(4) Countability Shifts
   a. MASS NOUN COMBINED WITH COUNT-SPECIFIC DETERMINERS
      We'd like **two waters** and **a juice**, please.
   b. COUNT NOUN COMBINED WITH MASS-SPECIFIC QUANTIFIER
      After the accident there was not **much car** left.

This dissertation investigates whether countability is a lexical property of nouns and how it is utilised during phrase comprehension. Three reaction time experiments were conducted that measured participants' reactions to nouns in three different environments:

(i) a congruent condition where the noun is preceded by a determiner that fits its countability class, for example *ein Bus, a bus*, and *viel Sand, much Sand*

(ii) an incongruent condition where the noun is preceded by a determiner that does not fit its countability class, for example **viel Bus, much bus,* and **ein Sand, a Sand*

(iii) a neutral condition where the noun is preceded by a segment of brown noise which does not hold any linguistic information

As trials of the congruent condition yielded shorter reaction times than trials of the incongruent condition, the experiments showed that countability is in fact a lexical property, that is in the mental representation of every German noun it is lexically specified whether that noun is a countable or an uncountable noun.

The experiments deployed different tasks: A lexical decision task, where participants decided if the noun in the noun phrase was a real German word or not, and a phoneme monitoring task, where participants decided whether a certain target phoneme was part of the noun. The lexical decision task is sensitive to the processing stages of lexical access and to post-lexical phrase integration processes, while the phoneme monitoring task is only sensitive to stages of lexical access.
The comparison of results obtained with the two tasks thus allowed for a precise location of the effect during noun phrase comprehension.
The first hypothesis how countability could be utilised entailed that the processing of a count-specific or mass-specific determiner would give rise to an expectation regarding the countability value of the upcoming noun. This expectation would be realised as the limitation of the search space (or cohort) of possible noun candidates that would be considered for lexical selection.

By observing a countability congruency effect with the lexical decision task, but not with the phoneme monitoring task, it could however be established that a different hypothesis prevailed. Rather than processes of lexical access, countability effects later, post-lexical stages of phrase integration. An incongruency is detected after the noun is lexically selected during the build-up of a noun phrase and a grammaticality monitoring or checking mechanism.

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der ‚Ordnung über die Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität' erstellt worden ist.

Sprockhövel, den 20.02.2023

Natalja S. Peiseler

Acknowledgments

The work on this dissertation has been a fascinating yet challenging journey that was made possible by the continuous support of many people.

My research has been part of the collaborative research centre CRC 991 “The Structure of Representation in language, Cognition and Science” at Heinrich-Heine-University Düsseldorf and was funded by the **Deutsche Forschungsgesellschaft** (DFG).
My work was part of Project C03 “Conceptual Shifts: psycholinguistic evidence” within the CRC 991 and I have greatly profited from the supportive infrastructure that was provided. I would like to thank **Prof. Dr. Alexander Ziem**, **Prof. Dr. Sebastian Löbner** and **Dr. Eva Gentes** for their continuous efforts in the Graduate Training Program StoRE. First and foremost, the weekly Graduate Research Seminar (“Doktorandenkolloquium”) has been a space of exchange, feedback and moral support which I greatly appreciated. I would like to thank all participants of the seminar for their active attendance and for the great sense of camaraderie we all experience to this day. The PhD candidates of the CRC have grown into a family and I will never forget the love, the support and the fellowship we shared and which connects all of us even now.
Further, the qualification program of SToRE provided me with additional courses and workshops on transferable skills that were an unforeseen source of resources that cannot be taken for granted.

The CRC in general provided me with a network of kindred spirits and excellent scientists, without whom this project would not have been the same. I would especially like to thank **Prof. Dr. Wiebke Petersen** for her collaboration on my first conference visit and the subsequent submission of my first publication. I am also grateful to **Prof. Dr. Hana Filip, Dr. Peter Sutton, Dr. Kurt Erbach** and the other members of Project C09 for the constant and fruitful exchange of ideas regarding countability and semantics. **Prof. Dr. Sebastian Löbner** and **Dr. Curt Anderson** are also to be mentioned here for their feedback and patients with my questions on theoretical semantics.

Mainly, this project would not have been possible without the help of my supervisor team: **Prof. Dr. Dr. Peter Indefrey, Prof. Dr. Ruben van de Vijver** and **Dr. Natalia Bekemeier**.
Especially **Peter Indefrey**, my “Doktorvater”, has my deeply felt gratitude. You have been my scientific mentor since my bachelor days and I owe a lot of who I am now to your guidance, support and shining example. Thank you for making me the scientist I am. Thank you for believing in me. Thank you for supporting me and thank you for teaching me. You are the person I want to be when I grow up.

I would like to thank all members of our psycho(-linguists) group for exchanging ideas, knowledge and experiences (and scripts); my colleagues **Kathi Sternke** and **Margit Scheibel**, but also our students assistants, especially **Sandra Bethke** for being at my side during this journey.

I thank **Benedikt Wischer** for using his radio training and charming voice for the recordings of my stimuli.

A very special thanks goes to **Frauke Hellwig** whose help in programming my experiments has been invaluable and whose constant friendship and encouragement as well as keen insight into the world of academia has made me a better scientist and a happier person.
I also thank **Dr. Alexandra Redman** and **Roman Gelzhäuser** for their help and guidance while

programming the scripts for the statistical analyses. There is not enough vegan chocolate in the world to ever properly repay you.

I also want to thank the CRC and HHU staff, especially **Tim Marton** and **Lena Hierl** for taking a weight of my shoulders in always being there for my questions regarding paper work, logistics, teaching and much more. Your friendliness, assurance and natural support has often been the sunshine that battled the clouds of doubt away.

Lastly, I thank my friends and family - especially my parents, **Dr. Udo and Deborah Beckmann**, my brother, **Dr. Dennis Beckmann**, and my husband, **Heiko Peiseler**, and my friend **Lauentia Schreiber** - for believing in me, for their encouragement, their love and their patients during this journey, which has been truly trying in parts. All of you have spent a lot of time and nerves in being my guinea pigs; discussing examples, filling out questionnaires and engaging in debate over my work and I am truly thankful; not only for taking an interest in my work but also for your way of ensuring I would not give up.

Content

1. Introduction 3

1.1 How some things can be counted – and others cannot 4

1.2 The things that can be counted – and the stuff that cannot 8

1.3 How some uncountable stuff is counted – and some countable things are not 12

2. Background 16

2.1. How we think about things and stuff or: Acquisition, conceptualisation and properties of objecthood 16

2.1.1. Object permanence, individuation and numerical identity 16

2.1.2. The properties of objecthood 17

2.1.3. How humans learn to count 20

2.1.4. Discussion 22

2.2. How we talk about things and stuff or: Effects of language on object construal – Lexical projection, similarity judgment and quantity judgment 22

2.2.1. Presentation of the stimulus 23

2.2.2. Native language of participants 26

2.2.3. Discussion 29

2.3. How linguists think we talk of things and stuff or: Linguistic Theories on Countability 30

2.3.1. Countability as a property of nouns 31

2.3.2. Countability as a property of noun phrases 32

2.3.3. Countability as a continuum 34

2.3.4. Discussion 36

2.4. Words that mean many things (or stuff) or: Polysemy, Metonymy and the Lemma-Concept Interface 43

2.4.1. Co-predication, anaphoric reference and types of polysemy 43

2.4.2. Metonymy, Shifts and Coercion 47

2.4.3. Discussion 51

2.5. Words that mean stuff together or: Combinatorial Restrictions and the Processing of Nouns in Phrases 52

2.5.1. Nouns in interaction – Theoretical approaches 52

2.5.2. Inside or outside the noun – empirical evidence 57

2.5.3. Online processing of noun phrases 62

2.5.4. Discussion 63

2.6. General Discussion 63

3. Empirical Investigation 67

3.1. Pretests 67

3.1.1. Pretest 1: Corpus study 68

3.1.2. Pretest 2: Sentence Production Study 70

3.2. Lexical Decision Experiment 1 71

3.2.1. Method 73

3.2.2. Results 77

3.2.3. Discussion 82

3.3. Lexical Decision Experiment 2 84

3.3.1. Method 84

3.3.2. Results 86

3.3.3. Discussion 89

3.4. Phoneme Monitoring Experiment 90

3.4.1. Method 93

3.4.2. Results 97

3.4.3. Discussion 100

4. General Discussion 102

4.1 What actually counts or Main findings 104

4.1.1 Countability in the mental lexicon 104

4.1.2. Countability during phrase processing 110

4.2 What's left to count or Limitations of the current research 114

4.3 Conclusion 116

1. Introduction

Counting is one of the most elementary cognitive skills that we teach our children. We consider it the very base of mathematics and to be on the low end of complex thought. Yet, when considered carefully, counting represents a highly elaborate system of interacting with the world; of mapping reality onto thought and of mapping thought onto language.

The work presented in this thesis attempts to shed light on the question "Is noun countability a lexical property of German nouns?" To answer this question from a psycholinguistic perspective, psychological issues must be considered that go beyond language processing. The necessary cognitive processes that precede the formation of a meaningful linguistic utterance are prerequisites of the linguistic processes discussed here and are thus an important part of this thesis. This dissertation aims to give an overview of the conceptual and linguistic procedures that are involved in counting; from individuating entities in the world conceptually and forming a mental representation thereof to formulating an appropriate linguistic way of describing this representation.

In the introduction, the complexity of counting as a mental task and the problems that go along with it are outlined. With the help of examples, the reader is familiarised with the linguistic landscape of German countability but also confronted with the challenges that a native speaker faces when attempting to produce accurate utterances denoting the things he or she wishes to quantify.

The subsequent background chapter presents previous psychological and linguistic research and theories related to counting and the cognitive processes it involves. It is divided into six sections. The first section is concerned with the cognitive process that is the prerequisite of linguistic counting: conceptual objecthood. The second section describes processes and issues regarding the interplay between the conceptual and the linguistic level when it comes to counting. The third section gives an overview of linguistic and psycholinguistic theories on countability and the language data they are based on. The fourth section discusses linguistic phenomena of noun usage that are of crucial importance when discussing noun countability. The fifth section is concerned with phenomena specific to noun phrases, most importantly how the syntactic environment of a noun influences its interpretation. Finally, the sixth section ties together all these considerations in a general discussion of the work previously done on countability that is relevant for the present research question.

The third chapter of the dissertation describes three behavioural comprehension experiments that directly address the question of whether noun countability is a lexical property in German. Carefully selected noun stimuli were used in two lexical decision experiments and a phoneme monitoring experiment. The results are compared and contrasted as the different tasks target different stages during linguistic processing.

Finally, the fourth chapter presents a general discussion of previous research and ideas and the results of the current research described in this thesis. The chapter addresses the research question at hand as well as surrounding issues of noun countability in German and gives a description of the processes involved in using count language.

1.1 How some things can be counted – and others cannot

Figure 1: "About Geese and Waterfowl"
from the website of The MSPCA–Angell Non-profit

How many birds are in this picture[1]? In order to count something, one first has to decide what the appropriate portion is that counts as "one". Do we count individual birds or groups of waterfowl? Do we count pairs of wings or the number of different species? What about the birds that are only partially visible in the picture? Do we count each one of them as 'indicating the existence of an entire animal' or will our end result be a fraction? How much of a given bird must be visible for it to count as one?

Do we even need a precise number or is it sufficient to say "There are many birds in this picture." or "Some swans, several ducks and a number of loons." or even "There is rather a lot of waterfowl." If all those descriptions are accurate, why can we not say that there are "many waterfowls" in the picture?

From this simple example alone we see that already at the language-end of the equation the situation may be rather complicated indeed. It is clear that some nouns, like *bird(s),* can be used in a count construction whereas others like *waterfowl* cannot, even if they refer to the same things in the world. Noun countability is the linguistic term that refers to the property of nouns being countable. Some nouns can be counted more easily than others. Consider the examples in (1).

(1) a. ein halber Apfel, ein Apfel, zwei Äpfel
half an apple, one/an apple, two apples

b. ein halber Liter Wasser, ein Liter Wasser, zwei Liter Wasser
half a litre of water, one/a litre of water, two litres of water

c. *ein halber Moment, ein Moment, zwei Momente
half a moment, one/a moment, two moments

d. *eine halbe Hoffnung, eine Hoffnung, *zwei Hoffnungen
half a hope, one/a hope, two hopes

e. *ein halbes Möbel, ?ein Möbel, zwei Möbel
half a furniture, one/a furniture, two furnitures

[1] from https://www.mspca.org/animal_protection/about-geese-and-waterfowl/, 06/17/2022

An apple can be divided, individuated and counted as in (1)a, whereas water is better measured than counted as in (1)b. Moments, apparently, can be individuated and counted, but not divided as in (1)c, while hope can be individuated but neither divided nor counted as in (1)d. Furniture curiously can be counted, but neither divided nor individuated as in (1)e.

Linguists – like Allan (1980) for example – have proposed tests which a noun can pass in order to determine whether it is countable or not:

(i) a(n) + N
Can the noun head a noun phrase of the type [indefinite determiner + singular noun]?

(ii) F + Ns
Can the noun head a noun phrase of the type [fuzzy denumerator + plural noun]?

(iii) EX-PL
Does the noun take NP-external elements with plural marking?

(iv) All + N
Can the noun head an NP of the type [all + singular noun]?

(Allan, 1980)

Nouns that are 100% countable pass the first three tests and don't pass the fourth. Others only pass the fourth test and are thus felicitous in what Allan calls 'uncountable all-contexts'.

(2) a. ein Auto
one/a car

b. einige Autos
several cars

c. Manche Autos sind teurer, als sie wertvoll sind.
Some cars are more expensive than they are valuable.

d. *Alles Auto ist schlecht für die Umwelt.
All car is bad for the environment.

(3) a. *ein Schlamm
one/a mud

b. *einige Schlämme
several muds

c. *Manche Schlämme sind klebrig. / Mancher Schlamm ist klebrig.
Some muds are sticky. / Some mud is sticky.

d. Aller Schlamm ist feucht.
All mud is moist.

As can be seen in examples (2) and (3), there are German nouns that pass only the first three tests and nouns that only pass the last test. We call the first class of nouns count nouns and the second class mass nouns. Count nouns are felicitous in count contexts: they pluralise and can be combined with a certain

set of determiners like *einige, several, viele, many,* and *wenige, few,* and the indefinite determiner. Mass nouns on the other hand do not pluralise[2] and combine with a different set of determiners like *etwas, some, kaum, barely any, viel, much* and *wenig, little*. They do not combine with the indefinite determiner but can form noun phrases[3] without any determiner.

Interestingly, there are also nouns that pass all the tests.

(4) a. eine Eiche
one/an oak

b. einige Eichen
several oaks

c. Manche Eichen wachsen schnell. / Manche Eiche wächst gerader als die andere.
Some oaks grow fast. / Some oak grows straighter than another.

d. Alle Eiche/ Eiche ist stärker als Buche.
All oak/ Oak is stronger than beech.

We call these nouns dual nouns because they can occur in countable as well as in uncountable contexts. They exhibit the behaviour of both count and mass nouns simultaneously.

Moreover, Allan points towards more examples of nouns which pass subsets of the four tests. Invariant plurals or 'pluralia tantum' don't pass the first and the last test, as can be seen for *Leute, people*, in (5), and that also holds for the German translation of *furniture, Möbel* (see example (1)e). The fact that they do not have a singular word form prevents their felicity with the indefinite determiner (*a, an*) and singular verb forms, see (5)d. Proper names can behave like count nouns when they are used as an appellative, that is when they are not fully-defining, as in (6). These are only infelicitous with All + singular, see (6)d, but some can be combined with the indefinite determiner, as shown in (6)a.

(5) a. *ein Leute / *ein Leut
a people.PL / a people.SG

b. einige Leute
several people

c. Manche Leute sind /*ist groß.
*Some people are/*is tall.*

d. *Aller Leut ist im Herzen gut.
All people.SG is good at heart.

[2] For *Schlamm, mud*, a possible plural form was created by the author for this example that is based on phonologically similar count nouns. Note that there is no existing plural form of *Schlamm* since it is a mass noun. Similarly, the singular form of *Leute, people,* in (5) was created for the example and is not part of the German vocabulary.

[3] ‚Noun phrase' (NP) here is meant as the general term for a constituent headed by a nominal element. It should not be taken as a reflection of commitment to a specific syntactic theory, e.g. as opposed to 'determiner phrase' or 'reference phrase'.

(6) a. Eine Lisa ist meist weiblich. / ?Ein London ist meist groß.
A Lisa is most often female. / A London is most often big.

b. Es gibt einige Lisas in meiner Klasse. / einige Londons auf der Welt.
There are several Lisas in my class. / several Londons in the world.

c. Manche Lisas / Manche Londons sind größer als andere.
Some Lisas / Some Londons are bigger than others.

d. *Alle Lisa ist klug. / *Alles London ist hektisch
All Lisa is smart. / All London is busy.

As mentioned above, some nouns are countable but not dividable. Those nouns in German include words like *Tropfen, drop, Feuer, fire,* and *Krümel, crum*, which when cut in half are better called "small [noun]" than "half a [noun]" because the shape and size of their referent is not relevant for it to qualify for denotation. For other examples of these 'unidirectionally countable nouns'[4] it seems important that their referents stay whole in order to qualify for denotation by their nouns, like *Moment, moment, Fehler, mistake,* and *Idee, idea*. Then there are nouns that can be combined with *halbe, half the*, in a sentence but it seems that this division refers to the verb or the adjective describing the result state more than to the noun, see examples in (7).

(7) a. Utgardloki sagte Donar, er habe das halbe Meer leer getrunken.
Utgardloki told Thor he had drunk half the ocean (empty).

b. Die halbe Seereise lang war mir übel.
For half the sea voyage, I felt nauseous.

c. Der halbe Apfel lag im Gras.
Half the apple lay in the grass.

So, it seems that noun countability not only depends on whether or not there is a singular and a plural word form available or only one of the two, but also what exactly the noun refers to. From the examples in (7) it can not only be seen that there are many examples of unidirectionally countable nouns, but also that the same formulation does not have the same interpretation in all cases. In (7)a and b – the examples with the unidirectionally countable nouns – the sentences cannot be interpreted as meaning that 'half the ocean' was empty or that 'half the sea voyage' made me nauseous similarly to the example in (7)c, but rather that Thor drank the ocean 'half empty' and that I was nauseous for one half of the duration of the voyage.

For examples like those in (6) it seems crucial how many Lisas or Londons there are and whether or not a specific Lisa or London is meant or whether one is referring to all the people called Lisa or any of the cities called London, that is appellative usage.

[4] This is not a term used by Allan (1980) and, to my knowledge, there are no other linguistic articles that point out this exact phenomenon.

In (4) one might argue that the noun *Eiche, oak,* refers to one or multiple trees in (4)a-c while the sentence in (4)d refers to the wood of oak trees; a substance that does no longer depend on the form of the plant it originated from. Such a reading, however, is not readily available for all the nouns that are felicitous in the countable contexts, as can be seen from the example on *Auto(s), car(s)* in (2).

1.2 The things that can be counted – and the stuff that cannot

The case of dual nouns like *oak* in (4) highlights the notion of whether or not a certain entity can be perceived as an individual thing that can be clearly distinct from other things of its kind, like it seems to be the case for cars, or whether it is a more substance-like entity, like mud. While at first glance that notion can appear to be pretty straight-forward, it presents interesting borderline cases once considered more thoroughly. Things certainly get more complicated when concerning mapping perceived objects and substances to nouns. Why, for example, should pebbles be seen as individuals but gravel as an undivided mass?

In the pebbles/gravel case, one might argue that the decisive property is the size of the individual stones. A collection of small stones is gravel while a collection of larger stones constitutes pebbles. But where is the threshold and is it the same for everyone? Similarly, it is not easy to explain why rice is a substance but peas are individuals since their individuals units (grains of rice and individual peas) can have the same size and are even interacted with in the same way (e.g. as a side dish).

Similar to the gravel/pebbles example, where one and the same stuff in the world seems to have two names, one count noun and one mass noun, some things in the world are denoted by a count noun in one language and by a mass noun in another, raising doubt to the view that whether something is a thing or a stuff has anything to do with the countability of its associated noun. For example, lentils are referred to by a count noun in German *Linse(n*), but by a mass noun in Czech *čočka*.

There are more examples of the same entity with two different nouns in the same language, just as the animals in Figure 1 can be referred to both as *birds* and *waterfowl*. Further examples include *shoes/footwear* in English; *meubles/meubilair, furniture*, in Dutch and *Teilchen/Gebäck, pastries,* in German. It has been argued by linguists that these mass nouns are atypical in that they refer to things and not stuff, as 'classic mass nouns' do. They have thus been called object-mass nouns because of their reference to (collections of) objects. Those nouns and the theoretical linguistic considerations surrounding them will be discussed extensively in the Background chapter.

Another interesting case where language and perception seemingly lead to contradicting insights are nouns that form part of lexicalised noun phrases where they exhibit properties different from their interpretation as nominal heads in isolation. Additionally, nouns can be used as names and then show linguistic behaviour different from that of their sortal kind-interpretation. Example (8) shows some of the ways the German noun *Erde* can and cannot be used in countable and uncountable contexts.

8) a. Seltene Erden werden für die Herstellung von Computerchips benötigt.
*Rare earth.PL (*meaning *'rare earth elements') are used in the production of computer chips.*

b. ?Diese Erde ist selten.
This earth is rare.

c. ?Hier gibt es verschiedene Erden für Hochbeete, Topfpflanzen und Rasen.
Here they have different soils for raised beds, pot plants and lawns.

d. Diese Erde ist am besten geeignet für Gemüsepflanzen.
This soil is best suited for vegetable plants.

e. ?Drei weitere Erden wurden in unserer Nachbargalaxie entdeckt.
Three other Earths were discovered in our neighbouring galaxy.

f. Eine Erde finden wir so schnell nicht noch einmal.
We will not find another Earth so easily.

In (8)a, *Erden* is part of the lexicalised noun phrase *seltene Erden*, which translates to *rare earth elements*. While here *Erde* is used in the plural form (*Erden*) it is not fully transformed into a count noun in this meaning, as can be seen from the infelicity with the singular demonstrative in (8)b.

Curiously, the very close meaning of *Erde* as soil cannot easily be used in plural contexts, see (8)c, while it is felicitous in singular countable contexts, like (8)d, thus exhibiting the opposite behaviour pattern. In the sense of *Erde* as the name for humanity's home planet – arguably an individual object rather than a substance-like mass – the noun allows combination with the indefinite determiner, see (8)f but is not so readily available for plural counting contexts where it refers to Earth-like (i.e. habitable) planets, see (8)e, like the names *Lisa* and *London* in example (6)b and c.

When discussing countability we thus not only need to be cautious about what sense of the noun we are talking about and what its reference is precisely but also about whether a noun is part of a phrase that has special properties because of its specific meaning as a combination. *Seltene Erden*, like *United Nations*, the example used in Allan (1980), does not simply refer to types of soil that are rare or nations that are part of some union. There is additional meaning to those phrases that has to be taken into account when considering what type of noun *Erde* or *nation* is.

Coming back to the physical side of things, one of the easier decisions should be to determine whether something is for instance a liquid substance or a solid body. It should thus be easier to classify noun referents then it is to classify nouns. A popular common-sense definition of aggregate states is based on the behaviour of entities. One could argue that there are qualitative differences between gaseous, liquid and solid entities when it comes to flow or bulk properties. That is, a solid is something that maintains a fixed volume and shape, while a liquid maintains a fixed volume but adapts to the shape of its container, and the matter of a gas expands to occupy whatever volume is available.

While those notions are difficult to apply to all entities – especially abstract ones like hope – thorough consideration makes them already problematic for entities that very clearly seem to belong to one category at first glance. Fardin (2014) shows the limitations of simple, common-sense definitions on

aggregate states – but also of more advanced ones from the field of rheology – by exploring the aggregate state and the dimensional control parameter flow rate of the species *Felis catus*. By referencing examples like the ones in Figure 2, Fardin most impressively shows that house cats can both keep and adapt their shape, complicating analysis based on conventional definitions of solidity.

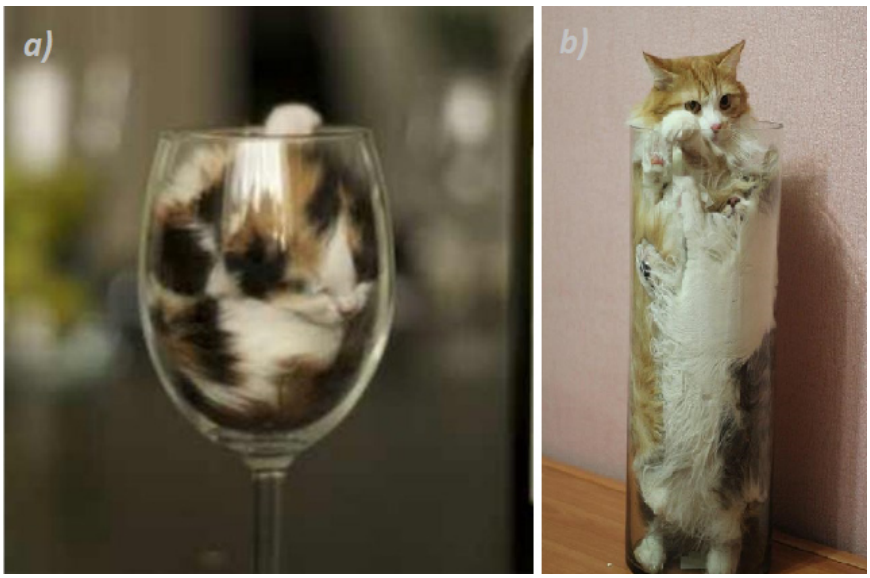


Figure 2: Cats in liquid state
(a) taken from Fardin (2014, p. 16), (b) from Daily Express[5]

For these examples, it seems more accurate to use a measuring description like “a glass of cat” (for 2a) and to ask “How much cat is there?”, rather than “How many cats are there?” (for 2b).

While the main aim of Fardin’s work was probably humorous in nature, it still shows that the definitions one works with determine one’s results and that even concepts that seem simple at first – like ’object’ or ’liquid’ – are not as simple once one tries to precisely define them.

Once concrete measurements like size and solidity do not apply anymore, properties of individuation are getting even harder to pinpoint. Consider the abstract count and mass nouns in the examples (9) and (10).

(9) a. Ein Risiko, wie Lisa es einging, könnte ich nicht bewältigen.
A risk such as the one Lisa took would be unbearable to me.

b. Einige Risiken sind es wert, erwogen zu werden.
Some risks are worth considering.

c. ?Ich bin heute schon drei Risiken eingegangen.
I already took three risks today.

d. Das ist zu viel Risiko für mich.
That is too much risk for me.

[5] Article „Cats pushing their nine lives? Felines squeeze into awkward places in hilarious pictures“, https://www.express.co.uk/news/weird/452565/Cats-pushing-their-nine-lives-Felines-squeeze-into-awkward-places-in-hilarious-pictures, 05/24/2021

(10) a. Von einer Flexibilität, wie Lisa sie an den Tag legt, kann ich nur träumen.
I can only dream of a flexibility such as the one Lisa is displaying.

b. *Einige Flexibilitäten sind kaum zu glauben.
Some flexibilities are barely believable.

c. *Während der Olympischen Spiele im Turnen gibt es viele großartige Flexibilitäten zu sehen.
During the gymnastic Olympics, many great flexibilities can be seen.

d. Für diese Yogastellung habe ich zu wenig Flexibilität.
For this yoga pose I have too little flexibility.

From the examples in (9) we see that the noun *Risiko, risk,* indeed has a plural form, referring to distinct situations that are risky. Therefore, the noun is also felicitous with the indefinite determiner. Yet it is not easily combinable with a numeral, as in (9)c. Combination with the mass-specific determiner *viel, much,* in the uncountable context in (9)d is unproblematic. Contrary, the plural form of *Flexibilität, flexibility,* seems odd and even though the sentences in (10)b and (10)c aim for a similar interpretation as those in (9)b and (9)c, i.e. there are multiple instances of flexibility displayed by different people, the usage with count-specific determiners seems infelicitous. However, (10)a with the indefinite determiner appears to be just as good as (10)d where *Flexibilität* is used in an uncountable context.

The difference is reminiscent of the different usages of *such* described in Constantinescu (2011). There, a 'kind-such' is distinguished from a 'degree-such' with the former being combinable with 'gradable nouns' as well as 'ungradable nouns' and clauses that pick out a certain instance of the kind (as in (11)), while the latter can only combine with 'gradable nouns' and indicates a specific intensity (like in (12)), among other properties. The same distinction is claimed to apply to the German word *solch-*.

(11) a. ?Er ist ein solcher Wissenschaftler, der empirisch arbeitet (KIND-SUCH)
He is such a scientist that works empirically.

b. Solche Frauen, wie wir gestern trafen, machen der Gesellschaft alle Ehre.
Such women as we met yesterday are a credit to society.

c. manche/wenige/?alle solche Hunde
some/few/all such dogs

(12) a. Sie ist ein solcher Wissenschaftler, sie verlässt sich nie auf bloße Intuition. (DEGREE-SUCH)
She is such a scientist, she never relies on pure intuition.

b. Diese Leute sind solche Idioten, dass man ihnen nicht trauen kann.
Those people are such idiots that one cannot trust them.

c. manche/wenige/alle solche Idioten
some/few/all such idiots

(examples adapted from Constantinescu (2011), p. 98, 99)

(13) a. Es ist eine solche Ehre, Sie kennenzulernen. (DEGREE)
It is such an honour to meet you.

b. *Solche Ehren sind nicht leicht abzulehnen. (DEGREE/KIND)
Such honours are not easy to refuse.

c. ?Einige solcher Ehren sind leichter zu erhalten als andere. (KIND)
Some such honours are easier to receive than others.

When the different *solch*- variations are combined with a gradable mass noun, like in (13), only the degree reading seems completely felicitous, at least in German, though it is possible that the feeling of ungrammaticality stems mainly from the unfamiliar plural usage of *Ehre, honour,* in (13)b and (13)c, even if distinct instances of honour should be no less conceivable than different instances of risk.

The important point to take away from this is that the same expression, e.g. *solch eine Ehre*, can be interpreted in more than one way. A kind reading (such as in (13)b and (13)c) can be analysed along the lines of a countability shift, where reference to a unbounded substance-like entity is shifted towards an interpretation where the noun refers to an individuated sort that is distinct from other sorts of the same kind. Whether the degree reading on the other hand represents a shift of the same magnitude is questionable. One might argue that in this case the noun still has the same reference.

Thus, not only does one noun have different interpretations and consequently shows different behaviour when it comes to count language in different constructions and linguistic environments (see examples of *Erden* in (8)), additionally the same construction with the same intended interpretation can be infelicitous for nouns of the same semantic category (see examples for *risk* and *flexibility* in (9) and (10)) or even the same noun (see (11) – (13)). These two pieces of evidence show that it is neither solely the noun nor yet the construction that fully determines interpretation and linguistic countability. To make the phenomenon of counting even more complex, even physical notions of entities in the world are not as easily applicable when it comes to distinguishing those entities that should be countable with language from those that should not, as exemplified by the liquid cats in Figure 2.

While the examples in this section may present extreme cases, the next section is concerned with a feature of language that is comparatively common, yet illustrates the same problem: One and the same noun on the one hand is categorised as a count or mass noun based on its linguistic behaviour and concomitant interpretation. However, both behaviour and interpretation can be changed purposefully and productively for many nouns. So, what is it then that makes a noun countable?

1.3 How some uncountable stuff is counted – and some countable things are not

The phenomenon of noun countability is further complicated by the epiphenomenon of countability shifts, also called countability coercions. These are cases where a noun that usually belongs to one countability class – say, count nouns – is inserted in an incongruent context, i.e. an uncountable

context, and as a result shifts in interpretation from individual entity toward undividable substance-like entity.

Most often cited examples come from the semantic domain of food and drink, where for example the name of an animal or fruit is also used to refer to its flesh, or where a mass noun in count-specific context refers to portions or sorts of a beverage. However, nouns from many different domains are found[6] to undergo countability coercion and interpretation shifts in many different ways.

(14) a. In der Wüste Gobi liegt ein Sand wie man ihn sonst nirgends anders findet. (SORTER SHIFT)
In the desert of Gobi there is a sand that one cannot find anywhere else.

b. Ich zahle jeden Monat eine Miete. (PACKAGER SHIFT)
Every month I pay a rent. (= one rent)

c. Ein Krieg ist das unnötigste auf der Welt. (EXISTENTIAL SHIFT)
A war is the most unnecessary thing in the world. (= wars in general)

d. Ich habe solch eine Angst noch nie in ihren Augen gesehen. (DEGREE SHIFT)
I have never seen such a fear in her eyes. (= this degree/ amount of fear)

(15) a. Viel Pflanze ist nach dem Unfall mit dem Blumentopf nicht mehr übrig. (GRINDER SHIFT)
There is not much plant left after the accident with the flowerpot.

b. In der BILD ist nicht wirklich viel Zeitung. (FUNCTION SHIFT)
There is not really much newspaper in the BILD. (= BILD does not have many properties of a newspaper, e.g. text, facts, results of journalism)

c. Die Blutlache gibt viel Hinweis auf ein Verbrechen. (DEGREE SHIFT)
The puddle of blood is much clue of a crime. (=a big/strong clue)

Four different shift types for mass-to-count shifts are exemplified in (14), and in (15) examples for three types of count-to-mass shifts are shown. Some of these shift types have been identified in the literature (see e.g. Wiese & Maling, 2005, for German), others where specified as such during the work on this dissertation, specifically the existential shift in (14)c, the function shift in (15)b and the degree shifts in (14)d and (15)c.

Interestingly, even when considering the most frequent shifts, the packager shift and the sorter shift from mass to count in the domain of beverages, there are limitations to which nouns can be used in such constructions. There is no obvious semantic or pragmatic reason that would explain why *Wein, wine,* in example (16) can be shifted and *Milch, milk*, cannot since they are both common beverages served in

[6] Example sentences presented in (14) and (15) were results of the sentence production study described in this dissertation, see Section 3.1 “Pretests”

portions like glasses and are both available in different sorts, for instance red and white wine, and regular and fat-reduced milk.

(16) a. ein Wein, zwei Wein (PORTION SHIFT)
a wine, two wine.SG

b. zwei Weine (SORTER SHIFT)
two wines

c. eine Milch, ?zwei Milch (PORTION SHIFT)
A milk, two milk.SG

d. *zwei Milche/Milchs/Milchen (SORTER SHIFT)
two milks

While it is acceptable to combine *Milch* with the indefinite determiner (see example (16)c), the combination of the singular form with a numeral seems odd and an attempt to form a plural fails completely (example (16)d)[7]. In the next paragraph, phonological factors that influence countability shifts are discussed. However, it is unlikely that the shift of *Milch* is prevented by the sheer oddness of the plural form, as firstly that shouldn't prevent the noun to undergo a shift of the type numeral + singular form (like in (16)a and (16)c) and secondly, there are several nouns in German with the same ending that are clearly count nouns and thus pluralise, though they are comparatively infrequent: *Knilch(e), boor* or *lout*, *Kilch(e*), *Lake Constance whitefish*, and *Bilch(e), dormouse*.

Some of the shifts in German proceed similarly to those in other (Germanic) languages but German is strikingly interesting in its mass-to-count shifts because especially when the resulting noun is interpretable as a portion ('packager shift') sometimes the singular word form can be directly combined with a numeral, as for instance in (16)a. While it has been proposed that those shifts present elliptic constructions (see example (17)a) that assumption does not seem to hold. Firstly, Wiese & Maling (2005) point out that this assumption doesn't conform to observed agreement patterns: when combined with the indefinite determiner that determiner agrees in gender with the shifted mass noun not with a hypothetical container noun, see (17)b. Secondly, some container phrases can also be used in singular form in these kind of constructions, as in the examples in (17)c, while others cannot (see (17)d).

(17) a. Drei [Gläser] Bier
Three [glasses of] beer.SG

b. Einen [*Tasse] Kaffee
a.MASC [cup.FEM] coffee.MASC
a [cup of] coffee

c. Drei Glas Bier, Drei Fass Wein
three glass.SG beer.SG, three barrel.SG wine.SG
three glasses of beer, three barrels of wine

[7] In the German dictionary "Duden", two plural forms for *Milch* are listed: *Milche* and *Milchen*. These are declared technical terms and most people don't know of their existence.

d. Drei Tassen / *Tasse Kaffee, Drei Messerspitzen / *Messerspitze Zucker
three cup.PL / *cup.SG coffee, three point.of.knife.PL / *point.of.knife.SG sugar
three cups of coffee, three pinches of sugar

This combination of numerals with singular nouns might have to do with the fact that German has so many different productive ways of forming the plural from the singular word form (compared to e.g. the one way in English: + *-(e)s*). While in the cases of *Tasse* and *Messerspitze* in (17)d, only an *–n* is added to the singular form, the pluralisation of the container nouns in (16)c includes the addition of *–er* and the change of the vowel from *a* to the Umlaut version *ä*, /æ/. It is possible that the singular form of the mass nouns is kept in some packager shifts in order to avoid the effort and insecurity that would come with the need to productively form a plural. A principle of minimising effort would explain why in cases with *Tasse/Tassen, cup/cups*, the plural is used, but in the cases with *Glas/Gläser, glass/glasses*, where more effort is needed to form the plural, singular forms are also acceptable.

Thus, apparently there are even phonological factors that come into play where noun countability and the limitations to its productive usage are concerned. Considering the complicated picture outlined above, German count language seems an ideal subject of research to investigate the fascinating interactions between world, mind and language that speakers and listeners have to engage in in order to effectively communicate.

In this introduction, it has been shown that the interplay between what is out there in the world, how humans perceive and mentally represent it and how these thoughts are translated into language is complicated extensively not only by the properties of (the German) language but also by our perception of entities, both concrete and abstract.

So, what is it that makes us count some things and not others? Physical properties seem to depend on our perception thereof, which appears to vary not only between people but also depending on the situation. And since also abstract entities without physical properties in the classical sense experience linguistic countability, these properties cannot be the whole answer.

In this dissertation, I thus want to focus on the nouns we count and don't count. Is countability determined for each noun in a given language, or do other linguistic features, such as the noun's environment determine its countability? How is countability interpreted during the online processing of language? How do we arrive at an interpretation of a perceived message that involves count language? As a first important step towards answering these questions and shed some light on the complex issue of counting as a cognitive and linguistic task, I ask: Is countability a lexical property of German nouns?

After an extensive overview of existing literature on human object perception and construal, German noun countability, semantic theories thereof and empirical findings on shifts and noun phrase congruence, three reaction time experiments are presented that explore how countability is stored in the mental representation of German nouns and how this property is utilised during the online processing of congruent and incongruent (shifted) noun phrases.

2. Background

2.1. How we think about things and stuff or: Acquisition, conceptualisation and properties of objecthood

Before we can turn to the way we speak about objects and substances, we need to establish how we think about objects and substances.

The conceptual construal of an entity as either an object of sorts or a substance of sorts is an ordinary everyday task that nonetheless has wide and immediate repercussions on our life. For the purpose of this dissertation the most important issue with object construal is the one once noted by Noam Chomsky: construal as either object or substance has systematic consequences for the way we think and, importantly, speak about entities.

2.1.1. Object permanence, individuation and numerical identity

In the field of cognitive psychology, it is assumed that object construal in humans involves at least two systems. The first and more rudimentary is the mid-level vision system that establishes mental representations of objects in the world, indexes them and tracks them through time and occlusion (also called mid-level object file system, e.g. Carey & Xu, 2001, p. 181).

This first system relies mostly on spatiotemporal information such as coherence and continuity through space and time for the establishment of individuation and numerical identity. A mental representation of an individuated object (or 'object file') may be formed and indexed based on FINST indexes (FINgers and INSTantiations) which serve a deictic function (like a finger pointing at an object) and provide the mental representation of the object with an "address" (Pylshyn, 2001, also see Kahneman, Treisman and Gibs, 1992, for the relation of FINSTs and object files). In a second step, the system indexes the mental file further with information that might play a secondary role in questions of number identity when spatiotemporal information is neutral, such as colour, shape and texture.

The second system is more sophisticated and is assumed to be the one that adult humans mostly rely on in their perception of objects in the world. It is fully conceptual in that it draws information about kinds for decisions about individuation and numerical identity, and hence is called the kind-based object individuation system. In cases where no spatiotemporal information is given, adults can still decide that, for example, the mug on the windowsill is the same we left there yesterday, but the cat on the windowsill is not the same as the mug (cf. Carey & Xu, 2001). Property changes are relevant on the conceptual level and can override spatiotemporal-based object construal (for instance when we perceive that a person 'no longer exists' after death, even though the body has not vanished) but don't stand on their own, since the relevance of such changes is kind-relative (e.g. humans grow old, change appearance and die, whereas mugs don't).

It is the first of these two systems that we can assume establishes object permanence; the first ingredient needed to think about things. Object permanence refers to the cognitive ability to realize that an object keeps on existing even though we may not perceive it anymore. While early and mid-twentieth

century scientists believed this ability to be mastered by humans at around 18-24 months of age, as indicated by the pioneering work of Piaget (1954), there have been more recent studies showing that children as young as four or five months are not only aware that an object does not cease to exist when it is out of sight, but also that this spatiotemporal discontinuity is utilized in object construal and number identity. Using children's gaze duration as indication of surprise or the perception of an unexpected outcome, several studies were able to show that infants track objects through space and time when they disappear out of sight (e.g. Spelke, Kestenbaum, Simons and Wein, 1995), and can conceptualize perceptually identical objects as different, distinct entities on the basis of spatial, directional information (e.g. Wynn, 1992). Other studies show similar abilities in children of only two months of age (e.g. Hespos & Baillargeon, 2001).

Object permanence, however, obviously is not limited to objects in our sense, but holds for substances as well. Thus, as a second step after a 'mental file' has been created to hold the representation of a given entity, its location and some properties in working memory, the perceiving mind has to decide whether it is dealing with an object or (an instance of) a substance.

2.1.2. The properties of objecthood

What properties make an entity likely to be an object or substance in the human mind has been subject of a multitude of investigations, dating back as far as the Aristotelian Aitia. Pustejovsky (1995, following Moraycisik, 1975) made this tradition of Aristotelian theory influential again by using four causes (or 'qualia' in Pustejovsky's terms) to characterise possible kinds of entities. In this account, any given kind can be specified by its material, efficient, formal, and final cause, that is what entities of this kind are made of, what brings them into being, what they essentially are, and what purpose they have. In Pustejovsky's approach, the four qualia taken together provide the lexical meaning of a noun denoting a kind, though it is possible for one quale to take a null value. Natural kinds for example are said to lack a 'telic quale', that is a purpose or function and can therefore be defined by three qualia.

Subsequent approaches of this kind add or remove certain qualia from the list. Del Pinal (2018) for instance added information about typical appearance, which not everyone agreed with.

With a series of experiments, Prasada, Ferenz & Haskell (2002) seek a more fine-grained and evidence-based view of what properties lead to an entity being construed as an instance of an object-kind. As previous researchers before them (see below) they created three types of novel entities: non-solids, regularly shaped solids and irregularly shaped solids and had participants indicate their construal decision by choosing a descriptive sentence for the novel stimulus entity in either count syntax ("There is a blicket in the tray.") or mass syntax ("There is blicket / a piece of blicket in the tray.").

Previous to the experiment, the stimulus items were rated according to the likelihood that "the structure of the entity was the result of a process directed at creating that structure" (p. 147). The high and low extremes of that rating scale were used for the description task.

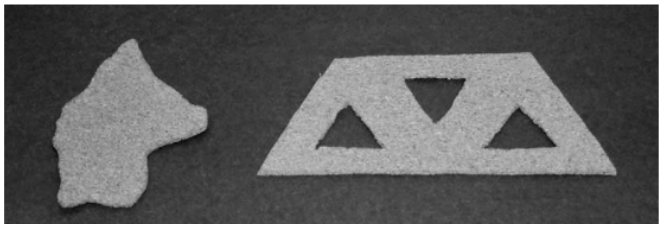


Figure 3: Examples of irregularly and regularly shaped non-solids, from Prasada, Ferenz & Haskell (2002, p. 148)

It was found that shape regularity is a decisive factor for object construal with the count-noun choice occurring significantly more often for regularly shaped solids than for irregularly shaped solids. Additionally, the rating task, which was repeated on the same participants after the description task, indicated the same trend with the regularly shaped solids being judged as more likely to be the result of a process directed at creating that structure than the irregularly shaped entities (Experiment 1A).

This factor however was found to be modulated by repetition. When the irregularly shaped items were presented among other entities of the same make-up they elicited significantly more count-responses and higher ratings of being intentionally created as such a structure than when they were presented in isolation (Experiment 1B). Crucially, the same irregularly shaped items elicited more count-responses and higher judgment values when presented among entities of identical structure than when presented among entities whose structure was different from the one of the standard item (Experiment 1C).

When the experiment was changed to a between-subject design, where participants saw the items in only one of the conditions (single regularly structured instance, single irregularly structured instance, multiple identical irregularly shaped instances, multiple non-identical irregularly shaped instances) the results remained the same: single regularly shaped items and irregularly shaped items presented among identical structures elicited more count-responses than irregularly shaped items presented in isolation or among different structures.

Introducing another factor into the design, the authors then tested items that had a shape that was rated to be medium-regular, i.e. that were in the middle of the rating scale between the regularly and irregularly shaped items of the previous experiments. For those items, Prasada and colleagues prepared short video clips in which the item either performed a “shape-dependent function” (perfectly fitting into a slot in a devise and turn a lever which causes a bell to ring) or a “shape-independent function” (hitting the bell on the device with the item).

In both the within-subject version (Experiment 3) and the between-subject version (Experiment 4) the results showed a significant main effect of this factor, with the items presented as having a shape-dependent function eliciting more count-responses than the items presented with a shape-independent function. The results of the rating task showed the same effect: a shape-dependent function increased the likelihood of an object-response.

Figure 4: Stimulus entity performing a shape-dependent function,
from Prasada, Ferenz & Haskell (2002, p. 155)

In a similar endeavour, Li, Duham & Carey (2008) employed a similarity judgment task, where participants were tasked to compare a novel entity with two other entities: one that matched the standard entity in shape, but not in material (shape-match) and one that matched the standard in material, but not in shape (material-match). Participants had

to indicate which of the two new entities was "the same" as the presented standard stimulus entity. The assumption behind this was that if something was conceptualised as an object the shape-match would be chosen, and if something was conceptualised as a substance the material-match would be chosen. The stimulus entities were non-solids, simple shape solids and complex shape solids. It was found that entities with complex shape elicited more shape-matches than material-matches compared to simple shape solids (Experiment 1).

The authors went on to cross the factor Complexity with results from a rating task that asked participants to rate each item for the relative "meaningfulness of its shape to its function" (p. 494), yielding four classes of solid items: neither complex nor functional, complex but not functional, not complex but functional and both complex and functional.

Functionality, as well as Complexity reached significant main effects in that again complex items had more shape-matches than simple items, and functional items had more shape-matches than non-functional items (Experiment 2).

In a more direct paradigm, the authors then asked participants to explicitly state whether a given stimulus was to be construed as an object-kind or as a substance-kind (Experiment 3). Here, too, main effects of Complexity and Functionality were found, representing that more complex items as well as items whose shape was important for their function elicited more object-kind responses.

The authors also repeated the similarity judgment task with children from two to five years of age (Experiments 4, 5 and 6). All but the two-year-olds showed significantly more shape-match responses for complex solids than simple solids and more for simple solids than for non-solids. That is, an effect of solidity could only be observed from three years on. Additionally, the main effect of functionality observed in the adult participants could be replicated in all age groups, with the exception of the two-year-olds. Four- and five-year-olds as well as adults showed more shape-matches for complex, non-functional solids[8] than for non-solids.

Adding to the scope of entity types, a study by Middleton et al. (2004) explored the influence of ease of perception of an individual unit in aggregates and the number of units used in typical interactions to investigate whether a given aggregate is construed as a (coarse-grained) substance or a collection of individual things (consider the case of pebbles and gravel mentioned in the Introduction). The first four experiments were rating studies on existing aggregates.

In experiments 1a and 2a, participants judged on a 1-to-9-scale how easy it was to see the individual units of 112 stimuli (42 referents of mass nouns and 70 referents count nouns in English) that were displayed orthographically either in the form of "a(n) [count noun]" or "a unit of [mass noun]" with the scale. The difference between answers to count and mass noun referents was obtained with a t-test that was highly significant, even though the respective means were both rather low and only differed by 1.04 points (count nouns: 3, mass nouns: 4.04).

Experiments 1b and 2b asked participants to judge on a 7-point-scale whether a given aggregate was used by most people in a way that "one or few units" (1) or "very many units" (7) were interacted with. Here again, the t-test revealed a highly significant difference between count noun referents (mean: 3.6) and mass noun referents (mean: 5.58).

[8] This class of items was chosen for the comparison in order to limit the confounding effect of functionality.

In Experiment 3, novel aggregates were presented in pairs, each pair differing either in the size of their constituents, in the spatial proximity of their constituents, or along both dimensions (see examples in Figure 5). Participants indicated their choice by picking one of the two aggregates as a more fitting representation of the novel noun in the sentence below. That novel noun was either presented in count syntax or mass syntax.

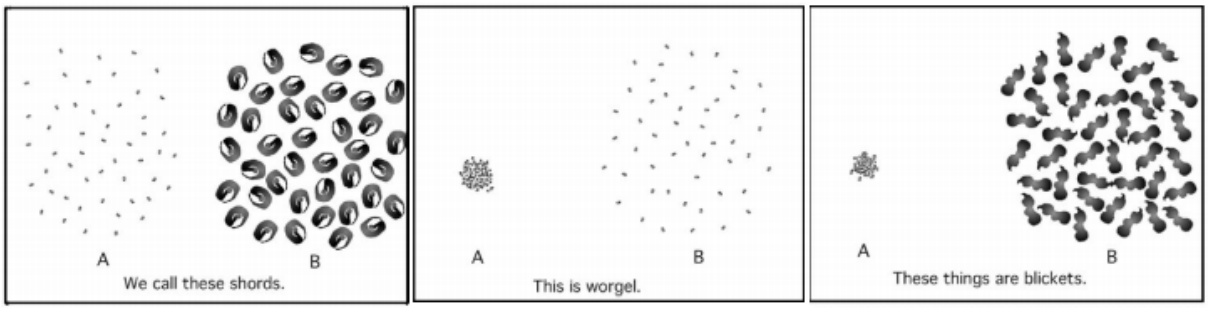


Figure 5: Examples of experimental stimuli from Experiment 3, from Middleton et al. (2004, p. 383)

In total, participants seemed to take spatial contiguity into account when making their choice of which aggregate represented the noun in the given sentence better and relied less on the size of the individual units. However, the results are not very straightforwardly interpretable, as the authors themselves point out (cf. p. 385).

In Experiment 4, participants were shown a novel aggregate (resembling yellow sugar) directly by the experimenter. In control trials, participants viewed the novel aggregate in a box and were asked to describe it by choosing a sentence containing either count or mass syntax. 60.7% of participants chose mass syntax, though this result did not significantly differ from chance level.

A different group of participants was presented with experimental trials where after viewing the aggregate in the box the experimenter used a tool to manipulate the aggregate grain by grain. Each individual grain was pushed through a hole slightly bigger than the biggest grain. Here, 69.2% of participants selected count syntax to describe the aggregate. This number did differ significantly from chance level.

The experimental evidence described in this section shows that solidity, complexity and functionality are deciding factor in the decision concerning objecthood. Like Pustejovsky and Aristotle have suggested, what entities are made of, what brings them into being and what purpose they have is considered by people when asked whether something is an object or a substance. From an early age on, people use these factors to decide what to count as “one” or if the entity in question can be counted at all.

The next section is concerned with the next step in the counting process: representing multitudes of entities and quantising them.

2.1.3. How humans learn to count

Carey (2009) argues for a sequence of events in children’s development of counting that goes as follows: From a very young age (earliest evidence found in two-months-olds) children can individuate

objects from their surroundings and from other objects, i.e. possess object permanence. After that, children master numerical identity in the sense of mentally representing a small number of different objects. That is however, Carey argues, not a case of analogue magnitude number representation, which is assumed to be applied only to greater numbers of objects in children and in adults. In fact, famous experimental outcomes such as Wynn (1992b) showing that infants expect to see exactly two objects behind a screen and not one or three when they saw two objects disappear behind that screen are said to be best explained by a representational system where there are no mental symbols for numbers at all. Rather, the child represents each object in working memory, not in the way of {two puppets} but in the way of [{puppet} {puppet}]. By means of this 'parallel individuation' infants are able to represent individual objects, sound bursts and events (Carey, 2009, p. 7).

Crucially, since working memory is limited, the system has its limitations, most impressively demonstrated by Feigenson and colleagues (Feigenson & Carey, 2005; Feigenson et al., 2002) who showed 10-12 months old infants' successful choice between 1 and 2, and 1 and 3, and 2 and 3 crackers, but their failure to distinguish between 4 and any of the other numbers. Studies like this present conclusive prove that infants' choices cannot be based on Weber-fraction differences[9] or other forms of analogue magnitude representation, as 1 vs. 4 should be a more favourable ratio than 2 vs. 3 according to those systems.

At any rate, very young children, similar to non-human animals, have the ability to differentiate singular from plural entities (see Barner, Wood, Hauser & Carey, 2008, for data on rhesus macaques; and Barner et al., 2007, for data on young infants), even though they may fail to distinguish between plural sets of different sizes when they move together as a united set or when the ratio for comparison is too small, that is in cases were parallel individuation is hindered and analogue magnitude representations are not salient enough.

Children's representations of number words evolve from a meaningless, but ordered list of "one, two, three, four, etc." (which they learn in a similar way as other meaningless ordered lists, such as "ene, mene, muh") to a stage where sets of different sizes exist in their minds, either as abstract symbols of individuals like {a} {b, c} {d, e, f} or possibly long-term memory models of particular sets like {my head}, {my hand, my hand} and so forth which can be mapped in 1-to-1 correspondence to newly encountered sets.

According to Carey (2009), children learn the cardinal principal based on the fact that the list is ordered, much like they learn linguistic quantifiers. They start out by mapping the numerals to representations of single entities, that is individuals (one), pairs (two), triplets (three), and so on and treat higher numerals in a placeholder fashion:

> "The bootstrapping process explains how children learn how the list itself represents number, which in turn explains how they assign numerical meaning to numerals like "five" and "seven." When children first become cardinal principle knowers, the meaning of "five" is exhausted by the child's mastery of counting. The counting principles ensure that the content of "five" is one more than four, and the meaning of "seven" is one more than six, which is one more than five, which is one more than four, and so on."
> (Carey, 2009, p. 19)

[9] Weber's law states that the bigger the difference in number between two multitudes the easier it is to decide which is higher in number.

2.1.4. Discussion

What have we learned so far? From a very young age on, children can make out objects and are aware of their existence (object permanence). They use the so-called mid-level vision system and parallel individuation to conceptualise an object they see as a thing with say a colour, a shape and a texture that moves in space over time. From the experiments by Li, Duham & Carey (2008) we know that complexity of shape and functionality are factors that influence object construal in children as young as four years old. Solidity was only found to be of significant influence from three years on, however it needs to be considered that the children did not see any of the stimulus items move or were allowed to touch them to acquire some spatiotemporal information about them.

Adults mostly utilise a kind-based object individuation system that relies more on what a thing is, i.e. what kind, rather than the individual information about shape, colour, etc. that are perceived online. Adults chose whether a given thing – say a sock – is a sock or a piece of cloth because they recognize it as an instance of the sock-kind and they know that socks are (individual) things. This is substantiated by the results of Middelton et al.'s (2004) rating studies. For existing aggregates that already have a name in participant's native language, entities are conceptualised as individual units more often if the noun denoting them is mainly used in count syntax. When inspecting Middleton et al.'s stimulus set, it becomes clear that the referents of those nouns differ greatly with respect to the individual unit's size and the number of units people typically interact with. For example, *ashes* and *screws* are both part of the count noun referent group, even though individual screws are rather big (as compared to units of typical aggregates such as rice) while the individual units of ashes are hard to imagine at all.

However, that does not mean that adults cannot use the previously employed method for object construal based on a thing's solidity, regularity, function and whether or not there are others like it, pointing towards the fact that the thing was specially designed as it is for a specific purpose, as shown by the results for novel entities by Prasada, Ferenz & Haskel (2002) and by the results of Experiment 3 and 4 by Middleton et al. (2004).

It can thus be concluded that in adults both systems are consulted when a novel entity is encountered or possibly even when people explicitly consider and evaluate a familiar entity's objecthood.

2.2. How we talk about things and stuff or: Effects of language on object construal – Lexical projection, similarity judgment and quantity judgment

For a long time now, but at least ever since Linguistic Relativity (also known as the Sapir-Whorf-Hypothesis; Whorf, 1956) and Quine's claims regarding the linguistic influence on conceptual object construal have entered the discourse, researchers have been interested in how language shapes the conceptual representation of objects and non-solid substances. Quine (1969) claims that early in language acquisition, children use nouns to refer to a portion of experiences they consider related to the referenced entity and have no ontological categories distinguishing e.g. objects from (portions of) substances, and only start to differentiate between different types of nouns as they acquire quantifying language such as plural morphology, numerals or quantifiers.

There is in fact empirical evidence that children up to 18 months of age overgeneralize nominal terms and use them in a variety of different situations, for example the word *paper* might be used to address the act of cutting, the act of drawing and to refer to pens, pencils and paper (Dromi, 1987), which appears to be in accordance with Quine's suggestions.

Additionally, Barner et al. (2007) report experiments that show a link between the ability of differentiating quantities of four from those of one with the acquisition of plural morphology. As mentioned in the last section, children of 20 months of age succeed in distinguishing quantities of one, two and three, but start experiencing problems when confronted with quantities of four and higher. In fact, it has been suggested that they do not even perceive quantities of four to be multitudes.

Barner and colleagues found that children of 20 months fail at this task, even when explicitly cued to the difference using instructions where plural morphology is used. However, 22- and 24-months-olds succeeded, irrespective of whether they were given linguistic cues. The authors state that the success of the children was predicted by whether or not they used plural morphology in their own speech, as the parents reported.

Thus, for counting itself, linguistic knowledge seems to be useful at least. Next to plural morphology, the ability to combine count and mass nouns with the correct determiners and quantifiers or classifiers is part of proficient language use and has been hypothesized to interact with – if not allow for – the abilities that are needed to conceptualise entities correctly.

With regard to how language might influence object construal, the lexical projection task, the similarity judgment task and the quantity judgment task have been used to address this question empirically. Using these tasks, two main factors have been identified that influence how humans perceive entities, i.e. whether as an object-kind or as a substance-kind: Presentation of stimulus and native language.

2.2.1. Presentation of the stimulus

First, the study by Li, Duham & Carey (2008) already outlined above was conducted with two versions of instructions (Experiment 1); a label condition, where the new entity that had to be compared to a shape- and a material-match was referred to by a novel noun ('lexical projection task', e.g. "Which one is the blicket?"), and a non-label condition, where the entity was referred to by pronouns ('similarity judgment task', e.g. "Which one is the same?"). The researchers however did not find any influence of the instruction on the choices participants made. Note that both versions of the instruction are neutral when it comes to syntactic countability (definite determination).

Thereby, the authors replicated findings by Imai and colleagues who also observed similar response patterns in all of their participant groups for studies with non-linguistic similarity judgment task (Imai & Mazuka 2003, 2007) and a study with a lexical projection task, where pseudowords were used as names for the stimulus entities (Imai & Gentner, 1997)

A previously not mentioned task that was used extensively in pursuit of the question whether something is perceived as a thing or some stuff is the so-called quantity judgment task. Most famously, Barner & Snedeker (2005) showed participants pictures with one big item of something and a collection of three smaller items of the same kind that in total did not exceed the volume or surface area of the

big item. The participants were told that the big item belonged to one person and the three small items belonged to another person and had to answer the question "Who has more X?"

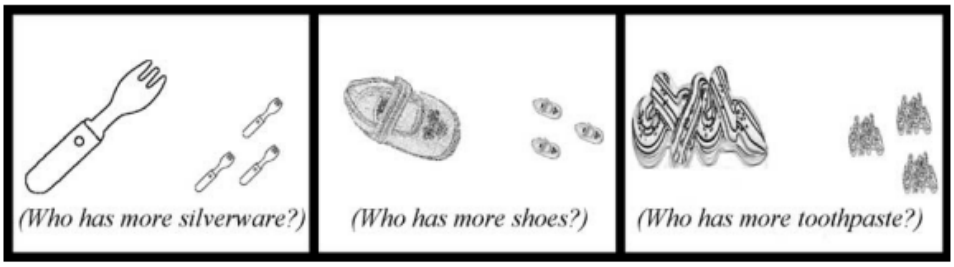


Figure 6: Example stimuli of Quantity Judgment Task
Note: Left: collective noun, middle: count noun, right: mass noun,
from Barner & Snedeker (2005, p. 50)

The stimuli consisted of three classes of objects: those whose linguistic label is a count noun in English (e.g. *shoes*), those that are referred to by an English mass noun (e.g. *toothpaste*) and those that are referred to by a collective or 'object-mass noun' in English (e.g. *silverware*).

The experiment showed that both adults and children of four years of age interpreted the big item as being "more" than the three individual smaller items for referents of mass nouns, such as *toothpaste*, but not for the referents of count nouns such as *shoes* and most interestingly, also not for the referents of object-mass nouns such as *silverware*. For the latter two stimulus classes, participants indicated the three individual instances as being "more" than the one big item, that is they based their judgment on cardinality rather than volume. In order to prevent the referents of object-mass nouns to be interpreted as referents of count nouns (i.e. *silverware*, rather than *spoon*) they repeated the experiment with stimulus pictures where the object-noun referents were presented in pairs. For example, the picture for *furniture* compared one large table and chair with three tiny table and chair sets (p. 52). The results remained the same. Children and adults alike based their judgment on cardinal number for the referents of count nouns and object-mass nouns, and on volume for the referents of mass nouns (Experiment 2).

The researchers also used the same paradigm to test referents of four dual nouns which are flexible enough to be grammatical with either count or mass syntax and whose interpretation as either an object-kind or a substance-kind depends on the syntactic environment they appear in: *string, chocolate, paper* and *stone*. With this kind of stimuli, judgment of both adults and children was significantly modulated by whether the question utilized mass or count syntax to refer to the item. If count syntax was used, e.g. "Who has more strings?", than the participants decided more often based on cardinality, while in cases where mass syntax was used, e.g. "Who has more string?", judgment was based on volume more often. These results were replicated with speakers of Brazilian Portuguese by De Lima (2015).

Barner & Snedecker (2006) found that – as for dual noun referents like *string* and *stone* – influence of syntactic countability of the noun used also holds for novel nouns. When the novel nouns were presented as count nouns (e.g. "This is a blicket.") participants construed the referent entity as an object, irrespective of its solidity and complexity. However, for entities that were introduced as mass noun referents, judgment was stronger modulated by entity type. The effects were found with both the lexical projection task and quantity judgment task.

In a later version of the experiment, Barner, Wagner & Snedecker (2008) found a similar pattern when verbs were used as stimuli instead of nouns. Here, two types of actions were distinguished: those that are judged as durative (e.g. *dance*) and those that were judged as iterative (e.g. *jump*). For the durative actions, quantity judgment (i.e. "Who does more dances/more dancing?") was affected by whether the

action word was presented in count syntax (*dances*) or mass syntax (*dancing*). The effect was not found for iterative actions like *jump* which were judged according to cardinality (i.e. many tiny jumps > one big jump).

Similar to the multiple/single difference and the observed functionality effect from Prasada, Ferenz & Haskell (2002), Grimm & Levin (unpublished manuscript)[10] performed experiments with a quantity judgment task on object mass nouns with results pointing in the same direction.

The first of these experiments aimed to establish the effect of heterogeneity. There, stimuli were a set of identical items within the denotation of the noun (e.g. 5 chairs for *furniture*) that had to be compared to a heterogeneous set of denotation targets (e.g. a sofa, two chairs, a coffee table, and a bookcase). They found that 20 out of 20 participants judged the second set – that is heterogeneous set – as representing "more furniture", adding to Barner & Snedeker's (2005) findings that object-mass noun referents are not just as well be perceived as heterogeneous collections, but better as heterogeneous collections than homogeneous ones.

In a second experiment, they provided a short context pointing toward the intended function of the mentioned items as participants of an event. Across all nouns, at least 5 out of 20 participants judged the more heterogeneous set more appropriate to fulfil the event of being "more [noun]". For the items *change* and *furniture,* at least 12 participants chose the more heterogeneous set as "being more". An example is given below:

(18) CONTEXT: Soldiers in the army.

— Soldier A has four bullets and two grenades.
— Soldier B has two bombs, two grenades and one bullet.

QUESTION: Who has more ammunition?

(Grimm & Levin, unpublished manuscript, p. 14)

Some participants even pointed out, that they based their judgment on the fact that "[t]wo bombs would be more lethal than the three extra bullets that Soldier A has" (p. 14), underscoring the importance of the specific event the noun referent takes part in.

To rule out that for this class of nouns cardinality is the preferred dimension of quantity judgment because of their superordinate nature, Grimm & Levin conducted a third study using countable superordinate nouns as stimuli. Here, most participants (at least 15 out of 20 for each noun) based their judgment on cardinality rather than functionality when asked what constituted more *weapons*, *vehicles* or *tools*. They conclude that, unlike a non-countable superordinate noun such as *furniture*, countable superordinate nouns only require one entity as event participant to satisfy their semantic requirements.

A fourth experiment tested the context that pointed participants toward the event the noun referents from Experiment 2 took part in against a neutral context. The effects were less straightforward, but do show a trend of context effect. That is, for five out of six nouns, participants showed a preference to judge the functionally more useable amount as "more [noun]" when presented with a context that pointed towards the function. Only the noun *change* (as in *pocket change*, i.e. coins) yielded more

[10] Manuscript received after personal correspondence with Scott Grimm. Title taken from http://www.sas.rochester.edu/lin/sgrimm/publications.html

judgements of 'functionally more' to be 'more change' in the neutral context. The authors suggest that the functional way of counting change (i.e. value rather than cardinality) is most salient for this concept, which may have overruled the context effect.

In the next section, studies investigating the effects of cross-linguistic differences are summarized.

2.2.2. Native language of participants

The first question with regard to native language influence must of course be whether it matters if subjects' minds can be influenced by language ability at all. Soja et al. (1991) tried to address Quine's claims that count-mass syntax is necessary for object construal directly and found no influence of linguistic knowledge of this kind. They tested the ability of two groups of 2-year-olds to generalize a newly learned word over another object or substance of the same kind, either with or without linguistic cues of mass-count syntax in the instruction. Even though many of the children did not show the ability to use count-mass syntax properly in the obtained production data, they all generalized words learned for solid objects over a shape-match rather than a substance-match and words learned for substances over other portions of the same material rather than samples having the same shape. Whether the word was learned as part of a count or mass-specific noun phrase or in count-neutral context did not influence children's choices (Experiment 1).

Another interesting contrast when it comes to the question of a possible language-on-thought influence in the Whorfian sense is the comparison of speakers of languages with count-mass syntax with speakers of languages that are assumed not to mark objecthood linguistically.

Classifier languages used to be the typical choice, for they do not have plural morphology in the same way as say English or German. Yet, in recent years more and more researchers have pointed out, that the classifiers used for nouns referring to individuated (object-like) entities differ systematically from those used with nouns referring to non-individuated (substance-like) entities and that even equivalents of object-mass nouns are present (e.g. Cheng, 2012).

Despite the growing realization that these languages, too, have linguistic countability in a way, there have been some interesting findings regarding the results of similarity and quantity judgment tasks. Li, Duham & Carey (2008) reported significant language effects in their similarity judgment data, with adult English speakers choosing the shape match more often than adult Mandarin speakers (Experiment 1) and more often than adult Japanese speakers (Experiment 2).

Li and colleagues were able to show the onset of these differences based on native language to be during the fifth year, in that English-learning children of four, five and older picked more shape-matches than Mandarin-learning children (Experiment 5) and Japanese-learning children (Experiment 6).

Contrary to this finding, the Imai and colleagues studies already mentioned above used the lexical projection task with complex solids, simple solids and substances on English and Japanese speakers of three age groups (young two-year-olds, late two-year-olds, adults). The results were comparable for all groups when it came to complex solids and substances. However, interestingly, for simple solids, English speakers of all ages projected the lexical label across shape-matches, while Japanese children chose both options equally and Japanese adults preferred material-matches. Thus,

these results pinpoint the age when the difference influenced by native languages arises at a much younger age (Imai & Gentner, 1997).

Imai & Mazuka (2003) used the same stimuli in a no-label similarity judgment task with English and Japanese adults and four-year-olds. They found no difference in response patterns to the earlier study, except for solids, namely the English four-year-olds matched them with the shape-match to a much lesser extent (see Figure 7). Note however, that the difference between the conditions is maximally 20% (Substances ≈ 35%, Complex Solids ≈ 55%) varying around chance level (50%) and that the English adults chose the shape-match even for non-solid substances in about 50% of trials.

These results are in fact reminiscent of results from Subrahmanyam, Gelman & Landau, (1999), who found English-speaking children to only employ ontology-based categories for their classification of novel entities when those entities where given names.

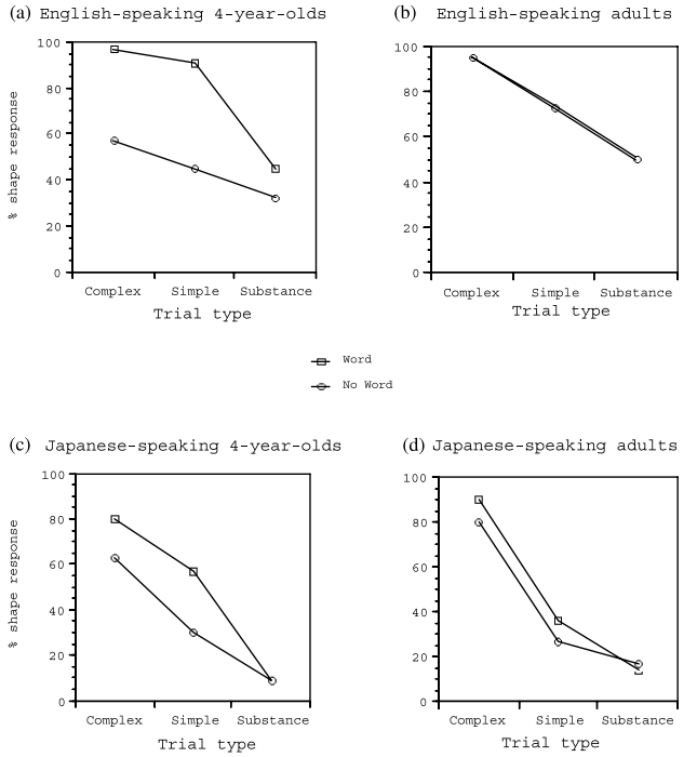


Figure 7: Age and native language effects in Lexical Projection ("Word") and Similarity judgment ("No Word") studies by Imai and colleagues, from Imai & Mazuka (2003, p. 443)

What seems clear is that Japanese-speakers chose the shape-match less (in only about 30% of trials) than English speakers, with the exception of complex solids, pointing towards an effect of native language at least for the adults.

Imai & Mazuka (2007) additionally used count and mass-syntax to introduce the standard entity to English-speaking adults and 4-year-olds. Within each age group, one half of participants was given instructions using count syntax ("Look, this is a X! Can you point to another X?"), whereas the other half was given instructions where the standard was referred to by a mass noun ("Look, this is some X. Can you point to some more X?"). In both age groups and across stimulus types, there was a highly significant effect of this factor, signifying a strong influence of syntax onto match choices.

Table 1: Mean proportion and standard deviation of the shape response as a function of language, age, and context, from Imai & Mazuka (2007, p. 393)

			Complex	Simple	Substance
Japanese	4-year-olds	No-word (Study 1)	0.63 (0.44)	0.30 (0.38)	0.089 (0.27)
	Adults	No-word (Study 1)	0.08 (0.36)	0.27 (0.35)	0.17 (0.29)
English	4-year-olds	No-word (Study 1)	0.57 (0.43)	0.45 (0.45)	0.32 (0.43)
		Count (Study 3)	0.89 (0.16)	0.84 (0.23)	0.41 (0.42)
		Mass (Study 3)	0.59 (0.48)	0.46 (0.44)	0.20 (0.26)
	Adults	No-word (Study 1)	0.95 (0.14)	0.73 (0.33)	0.50 (0.41)
		Count (Study 3)	0.93 (0.27)	0.79 (0.29)	0.50 (0.35)
		Mass (Study 3)	0.48 (0.43)	0.15 (0.26)	0.13 (0.35)

One thing is noteworthy about these studies, namely that most of the entities in the simple solid condition were made of soft or jelly-like materials, therefore crucially different from the entities in the complex solid condition, which were formed of plastic, wood, porcelain and ivory.

Figure 8: Examples of simple solid stimulus entities, from Imai & Mazuka (2007, p. 391)

This could be assumed to lead participants toward analysing the stimuli as instances of the material, thus pointing them towards choosing the material-match. In light of this, it is even more baffling that the English-speakers chose the shape-match for this condition to such a strong extend and also for the non-solid stimuli and lets the effect the English language has on object construal seem even stronger.

In a series of experiments on speakers of Yudja, a language spoken in Brazil that limits plural morphology to nouns referring to humans, Lima (2014) showed the effects of entity type observed in quantity judgment tasks for English speakers (e.g. by Barner & Snedeker, 2005) to be absent in Yudja speakers of all ages (children: 2-5 years old, children: 6-11 years old and adults).

De Lima (2015)[11] repeated the test with two groups of speakers of Brazilian Portuguese (BP). The first group consisted of adults who only spoke BP. For these participants, the number of cardinality-based quantity judgments increased with “natural individuatedness” of the entity (substances 21% < objects 86% and aggregates 97%). The second group of participants were L2 speakers of BP that learned Yudja

[11] Note that S. O. de Lima and S. Lima are the same person publishing under slightly different names.

as their first language. With those participants, there was no effect of entity type (see Table 2) when they were tested in Yudja, but the paper reports that when tested in Portuguese, substances are significantly less likely to be judged based on cardinality (cf. p. 89). It should be noted that the study employs the rather atypical statistics of "Helmert contrasts" and the reported percentages of cardinality judgments for BP speaking Yudja speakers do not always differ dramatically at first sight (see Table 2).

Table 2: Percentage of cardinality responses of Yudja speakers in L2 (BP) and L1 (Yudja), from Lima (2015), p. 88)

Noun category	Answers (Brazilian Portuguese)	Answers (Yudja)
Mass (farinha ′flour', água ′water', carne ′meat')	55 %	85 %
Count (cuia ′bowl', galinha ′chicken', colher ′spoon')	63 %	83 %
Collective (roupa ′clothes', cerámica ′ceramics')	65 %	79 %

Similar to the Yudja-speaker's results, Mandarin-speakers were found to allow cardinality and volume-based judgments for all types of nouns (Lin, Hacohen & Schaeffer 2018). However, it was found that when tested in their proficient L2 English, Chinese speakers (Yin & O'Brien 2018), Japanese speakers (Inagaki, 2014) and Korean-speakers (MacDonald & Carroll 2018, except for flexible nouns like *stone*) performed like the English-natives from Barner & Snedeker (2005), but different from late learners as found by Inagaki (2014).

2.2.3. Discussion

Interpreting this body of results is not easy, as some results contradict each other. Li, Duham & Carey (2008) report that in their data, there was no difference between the label condition (lexical projection) and the no-label condition (similarity judgement). Subrahamanyam, Gelman & Landau (1999), however, do report an effect of that difference. In the data from the Imai & colleagues studies (Imai & Gentner, 1997, Imai & Mazuka, 2003, 2007), this difference seems to be relevant for object construal of children but not for that of adults. Here, the difference is most salient for simple solids, which – as pointed out above – may in fact have been perceived as 'less solid' than the items of the complex solids condition. That confound, however, does not explain the large difference between the label ('Word') condition and the no-label ('No Word') condition found with English four-year-olds (see Table 1).

It is also strange that Li, Duham & Carey (2008) found the difference between Japanese and English speakers not to be modulated by entity type, whereas Imai and colleagues did find it more pronounced for solid entities. However, Barner, Wagner & Snedeker (2008) found that responses to durative actions (which are understood as the verbal equivalent to non-solid substances here) were more modulated by syntax of presentation than iterative (bounded, more 'object-like') actions. Also, viewed the other way around, responses to mass noun referents were modulated by entity type more than count noun referents (Barner & Snedeker, 2006).

Thus, it seems that the effect of native language does not affect object construal in the same way as language cues (i.e. count vs. mass-syntax) in the instruction. Additionally, the age at which the language cues in instructions are influential have been pinpointed by Li et al. (2008) at four years, while the effect of native language influences choice behaviour from two years on according to Imai and colleagues (Imai & Gentner, 1997, Imai & Mazuka, 2003, 2007).

Other results form a more coherent picture. Barner et al.'s (2007) findings that children that use plural morphology themselves have the ability to distinguish multitudes above three items, irrespective of whether plural morphology was used to instruct them, lines up nicely with the finding of Soja et al. (1991) that whether count-mass syntax is used in the instruction does not influence children's behaviour. That points towards an interpretation of the findings by Barner and colleagues as indicative of a thought-on-language effect, rather than a language-on-thought effect, in that one's own language abilities and practices show an influence on object construal more so than perceived language cues do.

It also seems straightforward that interpretation of both dual nouns (Barner & Snedeker, 2005) and novel nouns (Barner & Snedeker, 2006) in adults is heavily influenced by whether the noun is introduced as a count or a mass noun referent. However, it is far from clear why entity type was able to overrule that effect for mass noun referents, but not for count noun referents. That Grimm & Levin (unpublished manuscript) also found a difference in quantity judgement behaviour for superordinate plural count nouns like *weapons* when (indirectly) compared to object-mass nouns like *ammunition* fits these findings as well.

Thus, it was shown that (i) presenting a substance in individual portions is not enough to make participants construe them as individuated, countable objects and (ii) that entities which are labelled linguistically by object-mass nouns (e.g. *silverware*) can and will be construed as countable objects based on their properties and that (iii) the way one's native language encodes objecthood influences whether we focus our attention on an object's shape or its material when categorising it.

2.3. How linguists think we talk of things and stuff or: Linguistic Theories on Countability

Before exploring empirical results on countability further, this section summarises theoretical accounts on how count and mass noun word forms are connected to their respective conceptual representations and what properties those representations would have to have in order to be linguistically realised as a count or mass noun, respectively.

First, there are accounts that state that nouns by themselves don't commit to a countability class in any rigorous way. In that view, noun phrases but not nouns can be assigned countability, i.e. a noun phrase including the indefinite determiner for example is count and one including a quantifier like *much* is mass. Any noun (or almost any, see Pelletier 1975, 2012) can occur in any noun phrase and interpretation changes according to the (presence of a) determiner or quantifier, like it is the case for dual nouns.

2.3.1. Countability as a property of nouns

In theoretical semantics, there are mereological approaches to countability. Those assume individuals[12] and sums of individuals to be of the same semantic type and assume a domain of entities structured either by the sum operation or the ordering operation.

Link (1983) is generally assumed to be the first work suggesting a mereological theory of nominal semantics in which individuals and their sums are of the same semantic type with regard to countability. In mereological approaches, either the sum operation, ⊔, or the ordering relation, ≤, is assumed to give structure to the entities in a domain, which is to be understood as the realm of potential denotation targets of a word. Link (1983) assumes that there are different domains for count and mass nouns (also see Champollion, 2017).

The article discusses Quine's (1960) notion of cumulative reference for both plural count nouns and mass nouns, and points towards the verb *gather* that can combine with both classes of nouns.

(20) a. The children gathered around their teacher.

b. The water gathered in big pools.

(Link 1983, p. 127)

Both *the children* and *the water* have cumulative reference in the sense that a subset of *the children* (say, three out of seven) is still called *children* just like a subset of *the water* (say, 300ml out of 1l) is still called *water*. However, according to Link, mass nouns denote entities in a domain(-part) of non-atomic entities and count nouns (both singular and plural) denote entities in a domain of atomic entities, which distinguishes plural count nouns from mass nouns. This distinction enables an analysis that can explain how reference to entities occupying the same space can be achieved by different nominal expressions, like *ring* and the *gold* the ring is made of (cf. Link 1983, p. 135).

Krifka (1989, 1995) follows in this direction but assumes a single domain that is non-atomic. That means every potential denotation target of a given noun is non-atomic – like the denotation of mass nouns – and what makes counting possible is a function applied to that non-atomic concept. For count noun denotations that function is the 'NU (natural unit) extensive measure function' which provides the information about what can be called or interpreted as *one [NOUN].*

In his more recent works, Chierchia (2010, 2015) pinpoints the difference between count and mass nouns on their different denotations. According to him, count nouns – but not mass nouns – 'denote stable atoms'. That is, in any context it is unambiguously clear what counts as *one [count noun]*. Say for *boy* it is clear that *a boy* is the entirety of exactly one male child, not half of it and not two of them, but exactly one. Contrary, for *rice* it is not clear what exactly *a rice* would be, e.g. several grains, one grain, half a grain, some portion of ground-up grain dust. Rather, all those quantities are said to be part of the 'vagueness band' of *rice*'s denotation (Chierchia, 2010). Pluralisation, as well as counting (that is directly combining with numeral), Chierchia argues, requires unambiguous – that is not vague – atoms and therefore mass nouns can neither be pluralised nor counted.

[12] Note that 'individuals' here does not refer to object-like bounded things in the physical sense but to a certain kind of entity, namely individuals as opposed to e.g. events.

Rothstein (2010) builds on Krifka (1989, 1995) in rejecting Link's (1983) two-domain (atomic and non-atomic) ontology which gives nouns their denotation, but rejects Krifka's proposition of the natural unit measure function that states that nouns (e.g. *cow*) have cumulative predicate denotations that are quantized with an extensive measure function to arrive at a derived meaning of singular count nominal predicates where they denote entities that have no proper part falling under the same predicate. That is, a cow's tail cannot be called *a cow* or even *cow* at all. Rothstein (2010) criticises this idea by pointing towards nouns that she calls 'homogenous', such as *fence*, claiming that they fail to be quantised since they may have proper parts that fall under the same predicate. Generally, in her account the denotation of a given noun intersects with a subset of the domain called the 'counting context'. For default contextual interpretations, the context yields a set of entities that count as *one [noun]* by removing overlap. Overlap here refers to the phenomenon that it is sometimes unclear what counts as one, which can be nicely illustrated by envisioning a cup and a saucer and trying to make out how many items of kitchenware that is/those are. Mass nouns, so the account states, are predicates of individuals ("of type <e, t>") that may be naturally atomic, whereas count nouns are predicates of indexed individuals ("of type <e x k, t>", where k is the counting context) and are 'semantically atomic'. Semantically atomic nouns refer to individuals that are indexed for the context k, in which they are 'disjointed' and count as one and are thus countable. This way, it is claimed, it is possible to account for such situations like a square surrounded by fence, where one person might say that there is one fence, while another person might argue that there are four fences. *Fence*, thus, has semantic atoms, though they vary depending on the context, and is therefore a context-sensitive count noun, compared to say *boy*, which has the same semantic atom in all the counting contexts.

In Landman (2011) so-called generator sets are proposed that determine for a given nominal predicate what counts as *one [noun].* Crucially, counting is only possible if the generators in the set are not overlapping, as is the case in e.g. a set of cats [Trille, Sissa, Murre] but not for a set of water, that cannot be unambiguously divided into single 'waters' (say, [molecule1, molecule2, molecule1⊔molecule2, puddle1⊔molecule1]), just as a cup and a saucer might be one or two pieces of *kitchenware*. Instead of making a choice between all the different possibilities that the entities in the domain of *water* or *kitchenware* could be counted, Landman (2011) argues that counting goes wrong because of the overlapping generators in the noun's denotation.

As in his previous work, the refined account in Landman (2016) rests on the idea that lexical entries for nouns represent 'i-sets' with a base and a body. The notional entities for counting are formed by bases, and bases generate bodies under sum (⊔). If a noun's base is disjointed as is the case for *cats*, it is a count noun, if not it is a mass noun.

Next, a viewpoint is summarized that sees countability not as a binary factor but as a spectrum in order to preserve the idea of a correspondence relation between things and stuff in the world and the nouns referring to them without running into the same challenges that both extremes outlined above face.

2.3.2. Countability as a property of noun phrases

Authors like Pelletier and Allan (e.g. 1980), to name the most prominent ones, regard countability shifts not as exceptions or transformations of a bivalent property but claim that all nouns are flexible in their ability to be interpreted as either a bounded or unbounded entity.

It is not such an outrageous idea, considering that language is first and foremost a tool of communication and naturally bends to the communicative needs of its users. That is, if a speaker has the need to portion blood or mud or refer to an unbounded substance that usually occurs confined in the body of a bounded entity in the world then countability shifts (or the insertion of a noun in an atypical syntactic environment) gives this speaker an opportunity to do so.

Consider the sentences in Example (19)[13]. In a corpus, those nouns only occur in count-specific noun phrases (Examples (19)a and b) or mass-specific noun phrases (c and d), respectively. Yet, when native speakers are given enough time to come up with an appropriate context, they have no problem forming sentences with the nouns in incongruent environments that are perfectly interpretable. (Examples are taken from the sentence production study reported in this dissertation. For more information see Section 3.1 "Pretests".)

(19) a. Viel Pflanze ist nach dem Unfall mit dem Blumentopf nicht mehr übrig.
There is not much plant.SG left after the accident with the flowerpot.

b. In der BILD ist nicht wirklich viel Zeitung.
There is not really much newspaper.SG in the BILD. (= BILD does not have many properties of a newspaper, e.g. text, facts, results of journalism)

c. In der Wüste Gobi liegt ein Sand wie man ihn sonst nirgends anders findet.
In the desert of Gobi there is a sand.SG that one cannot find anywhere else.

d. Der Bauer jagte uns vom Hof, wenn wir auch nur eine Gerste berührten.
The farmer chased us off the farm when we just touched a barley.SG.

While other authors insist that count nouns, but not mass nouns have 'atomic reference' and mass nouns, but not count nouns experience 'divisibility' to reinforce their view of countability being an inherent property of nouns, Gillon (1999) advocates the view that these notions are insufficient to determine if a noun is count or mass, since in his opinion there are count nouns that experience divisibility and mass nouns that have atomic reference.

Atomicity generally refers to the phenomenon of a noun denoting exactly one bounded entity of the referenced kind. For example the word *chair* refers to exactly one chair in the world. That is, two chairs cannot be referred to by *a chair* and neither can a part of the chair, say a leg, be called *a chair*. Divisibility – a concept that's closely related to the term 'cumulative reference' – describes that a word like *sand* refers to any part or sum of sand[14]: One kilogram of sand can be called *sand*; if another kilogram is added to that the resulting sum is still *sand* and when 200 grams of that are taken away, both remaining portions are still called *sand*, too.

Gillon names words like *rope, rock, ash, cord* and *tile* as count nouns with cumulative reference and states that mass nouns like *furniture, silverware, jewellery, clothing, traffic* and *infantry* have individuals or 'atoms' in their denotation, arguing for the inadequacy of those properties to classify nouns for countability.

Similarly, many authors have rejected the idea of a one-to-one mapping of objects and substances on the one hand and count and mass nouns on the other, citing cross-linguistic and intra-linguistic

[13] Example (19)a-c are first shown in the introduction chapter as Example (15)a, b and (14)a.
[14] Divisibility is reference to a proper part of the whole, while cumulativity is reference to a sum or both a sum and a part.

differences in nouns referencing the same real-world thing or stuff (see Chierchia, 1998b, as a prominent example). For instance, the English word *hair* is mass, while it's French equivalent *cheveux* is an invariable plural (Palmer, 1981), and in German there is both a mass noun *Haar* and a count noun *Haar(e)* showing that even within one language, there is variability.

Before discussing whether those nouns really represent cases of 'regular count and mass nouns' denoting concepts opposite to the class they are supposed to denote, theoretical accounts are summarized that argue that count and mass nouns are semantically distinct from each other and that atomicity and cumulativity can be used to determine which noun belongs to which class.

2.3.3. Countability as a continuum

The first line of work I want to use to exemplify this is by Peter Sutton and Hana Filip (Sutton & Filip, 2016a, 2016c, 2018, 2019; Filip & Sutton 2017). They focus on cross-linguistic data and divide noun referents into five classes, with the first (prototypical objects) and the last (substances, liquids and gasses) being relatively stable across languages with regard to their countability, while the other classes are regarded as being more variable across and within languages. Table 3 displays the five classes with examples of several languages.[15]

Table 3: Noun Classes of different countability behaviour, from Sutton & Filip, (2016c, p. 14)

Noun Class	**Examples**
Prototypical Objects	chair$_{[+C]}$; tuoli$_{[+C]}$ ('chair' Finnish); Stuhl$_{[+C]}$ ('chair' German) dog$_{[+C]}$; koira$_{[+C]}$ ('dog' Finnish); Hund$_{[+C]}$ ('dog' German) boy$_{[+C]}$; poika$_{[+C]}$ ('boy' Finnish); Junge$_{[+C]}$ ('boy' German)
Collective Artifacts	furniture$_{[-C]}$; huonekalu-t$_{[+C, PL]}$ ('furniture', Finnish); meubel-s$_{[+C, PL]}$ ('furniture', Dutch); meubiliar$_{[-C]}$ ('furniture', Dutch); kitchenware$_{[-C]}$; keittiöväline-et$_{[+C, PL]}$ ('kitchenware', Finnish); Küchengerät-e$_{[+C, PL]}$ ('kitchenware', German) footwear$_{[-C]}$; jalkine-et$_{[+C, PL]}$ ('footwear', Finnish); Schuhwerk$_{[-C]}$ ('footwear', German) jewelry$_{[-C]}$; koru-t$_{[+C, PL]}$ ('jewelry', Finnish)
Homogenous Objects	fence$_{[+C]}$ vs. fencing$_{[-C]}$ wall$_{[+C]}$; Mauer$_{[+C]}$ ('wall', German); walling$_{[-C]}$; Gemäuer$_{[-C]}$ ('walling/masonry', German) bush$_{[+C]}$; Busch$_{[+C]}$ ('bush', German); shrubbery$_{[-C]}$; Gebüsch$_{[-C]}$ ('shrubbery', German)
Granulars	lentil-s$_{[+C, PL]}$; linssi-t$_{[+C, PL]}$ ('lentils', Finnish); lešta$_{[-C]}$ ('lentil', Bulgarian); čočka$_{[-C]}$ ('lentil', Czech) oat-s$_{[+C, PL]}$; oatmeal$_{[-C]}$; kaura$_{[-C]}$ ('oat', Finnish); kaurahiutale-et$_{[+C, PL]}$ ('oatmeal', lit. oat.flake, Finnish) bean-s$_{[+C, PL]}$; pavu-t$_{[+C]}$ ('beans', Finnish); bob$_{[-C]}$ ('bean', Bulgarian)

[15] Though, I do not agree with all of the classifications, especially regarding the German examples, most importantly *Gemäuer* and *Gebüsch*.

Substances, Liquids & Gasses	$mud_{[-C]}$; $muta_{[-C]}$ ('mud' Finnish); $Schlamm_{[-C]}$ ('mud' German) $blood_{[-C]}$; $veri_{[-C]}$ ('blood' Finnish); $Blut_{[-C]}$ ('blood' German) $air_{[-C]}$; $lenta_{[-C]}$ ('air' Finnish); $Luft_{[-C]}$ ('air' German)

Additionally to the typological motivation for the classification, the authors claim that 'granulars' and 'collective artifacts' resist mass-count coercion in the way that mass nouns referring to substances, liquids and gasses do not. The classic packaging shift that is felicitous for some mass nouns of the last class (*three waters* in the sense of three portions of water) is said to be impossible for those classes.

> "[G]ranulars and collective artifact mass nouns resist object 'packaging' readings: *Three furnitures* cannot be read as three ITEMS OF furniture, and *three rices* cannot be read as three GRAINS OF rice."
> (Sutton & Filip, 2016c, p. 42, italics and other markings adapted to current formatting).

They adapt the notion of counting context from Rothstein (2010) and the notion of overlap from Landman (2016) (here called 'precisification') and describe the behaviour of the different classes with regard to what can count as *one [noun]* given what context.

Grimm (2018) is the second work used here to exemplify such a viewpoint. Examining cross-linguistic data on grammatical number systems as well as psycholinguistic and philosophical work on what he calls 'individuation' (see Section 2.1 *How we think about things and stuff*), Grimm (2018) also rejects the idea of a direct one-to-one mapping between entities and nouns but proposes that it is possible to estimate the likelihood of a given entity to be referred to by a countable noun based on certain properties not only of its natural composition but also based on the way language users interact with it; much like the more 'agent-like' a thematic role is, the more likely it is to be realised as the grammatical subject of a sentence (cf. Dowty, 1991)[16].

In this, he highlights first the flexibility and changeability of a given word's countability status, nicely citing the example of *chad*, which (as he claims, p. 22) was mostly used as a mass noun describing the left-over paper from punch-card ballots, until it became necessary to count the individual pieces to verify proper election procedures, shifting the grammatical status of the noun to a dual use with both a mass and a count version.

Secondly, he explains the cross-linguistic variability of countability by the fact that different cultures – and therefore different language communities – interact with many forms of entities in different ways, focussing on different aspects of their composition or behaviour, such as the size of animals and whether they are domesticated, being used as food or usually occur in groups, like swarms or herds.

In their discussion of object-mass nouns, Grimm & Levin (unpublished manuscript) point out the artificial nature of their referents to be of crucial importance. They follow Nichols (2008) in arguing that artefacts – but not natural kinds – are "characterised by an associated event" (p. 3) and state that while object-denoting nouns like *pen* require a single entity to participate in that event, object-mass nouns allow for (and possibly need) more than one entity. That – so they argue – is what prevents the singular/plural contrast and therefore countability of this class, which they call 'artifactual aggregates'.

The emphasis on noun denotations being used as participants of an event is outlined in Levin et al. (2006) who states that this is especially the case for nouns denoting artefacts, referencing Brown (1999) among

[16] Though Grimm points out the crucial difference between those scales regarding truth conditions

others, who found that in Native American languages many[17] newly acculturated objects have received names that either directly refer to their function (e.g. the name for *money* equals something like "which we buy with") or draw parallels to already named artefacts (e.g. something equivalent to "wood sack" for *box*).

Grimm and Levin's account highlights (i) the functionality of artefacts and that that functionality may be reflected linguistically, (ii) that object-mass nouns – due to their superordinate nature – denote a multitude of objects and (iii) that this set of denotation targets consists of heterogeneous entities.

Taken together, their work explains crucial differences to other classes of nouns. Namely, that object-mass nouns (or 'artifactual aggregates') usually function as a group in the real world and therefore are treated as such linguistically. *Furniture* for example cannot be counted in the sense that one piece of furniture (e.g. a single chair) is called *a furniture* because one piece of furniture is not enough to furnish a room. Thus, a single chair might be acceptable as the only participant in a sitting event and can thus fulfil the requirements of the noun *chair*, but it cannot furnish a room (on its own) and therefor does not fulfil the requirements for being *a furniture*.

In conclusion, these works consider more (non-linguistic) factors than the ones outlined above and are thus better equipped to face the challenges presented by counterexamples that the community of researchers of linguistic countability has spent much of their time debating. Additionally, the account by Grimm & Levin (unpublished manuscript) honours the fact that language is continuously used and changed and adapted by human beings according to their needs.

2.3.4. Discussion

Applied to a psycholinguistic point of view, some of the notions in the accounts may be difficult to conceptualise, not in the least because the distinction between non-linguistic conceptual entities and linguistic word forms are not as explicitly made as it is typically done in empirical cognitive science.

What is common to all accounts is the assumption of a nominal domain – be it completely atomic / non-atomic or divided into sub-domains – which is the conceptual realm of non-linguistic entities that can be targeted by a noun's denotation. The crucial issue to understand here is that, despite those conceptual entities being called atomic (or non-atomic), what this property is often defined as is that if a noun does refer to them then its reference is atomic. So, in actuality, the conceptual entities in a domain are not atomic, but have the potential to take part in an atomic relation of reference or denotation between the noun and themselves.

From that point on, the accounts vary on how exactly the relationship between the noun and the conceptual denotation target in the domain is formed. While some accounts, like Chierchia's (e.g. 2010) and Link's (1983), state that there is just a certain class of concepts that can form denotation relationship with singular count nouns (namely those that "are atomic"), accounts like Rothstein's (2010) and Krifka's (1989, 1995) assume additional functions or contextual elements that need to combine with the concept in order to make it eligible for count noun denotation.

[17] about 2/3

So, to be absolutely clear, it is not the case that these accounts explain what kind of properties or form the conceptual entities must have, but rather in what way the denotation process takes place in order to combine a given concept with a given word form in a reference relationship.

Whether these accounts assume countability to be a lexical property is hard to make out, not only because it is unclear whether the additional measures needed to determine "what counts as one" such as the natural unit extended measure function or the precisification context is part of the mental lexicon. The main problem for the current purpose of this dissertation is that these accounts do not include any assumptions on processes of language production or comprehension (in the way psycholinguistic accounts do) that would allow one to derive testable predictions from them that could be confirmed or disproven by empirical testing.

These accounts attempt to explain the phenomenon of the mass/count distinction by means of the dissociation of count and mass nouns on the one hand and objects and substances on the other. While all of these accounts have no problem explaining the behaviour of classic examples, where count nouns denote physical objects and mass nouns refer to non-solid substances, problems typically arise in cases where a noun is analysed as 'grammatically mass' but 'physically an object'.

At this point a second type of misunderstanding is typically encountered namely the blending together of properties of real-world entities and their conceptual representations within the human mind. That point has led to much confusion, because so many properties of actual physical entities in the world are discussed (like the molecular structure of salt in Landman (2011) or the aggregate states of rice in Chierchia (1998)). Rather, the field of semantics is concerned with denotation; that is a relationship between a linguistic entity, such as a noun word form, and a conceptual entity in a speaker's mind, none of which has a molecular structure or an aggregate state.

The focus on real-world entities rather than mental concepts by most of these accounts (Scott Grimm's work (e.g. 2018) possibly being an exception because of the focus on human interaction with such entities) has steered the discussion towards cases that precisely not deal with classic things or stuff, but with entities that are classified as in-between cases and their denotation relationships. Some ideas and problems are discussed below.

Solid substances

One example of such an exception is the mass noun *gold* referring to a substance that is however mostly encountered in solid form (a brick or bar, a clump). Under the assumption that objects (i.e. entities that do not change form when moved through time and space) are typically referred to by count nouns and that *gold* is used to refer to a gold bar, the noun *gold* is said to constitute an outlier that does not follow the assumed pattern (see also Erbach, 2021, on that point). This problem is not to a small extent rooted in the fact that many of these accounts base their definition of what constitutes an object and a substance on the work of Soja et al. (1991) where only non-solid substances are considered, even though there is a vast body of psychological literature on the properties of objecthood (partially summarised in the section 2.1 *How we think about things and stuff*).

Some accounts thus aspire to extend the definition of substance. Rothstein (2010) for example, included solid entities that are used to construct artefacts (like gold and wood, cf. p. 343-344) in the category of substances arguing that the physical boundedness of a given gold sample does vary over time in that it depends on the artefact it is constructed into.

In a similar endeavour, it has been pointed out that while not technically artefacts, referents of nouns like *cattle* or *infantry* might be 'natural artefacts' since that is how humans treat them, and those nouns could therefore be included into the class of object-mass nouns (see e.g. Erbach, 2021, for a discussion of *wildlife*).

Divisible count nouns, dual nouns and nouns with overlapping reference

The dual nouns discussed in most of the semantic literature regarding countability are constrained to nouns like *rope(s), stone(s), cake(s)* and so forth. That is, these nouns have two word forms, singular and plural, like count nouns but the singular form is ambiguous with regard to countability, or put differently there are two realisations, one count with a singular and plural form, and one mass with only a singular form. That is, depending on the (syntactic) context they occur in, those nouns are interpreted either in the same way as classical count nouns or classic mass nouns, see examples (21) and (22) for this phenomenon in German.

(21) a. Johann warf zwei Steine auf mich; einer davon traf mich am Kopf.
John threw two rocks at me; one of them hit my head.

b. Diese Wand ist aus Stein gemacht.
This wall is made of rock.

(22) a. Ich nutzte zwei Seile um ihn festzubinden; Eines band ich um seine Handgelenke, das andere um seine Fesseln.
I used two ropes to tie him down; one I tied around his wrists, the other one around his ankles.

b. Wir banden etwas Seil um das Ende der Klinge, sodass wir das Messer halten konnten.
We wrapped some rope around the end of the blade so we could hold the knife.

Note that all the 'count nouns' used in Gillon (1999) to exemplify a subset of count nouns that experience divisibility or cumulative reference are actually dual nouns. Consider first the example nouns *rope, rock* and *cord*. While Gillon admits that those are dual nouns, he claims that the cumulative reference applies to their count word forms, but does not offer any data to strengthen his claim, other than the idea that one can have a cord which is cut producing two instances of a cord.

Here again, the unclear distinction between things and words becomes clear. One cannot cut *a cord* (noun phrase) but only a cord (physical entity), just as the noun *rice* cannot be classified by the size of individual grains, because nouns don't have grains. The fact that the same noun can occur in different noun phrases (of different countability) cannot be explained by physical operations on the material. As should be apparent from the lexical projection tasks outlined in the previous chapter, all physical entities have both a shape and a material. Whether a noun refers to one or the other is a matter of focus.

It is also important to understand that the dual-noun examples above do not describe the same phenomenon as the one that is exemplified by the class of 'homogeneous objects' of Sutton & Filip (2016c), where a given real-life entity can be referred to by different nouns of different countability, like *fence/fencing* or *Mauer/Gemäuer*. In this case, I assume that we are dealing with reference distributions

of several nouns that can not only be of very different frequency but also of very different age (of the noun, i.e. a noun's usage is part of an outdated, deprecated way of language usage). That is, different speakers use these terms differently, depending on their needs, age and vocabulary size, among other things.

A nice example of this is the term *Bein* in German, and variations thereof. Consider the examples in (23) that were inspired by the consideration of *ash(es)* in Gillon (1999). The noun *ash(es)* is a special case (of English, possibly inspired by French) as its singular form describes a classic substance and is realized as a mass noun whereas the plural form *ashes* is mostly limited to the reference of residue of specific remains (mostly human)[18], similar to *Gebeine* in German, which is a plurale tantum grammatically.

(23) a. Dass Spinnen keine Insekten sind, kann man daran erkennen, dass sie acht Beine haben und nicht sechs.
That spiders are not insects can be seen from the fact that they have eight legs, not six.

b. Die Gebeine des Heiligen Nikolaus wurden angeblich in der Türkei entdeckt.
The remains of Saint Nicholas were allegedly found in Turkey.

c. Dieser Kamm ist aus Bein gefertigt und gehörte mutmaßlich einer germanischen Stammesfürstin.
This comb is made of bone and allegedly belonged to a Germanic tribal leader.

While Example (23)a should be perfectly acceptable for any German native speaker, the noun *Gebeine* in (23)b is already much less frequent[19] and may be unknown to many speakers. The mass noun *Bein* referring to the material, in turn, is so infrequently used that even this linguistically well-trained native speaker had only been aware of it as part of *Elfenbein* (*ivory*, even if technically tooth not bone) until its existence was pointed out by an anthropologist.[20]

It's thus plausible to assume that *Bein* was once used equivalent to *bone* (also notice the phonological similarity) and only later shifted to mean *leg*, while the original meaning was preserved in the mass noun *Bein* and as part of *Gebeine*, which denotes all the bones of a (bare) skeleton and not just the legs.

Similarly, it could be the case that examples like "ashes to ashes" and "to rise from the ashes" where *ashes* refers to remains, reflect a special, rather outdated usage compared to the commonly used mass expression *ash* (= material remaining from burned wood/coal/cigarettes/etc.), that has however now been used more widespread in metaphoric usage of the noun. The example *ash(es)* can therefore not be considered parallel to *rope(s)* and *stone(s*), but more similar to *fence/fencing*; a class that experiences overlap as the source of its individuation difficulties rather than divisibility.

[18] In the Corpus of Contemporary American English, only 23 out of the first 100 occurrences of *ashes* (plural) are describing a substance, while 19 describe animal remains and 46 are used metaphorically as in *reduced to ashes* or *rising from the ashes*, (https://www.english-corpora.org/coca/, 02/17/2021).

[19] Leipzig Wortschatz on November 17, 2020: Bein(sg) 10068 occurences, Beine(pl) 21486 occurences, Gebeine 655 occurences; though note that Bein(mass) is part of the first counting Bein(sg); *© 2021 Abteilung Automatische Sprachverarbeitung, Universität Leipzig.*

[20] Personal correspondence with Dennis Beckmann, PhD, Museum der Deutschen Binnenschifffahrt

As for *tile*, it might be that in English a shift such as the one in Example (24) might be considered valid. However, it is doubtful that any native speaker would refer to the entities depicted in Figure 9 as *a tile*.

(24) I bought a lot of tile because I am remodelling my bathroom.

We are thus dealing with a (possibly lexicalized) shift in countability, similar to *three beers* meaning 'three portions of beer' and not with the divisibility of a count noun.

Figure 9: A sum (left) and a piece (right) of tile(s), not referred to by *a tile*

Collective nouns

The other case of objects being referred to by mass nouns is the one of collection denoting nouns. For homogeneous collections such as *rice* and *lentils* (also referred to as 'aggregates' or 'granulars') it is argued that their substancehood is not as straightforwardly determinable as it is the case for e.g. liquids. Therefore, for instance Sutton & Filip (2016c) argue the countability of the respective nouns referring to collections of that sort varies within and across languages (see Table 3 above).

Nouns referring to heterogeneous collections have received the most attention in the semantic literature, where they are called 'object-mass nouns' (alternative/earlier names include count-mass nouns, neat mass nouns, fake mass nouns, superordinate terms and (naturally) atomic mass nouns). These nouns refer to (mostly artificial) objects of the same function, such as *footwear, silverware, cutlery, jewellery* or *furniture* in English and *Besteck, cutlery, Schmuck, jewellery, Gebäck, pastries,* or *Schrott, rubbish/ scrab metal,* in German. Some authors also include words like *foliage* and *cattle* into this category. While some insist on the artificial nature of the individual members of the collection as a qualifying property, others argue that the class is defined through the apparent mismatch of physical boundedness of their referents and the fact that the nouns are grammatically mass.

What makes those nouns a challenge for semantic accounts of the count/mass distinction is that they are grammatically mass[21] but are consistently said to refer to objects in the classical sense, i.e. physical entities with invariable boundaries. As evidence for that, the study by Barner & Snedeker (2005) is cited most frequently, where the referents of such nouns were compared according to cardinality rather than volume. This irregularity is further complicated by the assumption that those nouns do not shift reference like other mass nouns, namely it is allegedly not possible for them to undergo the packer and sorter shift (Sutton & Filip, 2016a).

[21] Though it has been pointed out in e.g. Rothstein (2010) that e.g. *furniture* behaves differently from 'substance mass nouns' in that it combines with "stubbornly distributive predicates" like *round* and *big*, and with reciprocal operators like *each other*.

(25) a. *Drei Schmuck/Schmücker habe ich in der Schublade gefunden.
I found three jewelry/jewelries in the drawer.

b. *We can put another two furnitures in the truck.

c. *Wir haben drei Schmuck/Schmücker hier: Gold, Silber und Modeschmuck. [22]
We have three jewelry/jewelries here: Gold, silver and custom jewelry.

d. *There are two furnitures upstairs: bedroom furniture and sitting room furniture.

In this thesis, as in my previous work outlined in Beckmann, Indefrey & Petersen (2018)[23], I lay out an alternative view on these nouns. Namely, that it is a false assumption that they refer to objects, as they crucially refer to collections of objects, similar to plural count nouns. I don't precisely follow Chierchia (e.g. 2010) in that I don't view nouns as operating in a conceptual domain of possibly 'atomic' denotation targets nor do I argue that all noun denotations have to be analysed as participants of potential actions like Grimm & Levin (unpublished manuscript). Rather, I take the German and Dutch translations of *furniture* as a starting point to explain the following:

(i) So-called object-mass nouns do not refer to bounded, atomic objects at all but to collections thereof, which calls the entire premise into question.

(ii) Object-mass nouns like *Schmuck, jewellery,* and *Besteck, cutlery,* have count noun equivalents which refer to the individual members of the collection, e.g. *ring* and *necklace* for *jewellery* and *knife* and *fork* for *cutlery*, just like plural count nouns like *dolls* have singular forms referring to one individual doll, which block the shift of the mass noun's word form to refer to one individual collection member.

Conceptually, there is of course no argument about the fact that chairs and tables and ottomans and such are in fact objects in the same way dolls, apples and knifes are. Therefore, it is not surprising that when confronted with a linguistic cue targeting the function-aspect of a collective concept, such as "Who has more furniture?", participants determine the amount 'most useable'. Since a greater number of individual pieces of furniture can be 'better used' to furnish a room than a smaller number of giant items, participants consistently say that many small items is 'more furniture' than few big items. Moreover, it has been found that participants judge a heterogeneous set of five pieces of furniture (a sofa, two chairs, a coffee table and a bookshelf) to be 'more furniture' than a set of five chairs by Grimm & Levin (unpublished manuscript).

Especially the word *furniture* has been called upon as a classic example of that ominous class of nouns and while the English version is clearly a mass noun grammatically, other Germanic languages realize this word in other forms. In German for example, *Möbel* is a plurale tantum, i.e. an invariable plural with no existing singular form[24]. Individual pieces of furniture are referred to with the count noun

[22] It should be noted at this point that the noun phrase "dreierlei Schmuck" would be felicitous in this sentence.
[23] Note that Natalja S. Beckmann, one of the authors of Beckmann, Indefrey and Petersen (2018) is identical to the author of the current dissertation, Natalja S. Peiseler, who in 2019 changed her last name as a result of marriage.
[24] The Duden does offer a singular form, *(das) Möbel,* however I suspect this form to be referring to a secondary meaning of 'big, clumsy item' originally, influencing the potential acceptability of the NP *ein Möbel* in comprehension, or is a remainder of the usage of the original French word *(le) moeble*..
https://www.duden.de/rechtschreibung/Moebel, 06/23/2022

Möbelstück(e), while the mass noun *Mobiliar* is also a collective noun. Similarly, Dutch has three word forms, a mass noun *meubilair* and a count noun *meubel* with the plural form *meubels* or *meubelen*. This more complete view on the possible reference-targets in the furniture-concept opens up the perspective towards complex concepts and gives rise to the hypothesis that a word form can only denote a given aspect of the concept if

(i) that aspect meets the (semantic) requirements of the word form, e.g. with regard to plurality, and
(ii) the target aspect of the concept is not already realized by the language in another (straightforward) way.

That is, *furniture* and *cutlery* are mass nouns because they do precisely not refer to an (i.e. one) individuated object, and can be combined with predicates like *big* or *round* for the same reason plural count nouns can be combined with them.

(26) a. Frau Schmidts Katzen sind groß.
Mrs Smith's cats are big.

b. Frau Schmidts Besteck ist rund.
Mrs Smith's cutlery is round.

Both sentence a and b in (26) only hold true, if there is more than one cat or more than one item of cutlery, respectively, and all members of the sum of them are big or round.

The nouns in question, however, can't shift to mean 'piece of furniture' or 'item of cutlery' because for that concept other nouns like *chair* and *knife* are already available and there is no need to use a less precise label[25].

For portions of other sizes (compare the concept of 'overlap' mentioned above) mass nouns of that conceptual class behave just like mass nouns referring to substances, such as *water*, as is shown in the following example:

(27) a. Wir brauchen noch ein / einmal / zwei / zweimal Besteck für Tisch 7.
We still need one / once / two / twice cutlery.SG for table 7.

b. Wir brauchen noch ein / einmal / zwei / zweimal Wasser für Tisch 7.
We still need one / once / two / twice water.SG for table 7.

(28) a. *Nachdem Gabel und Messer unter den Tisch fielen, war nur noch ein Besteck zu sehen.
After fork and knife had fallen under the table, only one cuttlrey.SG could still be seen.

b. *Der Doktorand isolierte ein Wasser unter dem Mikroskop.
The PhD student isolated one/a water.SG under the microscope.

In Example (27), conventional portions like a set (knife + fork + spoon) of cutlery (see (27)a) and a glass of water (see (27)b) are referenced and thus the shift is acceptable. In (28), however, an unconventional portion is used, such as one piece of cutlery and a molecule of water and therefore the usage of the

[25] Remember Grice's maximes at this point, (Grice, 1975).

mass noun is infelicitous, i.e. the shift is unsuccessful. As can be seen, this holds for both 'classical mass nouns' (see (28)b) and those mass nouns that refer to collections (see (28)a).

Thus, in conclusion, most of the problems and challenges the semantic accounts to countability are facing are rooted in imprecision when it comes to the differentiation of things/substances, concepts and words, as becomes clear from the examples discussed above. Additionally, most accounts lack consideration of all the facets and aspects of language, be it typological, pragmatic or operational, and are therefore not well serviceable as a base for psycholinguistic hypotheses.

For the purpose of the current research question, namely "Is countability a lexical property of nouns?", unfortunately only few of the ideas and considerations outlined here give any direct insight.

In the next section, the idea of complex concepts, metonymic reference and polysemy are explored in order to gain a perspective into denotation relations of a more complicated nature that might help explain the convoluted, tangled and incoherent picture presented by noun countability.

2.4. Words that mean many things (or stuff) or: Polysemy, Metonymy and the Lemma-Concept Interface

This section aims to review some ideas on polysemy and complex concepts with regard to the main question of the thesis: Is countability a lexical property?

Assuming that countability is a lexical property then it must be the case that for the majority of count and mass nouns (namely at least those that are not lexicalized dual nouns) either an object-like denotation or a substance-like denotation is preferred, if not even exclusively available. If, however, as some theoreticians like Pelletier and Allan believe (see Section 2.3 *How linguists think we talk about things and stuff*), all nouns are basically dual and can switch more or less freely between those two interpretations depending on their syntactic environment, then it should follow that a given noun denotes both these meaning aspects. That in turn would necessitate a more complex view of the lexical entry and the noun's connection to the conceptual denotation targets.

To that end, theoretical and empirical literature on polysemy and metonymy is reviewed in order to find some insights about how words can denote multiple sometimes most different concepts and how their lexical entries are to be envisioned consequently.

2.4.1. Co-predication, anaphoric reference and types of polysemy

A most cited, classic example of a polyseme is the case of *book* or *letter* which on the one hand denotes a physical entity and on the other hand some informational content. Words like *letter*, just as the examples of *university* or *Brazil* below are felicitous in sentences with co-predication and anaphoric reference, where the same noun is used to denote different concepts or conceptual aspects, such as in the examples in (29).

(29) a. The letter from the Council is on the table. It sounds a bit threatening.

b. I have a meeting with Laura at the university which has signed the manifesto.

c. Brazil, the largest country in South America, is a republic and has won five World Championships.

(Examples adapted from Ortega-Andrés & Vincente, 2019)

Each of the usages of *letter*, *university* and *Brazil* here denotes multiple aspects of the word's meaning, or a different quale of the entity's nature, to put it in Pustejovsky's terms. It's the context that restricts the meaning of the noun (or pronoun) to those conceptual aspects that can fit the semantic requirements of the predicate, like that only a physical entity can 'be on the table' and only informational content can 'sound threatening' but not the other way around. The interesting question is, are all those aspects part of the same (complex) concept and if so, can the respective noun refer to all of them in the same way?

Weinreich (1964) differentiates homonymy from polysemy by stating that the different denotations of homonymous word forms are 'contradictory' and therefore easily to disambiguate given enough context, whereas the different senses of polysemous words are 'complementary' and one sense may simply be more appropriate or focussed on in a given context.

It is hard to say, however, if the reference to an individual entity or to a substance, as in countability shifts of words like *beer* in (30)b, is as 'contradictory' to the substance-sense in (30)a as using *bank* to refer to a financial institution on the one hand and a waterfront on the other; a classic example of homonymy. In many examples, the context – especially the syntactical environment of the noun phrase – disambiguates the two senses of the noun in countability shifts rather clearly (compare (30)a with b). Yet, an individuated entity and a substance can have very different, even contrary, properties.

(30) a. Ich habe gestern viel Bier getrunken.
I drank a lot of beer yesterday.

b. Ich habe gestern drei Bier getrunken.
I drank three beer.SG yesterday.

c. Gestern habe ich drei Bier fallen lassen.
Yesterday, I let three beer.SG fall.

Vieban (2020) distinguishes three types of semantic variability: (i) mere polysemy, where one expression has multiple senses that each have one semantic value, (ii) polysemy cum context-sensitivity, where an expression has multiple senses that each have multiple semantic values and (iii) mere context-sensitivity, where an expression has one sense that has multiple semantic values. A sense here is a closely related conceptual aspect, such as the temporal and the spatial sense of *long*. In contrast, a semantic value is a denotation target that varies with context, such as the referent of *I* or *that*.

Unfortunately, Vieban (2020) states that the way to decide whether a word is polysemous, context-sensitive or both is based on 'linguistic intuition' regarding the number of senses and semantic value candidates, which seems a highly subjective measurement. Intuitively, in the case of (30), different senses seem more likely than different semantic values, however when one considers the reference of *Bier* in (30)a and (30)b to be a substance and in (13)c to be a substance and its container (say, a bottle),

the situation may be different. Without clearer criteria to differentiate senses from semantic values an analysis will yield no clean-cut result.

In the following, different perspectives and supporting empirical evidence on polysemy and metonymy are summarized. The analysis is limited to those works that might help to shed light on possibly complex concepts which include representations of bounded and unbounded entities, i.e. denotation targets of count and mass word forms, respectively.

Ortega-Andrés & Vicente (2019) differentiate between 'conventional polysemy' or 'regular polysemy' (after Pustejovsky, 1995) where different senses are lexicalized and are retrieved by 'sense selection' and 'non-conventional polysemy' where senses are said to be generated online by pragmatic mechanisms. Pustejovsky himself further argues that some words are 'inherently polysemous', such as *letter* and *book,* which are both a physical object and an informational content by their very nature.

Ortega-Andrés and Vicente argue for a view where polysemous nouns 'link' to a 'knowledge structure' including some – but not all – world knowledge on the entity described by the noun, how it is realised and implemented.[26] In that knowledge structure the typical senses of the polyseme are stored. The senses together form what they call an 'activation package'. All senses are activated when the noun is encountered and activating one sense co-activates all the others.

Thereby, their account enables them to unambiguously distinguish between polysemes and homonyms. Polysemous senses co-activate each other, while homonymous senses compete in lexical selection. The authors cite empirical evidence for their claim, namely that senses of homonyms differ in frequency (Frisson, 2009), while polysemous senses can prime each other and do not experience lower levels of activation once one sense is selected (Pylkkänen et al. 2006; Klepousniotou et al. 2008). Similarly, psycholinguists usually assume one shared entry for polysemes and different ones for homonyms. However, Ortega-Andés & Vicente (2019) caution from this view, citing different results for the *book* and the *paper* example (as in liberal paper and shredded paper) by Klein & Murphy (2002) and Foraker & Murphy (2012).

In their assumed knowledge representation for "drinkables" they include a container, stating that it is needed for the drinkable to be "actualized" as such, just as informational contents need a physical host of some kind, like a book [or a computer file, or some neurological structure] (my addition).

The authors give an example of a content-for-container shift that is felicitous with co-predication and anaphoric reference.

(31) a. The beer Susan was drinking fell out of her hands.

b. Susan was drinking her beer. It fell out of her hands.

(Ortega-Andrés & Vicente, 2019, p. 2)

The German equivalent in (32)a and b seems felicitous as well, though naturalness increases with more context (see(32)c):

[26] For *school,* this could be: "information about the kind of entity a school is, as well as information about its physical and temporal realization or implementation, the kind of people that take part in it, and its organizational structure." (p. 5)

(32) a. Das Bier, das Susanne trank, fiel ihr aus der Hand.
The beer that Susan was drinking fell out of her hand.

b. Susanne trank ein Bier. Es fiel ihr aus der Hand.
Susan drank a beer. It fell out of her hand.

c. In der vollen Kneipe trank Susanne ein Bier. Als sie angerempelt wurde, fiel es ihr aus der Hand.
In the crowded pub, Susan drank a beer. When someone bumped into her it fell out of her hand.

They further cite an example from Schumacher (2013) said to show that co-predication is less acceptable in cases where the container is used to refer to its content:

(33) ?The bottle that Susan was drinking fell out of her hands.
(see Schumacher 2013)

Here, the German version (see (34)a) also sounds odd and even more context (see (34)b) does not seem to make it better. The sentence does however sound more natural when the second predication is more semantically related to the first sense of the noun, as is the case in (34)c, where *a big bottle* also contains more drinkable liquid. Interestingly, when the shift is marked syntactically by placing the substance-denoting noun in a count-specific NP (combined with the indefinite determiner) as in (34)d, the sentence sounds acceptable.

(34) a. ??Die Flasche, die Susanne trank, fiel ihr aus der Hand.
The bottle that Susan drank fell out of her hand.

b. ??Lisa und Susanne haben in der Kneipe Bier getrunken. Die Flasche, die Susanne trank, fiel ihr aus der Hand, als sie angerempelt wurde.
Lisa and Susan drank beer in the pub. The bottle that Susan drank fell out of her hand when someone bumped into her.

c. ?Lisa und Susanne haben beide in der Kneipe Bier getrunken. Die Flasche, die Susanne trank, war allerdings größer.
Lisa and Susan both drank beer in the pub. The bottle that Susan drank was bigger, though.

d. In der Kneipe hat Susanne ein Bier getrunken. Die Flasche, die sie trank, fiel ihr aus der Hand, als sie angerempelt wurde.
In the pub, Susan drank a beer. The bottle she was drinking fell out of her hand when someone bumped into her.

Thus, the fact that a word like *beer* is felicitous in sentences with co-predication and anaphoric reference points towards a view where countability shifts (i.e. the packager shift, mass-to-count, in (32)b and (34)d) should be analysed similarly to classic polysemes like *book* or *letter*.
It is important, however, to honour the authors' exact intentions here. Ortega-Andrés & Vicente (2019) state clearly that they do not subscribe to a view where each polyseme necessarily denotes all possible conceptual aspects at once but that

> "knowledge structures […] only offer possibilities for denotation, i.e., a variety of possible denotations from which the speaker has to select. That is, the denotation potential of a word type is not explained in terms other than the information stored in the knowledge structure associated with such a word-type." (p. 15)

The utterance in its entirety, so they claim, is what can be unambiguously interpreted[27].

2.4.2. Metonymy, Shifts and Coercion

Contrary to that, Frisson & Frazier (2005) assume that countability shifts in the food and drink domain are not cases of 'true polysemy' but that there is a default sense (individuals or substance, in their terms) and a derived sense, which would classify countability shifts as 'non-conventional polysemy' (see above). The authors state that they adopt the assumption that countability shifts are a form of shift based on lexical derivational rules from Copestake & Briscoe (1995). They hypothesise that without biasing information the 'underlying sense' will be interpreted, that is they assume "that an abstract conceptual difference distinguishes mass nouns (substances) from count nouns (individuals) in the domain of food and drink terms" (p. 278).

'Meat-grinder shifts', so they say, cannot undergo copredication, as they exemplify with sentences like in (35).

(35) a. ?Mary fed and carved the lamb.

b. *Mary fed and carved lamb.

(Examples from Frisson & Frazier, 2005, p. 279)

In German, however, the example sounds quite natural, especially with a little more context:

(36) a. Es war Marias Aufgabe, das Lamm zu füttern und nach dem Schlachten zu portionieren.
It was Mary's task to feed the lamb and portion it after the slaughter.

With an eye-tracking experiment, the authors aim to test this assumption of a derivational rule application against what they call 'exposure hypothesis' which holds that the more one hears a shifted noun, the easier it is to interpret it this way, even in neutral context. The materials consisted of nouns referring to edible or drinkable things or substances, which were classified as count and mass nouns based on the authors' intuition and a corpus search of 20 appropriate sentences for the singular word form. The mass nouns were used with 'a mass interpretation' 65-100% of times and the count nouns were used with 'a count interpretation' 60-95% of times in the corpus. All stimulus nouns were matched for length and frequency of the plural form.

In Experiment 1, 'Portioning', the stimuli were presented in plural form and embedded in sentences, either as bare plurals (e.g. "imported pears/beers", neutral context) or preceded by a numeral or count-specific quantifier (e.g. "several delicious yams/stews", helping context). There was only a marginally significant effect indicating that mass nouns were a little easier to process in plural form when a preceding numeral/quantifier already prepared the reader for its occurrence. In the spill-over region of the noun, there was an effect of countability (called "type" in the original work, p. 285) in that mass

[27] They insist on that, explaining that this way sentences can have true conditions, while other authors (e.g. Chomsky, 2000) use these examples of polysemy to argue against truth condition semantics.

nouns in neutral context took longer to read than count nouns in neutral context. Additionally, mass nouns in the condition with helping context were read faster than in neutral context.

In Experiment 2, 'Grinding', the count and mass nouns were presented as singulars either with a mass-specific element preceding them (e.g. "just a small amount of imported beer/pear", helping context) or without ("imported beer/pear", neutral context). Here, the main finding was that there were more regressions towards count nouns in neutral context than in helping context. Also, the second-pass duration was longer there. Interestingly, for the other variables (single-fixation duration and gaze duration) only significant context effects for the mass nouns could be found, that is those nouns for which the sentences actually presented a congruent environment.

Despite this rather small result, the authors conclude that the processing mind commits to a default sense in the absence of biasing information and that shifting towards the 'derived sense' involves rule application.

They contrast the results with those found in Frazier & Rayner (1990) for what they call 'systematically related senses'. With the same eye tracking paradigm, Frazier & Rayner found that sentences containing words with multiple meanings such as *pitcher* (animate/inanimate) and *records* (concrete/abstract) take longer to read if meaning-disambiguating information follows the target word than when it precedes it. This was not the case for sentences with words with multiple senses such as *paper* (concrete/abstract)[28]. However, readers took longer to read the actual ambiguous target word when disambiguating information followed the target word in all three cases, which does in fact mirror the results of Frisson & Frazier (2005).

Schumacher (2013) tested two types of metonymy: container-for-content shifts as in (36) and content-for-container shifts, i.e. countability packager shifts, as in (37).

(36) a. Was hat Heinz hastig getrunken? (container-for-content shift)
Er hat <u>den Becher</u> hastig getrunken.
He drank the cup hurriedly.

b. Was hat Rolf wie seinen Augapfel gehütet? (control)
Er hat <u>den Becher</u> wie seinen Augapfel gehütet.
He guarded the cup like his eye ball.

(37) a. Was hat Asterix an seinem Gürtel festgeschnallt? (content-for-container shift)
Er hat <u>den Zaubertrank</u> an seinem Gürtel festgeschnallt.
He fastened the magic potion on his belt.

b. Was hat Miraculix vor dem Eintreffen der Römer gebraut? (control)
Er hat <u>den Zaubertrank</u> vor dem Eintreffen der Römer gebraut.
He brewed the magic potion before the Roman's arrival.

from Schumacher (2013, p. 5)

In two judgment task studies (with different participants) both classes of sentences that included shifts were judged as plausible (1.3 – 2.7 on a 1-to-5 scale) and as significantly more plausible than implausible

[28] Pitcher: animate= the person throwing the baseball, inanimate= a large container for beer
Records: concrete= the phonograph disc record, abstract= any written or otherwise sustainable proof of events
Paper: concrete= one instance of a newspaper (actual pages), abstract= the company publishing the newspaper

filler sentences (4.8 – 4.9). The container-for-content shifts were rated as slightly less plausible than the content-for-container shifts, but not significantly so.

The results of a subsequent ERP study conducted on those materials yielded a late positivity component for the container-for-content shifts as in (36), but no significant effect for the content-for-container shifts in any time window up to 1000ms post critical phrase onset.

The absence of an effect for the packager shift sentences is in line with Ortega-Andrés & Vicente's result and points towards a view of countability shifts (at least the packager shift) as being analysed as conventional polysemy.

Note however, that in this experiment the countability shift is not morpho-syntactically realized since in both the shifted condition (see (37)a) and the control condition (see (37)b) the definite article is used. The substance-to-object shift inherent in those stimulus sentences is only of semantic nature and arises from the selectional restrictions of the verb in use (*festschnallen, fasten*, and *brauen, brew*) that demand a physical object and a liquid, respectively. Because the verb occurs in sentence final position, questions including the same predicate occurred before the target sentence. It is unclear, whether this method of sense selection is equivalent to a countability shift that is syntactically marked.

In Schulzek (2014) a similar container-for-content shift is cited as an example for a contextually triggered metonymic shift, defined as an instance where "the reference of a lexeme is shifted from the potential referents of the lexeme to something that is in the broadest sense part of, or thematically linked to, these potential referents" (p. 222). The example "Jimmy drank a glass in one gulp" (p. 223) is listed along with other cases that have been named as examples of polysemy in the works described above, namely institution-for-employees, as for *university* in (29)b, or author-for-work like "a real Picasso".

The account utilizes frames, which are recursive attribute-value-structures that can be used to represent conceptual representations and their relationship to words (or lemmata)[29]. According to Schulzek, a metonymic shift involves the shifting of the reference node of a noun-frame, that is the conceptual aspect the lemma is linked to by default (often a place holder for the entire network of knowledge associated with a word, see Schulzek, 2014, p. 223-224, and Petersen, 2007, for more). The prerequisite for this shift is that the relationship between the 'old' reference node and the to-be reference node is one of 'bidirectional functionality'. That is, all value nodes need to be linked to the reference node by functional attributes and two nodes have a bidirectional relationship if there is another attribute that goes back to the first node. For example, considering the sentence mentioned above found in Schulzek (2014) "Jimmy drank a glass in one gulp", the glass Jimmy drank of is represented by a node 'glass' that has an attribute CONTENT which has the value 'content' (say beer, in this case). This content node in turn has an attribute CONTAINER that takes the glass node as its value.

[29] Frames can in theory be used to represent many kinds of knowledge, up to whole utterances, see Gamerschlag et al. (2013) and subsequent works by the CRC 991 "Structure of Representations" at the Heinrich-Heine-University Düsseldorf.

Beckmann, Indefrey & Petersen (2018) adapt this approach to apply to countability shifts, like the packager shift where mass nouns are used to refer to portions of substances. Consider Figure 10 as representing (part of)[30] the lexical entry for *Wasser, 'water'*.

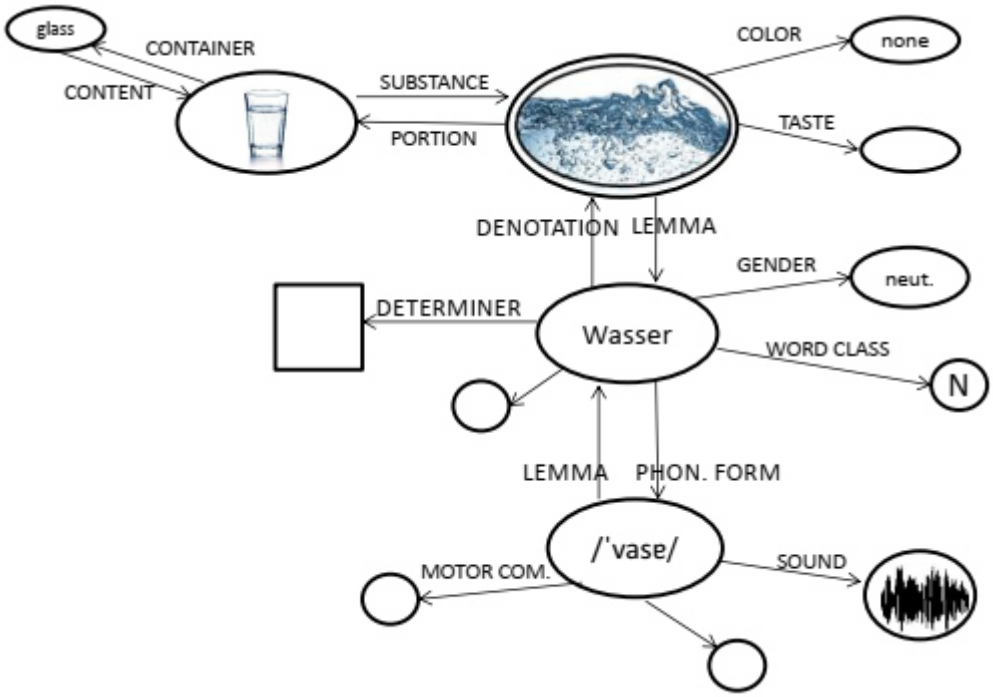


Figure 10: Frame representation of *Wasser*, '*water*'
adapted from Beckmann, Petersen & Indefrey (2018, p. 81)
Note: round node are value nodes, the double-circled node is the reference node, square nodes are empty nodes, arrows are attributes, values can represent abstract or concrete information, see Beckmann, Petersen & Indefrey (2018, p. 81) for a more detailed description

Now imagine a sentence with a packager shift like "Johann bestellte ein Wasser" (*John ordered a water)*. The noun *Wasser* here no longer refers to the conceptual aspect that represents an unbound entity, but rather to that representing a portion of water. In this context, we can envision the frame representation of *Wasser* to be something akin to Figure 11.

[30] As frames can represent everything one knows about the concept of water, all the information stored about the word 'Wasser' and so on, no frame representation depicted in an article is ever truly complete

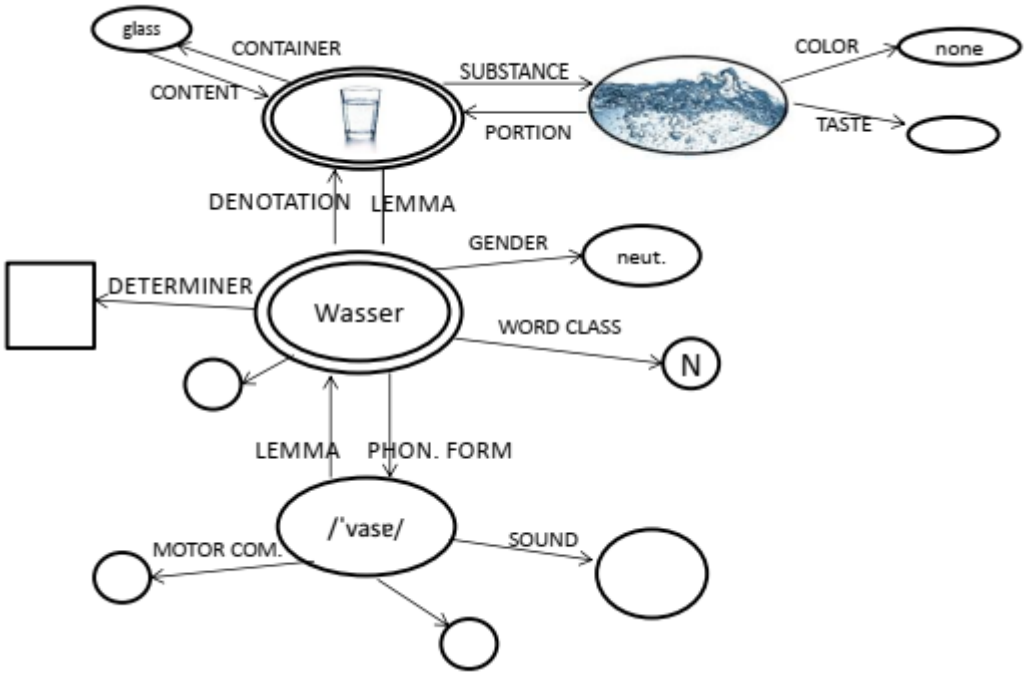


Figure 11: Frame representation of shifted *Wasser, 'water'*
adapted from Beckmann, Petersen & Indefrey (2018, p. 84)

Because the 'old' reference node in Figure 10 and the 'new' reference node in Figure 11 are linked bidirectionally (by the attributes SUBSTANCE and PORTION), the shift in reference is felicitous.

2.4.3. Discussion

This small sample of works already shows that classifying countability shifts as either cases of polysemy or as cases of a different – possibly related – phenomenon cannot be straightforwardly achieved. It seems that even when focussing on a single shift type, namely the packager shift, some evidence points in one direction, whereas other points into a different one. Additionally, cross-linguistic differences need to be taken into account, as can be seen from examples (35) and (36).

Viewing polysemy as a continuum rather than a binary distinction – as it was suggested by Viebahn (2020) – appears to be a better way. Then, however, a clearer method for placing individual example cases on the continuum is needed. Also, there doesn't seem to be a clear distinction of 'polysemy' and 'metonymy', as the same example cases are given for both phenomena (in e.g. Frisson & Frazier, 2005, and Schulzek, 2014, respectively). The data summarized above suggests an interpretation that places countability shifts on the polysemy-end of that continuum rather than the homonymy-end, with the data from Ortega-Andrés & Vicente (2019) and Schumacher (2013) supporting that both denotation targets – the default and the shifted one – are readily available to the interpreting mind. It is puzzling, in that respect, that Frisson & Frazier (2005) didn't find effects supporting that interpretation, especially because their stimuli occurred in the corpus in both interpretations (as individuals and substances). It needs to be considered, though, that the actual statistical results are comparatively weak considering the strong interpretation of the authors and that it's possible that not all of the individual noun stimuli showed the same syntactic behaviour, as the paper only gives the full range of all corpus results for each countability class. With regard to the shifted interpretations being 'context sensitive', it is difficult to differentiate if a context that precedes the critical noun is 'disambiguating between two readily available

forms' or whether it 'leads to the preparation of the mind for a shifted interpretation of the noun'. Frisson & Frazier (2005) employ a direct comparison and find a sign of processing difficulty for mass nouns in count-context only in the spill-over region. Whether that effect is a sign of re-interpretation of the noun or just a result of a mass noun occurring in plural form, is unknown.

In the end, more research into this very wide-spread phenomenon is needed with stimuli of a greater variety, as we are dealing with something very common in natural language. Linguists and cognitive researchers are realising more and more how ubiquitous and pervasive it is that words are used to refer to multiple (more or less different) conceptual entities.

Falkum (2011) rightly mentions that polysemy should be seen as a property of natural language that makes communication and acquisition easier and more efficient. Here, polysemy is seen as complying with the principle of communicative relevance, while maximizing effectiveness and minimalizing processing cost at the same time.

It is also fair to assume that many linguistic and non-linguistic factors interact with the interpretative ease of polysemous expressions. It might be the case that a sense that is far away from the original reference node is more difficult to access than one that is closer. That would be in line with the works by Murphy (Klein & Murphy, 2002, and Foraker & Murphy, 2012) where *book* (physical object → informational content) yielded different results than *paper* (physical object → company → management personal). But how would one unambiguously determine node distance? Another possibility is that there is a difference in processing between 'semantic sense selection' via selectional restrictions and morpho-syntactically marked shifts using different determiners or plural morphology. Further, the disambiguating strength of a given context, but also frequency of senses for individual hearers, number of senses in a given language and the general restrictiveness of the conversational topic at hand could be crucial for ease of processing; all things that are neither easily simulated in a lab environment nor straightforwardly quantized and factorized for an experiment.

2.5. Words that mean stuff together or: Combinatorial Restrictions and the Processing of Nouns in Phrases

This section outlines linguistic and psycholinguistic considerations on nouns and noun phrases, how their parts interact in processing and how a complete meaning for the entire phrase is achieved. This issue touches many fields and the focus here will lie on the syntax-semantic interface with regard to processing of noun phrases. Specifically, theoretical approaches and empirical results are collected that shed light on the question of how a given determiner or quantifier influences the grammatical and interpretative processing of the noun.

2.5.1. Nouns in interaction – Theoretical approaches

As a starting point, Pustejowsky's (1995) qualia analysis is taken as it represents the base for semantic interpretation in many linguistic theories, starting with the generative view on language. In this view, the focus on one or the other quale that constitutes part of the semantic representation of a noun is influenced by other linguistic items that can combine with it. For example, it can reflect demands

by selectional restrictions of verbs and adjectives in the sentence. A verb or adjective may require a particular type of argument (say a physical entity). In generative theory, this is represented by an annotation being part of e.g. the verb's argument position specifying a specific qualia type. An NP can then only fill that argument position if the noun that heads it has a quale compatible with the requirement specified by that annotation (Van Valin, 2001). The appropriate quale is in turn the focus of the noun's interpretation, that is the conceptual aspect of the noun's concept that is most dominantly referred to by the specific sentence.

With regard to the matter at hand, countability, Erbach (2021) among others nicely points out how so-called 'stubbornly distributive predicates' like *round* and *big*, and reciprocal operators like *each other* only combine with individuated entities, or collections thereof and are therefore typically thought only to combine with count nouns (and object-mass nouns).

Of course, there are also verbs that prefer or even demand nouns of a specific countability class or rather nouns that denote object-like, bounded entities, first and foremost the verb *to count*.

(38) a. Lisbeth counted the apples in the storage cupboard.
b. The greedy earl counted his gold.
c. Seymour counted the sugar but not the water. [31]

Note that the grammatical objects of the sentences in (38) – functioning as the undergoer of the verb *to count* – are more easily interpreted as a collection of bounded entities, irrespective of number marking (plural in a, singular in b and c) and the fact that the nouns in b and c most frequently refer to unbounded substances. In (38)b and c, the noun's interpretation is shifted in terms of boundedness and the mass nouns *gold, sugar* and *water* refer to portions, e.g. coins of gold, packages of sugar and bottles of water.

As for determiners, Löbner (e.g. 2011, see also Gamerschlag et al. 2013) developed a theory of concept types where he divides nouns into four classes based on two factors, uniqueness and relationality, and states that "[t]he properties that distinguish the types of nouns, […] correspond to types of determination and reference. Clearly, uniqueness is linked to definiteness, and relationality to possessive determination" (Löbner 2011, p. 287). Most straightforward examples are the noun phrases in Table 4.

[31] Example (38)c from Bale & Barner (2009), p. 221

Table 4: Löbner's Concept Types with determination, adapted from Brenner (2016, p. 149)
Note: combinations in green are said to be felicitous

		Uniqueness	
		non-unique [-U]	**inherently unique [+U]**
Relationality	**non-relational [-R]**	**SORTAL** ***Apfel / apple*** ein Apfel / an apple der Apfel / the apple sein Apfel / his apple	**INDIVIDUAL** ***Papst / pope*** ein Papst / a pope der Papst / the pope sein Papst / his pope
	Inherently Relational [+R]	**RELATIONAL** ***Arm / arm*** ein Arm / an arm der Arm / the arm sein Arm / his arm	**FUNCTIONAL** ***Mutter / mother*** eine Mutter / a mother die Mutter / the mother seine Mutter / his mother

Of course, there are a number of phrases that immediately come to mind that do not follow this pattern even for these exemplifying nouns. However, according to Löbner's theory, those are instances of shifts where a noun, e.g. an individual noun like *Papst, pope*, or *Mond, moon*, is used as a different concept type.

(39) a. A moon is an astronomical body that orbits a planet.

b. Sein Papst war er nicht.
His pope, he was not.

In phrases like *a moon* in (39)a, the noun *moon* is a sortal noun, referring to natural satellites in general rather than to Earth's moon. Similarly, *sein Papst, his pope*, in (39)b can mean the person who the speaker prefers in this role, parallel to *(not) my president*. Here, the noun *Papst* is relational and unique, i.e. shifted into a functional noun in Löbner's terminology.

So, for many of these semantic restrictions, it can be said that they hold for nouns in natural language use until they don't and the noun's interpretation shifts, even if only slightly. Whether all of these shifts represent irregularities or even cases of incongruency, or whether they are just part of the natural way nouns function, is a matter of discussion.

Next to semantic restrictions, there are also syntactic restrictions on NP construal that need to be obeyed in order to uphold grammaticality. In German, for instance, the items of a noun phrase must agree with the nominal head not only in number, but also in gender and case (though, see Schlenker, 1999, for a detailed description of exceptions). Additionally, there are certain restrictions as to what type of determiner or quantifier can combine with which type of noun. However, looking at countability shifts, the first exception to that rule already presents itself. In German packager shifts, the singular form of a mass noun is mostly directly combined with a numeral, as in (40).

(40) Drei Bier, bitte.
Three beer.SG, please.

Here again, violations of regulations are often used to signal a shift in the noun's meaning towards a non-canonical interpretation.

There are theoretical syntax frameworks that consider noun phrases to be complements of a higher syntactic projection, determiner phrases, and place a 'number phrase' above the NP as a separate functional head. In this view, that node not only makes the feature available to all syntactic operations such as agreement but due to its independence, it also has semantic content and is therefore a variable (see e.g. Ritter, 1988, 1993, for Government & Binding).

Bale & Barner (2009) aim to develop an account for why the syntactic environment seems to shift noun interpretation and state that nouns (in isolation or in the mental lexicon, what they call 'root nouns') are not marked for countability but that "nouns are formed by combining roots with functional heads, and that root nouns can be associated with denotations that either do or do not contain individuals as minimal parts" (p. 249). It is the functional head that can be count or mass and which maps onto denotations.

Another example is the account on 'restaurant talk' by Wiese & Maling (2005) that explains for instance a packer shift like *two coffees* as follows: *coffee* by nature is a substance in interpretation and syntactically behaves as a transnumeral (i.e. does not pluralise and can stand in isolation without a determiner or quantifier). When combined with a numeral the interpretation is enriched by "PORTIONS OF" and the noun behaves like any other plural nominal syntactically, i.e. combines with the quantifier to form a quantifier phrase like in (41):

(41) $[[\text{TWO}_{\text{numeral}}]\text{Q}^{0}_{\text{ind}}\ [\text{PORTION(COFFEE)}_{\text{plural nominal}}]\text{NP}\ ]\text{QP}_{\text{ind}}$
structure for *two coffees*

(adapted from Wiese & Maling, 2005, Figure 3, p. 8)

For German shifts, where the shifted mass noun remains singular morphologically, Wiese and Maling argue not for an elliptic construction with a container noun ellipsis (because the determiner agrees in gender with the beverage noun, not a container noun) but point out the similarity of shifted beverage nouns to object-mass nouns, that is they are said to be "object-denoting expressions, even though they remain transnumeral" in behaviour (p. 13).

(42) $[[\text{TWO}_{\text{numeral}}]\text{Q}^{0}_{\text{ind}}\ [\text{PORTION(COFFEE)}_{\text{transnumeral}}]\text{NP}\ ]\text{QP}_{\text{ind}}$
structure for *zwei Kaffee, two coffee.SG*

(adapted from Wiese & Maling, 2005, Figure 5b, p. 14)

In Role and Reference Grammar (RRG), all reference phrases including noun phrases, which are reference phrases with a nominal nucleus, have a layered structure (just like clauses and sentences) where different layers are modified by different operators. RRG considers countability a nominal aspect and thus a nuclear operator, whereas number and quantity are core operators and e.g. definiteness is an NP operator (Van Valin, 2005). Nominal aspect – of which countability is a part – determines whether the noun refers to an individual, a set of individuals, or a kind, similar to the telicity distinction in verbal aspect (see Jackendoff, 1990, for a discussion on that parallelism). For example *the scarf* would be represented as follows:

(43) $\langle_{\text{DEIC}}\textit{PROX}\ \langle_{\text{DEF}} +\ \langle_{\text{NEG}}\textit{Ø}\ \langle_{\text{QNT}}\exists\ \langle_{\text{NUM}}\textit{SG}\ \langle_{\text{NASP}}\textit{COUNT}\ \langle \text{scarf} \rangle \rangle \rangle \rangle \rangle \rangle$

(Van Valin, 2005, p. 53)

Thus, here countability seems to be a property of the noun itself. Unfortunately, there is no RRG-based account on countability shifts that could give insights in how the operators of the different layers influence each other in the processing of noun phrases.

For the RRG interpretation one can argue that countability, not unlike concept type, has to be in congruence with properties of – or expectations based on – the determiner and that incongruency might lead to a shift. Bale and Barner's (2009) account explicitly circumvents this by stating that count noun heads map onto denotations that contain individuals as well as on those that do not, and that mass noun heads are in fact identity functions[32]. Similarly, Wiese & Maling (2005) also decouple semantic interpretation from syntactic behaviour but explicitly state that given the syntactic environment both noun properties change (or are enriched).

From this perspective, these three phenomena – verbal restrictions on noun sense selection, concept type shifts and countability shifts – all represent examples in which a previously encountered element influences the interpretation of a subsequently occurring noun. In the latter two cases, it can be said that the semantic (re-)interpretation of the noun is triggered by an element that is classically thought of as grammatical i.e. syntactic in nature; a property embodied by a specific determiner, e.g. definiteness[33], and even more striking number agreement. Note that in all cases (except perhaps Bale and Barner's analysis, depending on the interpretation of 'functional heads') the underlying assumption is that the element combining with the determiner has a default interpretation (in the case of sense selection triggered by verbal restrictions) and/or a certain property (being sortal or singular or count) that influences its semantic interpretation and that 'goes together' with certain properties of the previous element. It thus appears that there is a 'normal' or 'default' way of which elements go together to form a noun phrase and that in 'abnormal' combinations something in the noun needs to change in order to allow the proper 'going together' of the phrasal elements again.

There are reasons to disagree with this strict view of default interpretation and (regular/productive/rule-based) re-interpretation in cases of incongruency. In the last section, it became clear that many (if not most) nouns simultaneously refer to different aspects of the concepts they are linked to and that more linguistic and extra-linguistic information may be needed to determine the precise meaning of a noun in a given utterance. Bale and Barner's (2009) account could be listed among these trains of thought in that they take nouns to map onto multiple denotations. From that point of view, it is not surprising that no small number of researchers claim that NPs (or functional heads therein), and not nouns, experience countability and that it may be the case that language itself cannot be analysed by asking whether a sentence is true or not.[34] If so many nouns refer to different entities in the world simultaneously, a mapping of language and reality becomes a difficult if not futile method of establishing meaning.

In order to find the best way of analysing countability shifts, we first have to understand whether countability is a property of the noun, that is a deeply rooted and unchangeable aspect of the nouns 'identity' without which the noun would 'mean something different', or whether it is a property of noun phrases; something that points us to the specific sense targeted by the noun in this case. A noun here

[32] In cases where "the root noun is mapped to a denotation with individuals, then the mass noun version of this root will also be mapped to such a denotation" (p. 249). This is assumed in order to explain object-mass nouns that denote 'atomic individuals', contrary to accounts that assume a purely semantic origin of the count/mass distinction.

[33] One might argue that definiteness – in contrast to e.g. number - is more of a pragmatic property than a syntactic one, since it is pragmatic information that informs the choice between the usage of the definite and indefinite article. However this is clearly not the view taken by Löbner (2011).

[34] In the former sections, Pelletier (2012) and Chomsky (2000) are named as examples for these views, respectively.

really must be thought of as referring to a complex web of concepts or conceptual aspects that cannot be unambiguously interpreted on its own.

2.5.2. Inside or outside the noun – empirical evidence

Let's consider two other nominal properties to help us form an understanding of how countability might relate to nominal entities in the syntax. It has been argued that gender, but not number, is a lexical property of nouns. Gender – grammatical gender at least – is invariable, unshiftable and in most cases an arbitrary syntactic property. However, contrary to grammatical gender one would assume that number is a property that influences the semantic level. Many nouns – that is count nouns – are used to refer to an individual entity or a set of them and it is number that tells us which it is, mostly through marking on several elements of a noun phrase, but sometimes only on those that are not the noun itself.[35]

Thus, number and gender are properties that seem clearly different from each other. Gender is seen as a lexical property of the noun, whereas number is more thought of as 'outside' of the (core) lexical representation of it and thus interacts with the other elements of the clause in a different way[36].

From the empirical side, the picture is less clear. Error rates regarding gender and number usually differ greatly within a language (e.g. Vigliocco, Butterworth, & Garrett, 1996, or Bock & Levelt, 1994). Primed lexical decision experiments have yielded results where the different agreement violations caused different effects (e.g. Fussart, Jakubowitz, and Costes, 1999) and where they didn't (e.g. Colé & Seguí, 1994, or see Lukatela et al., 1987). There may be various explanations for those different results. Fussart and colleagues used auditory lexical decision, while in the other two studies stimuli were presented visually in written form. Secondly, the stimuli from Fussart, Jakubowitz, and Costes' experiments were all animate. That is a gender violation here was not purely grammatical, as the grammatical gender of the nouns used reflected the biological sex of the noun's referent. Whether the same holds for the stimuli from Colé & Seguí and Lukatela et al. cannot be evaluated based on the stimuli descriptions in the articles. Colé and Seguí only provide one example *chat, cat,* which additionally to a grammatical gender does have a referent with a biological sex (*le chat, the.MASC tomcat* vs. *la chatte, the.FEM female cat*) but do not mention anything regarding this two-folded phenomenon in the text. Lukatela and colleagues also don't provide any description of their stimuli with regard to biological sex of the noun's referents. Peculiarly, all studies explicitly refer to the gender property as grammatical in nature.

Consequently, while it cannot be excluded that all the lexical decision studies (inadvertently) tested a semantic notion additionally to grammatical gender, the reason for the different results more likely stems from the difference in stimuli. Namely, in Fussart, Jakubowitz, and Costes' (1999) study, the semantic level was more involved than in the other two lexical decision studies by Colé & Seguí (1994) and Lukatela et al. (1987) due to the semantic difference in sex coinciding with the grammatical gender. That is, Fussart, Jakubowitz, and Costes' (1999) study is more similar to the observations of errors by e.g. Vigliocco, Butterworth, & Garrett (1996) or Bock & Levelt (1994) in that in both cases the processing

[35] This is the case for some German nouns where the singular form is (at least phonologically and orthographically) identical to the plural form, such as *Reiter, rider, Rahmen, frame,* and *Zettel, (piece of) paper*, for example, and of course shifted nouns.

[36] Though note that there are languages where subject and verb have to agree in gender, too; such as Arabic VSO constructions (Alkuhlani & Habash, 2001).

involved in the task included the semantic level. (For a more in-depth discussion of the processing levels targeted by the lexical decision task see discussions in Section 3.4 and Chapter 4.)

For countability, the picture the literature provides looks even more nebulous. Fieder et al. (2014) aim to provide an overview on experimental investigations of the semantic side of the count/mass distinction, thus here we are concerned with the nominal, invariable side of countability. Fieder and colleagues stated that

> "Bisiacchi et al. (2005) and El Yagoubi et al. (2006) investigated conceptual-semantic processing of concrete mass and count nouns in a semantic categorization task. Prior to Bisiacchi et al.'s experiment, participants were instructed about the semantic differences between mass and count nouns. During the task, participants were required to categorise visually presented words into mass and count by button press. The results showed that participants required longer processing times for the categorization of mass compared to count nouns (see also Mondini et al., 2008)." (p. 11)

Crucially, however, the cited articles do not mention the difference between count and mass nouns to be semantic in nature. In fact, Bisiacchi et al. (2005) provides no information at all about how participants were instructed about the count/mass distinction. Additionally, no behavioural results are reported that could justify a claim about processing difficulty of any kind. The reported results are limited to ERP data, which state that in an early time window (120-160 ms) a negativity was found that was lateralised to the left for count nouns, but that no lateralisation could be found for mass nouns. Further, they report the negativity in left anterior regions to be more negative for count nouns than mass nouns. And even this finding of an early left anterior negativity, that is supposedly stronger for count than for mass nouns, is called into question by the number of participants (n= 19), which is small for the current standard in ERP research.

Mondini et al. (2008) used a lexical decision task on 100 concrete count and mass nouns. Also here, it is not specified how the stimuli were categorized in the distinct countability classes. The article's introduction includes a form of definition of countability that encompasses next to morpho-syntactic restrictions a passage that – without explicitly mentioning it – states that count nouns apply to perceptual entities that are atomic ("…in combination do not yield another entity of the same kind", p. 1) whereas "[m]ass nouns , like 'water', 'sand' or 'oatmeal', denote instead, monadic entities whose boundaries are perceptually inaccessible: the samples to which they are applied are taken as constituting another sample of the same kind when in combination" (p. 1), i.e. are not atomic. Assuming that the mentioned atomicity and visibility of physical boundaries was indeed the author's decisive factor in stimulus categorisation, it can be said that a conceptual difference has been investigated. However, Fieder et al.'s interpretation of the results as "show[ing] that participants required longer processing times for the categorization of mass compared to count nouns" (p. 11) is puzzling since the original paper reports no statistically significant difference between count and mass nouns, be it in reaction times or error rates (see Mondini et al., 2008, p. 3). For the neurophysiological results, an early left negativity at posterior sites was found to be smaller for mass nouns than for count and pseudowords. Further, it is reported that count nouns elicited a significantly larger negativity at left posterior locations in comparison to mass nouns. The results, although not directly contradictory, differ from the left anterior response found by Bisiacchi et al. (2005) and therefore call the explanatory power of both findings and their interpretations into doubt. Here again, number of participants (n= 18) and proportion of analysed trials (mean: 51,5 % per condition per subject, see p. 2) were abnormally low.

Along with Bisiacchi et al. (2005) and Mondini et al. (2008), Fieder and colleagues name El Yagoubi et al. (2006) as having "investigated conceptual-semantic processing of concrete mass and count nouns in a semantic categorization task" (p. 11). In this case, again, conclusions provided by Fieder et al. are not directly obtainable from the original article. El Yagoubi et al. (2006) describe four experiments. The first two utilize a lexical decision task and found no difference between count and mass nouns in reaction times or accuracy (Exp. 1), not even when concreteness was treated as a separate experimental factor (Exp. 2). The passage of the article reporting Experiment 3 and 4 consists of exactly 105 words. Still, mismatches between the design reported in the original paper and the summary in Fieder et al. (2014) can be observed. In Experiment 3, participants (n= 15) were asked to perform a grammaticalilty judgment task for "240 sentences, half of which respected the mass/count syntax while the other half contained a syntactic violation" (p. 199). As the result, the authors single out the condition with concrete mass nouns in correct sentences as being "more difficult to accept" but fail to report the measured variable reflecting that processing difficulty. It is reported that Experiment 4 asked participants (n= 15) to judge words on concreteness, however, how concreteness was defined is not mentioned. Here again, the only reported result is that "[a] significant interaction was found: to decide that a count noun was abstract was more difficult than any other decision [$F(1,13) = 20.93$, $p < .01$]" (p 199). Between which experimental factors the interaction was present can only be speculated, as well as the variable reflecting this interpreted 'difficulty'.

The last article Fieder and colleagues claim provides evidence for a semantic influence is Gillon et al. (1999). Allegedly, here "concrete (atomic) mass nouns which are semantically more similar to count nouns have been shown to be processed differently to other (non-atomic, abstract) mass nouns (e.g., substances)" (p. 11). However, the original article reports no concern for concreteness whatsoever. With regard to atomicity, two out of the seven noun classes that were tested are distinguished by atomicity (MN and MA), with one additional class that could theoretically be interpreted as including only atomic noun referents (RC). The noun classes are summarised in Table 5.

Table 5: Noun Classes in experimental stimuli, adapted from Table 1 by Gillon (1999, p. 206)

	Singular	Plural
Regular count nouns (RC)	*table*	*tables*
Fossils (FS)	*man*	*men*
Invariant: -s (IP)		*trousers*
Mass: plural (MP)		*comics*
Mass: nonatomic (MN)	*water*	
Mass: atomic (MA)	*furniture*	
Dual (DL)	*rope*	*ropes*

The results of the two experiments reported in the article, interpreted as they may have been, show that the behaviour of the different noun classes clearly patterns along one distinction only: whether there exists only one word form (singular or plural) or two word forms (singular and plural).

Experiment 1 is said to "probe lexical features" (p. 207) impacting the speed and accuracy of a lexical decision, whether those are semantic or not is not mentioned. For singular nouns, a main effect of the factor Category (count vs. mass) is reported, that resolves to a significant difference in reaction times towards words of the class RC and words of the class MA. Without naming any numbers, it is claimed that "frequency was not found to be a determining factor in the above difference" (p. 208). The authors interpreted the delay as reflective of the extra effort to compute the feature "mass". The results of the

other mass noun class (MN) are not mentioned. For plural nouns, it was reported that reaction times towards words of the IP class were slower compared to all other classes. Further, they report significant differences between the class MP and RC and FS. How that difference is reflected in the reaction times is not mentioned. Also, the exact effects cannot be reconstructed due to a mistake in the reporting: "Fischer post hoc analyses revealed a significant difference between the category of MP nouns and those of RC and FS nouns (MP/RC, *p* , .0001; MP/RC, *p* , .0001)" (p. 209). Precise inspection of the noun classes prompts an alternative interpretation of the results. In the domain of singular nouns, the only class that shows behaviour significantly different from the others is MA. Crucially, words of the MA class do not have a plural form. From the text description of the results, it is possible to see that the words of the class MN also showed deviant behaviour (since the reason for that was assumed to be the presence of the mass-feature 'M'). Those words, too, only have a singular form. In the domain of plural nouns, it was again the classes that possess only one word form (here only plural) that showed behaviour aberrant from words of the other classes, namely IP and MP. No difference has been reported between reaction times towards those words that possess both a singular and a plural word form (DL, FS and RC).

In the second experiment reported in Gillon et al. (1999) all nouns were preceded by "a determiner or adjective to yield grammatical or ungrammatical combinations, e.g., *three babies-*three baby; much mud-*many mud"* (p. 209). No further detailed description of the critical stimulus material is provided. The main result of this experiment is that words of classes MA and IP show no effect of grammaticality. The authors interpret the outcome as showing that the noun's feature atomicity can be primed by the preceding determiner. The plural determiner-prime (supposedly *three* and *many*) "requires atoms for its interpretation, which fits the exceptional nature of the mass atomic nouns" (p. 211). The results on the IP nouns are interpreted accordingly: "[…] the singular determiner-prime facilitates the recognition of the word" (p. 211). Considering this explanation, it may be conceivable that perceiving "many" leads one to expect a noun referring to an atomic entity, and that confirmation of that expectation by the perception of e.g. "furniture" can overrule the grammatical mismatch. However, it is not clear what exactly the situation would be in the case of IP nouns. First of all, it is nowhere specified what qualifies as a singular determiner-prime. Assuming this to be a determiner that can only be combined with a singular noun, the only given example is the quantifier *much*. Assuming further that the reported 'PL ungrammatical' condition then consists of phrases like *much trousers*, it is hard to see how the expectation about the noun's atomicity value (NONATOMIC, in the case of *much*) would match the encountered atomicity value of words like *trousers*, which can denote single atomic entities despite their plural morphology (one pair of trousers). A feature that could be primed, if the 'singular determiner-prime' was in fact not a quantifier like *much* but something that was not exemplified in the paper, is '(not) referring to a multitude of entities'. If the encountered prime would give rise to an expectation such as "exactly one entity" and a noun like *trousers* would follow, that despite its plural morphology can denote exactly one entity, the match of features could be considered sufficient to overrule the morpho-syntactic mismatch. Note that similarly, the above mentioned case where *many furniture* is not significantly slower than *much furniture* can be explained the same way. Rather than atomicity, *many* leads to the expectation of a noun denoting more than one object, which is indeed the case with *furniture*.

In any case, neither experiment reported in Gillon et al. (1999) can be considered proof of a semantic influence on countability. If anything, the results of Experiment 1 could be taken as indicative that processing a word that only exists in either singular or plural form is somehow different from processing a word that occurs in both forms. Corbett (2000) theorises that each language has a number of words that are not "number differentiable" listing both pluralia tantum and mass nouns as examples, suggesting they might be differently processed from nouns having a singular and a plural word form.

Experiment 2 only shows that encountering words can possibly yield expectations regarding semantic properties of upcoming words, such as the number of referents in the denotation.

Thus, though it might qualify as technically true that some of the results reported in Bisiacchi et al. (2005), Mondini et al. (2008), El Yagoubi et al. (2006) and Gillon et al. (1999) could be interpreted as reflecting a semantic difference between mass and count nouns, the conclusion that these articles in fact provide experimental evidence for an impact of semantic factors on the online processing of count and mass nouns cannot be drawn. Additionally to the problems mentioned, it is worth noting that in none of the four articles it is stated how exactly the stimuli nouns were classified as count or mass.

A study that could give insights in the processing of count and mass nouns is reported in Steinhauer et al. (2001). Here, count and mass nouns were embedded in sentences providing either a plausible or implausible context for the specific noun. Example sentences are given in Table 6.

Table 6: Example sentences of experimental conditions, from Steinhauer et al. (2001, p. 1001)

Condition	Example
1a Plausible count	Yesterday, I translated Diane's story for the children.
1b Plausible mass	The detective shared Linda's information with the attorney.
2a Implausible count	The hiker put Kathleen's glacier on the floor.
2b Implausible mass	Yesterday, I donated Carmen's fog to the orphans.

Whether the plausibility of the combination was supposed to be determined by the countability of the noun, or any other semantic properties, is not mentioned and cannot be unequivocally determined from the example sentences. The article states as a main finding that contrary to the difference in N400 amplitude shown by the plausible/implausible distinction, the count/mass distinction is reflected in a different ERP response that shows a more negative frontal potential for count nouns as compared to mass nouns. The fact that "[t]his ERP effect was independent of any violations and proved to be functionally and topographically independent from the N400 component, [is taken as] challenging a primarily semantic explanation for the mass/count effect" (p. 1004). It is to say that the topological distinctiveness from the classic N400 component outlined in Steinhauer et al. (2001) is convincing. However, it is still under debate whether the classic N400 actually does reflect a violation of purely semantic nature (for an overview see Naumann & Petersen, 2019). Therefore, the results of this study cannot be taken as conclusive proof or disproof of a semantic impact on countability, either.

In summary, it seems that there is in fact a difference in processing of count and mass nouns. However, whether this difference stems from a semantic distinction cannot be evaluated. The results outlined in Basiacchi et al. (2005), El Yagoubi et al. (2006), and Mondini et al. (2008) are unreliable and those of Gillon et al. (1999) can be interpreted in different ways. Next to the confound variables of frequency and number of word forms in the design it is also important to point out that the lexical decision task does not necessarily target the conceptual level. Steinhauer et al. (2001) thus remains as the only source of a (lexical) countability effect and presents – as the only result – a rather scarce basis.

2.5.3. Online processing of noun phrases

In the following, the focus lies on how the individual parts of determiner-noun combinations act together during online processing. Regarding congruency within such noun phrases, there are two main points of view with empirical evidence differing depending on the phenomenon in question.

The first possibility is that the interpreting mind performs a verification that – once both parts of the noun phrase have been lexically selected – determines whether they 'fit together'. This 'well-formedness check' crucially is thought of as a post-lexical operation that establishes the grammaticality of the noun phrase.

The second option is that congruency operates on the lexical level. It is possible that once the first element (e.g. a determiner) is analysed it influences the lexical selection process of the upcoming nominal element. For example, it has been suggested that a prediction is being formed regarding the noun and that the list of candidates for lexical selection is limited to those that fit the determiner in the property in question. Cohort models involve the notion of multiple candidates competing with each other until one is unambiguously identified as matching the input (see e.g. Marslen-Wilson, 1987; Norris, 1994, for the 'short list model' or McClelland & Elman, 1986, for the connectionist model TRACE) while their predecessor, the Logogen Model, assumed only one logogen (the mental representation of a word, similar to the notion 'lexical entry') to be activated to a degree sufficient to 'cross the threshold' of lexical selection (Morton, 1964, 1968, 1969, 1979). Here, the effort of logogen selection is decreased with more top-down information from the context. That is, the more contextual information input to a logogen, the less sensory information is necessary to bring the logogen above the selection threshold (see Pisoni, et al. 1985, for a discussion)

Note that for both these accounts it is imperative that both the determiner as well as the noun are somehow lexically marked for the property under observation. A well-formedness check needs to check two properties for congruence, one on the first element and one on the second. That means that both elements need to have that property and that it is observable by the checking mechanism in some way or another with regard to the individual element in isolation. It is not possible to determine whether two things fit together based on just the properties of the combination in this case. Similarly, a given property must be detected in the determiner in order to limit the lexical search for the second elements. That element, too, must bear that property in a salient enough way in order to be selected or discarded in the selectional process.

Fussart, Jakubowitz & Costes (1999) argue that all grammatical agreement is not used for lexical selection but only during a post-lexical integration process. In their view, there are three stages of word retrieval: lexical selection, retrieval of lexical information and integration. If a gender agreement error occurs during the last stage, the system is assumed to return to the initial stage to check whether the correct entry has been selected since gender is a 'stem inherent feature'. Number, however, is not considered such a feature and thus a number agreement error would cause the processor only to repeat the final stage, not return to the initial one.

De Vincenzi (1999) and De Vincenzi & Di Domenico (1999) have contrasted gender and number agreement in whole sentences and found effects of both from a pronoun on the activation of its sentence-final antecedent in lexical decision, showing that this task is sensitive to both those properties. The authors interpret the result as in support of a view of co-reference processing in which gender and number – due to their different syntactic representation – are used at different stages of processing.

In ERP research, there have been findings of gender and number agreement violations on the sentence level, influencing amplitude in the P600 window (Osterhout & Mobley, 1995) and in the N400 window for adjective-noun pairs (Barber & Carreiras, 2003). Barbar & Carreiras (2005) also found effects of agreement violations in the N400 window. They tested article-noun pairs and also found no difference in the effects caused by the two agreement violation types there. However, in the P3 component (posterior regions between 450-750ms) there was a significant difference in latency between the two violation types.

Bölte & Connine (2004) concluded from their research on gender in auditory recognition of German nouns that it is not utilised during lexical access but rather that "the presence of a valid article increases the initial familiarity of a word, facilitating subsequent responses [in lexical decision]" (p. 1018).

For concept type congruency, Brenner (2016) was able to pinpoint the effect on an early post-lexical stage of processing. Using the same paradigm as Bölte & Connine (2004), her results show that the measured effects reflect an influence during the construction of the noun phrase subsequent to the lexical access of the noun.

2.5.4. Discussion

The evidence gathered here seems to be in line with Fussart, Jakubowitz & Costes' hypothesis: effects of number, gender and concept type must be attributed to a late stage of phrase processing, at least 350ms after word onset (that is the start usually assumed for the N400 window), if not even later, as violation-triggered ERP responses were found in the P600 window and the P3 window as well. This late timing speaks against an influence during lexical access, or at least against the limitation of the effect to the stage of lexical access, as these late components (especially the P600) are classically interpreted as reflecting effects on (syntactic) integration stages. The behavioural data summarised above also supports that view.

In sum, there is ample evidence that somewhere in the lexical entry of a noun, be it on the semantic/conceptual level or as part of a grammatical/morpho-syntactic representation, there is information stored regarding gender, number, concept type and possibly countability since for all these properties effects have been found.

2.6. General Discussion

From the literature summarized in this chapter, some interesting hypotheses can be drawn about how humans construe mental representations of bounded and unbounded entities, how that in turn is reflected in our way of talking about such entities and lastly how countability and shifts thereof are processed by the language-competent mind.

Firstly, it can be assumed that adults mostly use the kind-based system for object construal, but can still use the primitive mid-level vision system that young children use for novel entities or in cases of uncertainty. It might even be the case that when people are explicitly asked to judge an entity for

boundedness, factors that were originally employed by the mid-level vision system for object construal are considered in addition to kind-based information.

The empirical data summarised in the section *How we think about things and stuff* (2.3.) has revealed the following factors to be of importance for judging whether something is thought of as an individuated entity or not: From a very early stage of development on humans let their decisions be guided by the shape and the apparent functionality of stimuli, with entities of complex shape and high likelihood of functionality being associated more with bounded entities than with unbounded ones. In older participants, solidity, regularity and uniqueness have been found to play a crucial role as well. This last factor, uniqueness, appears to be intertwined with the notion of whether or not an entity might have been artificially designed or crafted to appear the way it does, rather than it being 'found like this' in nature.

When language comes into play, that is once the linguistic tools for expressing and understanding the perception of boundedness as well as multiplicity are acquired, these factors that are used by the mid-level vision system start to interact with them and a constant interplay occurs between the two sources of information. The outlined evidence suggests that the way our native language linguistically marks boundedness (e.g. the possibility of plural marking) influences object construal in children as young as two years, but within languages that have linguistically marked countability, the presence or usage of it during the experiment shows effects only in participants of four years and older. The work by Barner et al. (2007) on children's acquisition of plural morphology suggested that whether or not the children have fully acquired a linguistic skill (i.e. use it in their own speech) is of more crucial influence on their performance in object construal tasks than whether or not a linguistic ability is exemplified by someone else during an experiment, for example when count/mass syntax is used as a cue during stimulus presentation.

Another interesting observation with regard to the interplay of the two information sources for object construal is how it seems to reflect the inter- and intra-language variability of countability proposed by Sutton & Filip (2016c). In cases where information from both systems (mid-level vision and kind-based) provides the same conclusion, countability appears to be comparatively stable, e.g. for 'prototypical substances' that are non-solid, simple in shape and don't appear to be functional most languages and speakers prefer mass nouns.

Parallel individuation has been identified by Carey & Xu (2001) as one of the first steps towards adult-like counting ability. That means that the ability to individuate entities by means of mental representation (or the perception of individuality of physical objects) is very crucial indeed, but also that it is an ability that has to be acquired. Considering this might help to understand the result of count/mass syntax being able to influence the judgement of novel entities when it comes to count noun referents, but not for mass noun referents. Possibly, a stimulus being unbound perceptually (i.e. liquid and shapeless and non-functional, for example) prevents the acceptability of a count noun whatever the linguistic cue because individuation is just not possible. However, the finding by Barner, Wagner & Snedeker (2008) on verbs contradicts this assumption. Here, durative actions were judged depending on syntactic presentation (more dances/more dancing) while iterative actions (e.g. jumping/jumps) where always judged according to cardinality, irrespective of linguistic presentation.

In less clear cases, however, countability within and across languages seems to vary. It seems that in these cases effects of entity type can overrule linguistic cues (i.e. kind-based information) and vice versa. For dual noun referents and in cases where novel nouns were used to describe a novel stimulus entity,

participants judged according to the syntactic cue they received during stimulus presentation. The observational data on co-predication from Ortega-Andrés & Vicente (2019) and Schumacher (2013) falls in line here, supporting that both denotation targets (bounded and unbounded) are readily available to the interpreting mind. Contrary, Frisson and Frazier (2005) found a difference between count and mass environment for what could be interpreted as dual nouns. However, as pointed out earlier, the actual statistical effects were comparatively small and it is unclear in what distribution the nouns appeared in count and mass environment in natural language (i.e. the corpus used for stimulus selection).

For aggregates an influence of native count/mass syntax could be shown by Middleton and colleagues, though as pointed out above the results are not unequivocally straightforward. Note that this data mostly stems from lexical projection tasks where people had to indicate the best reference relationship between word form and entity. In the last study reported in Middleton et al. (2004) participants saw an experimenter interact with individual units of the aggregate and description changed from mainly mass to mainly count syntax.

For heterogeneous collections that are referred to by object-mass nouns, though, syntactic cues have been shown to be unable to overrule the perceptions of entity type. Crucially, the amount of these entities was judged according to cardinality, presumably because the referents of these mass nouns 'denote objects'. Thus, here it seemed that entity type overruled linguistic information. The case of heterogeneous collections in quantity judgement tasks, as pointed out above, however, is a special one that has caused much confusion and misconception, especially in the linguistic literature. It is dangerous to derive too strong conclusions from the data presented and when interpreting psycholinguistic studies in general, e.g. for the purpose of semantic analysis of word forms (in isolation), it is always crucial to consider possible task effects. I argue that the complexity of the interplay of conceptual, semantic and morpho-syntactic processes employed during this task has not been fully appreciated by authors who take it as straightforward evidence for the idea that collective nouns refer to objects. That is, while from a world-view perspective pieces of furniture are clearly objects – not substances – and are treated as such by naïve participants, from a linguistic perspective it simply cannot be argued that the word *furniture* refers to an object the same way a count noun like *shoe* does. I hypothesize that, in this particular task, participants could have interpreted the noun phrases as referring to kinds in that the question "Who has more [plural count noun] / [singular mass nouns]?" is understood as "Which of these pictures shows a greater amount of useable real-life entities of that kind?". That way, it is conceivably better to count furniture according to cardinality rather than volume since firstly, unlike drops of water or shaving cream, "portions" of furniture don't physically merge when put together and secondly, a bigger chair is still just one chair functionally, i.e. not "more to furnish" than say three smaller chairs.

Collections of things seemed to be a special case for perception, irrespective of task effects. Again, it is work by Barner and colleagues (Barner, Wood, Hauser & Carey, 2008; Barner et al., 2007) that provides the interesting finding that sets of multiple objects cannot be distinguished in size by young infants and non-human primates when they move together as a set. Perhaps, something about individual entities that are grouped together, either functionally or spatial-temporally, just provides our minds with contradictory information about whether we are dealing with one entity (a set) or many entities (individual objects). Here, conceptualisation may depend more on functionality, i.e. on what we want to do with them, than it is the case for other entities.

Leaving physical and conceptual issues aside and turning to the processing of linguistic elements, an interesting parallel between theory groups in theoretical semantic literature on countability and psycholinguistic evidence on noun phrase agreement can be observed that concerns the question of

where and when exactly a given property should be assumed to take effect on phrase processing. In both fields of research there seem to be three possibilities considered:

(i) The property in question is an integral part of the noun and is considered as soon as processing starts. Consider here for example Chierchia's and Link's accounts on countability and the hypothesis that agreement properties might be used during a noun's lexical access.

(ii) The property is located outside the noun's core representation and there is an additional element or mechanism that determines it for the noun. Consider here Rothstein's and Krifka's proposals of e.g. the natural unit extensive measure function and the considerations regarding additional functional heads between N^0 and NP from generative grammar accounts. Wiese and Maling's suggestions regarding an 'enriched interpretation' may also fall into this category.

(iii) The property is really to be considered a feature of noun phrases and not nouns and is utilised during processing only at this stage, for instance during a well-formedness check. This far end of the spectrum also falls in line with e.g. Pelletier's view on countability in that both elements, article and noun, are needed to establish countability.

From an empirical point of view, all reviewed evidence points towards a post-lexical locus of congruency effects during noun phrase processing, be it for gender, number or concept type. With regard to countability, the database alas is very slim. After dismissing unreliable evidence, only one study on online processing of count and mass nouns remains, Steinhauer et al. (2001), and the results only indubitably signal that countability as a property is somehow captured during the comprehension of nouns in sentence context. While that hints towards an interpretation where stored representations of nouns are at least marked for countability, it provides little insight in how exactly countability is utilised during phrase processing and thus how incongruency and shifts are to be analysed.

In the following, empirical evidence from three reaction time experiments will be presented that aim to shed light on this very question.

3. Empirical Investigation

In the following, a line of empirical research is presented that aims to shed light on the questions of how countability is stored in the mental lexicon and how it is utilised during language processing. The specific research question is: Is countability a lexical property that is utilised during the comprehension of determiner-noun combinations?

In order to answer this question, first an investigation was conducted that classified a large number of nouns for several properties that relate to countability (Pretest 1 and 2). Next, the first lexical decision experiment investigated whether the language perceiving brain registers differences in countability as a restriction on the felicity of determiner-noun combinations and which previously measured noun properties also influence processing time. The following two experiments, Lexical Decision 2 and Phoneme Monitoring aim to clarify the precise role of countability during the process of noun phrase comprehension. Specifically, whether an effect of countability congruency acts on lexical access of the noun after processing the determiner or whether the locus of the effect should be interpreted as post-lexical.

In the three reaction time experiments, nouns were presented either in a congruent, an incongruent or a neutral condition. The congruent condition presented count nouns with the indefinite determiner *ein(e)* and mass nouns with the mass-specific quantifier *viel, much*. In the neutral condition, the noun was preceded by a non-linguistic sound. This condition was added in order to have a baseline measurement; a situation in which the processing of the noun would occur without any influence of previous linguistic information. With regard to countability, the definite determiner would have been neutral as well, yet since the definite determiner is gender-marked in German (just as the indefinite determiner), using it would have introduced another form of influence on the processing of the noun. Using a non-linguistic sound has been proven to be a viable method in this type of research. Schriefers, Meyer & Levelt (1990) for example found a noise baseline to be perceived similar to silence and to be less disruptive than linguistic stimuli that were unrelated to the experimental manipulation by employing a language production interference paradigm. (For an extensive discussion see also Bölte & Connine, 2004.)

Other lexical decision studies reviewed in the previous chapter (especially Fussart, Jakubowitz, and Costes, 1999, and Brenner, 2016) have shown more reliable effects for auditorily presented noun phrase stimuli than for written stimuli, which is why for these experiments spoken language was chosen as well.

3.1. Pretests

The stimuli used in the reaction time experiments were controlled for two factors that were established for each noun by pretests. The first pretest was a corpus study aiming to classify nouns in terms of the grammatical properties concerning countability.

As mentioned in the introduction, nouns can first and foremost be classified as count or mass nouns by means of their occurrence in noun phrases. That is count nouns occur in combination with different determiners and quantifiers than mass nouns do. The first pretest yielded a large pool of nouns specified for this 'syntactic countability'.

The second pretest was a sentence production study that aimed to categorise nouns according to the ease of interpretation in incongruent phrase context. This more semantic factor was determined by eliciting a sentence containing an incongruent noun phrase and subsequently analysing the noun phrase's denotation in that sentence.

3.1.1. Pretest 1: Corpus study

The nouns entering this pretest were chosen to reflect a wide variety of semantic properties. The conceptual property that is most discussed in the theoretical semantics literature is the distinction between nouns referring to an object in the broadest sense ('atomic' entity) or a substance in the broadest sense ('non-atomic' entity). The main problem[37] with the conception of noun countability as reflecting the atomicity of an entity is its limitation to concrete nouns. Abstract nouns experience the same syntactic restrictions when it comes to the grammaticality of a determiner-noun combination as concrete nouns. However, it is much more difficult to think about the atomicity or boundedness of an abstract concept like *democracy* than of a concrete concept such as *apple*.

Therefore, the range of semantic diversity in the pool of testing material included nouns of three concreteness levels. Nouns were categorized for concreteness taking into account different criteria; among others concreteness, imagery and meaningfulness as defined in Paivio, Yuille, Stephen & Madigan (1968). There, concreteness and imagery refers to the mental connection of nouns to sensory and especially visual experiences. Meaningfulness is defined as in Noble (1952) and relates to the amount of associations a noun evokes. Further, the notion of specificity from Kammann & Streeter (1971) was entertained.

Taking all these notions into consideration resulted in the following concreteness values: the noun referent is either (a) a single kind and perceivable with more than one sense (concrete), (b) not a single kind and/or perceivable with only one sense or introception (semi-concrete) or (c) not perceivable by senses or introception (abstract). Example nouns are given in (44).

(44) a. Concrete: *Car, Sand, Athlete*
b. Semi-Concrete: *Night, Cutlery, Region, Anger*
c. Abstract: *Function, Risk, Illusion*

Moreover, special care was taken that the pool included nouns denoting collections (homogeneous collections like *Reis, rice*, and heterogeneous collections like *Schmuck, jewellery*), nouns denoting events and nouns of different animacy. The goal of this was to make sure the pool included a diverse range of nouns so as not to miss possible influencing semantic factors. As mentioned in the Background Chapter (see especially section 2.1 and 2.2), there is little empirical evidence on which semantic factors have an influence on noun countability.

[37] But by far not the only problem, see Chapter 2.3 *How linguists think we talk about things and stuff*.

The resulting collection of nouns thus contained a large semantic variety. All nouns in the sample could be assigned to one of the semantic categories exemplified in Table 7 that I have called 'meaning classes'.

Table 7: Meaning classes present in the stimulus pool that entered the first pretest

Meaning class	**Example**	**English Translation**	**Concreteness**
Object/Substance	*Auto/Sand*	*car/sand*	Concrete
Human	*Athlet*	*athlete*	
Event	*Nacht*	*night*	Semi-Concrete
Collective	*Besteck*	*cutlery*	
Space	*Region*	*region*	
Inner State	*Angst*	*fear*	
Abstract	*Funktion*	*function*	Abstract

Methods

For the first pretest, nine determiners, numerals and quantifiers were chosen that can be used either only with count nouns or only with mass nouns, see Table 8.

Table 8: Noun phrase elements used in corpus study with English translations.

Count-specific	**English translation**	**Mass- specific**	**English translation**
ein(e)	*a(n) / one*	kaum	*barely some*
zwei	*two*	etwas	*some*
einige	*several*	wenig	*little*
wenige	*few*	viel	*much*
viele	*many*		

In a collection of German newspaper articles (DWDS, corpus of 'Berliner Zeitung') each noun was tested for co-occurrence with all the noun phrase elements listed in Table 8. Additionally, the frequency of the noun's singular word form was determined, and it was determined whether the noun had a plural word form. This corpus was chosen because first, it allows one to search for combinations of words (e.g. a certain determiner combined with a certain noun, such as *ein Bus, a bus*) and second, the 'Berliner Zeitung' is written in medium register, that is it contains some colloquial language while still preserving the main grammatical structure of German.[38]

A noun was classified as a count noun if it never occurred with any of the mass-specific determiners and had a frequently used plural form. It was classified as a mass noun if it never[39] occurred with any of the count-specific determiners. Nouns occurring with determiners of both classes to a representative extend (at least 10%) were characterized as dual nouns. Depending on the respective distributions, dual nouns were classified as dual-count (more than 50% occurrence with count determiners) or dual-mass (more than 50% occurrence with mass determiners).

In addition to syntactic countability, concreteness and meaning class, the nouns where specified for length in syllables and frequency (of the noun in singular form without considering a specific determiner). None of the nouns were compounds or recently (<10 years) introduced loanwords.

[38] For example, headlines may contain countability shifts of a radical, bold nature, such as *viel Auto für wenig Geld, much car for little money,* that are usually only found in spoken German, but nevertheless represent an important part of the current way the German language is used by its native speakers.

[39] Except for uses of the indefinite determiner in emphatic statements, e.g. *so ein Chaos, such a mess*.

3.1.2. Pretest 2: Sentence Production Study

The second pretest was conducted to assess native speakers' willingness to interpret the noun in an incongruent environment. I have called this factor "Shiftability".

The stimuli for this study (n= 228) consisted of nouns that were previously classified in the corpus study as either (pure) count, dual-count, dual-mass or (pure) mass. The three levels of concreteness were represented in the stimuli to a comparable amount.

The study was conducted online on the German version of the SoSci Survey platform (www.soscisurvey.de)

Methods

A total of 522 native speakers of German participated in the online survey. They were presented with 30 nouns in a specific syntactic environment. Each participant was asked to produce one sentence for each of the presented stimuli. Syntactic environments were distributed between subjects, i.e. a given participant would receive all stimuli in one of the following environments:

(i) bare noun, e.g. *Puppe, doll* (n=154)
(ii) preceded by the indefinite determiner, e.g. *eine Puppe, a doll* (n=179)
(iii) preceded by the mass-specific determiner "viel", e.g. *viel Puppe, much doll*. (n=189)

Five participants and 14 stimuli (that is a noun in one specific environment) had to be excluded from analysis due to an insufficient number of responses (less than 8 answers/ less than 10 participants).

The remaining responses were judged independently by two German native speakers for the following criteria: presence of sentence[40], grammaticality of sentence, interpretability of sentence, noun usage specific or generic and shift type. Identified shift types are listed in Table 9.

Table 9: Shift Types and non-shift readings observed in sentence production study.
Note: Examples with glosses. DEF= definite article, PROPNOUN= proper noun/name), INDEF= indefinite article, SG= singular, PAST= past tense, PRO= semantically empty pronoun

Count noun in "viel"- context	Mass noun in "ein(e)"- context
Grinder Shift: Noun refers to (main) substance of default referent Da ist **viel** **Hühnchen** in der Suppe. There be.3SG much chicken.SG in the soup. *There is a lot of chicken in the soup.*	Packager Shift: Noun refers to a portion of default referent Ich habe **ein** **Wasser** bestellt. 1.SG have.1SG INDEF water.SG order.PAST *I ordered a water.*

[40] While the instructions did not explicitly mention the possibility to refuse an answer, participants did so by typing answers like "X", "…", "Ich weiß nicht" *I don't know*, or the like.

Function Shift: Noun refers to functional property of default referent In der BILD ist nicht **viel Zeitung**. In DEF PROPNOUN be.SG not much newspaper.SG *There is not much newspaper in the BILD.* *(meaning the journalism has a very low quality.)*	Sorter Shift: Noun refers to a sub-kind of default referent So **einen Sand** gibt es nur hier. Such INDEF sand.SG exist.3SG PRO only here *Only here, there is sand like this.*
Degree Reading: Usage of quantifier to indicate extend/amount Die Blutlache gibt **viel Hinweis** The puddle.of.blood exist.3SG much clue.SG auf ein Verbrechen. on INDEF crime *The puddle of blood gave much (of a) clue for a crime.*	Instance Reading: Noun refers to an instance of default referent Das hier ist **eine Reklame**. DEF here be.3SG INDEF advertisement *This right here is an advertisement.*
	Degree Reading: Usage of indef. det. to indicate extend/amount Niemand sah was für **ein Fleiß** Nobody see.3SG.PAST what for INDEF diligence nötig war… necessary be.3SG.PAST *Nobody saw what kind of diligence it took…*

The two most frequently used shift types (Grinder and Packager) were used for further testing. Among the stimuli of this pretest, 15 count nouns were identified that were most frequently used in a Grinder Shift construction and the 15 count nouns that were barely used in a shift construction. Similarly, 15 packable and 15 unpackable mass nouns were identified. These 60 count and mass nouns, together with 30 dual nouns, were used in the first lexical decision experiment.

3.2. Lexical Decision Experiment 1

This first experiment was conducted in order to investigate a possible congruency effect for countability during noun phrase processing. As laid out in the Background Chapter (see sections 2.3. and 2.6. in particular), some researchers believe countability restrictions only to apply to a few nouns but most nouns to be relatively free in combining with determiners. Interpretation of the noun would then be purely based on the syntactic environment. According to this view, the noun *apple* itself for example should not be viewed as a count noun referring to a bounded, individuated entity but as not specified in that regard. Rather, a noun phrase like *an apple* is count and referring to a bounded object while a phrase like *much apple* is mass and referring to an unbounded substance. That would mean that as long as a given phrase is interpretable, processing of one version (count or mass) should not be more difficult or effortful than processing the other.

Since a lexical decision task does not directly target participants' ability to interpret a phrase but only the lexical access of the noun and its integration into the noun phrase on a grammatical level, interpretability was investigated for each stimulus phrase prior to the lexical decision experiment.

Ease of interpretation can be broken down to several factors including whether a noun phrase is familiar to the participant, i.e. whether the phrase is part of the general usage of the noun in natural language. Another factor targets the production side of language and concerns whether a noun phrase can be productively used in a sentence. That ability reflects naturalness of usage but also consistency of reference, in that different participants use the noun in the same way referring to similar entities. As pointed out above, all stimulus nouns were evaluated for these factors (Countability and Shiftability) by means of pretests as only phrases that are part of the general vocabulary of German can be found in newspapers to a representative extent (Pretest 1: Corpus Study) and all nouns were checked for use in sentences with congruent phrases that were produced by multiple native speakers during the sentence production study (see sections 3.1.1. and 3.1.2 for details).

In the previous chapter, empirical results have been summarized that targeted other properties of nouns, such as gender, number and concept type in lexical decision experiments and have found them to influence reaction times. The current experiment aims to broaden our knowledge base and investigate whether countability should be included in the list of properties that have to be stored in a noun's lexical entry.

Firstly, if a congruency effect is observed, that is a difference in reaction time towards the noun depending on the element preceding it, it can be concluded that somewhere in the mental representation of the noun a value for countability is stored, disproving the claim that all nouns are basically (or lexically) dual in nature and freely combine with count- and mass-specific determiners and quantifiers.

Secondly, it could be the case that this congruency effect is modulated by any of the other properties the current stimulus nouns are classified for, for example the exclusiveness of their occurrence with certain determiners, i.e. whether something is a (pure) count noun or mass noun or a dual noun. Moreover, it is possible that one or more of the semantic factors that were taken into account during stimulus selection interact with the congruency effect. If, for example, the congruency effect was only to be found for unshiftable nouns or for concrete nouns it would hint towards an analysis where countability is strongly guided by the interpretability of a given noun phrase. If a congruency effect could only be found for concrete nouns but not for abstract nouns such a result would point towards a strong influence of physical factors of the noun's referent entity, like those summarized in Section 2.1.2 "Properties of objecthood". Analogously, if only unshiftable nouns were to show a reliable congruency effect the idea of a strong semantic influence would have to be entertained. In that case, the results would show that language comprehenders are influenced not only by how often a noun occurs in a given phrasal context (i.e. 'syntactic countability') but that whether or not a given phrase is interpretable at all can overrule such familiarity notions and directly influence processing speed in a lexical decision paradigm.

Moreover, with the neutral condition lacking any linguistic information that might influence processing of the noun, it is possible to distinguish whether any observed differences between conditions reflect an effect of facilitation by the congruent article or rather an effect of inhibition by the incongruent article. Such a distinction might be meaningful. In discussing gender congruency, Neely (1991) argued that facilitation effects should be interpreted as reflecting automatic, unconscious processes, but inhibitory effects should be taken as a sign of attention-driven, controlled processes. In line with that claim, Bates et al. (1996) states that facilitative effects have traditionally been interpreted

as possibly caused by automatic processes (e.g. automatic forms of spreading activation) while a combination of facilitation and inhibition effects would point towards a more controlled or strategic mechanism. That is, if any inhibition is observed it has been taken as a sign of a strategic process taking place (Bates et al. 1996, p. 995).

There has however been evidence of probably automatic (i.e. fast, unconscious and rapidly decaying) processes which show inhibition effects, e.g. in colour priming (Di Pace, Marangolo, Pizzamiglio, & Burr, 1994; Marangolo, Di Pace, & Pizzamiglio, 1993) and picture-word Stroop tasks with very short SOA (Glaser, 1992).

This discrepancy can be explained when considering grammatical gender information as affecting multiple levels of phrase processing, a view put forward by Friederici & Jacobsen (1999). It thus seems prudent not to draw too strong conclusions about the precise nature of the influence one's effect reflects when a task such as lexical decision is used that does not distinguish between e.g. lexical and post-lexical effect loci.

With this current experiment, a first step will be taken in gathering a larger pool of experimental evidence regarding the role of countability in noun phrase processing. In combination with the other two reaction time studies, results may be able to shed sufficient light on the precise nature of this property and how it is utilised in language comprehension.

3.2.1. Method

Materials

As a result of both pretest studies, 90 nouns (henceforth called Stimulus Set 1) could be classified according to Countability Class and Shiftability (with regard to the Grinder and Packager Shift) yielding the experimental stimuli for Lexical Decision Experiment 1.

Countability Class as a factor divides into the levels: count nouns, dual-count nouns, dual-mass nouns and mass nouns. Level assignment was determined by means of the frequency of co-occurrence of a given noun with count-specific and mass-specific determiners in the corpus study (Pretest 1) and by the results of the sentence production study (Pretest 2). Specifically, care was taken that all the sentences produced with the congruent noun phrases (that is combined with the determiner that was part of the more frequently associated class in the corpus study; count-specific or mass-specific) where interpretable and using the noun in the intended manner. For example, a count noun like *Bus*, *bus,* only occurred with count-specific determiners in the corpus study and all the sentences with *ein Bus, a bus,* that resulted from the sentence production study where interpretable in a way that the noun referred to an individuated, bounded object. The three concreteness classes were approximately equally distributed within the countability classes.

Shiftability, as can be seen in Table 10, divides into shiftable nouns (grindable count nouns and packageable mass nouns) and not shiftable nouns and reflects a reader's ability to interpret a noun in incongruent syntactic context according to well-studied shift types; the Grinder shift and the Packager shift.

All nouns were similar in length (1-3 Syllables). Lexical frequency was established using the corpus *Projekt Deutscher Wortschatz*[41] and stimulus nouns were chosen so that all countability classes had a similar frequency distribution (shown in Table 10).

Table 10: Mean frequency and standard deviation of stimulus nouns per countability class.
Note: The frequency is given in Häufigkeitsklassen (engl. class of frequency) as used by the corpus. Calculation is based on Zipf's law. For more details see the project's Wikimedia page[42].

Count nouns - shiftable		**Count nouns - not shiftable**		**Dual count nouns**		**Dual mass nouns**		**Mass nouns - shiftable**		**Mass nouns - not shiftable**	
Mean: 11,60	SD: 2,09	Mean: 11,20	SD: 3,08	Mean: 10,40	SD: 2,44	Mean: 10,33	SD: 2,12	Mean: 12,07	SD: 2,18	Mean: 12,13	SD: 2,23

Based on these existing nouns, 90 pseudowords were created by changing the onset or the onset plus the nucleus of the initial syllable of the noun. Care was taken that German phonotactic rules were obeyed, so as to not create non-words.

In the congruent context, a given noun was preceded by a determiner of the countability class the noun occurred most often in, that is the indefinite determiner *ein(e)*[43]*, a(n),* for count nouns and dual-count nouns and the mass-specific quantifier *viel, much,* for mass nouns and dual-mass nouns. In the incongruent condition, count and dual-count nouns were preceded by *viel, much,* and mass and dual-mass nouns were preceded by *ein(e), a(n).*

Every pseudoword was presented in congruent, incongruent and neutral context, preceded by *ein(e)*, by *viel* or by a length-matched noise segment (LMNS). Countability Class and gender were taken from the existing noun the pseudoword was based on. These variables were controlled for in the same way as it was done for the existing nouns, e.g. during list creation, so as not to miss a possible influence of the original noun's properties during the processing of pseudo words.

All noun phrase stimulus items with real and pseudowords were recorded by a male German native speaker in the two different syntactic contexts, congruent and incongruent, using Marantz PMD620 portable stereo audio recorder and digitized with 16 bits precision and 44.1 kHz sampling rate.

The neutral condition was created by replacing the determiner with a length-matched signal of brown noise (LMNS). To ensure that the length of the noise segment was not indicative of the countability class of the upcoming noun, the neutral environment was created based on the recording of the congruent condition in half of the cases, and based on the recording of the incongruent condition in half of the cases.

An overview on the stimulus conditions with concrete examples is given in Table 11.

[41] *© 2020 Abteilung Automatische Sprachverarbeitung, Universität Leipzig.*
[42] https://meta.wikimedia.org/wiki/H%C3%A4ufigkeitsklassen
[43] German inflects the indefinite determiner for gender of the upcoming noun. *Ein* is used with masculine and neutral nouns, *eine* is used with feminine nouns. This does not influence countability or any other grammatical or semantic property.

Table 11: Overview of experimental conditions
Note: The LMNS is represented by [xxx]

	Count		**Dual Count**	**Dual Mass**	**Mass**	
	grindable	not grindable			packagable	not packagable
Congruent	eine Pflanze	ein Bus	ein Tuch	viel Pelz	viel Müll	viel Laub
	a plant	*a bus*	*a cloth*	*much pelt*	*much garbage*	*much foliage*
Incongruent	viel Pflanze	viel Bus	viel Tuch	ein Pelz	ein Müll	ein Laub
	much plant	*much bus*	*much cloth*	*a pelt*	*a garbage*	*a foliage*
Neutral	[xxx] Pflanze	[xxx] Bus	[xxx] Tuch	[xxx] Pelz	[xxx] Müll	[xxx] Laub
Pseudoword	viel Plonze	ein Dus	[xxx] Düch	viel Bulz	ein Nill	[xxx] Rob

Using Adobe Audition 3.0 software the resulting 540 audio files were normalized for volume (i.e. the RMS amplitude of the samples was normalized to 70%). Additionally, all phrase-initial and phrase-final segments (circa 35ms) were edited with a smoothing kernel ('fade envelope' function) to ensure naturally sounding transitions.

The noun phrase items were distributed over three lists, so that each list contained each noun and pseudoword only once. Care was taken that each list contained the same number of noun phrase items of each condition, Countability Class and Shiftability. Lists were ordered in a pseudo-randomised fashion, so that no more than three repetitions of Countability class, Shiftability, determiner/quantifier/LMNS and condition followed directly upon each other.

Subjects

52 native speakers of German between the age of 18 and 36 years participated in the study. Each reported to have no language-specific, psychological or neurological disorders. Every participant had normal or corrected to normal vision and hearing ability and was able to move his or her arms in a normal fashion (e.g. no casts or paralysis).

None of the participants reported to have an additional native language without German-like morpho-syntactically realised countability (e.g. Japanese) or to have completed more than 2 semesters of linguistic studies at university.

Procedure

The experiment ran on a Windows 7 PC with a SyncMaster P2270 display (Refresh Rate: 120hz, Resolution: 1920*1080, Color depth: 32bit) using Presentation Software Version 19.0 (Neurobehavioral Systems Inc., www.neurobs.com). Stimuli were auditorily presented via Sennheiser HD 437 headphones.

All 52 participants were tested in a reaction time lab on the campus of Heinrich-Heine-University, Düsseldorf.

When the participants arrived in the lab, they were greeted and asked to take a seat at a table to fill out some forms. After filling out informational forms about their age, education and language experiences, and giving informed consent on their participation, participants were verbally instructed about their task and on what to expect from the stimuli. They were told that the experiment tested their language

perception, that they would hear noun phrases and that they had to press a button as fast as they could to indicate whether the noun, i.e. the second word of the noun phrase, was an existing German word or not. An example of a noun phrase containing a real word and one containing a pseudoword was given. It was explained that some of the noun phrases contained a noise segment instead of a determiner/quantifier but that participants should perform the task normally on these items as well.

All instructions were read from a prepared sheet by the experimenter so that each participant received the same instructions and to ensure that they would not miss something when reading it themselves.

After being asked to leave their phone, smart watch and any other devices that might distract them during the experiment behind, participants were individually seated in a dimly-lit, sound-attenuated chamber and asked to adjust the chair there so that they could reach the button box on the table in front of them comfortably. Once seated, they were familiarized with the button box, especially the keys they had to use during the experiment, namely the ENTER button and the buttons they had to press to indicate their answer to the lexical decision task. They were informed that the experiment would consist of individual blocks interrupted by pauses, the length of which they could determine themselves via button press. Finally, they were informed that the experiment would start with a training phase so that they could familiarize themselves with the task and instructed to put on the headphones.

The training phase started with written instructions appearing on the screen informing the participant again that s/he would hear recordings of nouns preceded by either an article or a non-linguistic sound and asked to decide for each noun as fast and as accurately as possible whether it was an existing German noun by means of a button press. Position of buttons was programmed so that the button for real-word answers was operated with the dominant hand and the button for pseudoword responses was operated with the non-dominant hand. Dominance was determined informally based on participants' self-report. The button for real words was green and the button for pseudowords was red. During the instruction, squares in the same colours labelled "Wort", *word*, and "Kein Wort", *not a word*, were shown on the screen, as depicted in Figure 12.

Figure 12: Buttons as depicted on participant's screen
for right-handed participants (left side) and left-handed participants (right side)

The training phase consisted of eight trials that were not part of the actual experiment and which contained all conditions and real words as well as pseudowords. Each trial started with a fixation cross accompanied by a beep sound. While the recording of the noun phrase item was played the coloured squares representing the buttons appeared on the screen. Once the participant had pressed one of the buttons, the boxes disappeared from the screen and the next trial began after a black screen was presented for 1000ms (Inter stimulus interval, ISI). If the participant did not press either button, the next trial would start automatically after 2000ms and the ISI (see Figure 13).

After the training phase, the participants were asked whether they had understood the task. This was also monitored by the experimenter present by means of checking the obtained score that showed the

accuracy of the participant's answers. They were asked whether the volume of the headphones and the level of light were comfortable to them, which was adjusted if they wished.

In cases of comprehension problems, the participant's questions were answered and the training phase was repeated. They were then instructed to place the index fingers of their hands on the buttons they had to press to indicate their answers and only remove them during the presentation of instructions or during pauses.

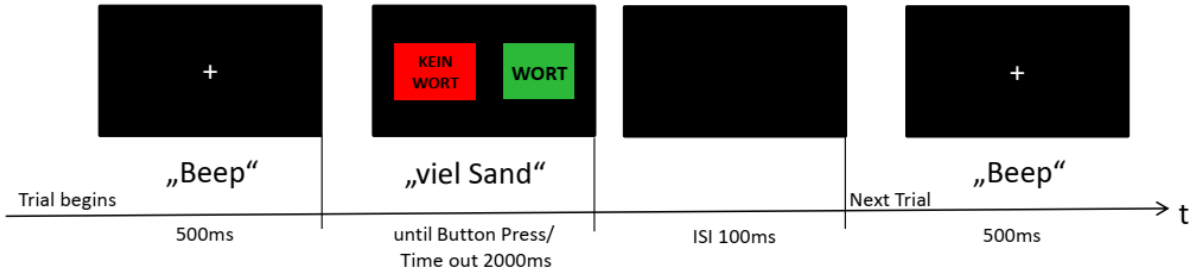


Figure 13: Representation of trial procedure
Note: Squares for buttons as for right-handed participants

The trials of the experiment proceeded in the exact same way as the trials of the training phase. There were three build-in breaks equally distributed over the duration of the experiment. The participant decided each time individually when to continue with the experiment. During the breaks, the obtained score of the current block (correct responses in percent) was displayed.

After the last block was completed, a screen thanking the participant occurred. The experimenter then opened the door of the sound-attenuated chamber and let the participant exit. Each participant was reimbursed with 8€ and candy and thanked again for their participation.

Before leaving the lab, each participant was asked what they thought the experiment was about. None of them reported to have been aware of the critical manipulation (i.e. countability congruence).

The entire procedure took approximately 45 minutes.

3.2.2. Results

R version 3.4.2 (2017-09-28, The R Foundation for Statistical Computing, 2017) was used to carry out all pre-processing and analysis on the data. The lme4 package (Bates, Mächler, Bolker & Walker, 2014) was used to perform linear mixed effects analyses for the reaction time data and mixed logistic regression analyses for the error data. P-values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question.

All statistical analyses were carried out on trials for real words only. Three items had to be excluded from analysis because of technical difficulties during the measurement.

Error Analysis

An erroneous trial was defined as a trial where the lexical decision for words was not answered with the word-button or the decision for pseudowords was not answered with the button assigned to pseudowords ("Kein Wort", *not a word*). Error rates for all items and all subjects were calculated. Four

Items were excluded from analysis because of an error rate higher than 25%. No subjects made more than 25% errors.

After this, analysis preceded with real-word items only.

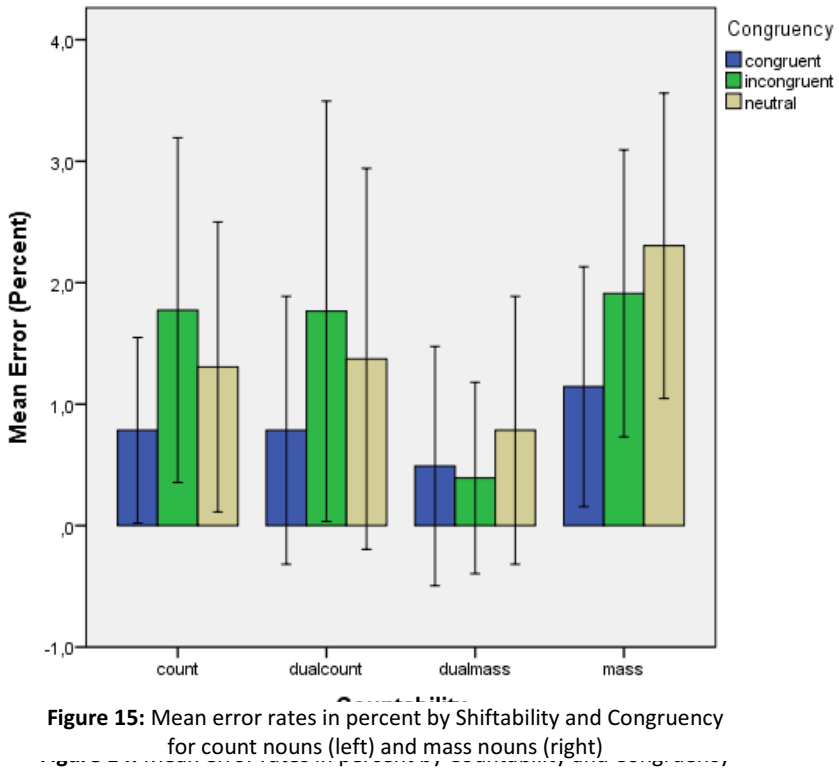
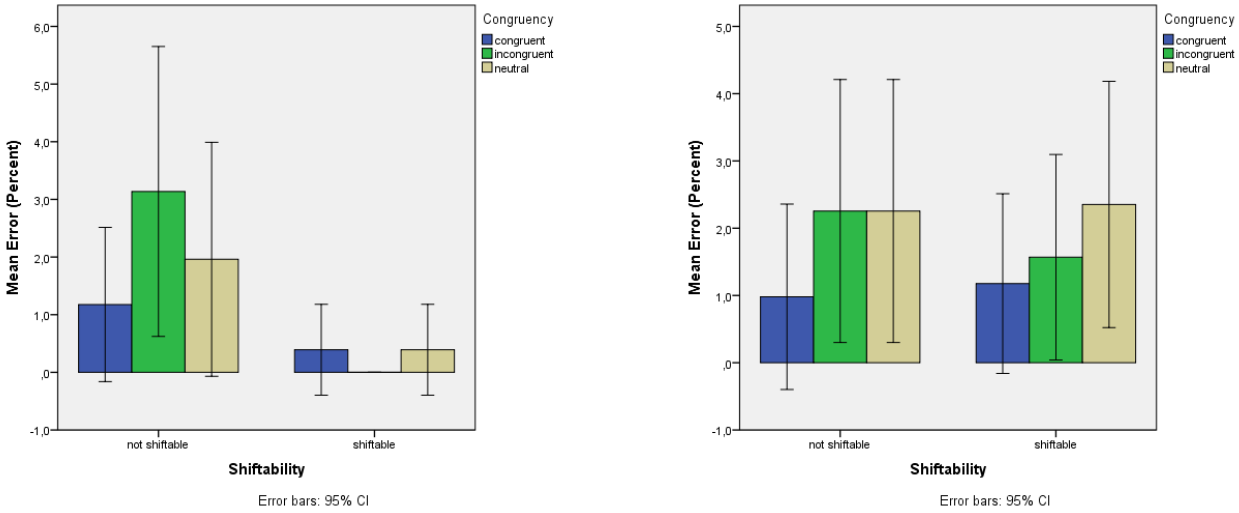


Figure 15: Mean error rates in percent by Shiftability and Congruency for count nouns (left) and mass nouns (right)

Countability Class, Congruency, Shiftability and Concreteness were entered into the mixed logistic regression model as fixed factors. As random effects, intercepts for subjects and items were included as well as by-subject and by-item random slopes for the effect of Congruency.



Statistical analyses revealed no significant main effects or interactions for any of the factors. However, as can be seen in Figure 14, numerically the error rates are higher for incongruent trials than for congruent trials and neutral trials for count nouns and dual-count nouns. For mass nouns and dual-mass nouns, most errors were made in trials of the neutral condition, and for mass nouns there were also more errors for incongruent trials than for congruent trials.

As for count nouns, trials including non-shiftable nouns had a numerically higher error rate in the incongruent condition than trials of the incongruent condition with shiftable nouns. This also holds true for mass nouns, though the difference is not as big (see Figure 15).

Reaction time Analysis

The reaction time was measured from noun onset on until the moment of the button press. Before analysis, all trials with incorrect answers to the lexical decision were excluded (122 trial, 2,69%). Additional 192 trials were excluded from analysis because the reaction deviated from the participants' mean reaction time per condition by two standard deviations or more (4,36%).

For statistical analysis, the natural logarithm (base-e logarithm) of the reaction time was used as dependent variable in linear mixed models.

First, Countability Class, Congruency and Concreteness were entered into the model as fixed factors. As random effects, intercepts for subjects and items were included, as well as by-subject and by-item random slopes for the effect of Congruency.

The analysis revealed a highly significant main effect for Congruency ($\chi^2(2)= 27.77$, $p< 0.0001$). Furthermore, there was a significant interaction of Congruency with Countability Class ($\chi^2(6)= 17.65$, $p= 0.007$). When the data was divided for the different levels of Countability Class, a significant main effect for Congruency was found for count nouns ($\chi^2(2)= 21.32$, $p<0.0001$), for dual-count nouns ($\chi^2(2)= 21.38$, $p<0.0001$) and for mass nouns ($\chi^2(2)= 6.98$, $p= 0.03$), but not for dual-mass nouns. No significant effect or interaction could be found with regard to Concreteness.

Pairwise comparison using the Satterthwaite method with Bonferroni correction revealed that for count nouns the difference between (the logarithm of) reaction times in the congruent condition ($M= 6.84$, $SD= 0.19$) and the incongruent condition ($M= 6.90$, $SD= 0.20$) was significant ($t(27)= -3.12$, $p= 0.005$). The difference between the congruent condition and neutral condition ($M= 6.94$, $SD= 0.19$) was significant ($t(25)= -5.71$, $p= 0.02$) and the difference between the incongruent and the neutral condition was significant ($t(29)= -2.68$, $p<0.0001$).

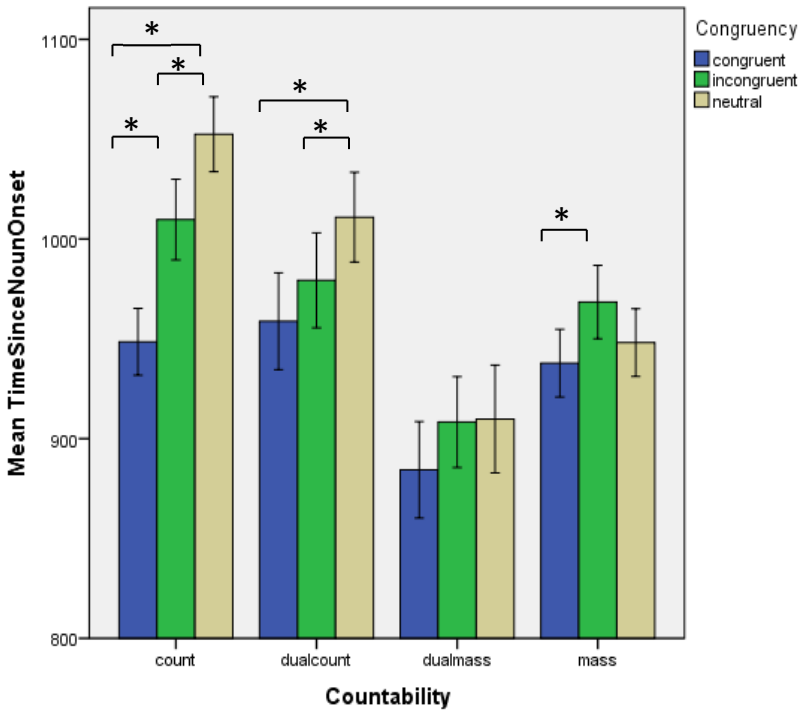


Figure 16: Mean reaction time by Countability Class and Congruency

For the dual-count nouns, the same pairwise comparisons[44] revealed a significant difference between the congruent condition (*M*= 6.85, *SD*= 0.19) and the neutral condition (*M*= 6.90, *SD*= 0.17) (t(666)= -4.59, *p*<0.0001) and a significant difference between the incongruent (*M*= 6.87, *SD*= 0.18) and the neutral condition (t(667)= -2.99, *p*= 0.008). Since for the dual-mass nouns no significant main effect of congruency was found, no pairwise comparison was performed. For the mass nouns, pairwise comparison revealed a significant difference between the congruent (*M*= 6.85, *SD*= 0.19) and the incongruent condition (*M*= 6.87, *SD*= 0.19) (t(30)= -2.57, *p*= 0.03).

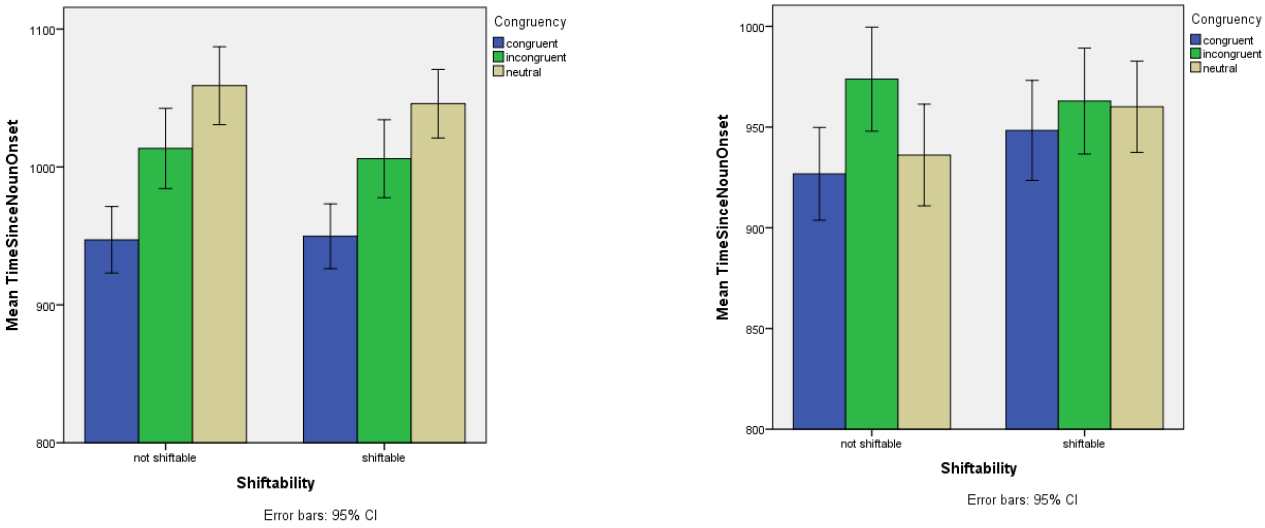


Figure 17: Mean reaction time by Shiftability and Congruency for count nouns (left) and mass nouns (right)

As Shiftability was only defined for pure count nouns and mass nouns, separate analyses were conducted within those two noun classes including the fixed factors Congruency and Shiftability. There was no

[44] Note that the model for the dual-count nouns was only possible without the random slopes. Random intercepts, however, were included. This procedure is recommended in Bates et al. (2014).

significant effect of Shiftability. The model including the factors Congruency and Shiftability did not explain the data better than the model that only included Congruency ($\chi^2(5)= 1.144$, $p= 0.95$) and neither did the model including main effects of Congruency and Shiftabilty plus their interaction ($\chi^2(7)= 1.407$, $p= 0.99$).

In order to shed more light on whether countability should be seen as an abstract property of lemmata in the sense of a categorical feature with two or three levels, like count/(dual)/mass, or whether it is more plausible to view countability as a more or less direct representation of the frequency with which native speakers encounter a noun in a given noun phrase, further testing was conducted that compared these different versions of the factor Congruency.

For the count nouns and dual-count nouns, likelihood ratio tests were performed of the model including Congruency against a model that replaced the categorical value Congruency (1= congruent, 2= incongruent) with the natural logarithm (base-e logarithm) of the actual corpus occurrences with (a) all count-specific determiners from Pretest 1 and (b) only the indefinite determiner (i.e. the determiner used for the congruent condition in the lexical decision experiment). Compared to a null model without any fixed factors (log(L)= 1147.1), the model with all count-specific determiner occurrences (log(L)=1158.6) was significantly different ($\chi^2(3)= 23.04$, $p< 0.0001$). The model with the indefinite determiner occurrence (log(L)= 1159.4) was also significantly different ($\chi^2(3)= 24.7$, $p< 0.0001$) from the null model. As mentioned above, the model in which Congruency as a nominal factor was used (log(L)= 1159.4) also was significantly different ($\chi^2(2)= 21.32$, $p<0.0001$).

Comparing these models with each other did not yield significant differences. However, it can be seen that the log data likelihood "log(L)" differs across models. According to Pinheiro & Bates (2000), this measure shows how likely it is to arrive at the present data given the model, with the models showing the higher value of likelihood being a better fit. Here, the model using the nominal Congruency factor and the model using the indefinite determiner occurrences both reached a data likelihood value of 1159.4, which is higher than the value for the model with all count-specific determiner occurrences (1158.6). It can therefore be assumed that this last model does not explain the data as well as the other two.

Similarly, the possible influence of (a) all mass-specific determiners and (b) only *viel, much,* (i.e. the determiner used for the congruent condition in the lexical decision experiment) was tested in the same way. Here, again, the three models with fixed factors were all significantly different from a null model without a fixed factor: The null model (log(L)= 1075.8) and the model including the nominal factor Congruency (log(L)= 1079.8) were significantly different ($\chi^2(0)= 7.90$, $p< 0.0001$). The model including all mass-specific determiner occurrences (log(L)= 1079.6) was significantly different from the null model ($\chi^2(1)= 7.51$, $p= 0.001$), as was the model including the occurrences of *viel* (log(L)= 1079.6): $\chi^2(1)= 7.50$, $p= 0.001$).

Again, comparing the models with each other did not yield significant differences. The model including the nominal factor Congruency has the highest data likelihood value. It also yielded the highest Chi² value and lowest p value when compared to the null model.

3.2.3. Discussion

The results of the reaction time analysis show that there is a clear congruency effect for count nouns and mass nouns. In both cases, the noun was processed significantly faster when it was preceded by the congruent article (*ein, a(n),* for count nouns and *viel, much,* for mass nouns) than when it was preceded by the incongruent article. It can thus be concluded, that countability as a property is utilised during phrase processing.

Trends in the same direction were also observable for the dual-count and dual-mass nouns. However, here the difference was less pronounced and did not reach significance. The congruency effect in the reaction time data therefore reflects the results of the corpus study: the more often a noun was found to occur with a certain set of determiners or quantifiers the easier it is to process the noun in a noun phrase where it is combined with an article specified for countability. The effect can however not be seen as a direct reflection of the categorisation of the noun stimuli into the countability classes since the classification was based on a noun's co-occurrences with the entire set of determiners whereas in the experiment only one of the determiners was used to manipulate congruency.

The results of the error rate analysis confirm this view. Here, too, more effortful processing of the incongruent condition is reflected by higher error rates for this condition than for the congruent condition (except for the dual-mass nouns, where error rates are similarly low in both conditions). It is thus not the case that a trade-off has taken place between speed and accuracy, i.e. that noun phrase stimuli were processed and the subsequent lexical decision was made either fast or accurately. Note however that this difference was not significant in the error rates and should thus be considered with caution. In general, the error rates were very low (highest for mass nouns in neutral condition: 11 errors out of 480 trials, 2.29%).

Reaction times for the neutral condition were significantly higher than for the congruent and the incongruent condition for count nouns and dual-count nouns. For mass and dual-mass nouns, the neutral condition yielded reaction times faster or more or less equal to those yielded by the incongruent condition, and the difference was not significant. The main reason for this is probably rooted in the fact that the noise segment was actually perceived by the participants as equivalent to silence. As pointed out in the introduction, singular mass nouns but not count nouns can form a grammatical noun phrase on their own, that is they do not need to be preceded by an article in German. From that perspective, the neutral condition (perceived as a bare singular noun) is actually congruent for mass nouns but not for count nouns, which explains the different patterns. However, again this difference did not reach significance and can therefore not be interpreted straight-forwardly.

For both the pure count nouns and the dual-count nouns, the differences between the neutral condition and the other two conditions were significant. In both cases, reaction times were longest for the neutral condition. Even though the noise segment had the exact same length as an article (see Section 3.2.1 "Methods" above) and therefore the gaps between noise segment and noun were identical, it is possible that irrespective of the article (congruent or incongruent) the noun phrases in the neutral condition were perceived as less natural with regard to the noun and that the noise segment had an interfering effect on noun comprehension in these trials. It is however more likely that participants' speed was affected by grammatical processing, i.e. the unfamiliarity of hearing a count noun (or dual-count noun) in isolation. As stated above, mass nouns but not count nouns can form grammatical NPs on their own without an article. Thus, should the LMNS in the neutral condition have indeed been processed similar

to silence, as suggested by previous findings , it presents an incongruent condition for the count nouns and the dual-count nouns as the bare noun is considered an uncountable context syntactically.

Curiously, in the case of the neutral condition the pattern yielded by the error analysis is reversed. Here, the mass and dual-mass noun trials of the neutral condition yielded higher error rates than it was the case in count and dual-count noun trials, suggesting more effortful processing. The possibility of a speed-accuracy trade-off can thus not be excluded in this case. Why this should be different for the different countability classes, however, lacks an obvious explanation.

The reaction time data for the neutral condition patterns differently for the levels of Countability Class. The neutral condition did not yield either uniformly longer or shorter reaction times than the incongruent condition. The present data therefore does not offer a clear indication as to whether the congruency effect is of a facilitating or an inhibitory nature.

None of the semantic factors yielded a significant main effect or interaction. Concreteness was not found to significantly affect reaction times or error rates and thus properties of the noun's referents do not seem to influence noun processing during this task.

There was also no significant main effect or interaction of Shiftability, however as can be seen in Figure 15, participants made more errors during trials of the incongruent condition with non-shiftable count nouns than in the incongruent condition with shiftable count nouns. This suggests a more difficult processing of non-shiftable count nouns than shiftable count nouns when they are combined with the incongruent determiner. The difference, however, did not reach significance and is neither clearly present in the mass noun error data nor reflected in the reaction times (see Figure 17).

The analysis that compared a countability factor based on an abstract category with a countability factor based on the actual co-occurrence data from the corpus study unfortunately did not yield clear results. No significant difference could be found between the models using the two kinds of factors and the comparison between the log data likelihood did not show any strong differences either.

In conclusion, it can be said that countability certainly does affect the processing of nouns in noun phrases, as a significant main effect of congruency was found in the reaction time data. This disproves accounts that place the countability feature completely outside the lexical entry of a noun and strongly suggests that a noun's mental representation is marked for countability as well as the determiner's mental representation, as congruency reflects the goodness of fit between the two elements.

It can also be said that during the processing of the noun phrase in order to be able to make a lexical decision, syntactic factors such as co-occurrence with determiners (as determined by the corpus study) affect noun comprehension, whereas semantic factors like Shiftability and Concreteness do not. The lexical decision task mainly targets the lower levels of a lexical entry (i.e. the word form and possibly the lemma level), since the decision of whether or not a given sound string represents a noun in German can be made without reaching the conceptual level and thinking about what the noun or even the complete noun phrase refers to. Clearly, such additional processing cannot be excluded, yet most semantic effects found with lexical decision studies can be explained in terms of spreading activation between mental representations, a process that is thought to be automatic and uncontrolled in the sense that it is distinct from e.g. syntactic processing. These effects are thought to reflect the

architecture of semantic memory rather than the processes of language comprehension. It is thus not clear whether the source of these effects is actually conceptual in nature as defined here (i.e. non-linguistic or extra-linguistic, see discussion on metonymy in Section 2.4 “Words that mean many things (or stuff)”). Attempts to distinguish types of semantic priming effects by conceptual properties (e.g. subordination, such as *animal-dog* vs. association, such as *cat-dog*) have been found to be unreliable for the amount of data usually collected during psycholinguistic behavioural studies. In fact, it has been suggested that when a study rigorously controls for non-semantic factors (such as morphological/orthographical overlap of the prime and the target, or the possibility to strategize) semantic priming effects are even less reliable (Heyman et al. 2018).

Thus, to shed more light on how and when exactly the effect found here occurs, additional data is needed. The two experiments presented next will provide more insight into this question.

3.3. Lexical Decision Experiment 2

This second lexical decision experiment was conducted as a replication of the previous lexical decision experiment and as a pretest for a second set of stimuli, which was created for the phoneme monitoring experiment. The paradigm of using both lexical decision and phoneme monitoring on the same material has been previously employed by Bölte & Connine (2004) in their research on gender in German as well as by Brenner (2016) in her research on concept types and it is well suited to differentiate between an effect locus during the stage of lexical access and a post lexical locus (see Section 3.4. “Phoneme Monitoring” below for more details).

The nouns used in this set differ from the ones used in Lexical Decision 1 in two crucial points. First, the nouns of the current stimulus set are longer, mostly three syllables long. Second, the current noun stimuli are no longer classified for semantic factors such as Concreteness and Shiftability, since for those factors no effect on processing during a lexical decision task could be observed in the first experiment.

3.3.1. Method

Materials

Classification for Countability Class of the stimulus set used in Lexical Decision Experiment 2 was almost identical as for the stimulus set used in Lexical Decision Experiment 1.

48 count, dual-count, dual-mass and mass nouns (12 of each) were selected as critical items. For a few nouns, where not enough occurrences with these specific determiners/quantifiers could be found in the corpus used in Pretest 1, the first 100 occurrences of the noun were extracted from the corpus and classified sentence by sentence to see whether the noun was used with count or mass syntax. Based on the resulting distribution, the noun was classified for Countability Class with the same threshold parameters as in Lexical Decision Experiment 1 (100% count-specific = pure count noun, 90-50% count-specific = dual-count noun, 50-10% count-specific = dual-mass noun, 0% count-specific = pure mass noun).

These nouns were all 2-4 syllables in length. One quarter of those nouns ended in l, n, k and t respectively and did not have the respective phoneme at any other point in the word. Because German experiences final devoicing, nouns ending in d, which would be pronounced /t/ as well, were excluded. The nouns were all matched for frequency by Countability Class using the corpus *Projekt Deutscher Wortschatz* (see Table 12).

Table 12: Mean frequency and standard deviation of stimulus nouns per Countability Class
Note: The frequency is given in Häufigkeitsklassen (engl. class of frequency) as used by the corpus. (cf. Table 10)

Count nouns		**Dual count nouns**		**Dual mass nouns**		**Mass nouns**	
Mean: 13,00	SD: 0,70	Mean: 13,50	SD: 1,71	Mean: 12,83	SD: 1,28	Mean: 13,83	SD: 1,72

For each of these nouns (e.g. *Kartoffel, potato*) one minimally different pseudoword was formed by changing the first phoneme in one phonetic feature (place or manner of articulation). Thus the resulting minimal pseudoword started with a phoneme that slightly differed from the original initial phoneme (eg. *Gartoffel* as minimal pseudoword for *Kartoffel*). Further, for each word a maximally different pseudoword was created by changing the initial phoneme by more than one phonetic feature (e.g. *Nartoffel* as maximal pseudoword for *Kartoffel*). The systematicity was maintained whenever possible.[45] In cases were the original noun started with a vowel, the second phoneme was changed instead of the first one, eg. *Übelkeit, nausea*, was changed to *Üpelkeit* (min) and *Ükelkeit* (max). This way of creating stimulus material to be used with a phoneme monitoring task was inspired by Bölte & Connien (2004) and used in Brenner (2016) as well.

Each of the words and pseudowords was presented in three noun phrase environments representing the congruent, incongruent and neutral condition, yielding 432 critical noun phrase items. Additionally, 48 words that didn't include any of the target phonemes (/l/, /n/, /k/ and /t/) were included as fillers, also in three conditions, that is preceded by either *ein(e)*, a(n), *viel,* much, or a length-matched segment of brown noise. Thus, there were 144 filler noun phrase items.

The entire experimental stimulus set (576 noun phrase items) contained as much noun phrase items including real words as it did noun phrase items including pseudowords.

The procedure of stimulus recording and editing as well as the creation of stimulus items for the neutral condition was identical to the one used for the stimuli for Lexical Decision Experiment 1 (see Subsection 3.2.1)

The 144 critical items were distributed across nine experimental lists, so that each of the critical words and pseudowords would be used only once for each participant (see Table 13).

Table 13: List distribution of critical noun phrase items

Condition	**Base Word**	**List**	**Min Pseudo**	**List**	**Max Pseudo**	**List**
Congruent	eine Kartoffel	L1	eine Gartoffel	L2	eine Nartoffel	L3
Incongruent	viel Kartoffel	L2	viel Gartoffel	L3	viel Nartoffel	L1
Neutral	xxx Kartoffel	L3	xxx Gartoffel	L1	xxx Nartoffel	L2

[45] In cases where changing one phonetic feature yielded an existing word or a sound sequence that violated German phonology, more than one feature was changed.

The fillers were distributed so that each list contained 16 words preceded by *ein(e)*, *a(n),* 16 words preceded by *viel, much*, and 16 words preceded by a length-matched segment of brown noise. The order of each list was pseudo-randomized, so that not more than 3 repetitions of determiner/quantifier/LMNS, condition, or criticality (critical item/filler item) followed each other.

Subjects

39 German native speakers between 18 and 36 years of age participated in the experiment. Each reported to have no language-specific, psychological or neurological disorders. Every participant had normal or corrected to normal vision and hearing ability and was able to move his or her arms in a normal fashion (e.g. no casts or paralysis).

None of the participants reported to have an additional native language without German-like morpho-syntactically realised countability (e.g. Japanese) or to have completed more than 2 semesters of linguistic studies at university.

Procedure

This experiment, too, was programmed using Presentation software Version 19.0 by Neurobehavioral System Inc.

The procedure was identical to the procedure described for Lexical Decision Experiment 1: After filling out some informational forms about their age, education and language experiences, and giving informed consent on their participation, participants were verbally instructed about their task and on what to expect from the stimuli, i.e. that some of the items were not real words and that some of the noun phrases contained a noise segment instead of a determiner/quantifier.

Each participant then sat individually in a dimly-lit, sound-attenuated chamber and listened to the stimuli via headphones. After some short written instructions and a training phase of 7 trials that were not part of the actual experimental stimulus material, the experiment started. Participants pressed one out of two buttons to indicate whether the noun in the heard noun phrase was an existing German word or not. Buttons were assigned so that the button for real words was always operated with the dominant hand.

For further details about the technical specifications of the experimental set up, inter-trial durations, block length and the like, see section 3.2.1. “Lexical Decision Experiment 1; Methods”.

3.3.2. Results

As for Lexical Decision Experiment 1, all steps of pre-processing and analysis were conducted using R version 3.4.2 (2017-09-28, The R Foundation for Statistical Computing, 2017). Package lme4 (Bates, Mächler, Bolker & Walker, 2014) was used to perform statistical analysis. P-values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question. All statistical analyses were carried out on trials for real words only (c.f. Subsection 3.2.2).

Error Analysis

Again, an erroneous trial was defined as a trial where the lexical decision for words was not answered with the word-button or the decision for pseudowords was not answered with the button assigned to pseudowords ("Kein Wort", *not a word*). All 39 participants answered 96 trials each during the experiment. Data from one participant had to be excluded because he/she had answered more than 25% of all lexical decisions incorrectly. With the same threshold (25% of trials) individual items were checked for too many mistakes, but no item had to be excluded.

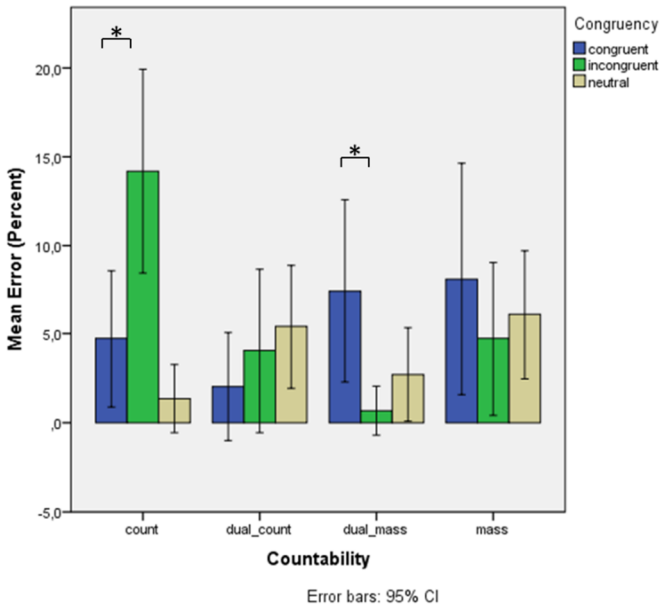


Figure 18: Mean error rates by Countability Class and Congruency

Countability and Congruency were entered into the model as fixed factors. As random effects, intercepts for subjects and items were included as well as by-subject and by-item random slopes for the effect of Congruency. Generalised linear mixed model analysis revealed no significant main effect for either factor, but a significant interaction of Congruency with Countability Class ($\chi^2(6)= 37.93$, $p<0.0001$). When the data was divided for the different levels of Countability Class, a significant main effect for Congruency was found for count nouns ($\chi^2(2)= 21.77$, $p<0.0001$). Pairwise comparison (Bonferroni-corrected) revealed that the incongruent condition (14,2%) had significantly more errors than the congruent condition (4,7%): z = 2.7, p= 0.018). There was no significant main effect of Congruency for the dual-count nouns. For the dual-mass nouns, however, there was a significant main effect of Congruency ($\chi^2(2)= 11.6$, $p<0.01$). Pairwise comparison (Bonferroni-corrected) revealed that the congruent condition (7,4%) had significantly more errors than the incongruent condition (0,7%): z = -2.47, p= 0.04). For the mass nouns, again, the difference between Congruency levels did not reach significance.

Reaction time Analysis

Reaction time was measured from noun onset on. 171 trials were excluded from analysis because the lexical decision was answered incorrectly (4,69% error rate). The outlier correction was conducted per condition. The threshold for exclusion was set to more than two standard deviations from the mean (for subjects and Items). For the congruent condition 27 trials were excluded, for the incongruent condition 30 trials were excluded and for the neutral condition 24 trials were excluded.

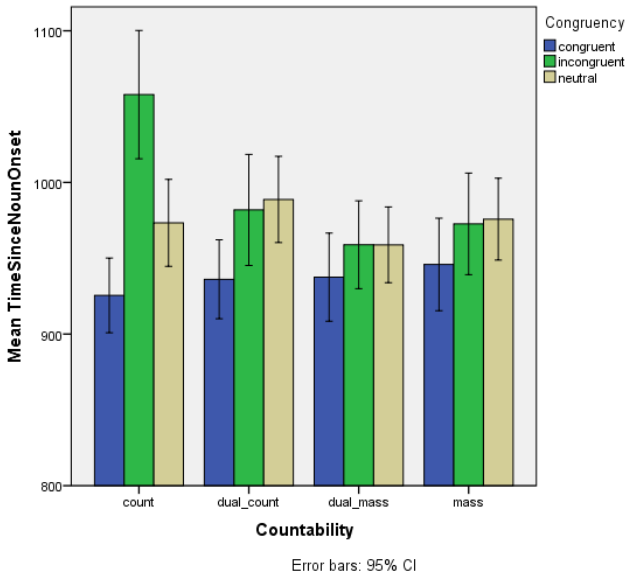


Figure 19: Mean reaction time by Congruency and Countability Class

For statistical analysis, Countability Class and Congruency were entered into the model as fixed factors. As random effects, intercepts for subjects and items were included, as well as by-subject and by-item random slopes for the effect of Congruency. A significant main effect for Congruency was found ($\chi^2(2)$= 13.78, p<0.01). There was no significant main effect of Countability Class and the interaction Congruency*Countability Class was also not significant.

Pairwise comparison using the Satterthwaite method with Bonferroni correction revealed a significant difference between trials of the congruent condition (*M*= 936,11ms, *SD*= 160,35) and trials of the incongruent condition (*M*= 990, *SD*= 206,24): t(16,7)=-3,07, *p*<0-01. Further, it revealed a significant difference between trials of the congruent condition and trials of the neutral condition (*M*= 973,87, *SD*= 160,94): t(11)=-3,79, *p*= >0,01. The difference between trials of the incongruent and trials of the neutral condition was not significant

.

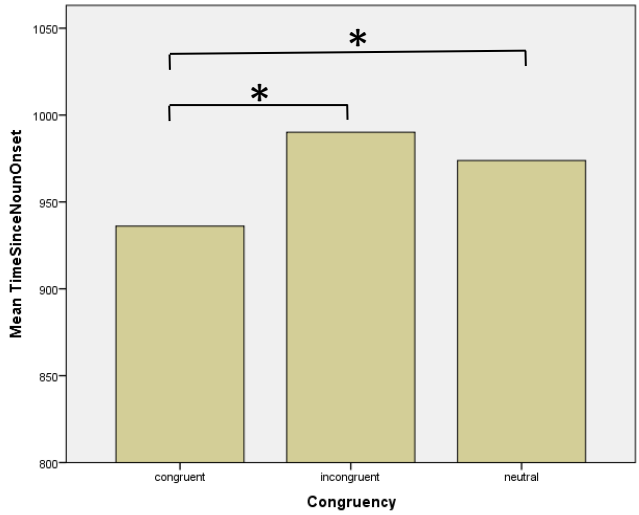


Figure 20: Mean reaction time by Congruency

3.3.3. Discussion

The two main goals of the experiment were achieved. The congruency effect of Lexical Decision 1 was replicated and the same effect licenses the stimuli for the lexical decision-phoneme monitoring paradigm.

Despite the significant main effect of congruency found in both experiments, the data patterns a little differently. Firstly, in the reaction time analysis of the first experiment, a significant interaction between Congruency and Countability Class was found in addition to the main effect of Congruency. It revealed that the difference between congruent and incongruent condition was only significant for pure count nouns and mass nouns, but not for dual nouns. This modulation of the congruency effect is absent in the reaction time data of the second experiment.

In Lexical Decision Experiment 2, the highest reaction times were obtained for incongruent count nouns. Here, reaction times were far higher than in any other condition (see Figure 19). Also, this condition yielded the most erroneous responses (see Figure 18). In Lexical Decision 1, half the count nouns were shiftable – that is interpretable with the incongruent determiner – which was not the case for the stimuli in the current experiment. While this manipulation did not turn out to be significant in the first experiment, it could have caused a slight difference to the processing of the current stimuli. It is possible that the currently used count nouns were slightly less variable in there countability in terms of interpretation. If however the stronger congruency effect had been due to reduced shiftability one would have expected the same for the mass nouns and there, the difference between congruent and incongruent trials was not as high. Additionally, count nouns in the second experiment were mostly concrete (8 out of 12). It is unlikely, though, that this could explain the higher difference between Congruency levels because that should have actually made the count nouns easier to undergo the Grinder shift and thus they would suffer less from incongruency.

A more crucial difference between the two experiments is the number of participants. Lexical Decision Experiment 2 had fewer participants than Lexical Decision Experiment 1 and it is conceivable that the one-sidedness of the congruency effect might have washed out somewhat with more data points. That is, the limited number of data points in Lexical Decision Experiment 2 could have led to differences caused by specific properties of individual items or the processing of individual participants being more prominently represented in the data, concealing more general patterns. The patterns however show similarities. In fact, as can be seen in Figure 16, the difference in reaction times between congruent and incongruent trials was also higher for count nouns than for mass nouns in Lexical Decision Experiment 1.

Another curious feature of the result of Lexical Decision Experiment 2 is the high error rate of dual-mass and mass nouns in the congruent condition. When one considers this in addition to the reaction time data, a speed-accuracy trade-off does not appear to have taken place. From the individual noun stimuli there is no apparent reason for this unusual result; none of the stimuli seems infelicitous with the article *viel, much,* see Appendix 2 for all nouns used.

With regard to the neutral condition, the finding of Lexical Decision Experiment 1 that reaction times for these trials were higher than for the incongruent condition for count nouns but shorter than for the incongruent condition for mass nouns could not be replicated. When the data in Figure 19 is

observed closely, one can see that for the dual-count nouns reaction times for the neutral condition were a little higher than for the incongruent condition. For the count nouns, this pattern is absent, mainly because the reaction times towards the incongruent condition were so high. For dual-mass nouns, the difference between incongruent and neutral condition is virtually non-existent, just like in Lexical Decision 1. Only the mass nouns thus show an unexpected pattern in that here the neutral condition yielded slightly higher reaction times than the incongruent condition. This difference however is very small indeed.

Since none of these differences are significant, however, they should be interpreted with great caution since it is possible that they simply reflect incidental properties of the data rather than indications of real effects.

In conclusion, the main finding of the experiment was the significant congruency effect in the reaction time data. It shows that also for stimuli that are chosen purely based on syntactic properties (i.e. co-occurrence with count- and mass-specific determiners) a significant difference between the congruent and incongruent condition can be obtained (see Figure 20).

3.4. Phoneme Monitoring Experiment

Phoneme monitoring is widely accepted as a task that – contrary to lexical decision – does not target post-lexical integration processes but only lexical access. In primed lexical decision of the kind reported in this dissertation, there are two stages of phrase processing that are possibly affected by the stimulus manipulation.

The first one is the stage of lexical access. This processing stage includes a search for an entry in the mental lexicon that matches the perceived input. Multiple candidates are activated during this search, which compete with each other for selection until finally the best candidate is selected.

There is ample evidence that word processing is influenced by previously perceived elements during the comprehension of complex linguistic units beyond individual words, such as phrases, clauses and sentences. This phenomenon is called priming and the assumption is that parts of upcoming linguistic material are pre-activated because of previously encountered input. This 'anticipatory' or 'predictive' processing is assumed to be part of many cognitive processes, and while it may not be strictly necessary for spoken language perception, it is a fundamental part of it (see e.g. Huettig, 2015, for a review).

For example, there is structural priming that allows the listener to make deductions on grammatical properties such as word class. Since part of our linguistic knowledge is knowing which syntactic structures are possible in our language and which are not, we know that for example after the input *She saw a long…* a noun or an adjective is most likely to follow, while a verb is rather unlikely since the resulting language segment would not follow the grammatical rules of English.

Priming is also assumed to affect the semantic side of processing. Previous context affects the likelihood of a particular word to occur, a phenomenon referred to as 'cloze probability'. A word is easier processed when the previous context 'builds up towards it' and limits the concepts that could plausibly be part of the state of affairs described in a given clause or sentence. For example, if one hears *Jimmy was buttering his bread with a…* the word *knife* has a far higher probability to occur than say the word *spoon*

and an even higher probability than the word *car*, because from world knowledge we know that one usually uses a knife to butter bread rather than a spoon or a car.

This sort of predictive language processing has been observed for many linguistic phenomena and in speakers of many different languages. Some researchers interpret the priming effects they have found as a sign that the search space of candidates for lexical selection is limited based on previously obtained information that can be used to deduce properties of the upcoming element. This is especially the case for phonological priming (see e.g. Mani & Plunkett, 2011; or De Long et al., 2005, for priming effects that cannot be explained by integration processes). However, there have been works that attempt to place (even phonological) priming effects at a post-lexical stage.

In light of the current research issue, the crucial question is: Does countability-specific information from the article limit the search space for the upcoming noun so that only nouns are considered that possess the expected countability value? That is, if we hear *ein…* do we predict that a count (or dual-count) noun will follow and therefore only consider lexemes marked for count (or compatibility with the indefinite determiner) as candidates during lexical access?

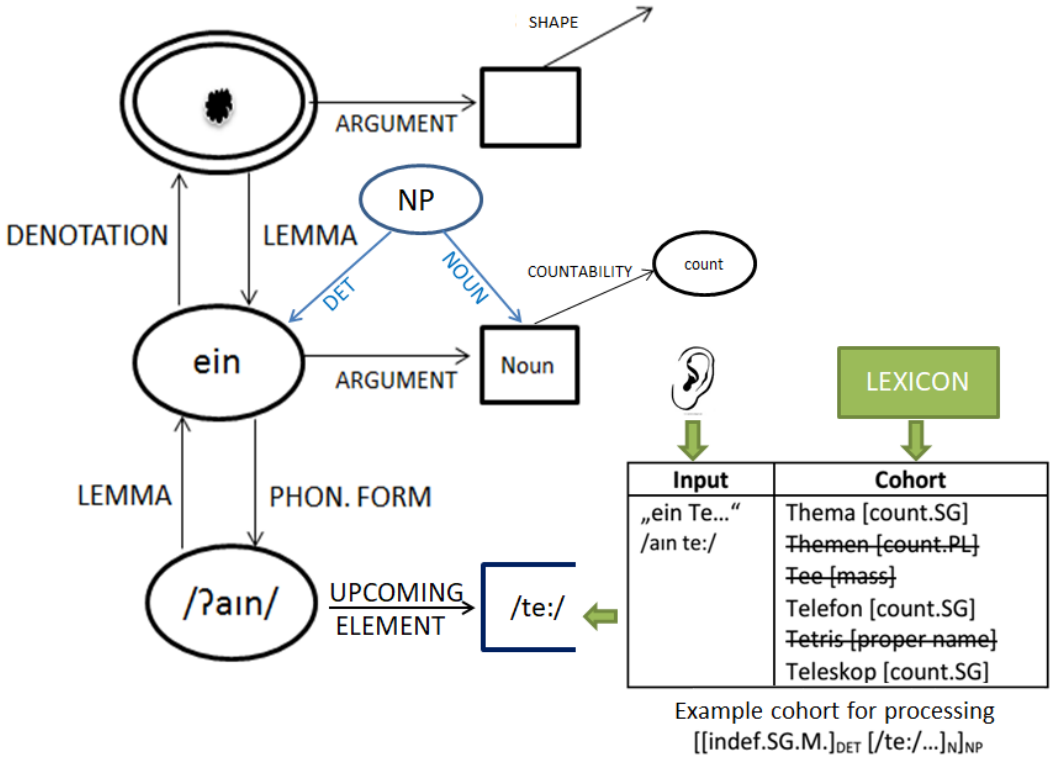


Figure 21: Lexical access modulated by countability
Note: round nodes = values, arrows = attributes, square nodes = empty values,
phon. form = phonological form, NP= noun phrase, DET = determiner, SG = singular, PL = plural, M = masculine

Figure 21 depicts the process of noun phrase comprehension under the hypothesis that countability information affects lexical access in the manner of predictive processing outlined above. On the left side the lexical entry of the determiner *ein, a(n),* is shown which raises the expectation of an upcoming argument on all three levels (from top to bottom: conceptual, grammatical and phonologically, see Subsection 2.4.2 “Metonymy, Shifts and Coercion” or Beckmann, Indefrey & Petersen, 2018, for more details on the assumed structure of lexical entries). These predicted elements are represented by square nodes and can themselves have attributes with fixed values that represent their predicted (or ‘deduced’)

properties. For instance, on the lemma level, the expected argument noun (square node) has an attribute COUNTABILITY with the value "count", meaning that a (syntactically) count noun is predicted as the argument of *ein*. On the bottom of the right side the process of lexical access of the noun is shown with a representation of the current (incomplete) input and a list ('cohort') of possible noun candidates from the mental lexicon. This list is restricted in terms of deduced properties based on the information gained from the determiner, namely that the upcoming noun is a (sortal) singular count noun.

The congruency effect found in the lexical decision experiments may reflect a mismatch between the expected noun (properties) and the noun that is actually perceived. That is, in the incongruent condition, the article *ein, a(n),* gives rise to the expectation of a countable noun but a mass noun follows, or the article *viel, much*, lets the participant predict that a mass noun will follow but then they hear a count noun. Because the cohort of possible candidates was limited to nouns with a congruent countability value, the process of lexical selection has to be corrected or repeated (or proceeds in a suboptimal way in any case) once it becomes clear that the perceived input cannot possibly be matched to one of the expected candidates. It might be this repetition or hindrance that is responsible for the longer reaction times (and the higher error rate) found for trials of the incongruent condition as compared to the congruent condition.

However, an alternative explanation for the congruency effect cannot be discarded yet because the lexical decision task also targets post-lexical operations and the effect could thus be a reflection of a problem with integration or a well-formedness check of the noun phrase that occurs after the noun has already been lexically accessed. After the lexical selection process is completed, information from the lexical entry of the selected lemma is retrieved and the noun is integrated into the noun phrase, that is combined with the determiner into one grammatical constituent (see e.g. Faussart, Jakubowitz & Costes, 1999). It is assumed that this process – like all of language processing – is subject to a self-monitoring process. While most of the early literature is concerned with monitoring one's speech production, there is convincing evidence that the monitoring process is in fact domain general, i.e. also applies to speech comprehension. Gauvin & Hartsuiker (2020), for instance, propose a model of self-monitoring where the constant interplay between the production system and the perception system automatically creates a monitoring system for perceived language input (among other things) which is checked for 'conflicts' by a domain-general checking mechanism.

It is thus entirely possible that countability does not affect lexical access – that is, nouns of all countability classes are considered and the cohort of possible candidates is not limited to only those that are countability-congruent – and incongruency is detected during a later stage of phrase processing. Namely, when the article and the noun are combined to (or 'integrated into') a noun phrase, a monitoring process could detect a mismatch between the countability value of the determiner and that of the noun (e.g. *ein*$_{\text{COUNT}}$ + *Wasser*$_{\text{MASS}}$, **a water*). This mismatch detection could lead to delayed processing, either because of a repair attempt including the repetition of (some stages) of the lexical and post-lexical processes or the detection of an error in itself could slow down integration.

In order to differentiate between those two possibilities, the stimuli of Lexical Decision Experiment 2 were also tested with a phoneme monitoring task, because – unlike lexical decision – phoneme monitoring is sensitive to lexical access processes but not to processes of integration or well-formedness checking mechanisms. In the phoneme monitoring experiment the task is to decide whether a certain target phoneme is part of the noun. It is therefore not necessary to actually perform an integration of the noun into the noun phrase.

The noun is lexically accessed in order to speed up the decision making process. That is, the uniqueness point of the noun – the phoneme from which on the noun can be unambiguously identified – may occur before the target phoneme occurs. Therefore to lexically select the noun as fast as possible is advantageous during this task. In the current experiment, all the target phonemes always occurred as the last phoneme of the critical stimulus nouns, thus per definition either after (or simultaneously with) the uniqueness point of the word.

During lexical access, activation of each cohort candidate is modulated by (among things like frequency and family size, etc.) how well the candidate matches the perceived input. That is, even if the input is not a perfect match (as is the case with the pseudowords here) an existing word form representation is activated and might even be selected (for an overview on evidence for multiple activation and the activation of words based on fragments see McQueen, Dahan and Cutler, 2011).

The current experiment employs three types of critical stimuli: real nouns, minimal pseudowords that differ from the existing noun a little and maximal pseudowords that differ somewhat more. The underlying assumption is that the pseudowords resemble the existing word enough for it to take part in the competition during lexical access. That is, the representation of the existing noun will receive enough activation during the comprehension of the pseudoword for it to influence the recognition process and therefore the task performance. I predict that the performance speed of the phoneme monitoring task will be mediated by how similar a given pseudoword is to its existing base noun, as was previously found (e.g. by Bölte & Connine, 2004) and consider this similarity effect as indicative of the sensitivity of the task to the lexical stage of language processing.

Thus, while in lexical decision phrase integration takes place as evidenced by the congruency effects found for e.g. gender and number agreement (see Subsection 2.5.3 “Online processing of noun phrases”), phoneme monitoring is assumed not to target post-lexical processing stages (see also Bölte & Connine, 2004, and Brenner, 2016, for a discussion of these task differences).

Whether or not the congruency effect found in Lexical Decision Experiment 2 is also found in the phoneme monitoring experiment will therefore tell which of the two processes described above is influenced by countability. If the congruency effect is found with lexical decision as well as phoneme monitoring then a lexical locus of the effect must be assumed. If, however, the congruency effect is not found with the phoneme monitoring task, then this result is indicative of a post-lexical locus of the effect.

3.4.1. Method

Materials

For the phoneme monitoring experiment, all stimuli from Lexical Decision Experiment 2 were used. That is, 48 critical nouns, containing equal numbers of count nouns, dual-count nouns, dual-mass nouns and mass nouns (n=12, respectively) were included and used as basis to form corresponding minimal pseudowords (differing a little from the original) and maximal pseudowords (differing somewhat more). All of these words and pseudowords ended in one of the target phonemes: /l/, /n/, /k/ or /t/ (¼, respectively).

Additionally, 32 words and 32 corresponding pseudowords that had one of the target phonemes in initial position and further 16 words that had one of the target phonemes in final position served as fillers. 64 words and 64 corresponding pseudowords that did not contain any of the target phonemes were included as fillers as well.

The entire stimulus set (n=352) therefore consisted of 160 words and 192 pseudowords. 96 words and 128 pseudowords contained the target phoneme (224 in total) while 128 words and pseudowords did not contain any of the target phonemes. All groups of words and pseudowords with a target phoneme (critical/filler, target phoneme initial/final) contained equal amounts of each of the target phonemes /l/, /n/, /k/ and /t/.

Each word and pseudoword was presented in 3 noun phrases, one for each condition (congruent, incongruent and neutral), that is being preceded by either *ein(e), a(n), viel, much,* or a length-matched noise segment (LMNS).

Recording parameters, specifications and subsequent sound editing (including creation of the LMNS) were identical to those described for the stimulus set for the lexical decision experiments. For details see the Methods section for Lexical Decision Experiment 1 (Subsection 3.2.1).

All stimuli were distributed over 9 lists in a way so that each noun phrase item with a critical noun and its corresponding minimal and maximal pseudoword only occurred once per list, as described in Table 14.

Table 14: List distribution of critical noun phrase items
Note: The LMNS is represented by [xxx]

Status	Condition	Noun Phrase Item	List
Word	congruent	eine Kartoffel	L1
Word	incongruent	viel Kartoffel	L2
Word	neutral	[xxx] Kartoffel	L3
Min Pseudo	congruent	eine Gartoffel	L4
Min Pseudo	incongruent	viel Gartoffel	L5
Min Pseudo	neutral	[xxx] Gartoffel	L6
Max Pseudo	congruent	eine Nartoffel	L7
Max Pseudo	incongruent	viel Nartoffel	L8
Max Pseudo	neutral	[xxx] Nartoffel	L9

Additionally to the resulting 48 critical noun phrase items per list, all 208 filler noun phrase items were included in every list, containing *ein(e), a(n), viel, much,* or a length-matched segment of brown noise in equal amounts. Consequently, each list contained 256 noun phrase items in total.

Each list was divided into four blocks; one for each of the target phonemes (/l/, /n/, /k/ and /t/). There were 24 versions for each list, so that the order of the blocks was not the same for every participant. The order of each block was pseudo-randomized, so that not more than three repetitions of determiner/quantifier/LMNS, condition, or criticality (critical item/filler item) directly followed each other. Care was also taken that the filler words and their corresponding pseudowords did not occur in the same block.

Subjects

90 German native speakers between 18 and 36 years of age participated in the experiment. Each reported to have no language-specific, psychological or neurological disorders. Every participant had normal or corrected to normal vision and hearing ability and was able to move his or her arms in a normal fashion (e.g. no casts or paralysis).

None of the participants reported to have an additional native language without German-like morpho-syntactically realised countability (e.g. Japanese) or to have completed more than 2 semesters of linguistic studies at university.

Procedure

Participants of the Phoneme Monitoring Experiment were tested in a reaction time lab on the campus of Heinrich-Heine-University, Düsseldorf. The technical equipment and set up was identical to that of Lexical Decision Experiment 1 and Lexical Decision Experiment 2, see Subsection 3.2.1 "Lexical Decision Experiment 1, Methods" for details. This experiment was programmed using Presentation software Version 19.0 by Neurobehavioral System Inc. as well.

When the participants arrived in the lab, they were greeted and asked to take a seat at a table to fill out some forms. After filling out informational forms about their age, education and language experiences, and giving informed consent on their participation, participants were verbally instructed about their task and on what to expect from the stimuli. They were told that the experiment tested their language perception, that they would hear noun phrases and that they had to press a button as fast as they could whenever they heard a specific sound in the noun that would be displayed on the screen. They were given an example like in (45)

(45) Example instruction of phoneme monitoring task for participants

"Stell dir vor, der gesuchte Laut ist L und du hörst ‚eine Hantel', ‚viel Laub' und ‚viel Sand'. Dann drückst du den Knopf bei ‚eine Hantel', weil in ‚Hantel' ein L ist und bei ‚viel Laub', weil in ‚Laub' ein L ist, aber nicht bei ‚viel Sand', weil in ‚Sand' kein L ist."

English translation – example phrases altered to maintain logical consistency

"Imagine the target sound was M and you heard 'a storm', 'the muffin' and 'much sand'. Then you press the button for 'a storm' because 'storm' has an M in it and for 'the muffin' because 'muffin' has an M in it, but not for 'much sand' because 'sand' doesn't have an M in it."

It was explained that some of the items were not real words and that some of the noun phrases contained a noise segment instead of a determiner/quantifier but that participants should perform the task normally on these items as well.

After being asked to leave their phone, smart watch and any other devices that might distract them during the experiment behind, participants were individually seated in a dimly-lit, sound-attenuated chamber and asked to adjust the chair there so that they could reach the button box on the table in front of them comfortably. Once seated, they were familiarized with the button box, especially the keys they had to use during the experiment, namely the ENTER button and the button they had to press to indicate that they had perceived the target phoneme.

They were informed that the experiment would consist of four blocks interrupted by three pauses, the length of which they could determine themselves via button press. Finally, they were informed that the experiment would start with a training phase so they could familiarize themselves with the task and instructed to put on the headphones.

All instructions were read from a prepared sheet by the experimenter so that each participant received the same instructions and to ensure that they would not miss something when reading it themselves.

The training phase consisted of five trials that did not contain critical items of the actual experiment. All conditions and real words as well as pseudowords were included in the training phase. After the training phase, participants were asked whether they had understood the task. This was also monitored by the experimenter present by means of checking the obtained score that showed the accuracy of the participant's answers. They were asked whether the volume of the headphones and the level of light were comfortable to them, which was adjusted if they wished. In cases of comprehension problems, the participant's questions were answered and the training phase was repeated. They were then instructed to place the index finger of their dominant hand on the button they had to press to indicate that they had perceived the target phoneme and only remove it during the presentation of instructions or during pauses.

The actual experiment started with another screen that quickly described the task again. The participant had to press the ENTER button to indicate he/she had understood the instruction and was ready to begin. After that, the first block started with a screen announcing the target phoneme and displaying the letter corresponding to the target phoneme for the block in a bright green square that matched the colour of the corresponding button on the button box. Again, the first trial did not start until the participant pressed ENTER. Then the first 64 trials, containing items including the first target phoneme and items that did not include any of the target phonemes were presented in Block 1. Between the trials, a fixation cross was displayed in the middle of the screen for 500ms that was accompanied by a short beep sound. Then, the noun phrase item recording was auditorily presented and the green square containing the letter representing the target phoneme was displayed on the screen. After the audio recording had finished the letter remained on the screen for another 2000ms, then a black screen was presented for 1000ms and the next trial began. After the first block was finished, the participant had a break, which was announced by a written instruction on the screen together with the score obtained during the block (correct answers in percent). After the participant finished the pause by pressing ENTER, the next block preceded in identical fashion.

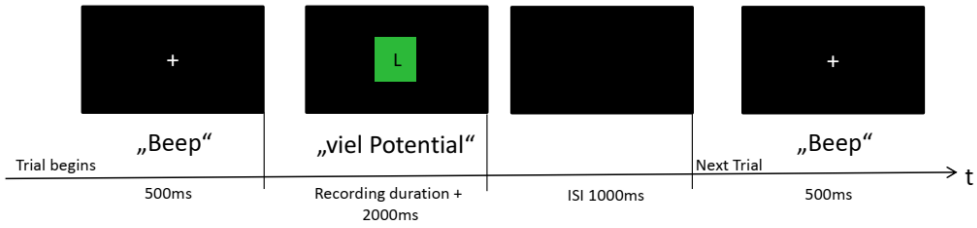


Figure 22: Representation of trial procedure

After the last block was completed, a screen thanking the participant occurred. The experimenter then opened the door of the sound-attenuated chamber and let the participant exit. Each participant was reimbursed with 4€ and candy and thanked again for their participation. Before leaving the lab, the participant was asked what he/she thought the experiment was about. None of them reported that he/she thought the experimental manipulation concerned countability. No participant reported to be consciously aware of morpho-syntactically realised countability.

The entire procedure took approximately 30 minutes.

3.4.2. Results

All analyses and preprocessing were performed using R version 3.4.2 (2017-09-28, The R Foundation for Statistical Computing, 2017). The lme4 package (Bates, Mächler, Bolker & Walker, 2014) was used to perform a linear mixed effects analysis. P-values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question.

Error Analysis

In the complete data set (including pseudowords, fillers, etc.) the error rate was 9,23%, that is trials where the button was either pressed even though the target phoneme was not part of the noun, or trials where the target phoneme was part of the noun but the button was not pressed.

The following pre-processing steps were conducted on data containing only trials including critical items, i.e. critical nouns and their corresponding minimal and maximal pseudowords in the three conditions of Congruency. Three participants were excluded from analysis because they made errors on more than 25% of trials. Additionally, 11 items were excluded from analysis because errors were made on more than 25% of trials. The remaining data had an error rate of 7,57%.

For the first test, Similarity was entered into the model as a fixed factor with the levels real words, minimal pseudowords and maximal pseudowords. As random effects, intercepts for subjects and items were included, as well as by-subject and by-item random slopes for the effect of Similarity. Statistical analysis revealed a highly significant main effect of Similarity ($\chi^2(12)= 37,55$, $p<0.001$).

Pairwise comparison (Bonferroni-corrected) showed that the difference between words (4,2%) and minimal pseudowords (10,4%) was significant ($z = 2.87$, $p= 0.008$) and that the difference between words and maximal pseudowords (11,6%) was also significant ($z= 3.52$, $p< 0.001$). The numerical difference between the two types of pseudowords (see Figure 23) did not reach statistical significance.

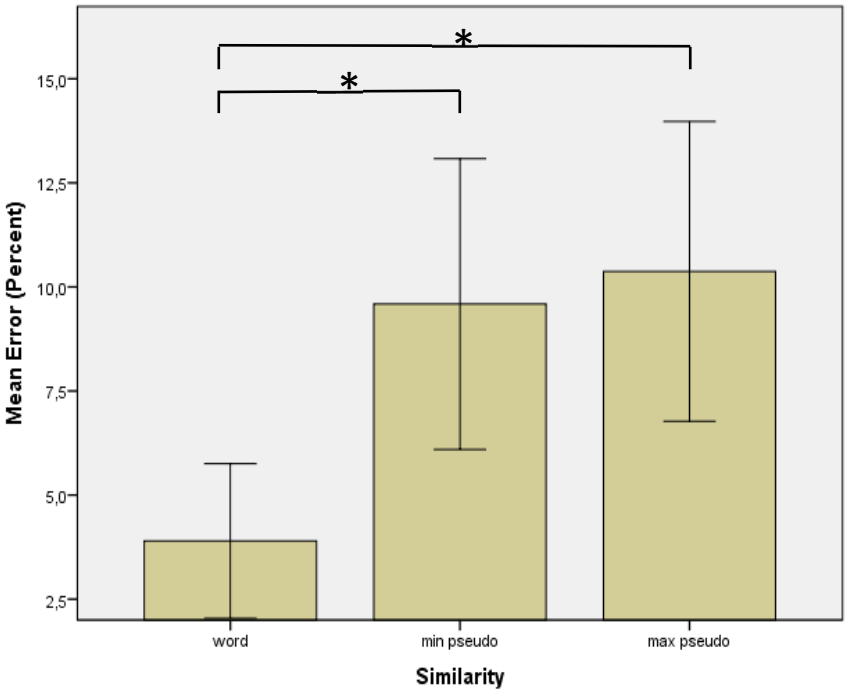


Figure 23: Mean error rates in percent by Similarity

The interaction of Similarity and Congruency was not significant. There was no main effect of Congruency.

Next, Congruency and Countability Class were entered as factors and intercepts for subjects and items were included as random effects as well as by-subject and by-item random slopes for the effect of Congruency. There were no significant main effects and the interaction Congruency*Countability Class also did not reach significance. That is, neither the models predicting main effects for Congruency and Countability Class, nor the model predicting an interaction between Congruency and Countability Class was a significantly better fit for the present data (words plus minimal and maximal pseudowords) than the null model, which didn't include any main effects or interactions.

There were also no significant main effects or interactions found when analysis was limited to trials including words only and repeated with the same models.

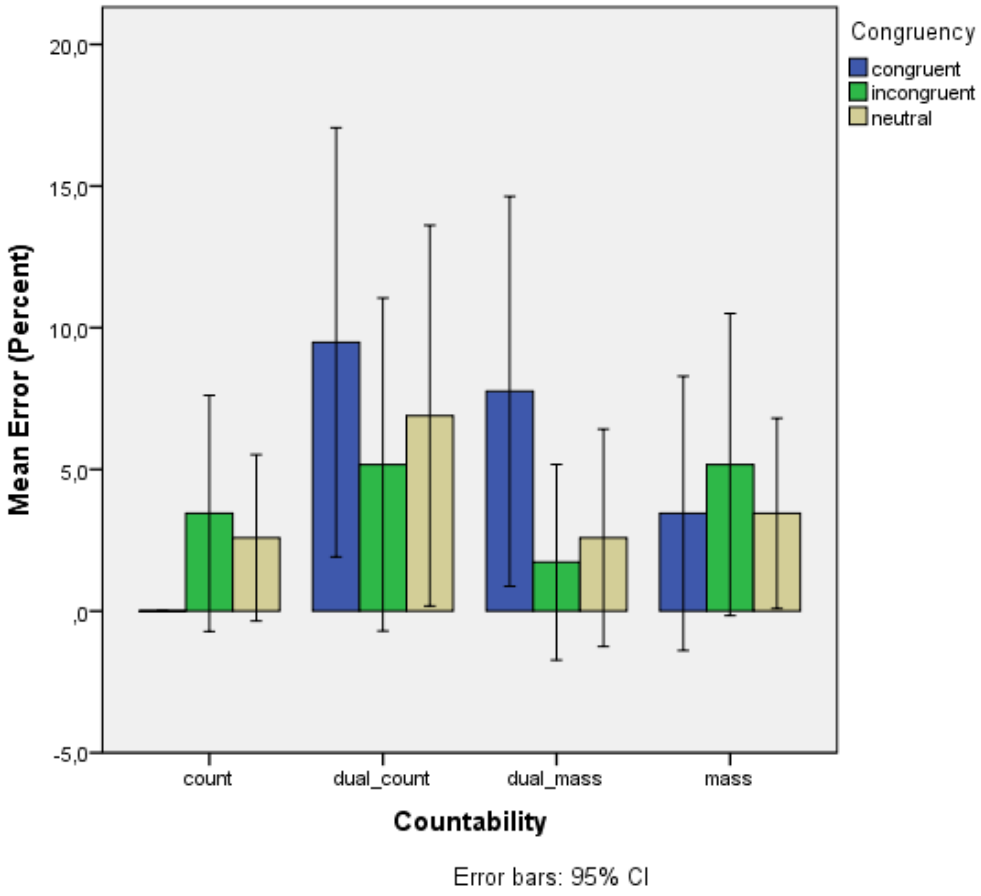


Figure 24: Error rates of real word items in percent by Congruency and Countability

Reaction times

All trials with erroneous responses were excluded (299, 7,57%) from further analysis. Further, trials with reaction times that deviated more than 2 standard deviations from the mean (subject's mean or item's mean) were considered outliers and excluded from analysis as well. The means were computed per Congruency condition. For trials of the congruent condition 36 trials were removed, for trials of the

incongruent condition 21 trials were removed and for trials of the neural condition 12 trials were removed as outliers.

For statistical analysis, the natural logarithm (base-e logarithm) of the reaction time was used. Reaction time was measured from noun onset on.

First, Similarity was entered into the model as a fixed factor with the levels real words, minimal pseudowords and maximal pseudowords. As random effects, intercepts for subjects and items were included, as well as by-subject and by-item random slopes for the effect of Similarity. Analysis revealed a highly significant effect of Similarity: $\chi^2(2)= 59{,}10$, $p<0.0001$.

Pairwise comparison with Bonferroni correction showed that the difference in reaction time between real words (M= 1166,06, SD= 341,47) and minimal pseudowords (M= 1351,23, SD= 378.38) was significant: t(133)=-6.87, p < 0.0001. Also, the difference in reaction time between words and maximal pseudowords (M= 1395,02, SD= 423,19) was highly significant: t(134)= -7,73, p < 0.0001. The difference between minimal pseudowords and maximal pseudowords did not reach significance, but as can be seen in Figure 25, the reaction times for maximal pseudowords were in fact longer than those for minimal pseudowords.

The Interaction between Similarity and Congruency was tested next. There was no significant main effect of Congruency and the interaction Similarity*Congruency was also not significant.

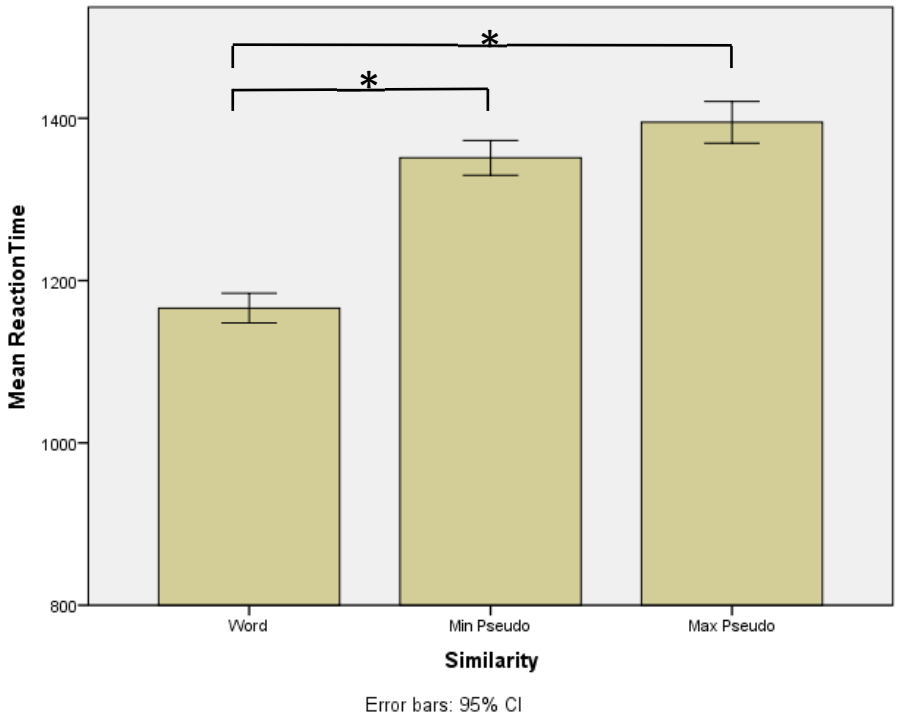


Figure 25: Mean Reaction time by Similarity

In a second analysis, phoneme monitoring data containing only trials including real words was used. Congruency and Countability were entered into the models as fixed factors. As random effects, intercepts for subjects and items were included, as well as by-subject and by-item random slopes for the effect of Congruency. Statistical analysis revealed a significant effect of Countability ($\chi^2(3)= 8{,}87$, $p<0.05$). There was no significant main effect of Congruency and the interaction Congruency*Countability was not significant either.

Pairwise comparison with Bonferroni correction revealed that only the difference in reaction time between the dual-count nouns (M= 1310,21, SD=405,03) and the dual-mass nouns (M= 1253,07, SD= 377,35) was marginally significant: t(50,3)=2,59, p = 0,075.

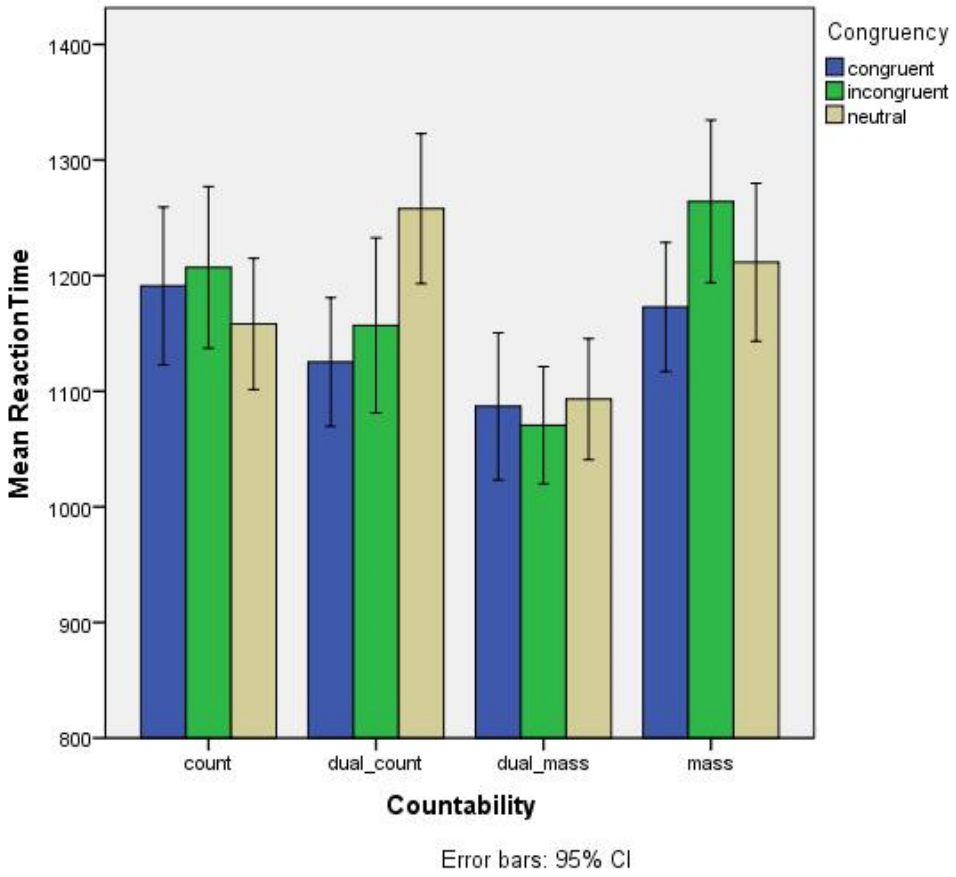


Figure 26: Mean reaction times for real word trials by Countability and Congruency

3.4.3. Discussion

The results of the phoneme monitoring task revealed two very important findings. First, there was a significant main effect of Similarity both in the error analysis and in the reaction time analysis that was not modulated by the element preceding the noun (no significant interaction with Congruency). This result not only replicates the findings of Connine et al. (1997), Bölte & Connine (2004) and Brenner (2016) and therefore adds to the evidence that the phoneme monitoring task in general is sensitive to effects of the lexical stage of phrase processing, it also shows that the lexical level was targeted in this specific case, i.e. in this experiment with these stimuli. It can thus be seen as proof of concept that the task was sensitive to the stage of lexical access. Had the stimulus words not been lexically accessed, there would have been no advantage for the real word items with this task and the reaction times and error rates to those stimuli would not be significantly different from those to the two classes of pseudowords. As outlined in the introductory text of this subsection (Subsection 3.4), the detection of the word-final target phoneme can occur prior to processing the complete word form in the case of real words because the word form can be accessed as soon as it is unmistakably identified as the input string perceived (uniqueness point). The accessed word form can then be searched for the target phoneme

and it is not necessary to wait and listen to the end of the auditorily perceived input. This advantage is not present for the pseudowords or at least there is another processing step involved, namely determining the original real word from which the pseudoword was generated which then can be lexically accessed. The fact that the minimal pseudowords yielded faster reaction times and less errors than the maximal pseudowords suggests that this was indeed the case. However, since this difference is not statistically significant no strong conclusions should be drawn from this result.

The second important finding is the absence of a significant congruency effect (also both in the error analysis and the reaction time analysis). With this result, there is strong evidence to place the influence of countability congruency not at the stage of lexical access but at a later stage of processing, namely integration. This finding is in line with the finding from Steinhauer et al. (2001) who found a significant difference in ERP responses to count nouns and mass nouns between 300 and 600ms post noun onset, the time window where integration processes are assumed to take place. Moreover, the result is further evidence for the claim made by Faussart, Jakubowitz & Costes (1999) that all grammatical agreement has a post-lexical locus.

Another less crucial but more baffling finding is that the dual nouns yielded such high error rates in the congruent condition (see Figure 24) which did not result from a speeded response to them, as evident by the reaction times (see Figure 26). The lexical decision experiment on these stimuli (Lexical Decision Experiment 2) also showed high error rates for dual-mass nouns in the congruent condition. Clearly, something about these noun phrase items disturbed the process of comprehension. As described in Subsection 3.4.1 “Methods”, all the dual nouns were found to occur more frequently with the respective count- or mass specific determiners in the corpus. However, closer inspection of the corpus results revealed that some of the dual-mass noun did not frequently occur with the determiner *viel, much*, specifically, but rather their classification was based on their co-occurrence with the other tested mass-specific determiners. This could have let to the often erroneous processing of noun phrases of the type “viel + mass-noun”. However, as mentioned in the section on the results of Lexical Decision Experiment 1 (Subsection 3.2.2), no direct evidence could be found that the specific co-occurrences with *viel* would be a better indication of ’mass-nounness’ than a value of a more abstract nature.

In conclusion, the main finding of the experiment is that there was no congruency effect or an interaction with the factor Congruency, while at the same time there was a significant effect of Similarity. This means that the locus of the previously observed congruency effect cannot be interpreted as reflecting an influence of countability on lexical access processes but must be seen as having a post-lexical locus.

4. General Discussion

This dissertation explores how the conceptual property of objecthood interacts with the linguistic phenomenon of noun countability and how countability as a noun property is utilised during the comprehension of German noun phrases.

The Introduction presented the different constructions German offers to linguistically realise the difference between bounded, countable entities, collections thereof and entities that are better quantified or measured in other ways. Syntactically countable and uncountable contexts were introduced which are used by means of determiner-noun combinations to signal whether a speaker intends to communicate the object-like qualities of an entity, such as the ability to move through space and time as a whole or the distinctiveness compared to other entities of its kind, or other qualities such as its nature or composition. It was shown that many entities can be referred to using both ways, depending on what aspect of the concept one chooses to focus on. Whether a group of animals is described as *waterbirds* or *waterfowl* depends less on properties of the animals themselves than on the communicative goals of the person describing them.

In fact, it was shown that even the most object-like entities can be found in situations where they are better described by uncountable, that is mass-specific language and even defy common-sense definitions of their physical nature as a bounded object. Where liquid cats are an extreme example of objects that can be denoted by mass language, they are by far not outliers of an otherwise straight-forward, rule-obeying system. There are also entities that – under the right circumstances – experience unforeseen objecthood or boundedness despite their apparent substance-like nature; countable risks and beverages that can be tied to one's belt or fall on the floor are only two of the examples presented in this thesis. Further, the existence of dual nouns as well as countability shifts must lead to the conclusion that there is not one unchangeable countability value per lexical item, but that the phrase environment the noun is presented in signals – in cases where the phrase is felicitous – whether reference is made to a bounded or an unbounded entity.

So, is it really the language that determines how an entity is perceived or which aspect of it is put into focus? Are speakers free to count or measure whatever entity they choose? Not completely, as counterexamples of unportionable milk and flexibility, just as oceans that cannot be parted, busses that cannot be measured and people that cannot be individuated, have convincingly exemplified the limitations that are posed on the freedom with which speakers can use linguistic tricks to bend entities and the nouns that denote them to their communicative intentions.

In the Background chapter, works are summarised that attempt to find the underlying property that decides which noun can be counted and which noun cannot. While some researchers have examined different physical properties of real-world objects, i.e. the referents of nouns, others have focused on the language-end of the equation and tried to identify properties of the noun constructions that enable a speaker to communicate the notion of boundedness.

Comparing the results of the different ways to investigate noun countability has led to the conclusion that the source of often seemingly contradictory findings in psychological and linguistic research is the

fact that adult humans most likely base their behaviour on clues from two cognitive systems that, too, offer occasionally contradictory information. These are the mid-level vision system that focusses on physical properties of entities or rather the perception and conceptualisation thereof on the one hand, and the kind-based system that is heavily influenced by a person's native language and that person's experience with a given noun on the other hand.

It was concluded, therefore, that in cases where the two systems offer converging evidence, noun countability is in accordance with the physical properties summarised under the term boundedness but in cases where the two systems offer a less clear result, countability within and across languages varies.

Additionally, this chapter offered insights into the notion of complex concepts by discussing phenomena of lexical items that denote multiple conceptual entities or notions thereof. These insights are crucial when discussing countability, especially for the discussions on dual nouns and countability shifts, since the ability to use language to focus on one or the other aspect of a concept demands a certain complexity. It was shown that complex concepts are far from uncommon and are a well suitable way to explain cases of polysemy and metonymy, as well as countability shifts.

Finally, research on the ways in which nouns are influenced by their syntactic environment during phrase processing was reviewed. This, too, is of the highest importance for the current discussion as one of the central questions this dissertation aims to answer is: How is noun countability utilised in noun phrase processing and how does it relate to other properties of nouns, such as gender, number or concept type. While there is disagreement about whether these properties should be seen as a core part of the noun's lexical entry or rather as an outside property applied to a noun in the theoretical literature, the empirical evidence seems to converge on an interpretation where nouns are lexically specified for gender, number and concept type but those properties are used in a comparatively late stage of phrase processing.

All of the results gathered in this chapter form the basis for the current considerations in that they provide an evidence-based framework of the syntax-semantic interface with regard to nouns both in the mental lexicon and during processing.

The experiments reported in the third chapter of this thesis aimed to contribute to our understanding of the issue of countability by focussing on the language-side of the issue. The experiments investigate the question whether countability is a lexical property of nouns – whether that property may stem from the perception of physical characteristics or not – and how this property is utilised during noun phrase comprehension.

In order to achieve a result as clear and objectively interpretable as possible, the independent variables of the experiments, namely the properties of the stimulus nouns, were meticulously measured in advance to the experiments. This was done to limit circularity in the argumentation of the research; a weakness of previous attempts to investigate noun countability which easily arises since countability touches so many linguistic and extra-linguistic levels. Here, an attempt was made to base classification of the stimuli for the experiments not on the researcher's intuition of noun countability, which may be based on syntactic, semantic or physical properties or a combination thereof, but to establish clear definitions for what counts as count, i.e. what makes a noun a count noun or a mass noun.

By means of distinguishing syntactic countability, analysed as determiner-noun co-occurrences and determined during the corpus study (Pretest 1, see Subsection 3.1.1.), from the flexibility that a noun experiences during interpretation, i.e. Shiftability as established by means of the sentence production

study (Pretest 2, see Subsection 3.1.2.), I tried to honour both sides of the coin that is language comprehension; the side that is the language and the side that is the comprehending mind.

After a solid base was established, two standard paradigms were employed to investigate how and when countability takes effect when a noun phrase is processed. The lexical decision experiments (Sections 3.2. and 3.3.) established countability as a lexical property of nouns and thus determined that there is in fact information regarding countability stored in the noun's lexical entry and it is therefore not possible to simply (that is without a more or less effortful reinterpretation) combine any noun with any syntactic context and bend its interpretation towards a bounded or unbounded entity. This was established by the observed congruency effect. Count nouns and mass nouns that are combined with a determiner that is not of the congruent class, i.e. count-specific determiners for count nouns and mass-specific determiners for mass nouns, are processed slower compared to nouns combined with a congruent determiner. Therefore, it must be that the lexical representation of a count noun or mass noun is specified for countability in the sense of the noun's ability to combine with certain determiners.

By contrasting the results of the Phoneme Monitoring Experiment (Section 3.4.) with those of the (second) lexical decision experiment it was shown that countability – like other lexical properties – is used during a stage of phrase integration that is subsequent to lexical access. As the congruency effect observed with the lexical decision task could not be found with phoneme monitoring, which is a task targeting only lexical access but not post-lexical stages of processing, countability cannot be interpreted as influencing processing stages that occur before the lexical selection of the noun is completed. This finding is not only crucial for researchers interested in countability but contributes to our understanding of language comprehension in general by adding to the body of research on the syntax-semantic interface both as part of stored lexical entries and during online comprehension.

In the final chapter of this dissertation, the main results of the literature review in combination with the results of the current experiments are discussed. Firstly, the results are evaluated in light of the question how noun countability is stored in the noun's lexical entry in the mental lexicon. Secondly, a proposal of the processes of noun phrase comprehension and the role noun countability plays during them is presented. Finally, the limitations of the current results are outlined and further research is proposed that would answer remaining questions with regard to the issue of noun countability and its importance during language comprehension.

4.1 What actually counts or Main findings

4.1.1 Countability in the mental lexicon

The main research question of this dissertation is whether or not countability is a lexical property of nouns. By means of the significant congruency effect found in Lexical Decision Experiments 1 and 2 it has been shown that lexical entries of nouns must be specified for (syntactic) countability, in the sense of combinatorial information on determiner-noun combinations. Countability thereby joins the list of

properties for which such an effect has been found and which have to be assumed to be properties that are part of the grammatical information stored for each noun in the mental lexicon, like gender, number (see Faussart, Jakubowitz, and Costes, 1999, among others) and concept type (Brenner, 2016).

It is most likely that the observed effects reflect a property of the lemma, that is the part of the lexical entry where grammatical information is stored. The lemma level is a well-established theoretical construct most famously discussed in Levelt, Roelofs & Meyer (1999). There is ample empirical support for its existence (see for example Roelofs and Ferreira, 2019, and references therein). Together with a conceptual level and a word form level, the lemma representation forms the lexical entry of a word.

It cannot be fully excluded that the observed effect of congruency reflects speeded processing because of the familiarity of the elements of the congruent noun phrases in contrast to the combination of the noun with an incongruent determiner or quantifier, which would be unfamiliar, or at least less familiar to the participants. That interpretation could place the congruency effect on the word form level as a reflection of subjective frequency of encounter of a specific noun with the indefinite determiner *ein* and *viel, much,* respectively. However, if that were the case the testing of the countability factor as reflecting the countability classes created here against the countability factor encoded as the actual statistic information of combinatorial frequency of the specific determiner with each noun should have revealed a better explanatory power by the model that reflected the actual frequency of co-occurrences of each noun with *ein* and *viel*. A familiarity effect would have predicted that how familiar a person is with hearing a noun with either *ein* or *viel* was in fact the decisive factor to explain which noun phrase could be processed easier (and thus faster). This, however, was not the result of the statistical analysis. Rather, there was no benefit of encoding countability as the actual co-occurrence frequency as compared to the classes based on co-occurrence frequency of all count-specific and mass-specific determiners. There is thus no direct indication of the congruency effect reflecting a familiarity advantage and hence no indication of the effect operating on the word form level. It therefore seems more plausible to analyse the effect as operating on the lemma level, where countability, along with e.g. gender and number, form grammatical information stored for the noun and utilised during the integration and grammaticality check of the noun phrase.

The congruency effect found in the current experiments is analysed as operating at the lemma level and not on the conceptual level mainly because there is no reason to assume that the tasks utilised (lexical decision and phoneme monitoring) target the conceptual level in any meaningful way. That is, while there have been semantic effects found with lexical decision experiments, those are most likely the result of spreading activation – an automated, unconscious process that does not reflect a level of cognitive regulation that would be assumed necessary for reinterpretation or shifts of reference, i.e. the processes involved in countability shifts during the comprehension of incongruent stimulus items (see Subsection 3.2.3. for a more detailed explanation and discussion of empirical evidence). While the processing of the incongruent phrases does not necessarily include a shift of reference (especially for the unshiftable nouns in Lexical Decision Experiment 1) a delay in reaction time caused by semantic factors in these cases would at least be reflective of an attempted shift. However, since there is no need to interpret the phrase in order to perform the lexical decision task accurately, it is unlikely that such an attempted shift operation is the (main) cause for the observed differences in reaction time towards congruent and incongruent noun phrase items.

Furthermore, no significant influence of any of the tested semantic factors could be found in Lexical Decision Experiment 1. Both Concreteness and Shiftability did not reach significant main effects or

interactions in either the error analysis or the reaction time analysis (see Subsection 3.2.2). There has indeed been no evidence found for a difference at the conceptual level between count nouns and mass nouns in the experiments reported here. The current evidence for countability stored as a lexical property in the mental representations of nouns is limited to a reflection of syntactic countability, that is the syntactic behaviour of nouns in the German language with regard to their frequency of co-occurrence with a given set of articles. Properties such as boundedness were not considered during the classifications of nouns. While they probably did play a role in establishing the factor Shiftability – in that participants of the sentence production study (Pretest 2, see Subsection 3.1.2) may have preferred to let bounded entities undergo the grinder shift as opposed to abstract entities – this factor was not found to have a significant effect on the processing of the noun phrases during the experimental tasks employed here.

That is not to say that such a difference on the conceptual level does not exist. If fact, intuitively, it is most conceivable that such a difference exists when dual noun reference and the interpretation of countability shifts are considered. Additionally, the results gathered with similarity judgement tasks and lexical projection task summarized in Section 2.2. "How we talk about things and stuff" provide a strong and extensive base of empirical results for the claim that linguistic countability does interact with the conceptual difference between individuated and non-individuated entities.

The currently employed behavioural experimental paradigms are just not sensitive enough to the conceptual level to detect such effects. To make the decision whether or not something is a real word or whether a certain phoneme is part of a given word form it is not necessary to interpret the perceived input, be it just the noun or indeed the entire noun phrase. The conceptual level does not need to be engaged during such a task.

However, even lexical (that is grammatical) countability of nouns has been in dispute, as some researchers suggested countability to be a property of noun phrases rather than nouns. Thus, the empirical proof of countability as a lexical feature of nouns is important for any further discussion of the issue.

So, if countability is a property stored in lexical entries, specifically the lemmata of noun entries, what can be said about the nature of this property? Interestingly, unlike gender and number, countability must be interpreted not as discrete categories or classes but as non-binary information, ranging from pure count-nouns via dual-nouns to pure mass-nouns. By means of the corpus study reported here ("Pretest 1", Subsection 3.1.1.) it was established that the frequency of co-occurrences observed for count- and mass-specific determiners and quantifiers was distributed over the entire range of possibilities; from 100% count-specific to 100% mass-specific with all intermediate distributions. Nouns that occur in both count-specific and mass-specific environments to a representative degree are thus to be categorised as dual nouns, where the same word form is found behaving like both a count noun and a mass noun.

In the experimental stimuli, the four classes of countability (count, dual-count, dual-mass and mass) were artificially created in that nouns were specifically selected to represent parts of the continuum. These discrete classes do not however reflect the reality of the German vocabulary and should thus be treated as an experimental manipulation not an observation of linguistic characteristics.

The question remains, whether countability is stored as an abstract category or as a form of statistic information about the likelihood or familiarity of a given determiner-noun combination. During the analysis of the data acquired in Lexical Decision Experiment 1, an attempt was made to answer that question by comparing a statistical model that encoded countability by means of the four countability classes used to classify the stimulus nouns to another model where countability was encoded reflecting the actual frequencies of occurrence of determiner-noun combinations as measured during the corpus study. Unfortunately, no clear result was obtained by this comparison as both models showed very similar explanatory power for the reaction time data.

What is clear, however, is that pure count nouns and mass nouns (i.e. those nouns that never occur in the respective incongruent syntactic environment) must be lexically specified as such. Otherwise, there would have been no detectable additional effort during the comprehension of items of the incongruent condition in the lexical decision experiments. Those nouns must somehow be licenced for combination with one set of determiners (or no determiner in the case of mass nouns) but not licenced for combination with the other.

The result for dual nouns presents an interesting challenge when it comes to interpretation. Those nouns are licenced for combination with both determiner sets, be it more so for one than the other, as evident from the way the reaction time data showed the same pattern as for the respective classes of pure count nouns and mass nouns, just less so. What posits a challenge for interpretation is the observation that a given noun is interpreted denoting a bounded entity in countable contexts and an unbounded entity in mass context. How should a lexical entry for such a noun be envisioned, considering that we are dealing with a single word form but assume that lemmata are specified for countability, and must be so in order to enable linking to the appropriate conceptual aspect?

To answer this question, I want to consider phenomena which are similarly challenging for theoretical accounts on lexical entries concerning other noun properties, namely number and gender. There are – as pointed out in the Background chapter (Subsection 2.5.2. “Inside or outside the noun – empirical evidence”) – German nouns where the singular and the plural word forms are homophones, such as *Reiter, rider*. In these cases, like in the case of dual nouns, only the syntactic environment signals unambiguously whether a singular or a plural entity is referenced, e.g. by means of articles or noun-verb-agreement. Furthermore, there are German nouns for which speakers do not uniformly agree on the appropriate grammatical gender that should be assigned to the noun. This is mostly the case for neologisms, loan words and proper names, such as *Nutella*, but there are also homophones that exhibit different genders and different referents, such as Tau_{MASC}, *dew*, and Tau_{NEUT}, *rope*. While the first phenomenon is often a case of regional (that is dialectal) or individual differences, the second can be seen as systematic in the German language; a way to lessen processing effort during the establishment of correct reference for homophones. That is, words like *Nutella* usually get assigned one gender per speaker and the gender just differs across speakers, whereas both versions of *Tau* exist simultaneously inside the mind of one and the same speaker and are usually considered ‘two different words’ (meaning true homophones, as opposed to e.g. polysemes). $Reiter_{SG}$ and $Reiter_{PL}$, however, share a common reference with the exception of the amount of people the noun refers to – arguably a less severe difference conceptually than that between dew and rope.

How does that compare to dual nouns like *Tuch, cloth/fabric*, or *Pelz, pelt/fur,* where the count-noun version refers to a discrete object and the mass-noun version to a material? It seems obvious that the two versions of each of these nouns are not mere homophones like the two versions of *Tau*. It is also

not the case that each speaker only accepts one version of syntactic marking (combination with count- or mass-specific determiners) for the noun, as for *Nutella*. Rather, it seems like – similarly to the case of *Reiter* – a specific conceptual property changes for the reference of the noun depending on the syntactic environment. Contrary to the case of *Reiter*, though, most German nouns are marked for number on the noun word form and nouns that lack that marking are the exception. Countability on the other hand is never marked on the noun itself in German (unless one considers plural markers, which I however see as an indication of countability not a direct grammatical marking of it).

Whether – as a psycholinguist – one concludes from this that dual noun versions should be represented by two distinct lexical entries depends on one's view on the form of lexical entries. The framework underlying the current position outlined in this dissertation is a frame formalisation of the mental lexicon inspired by the popular and widely-applied model of language by Roelofs & Ferreira (2019) (which heavily builds on the model by Levelt, Roelofs & Meyer, 1999). Outlined in more detail in Beckmann, Indefrey & Petersen (2018), the current approach assumes three distinct levels for each lexical entry: a conceptual frame, a lemma frame and a word form frame. The lemma frame is the one specified for countability in that here the information about combinatorial restrictions with articles is stored. Consequently, dual nouns would have to be modelled as having two lemma frames (one specified for count, one specified for mass) connecting to the same word form frame. The lemma specified for count connects to a part of the conceptual frame that represents an individuated object, while the lemma specified for mass connects to a conceptual node representing an unbounded material. An example for the dual-mass noun *Pelz, pelt/fur,* is shown in Figure 27.

All (sub-)frames of the three levels together form the lexical entry of the dual noun. This is made possible by the inherent recursiveness of frames. A frame is a structure of attributes and values and each value-node in turn can form a frame on its own. Therefore, the question of a dual noun being 'one word' or 'two words' is really not applicable, or rather depends on the definition of 'word' used.

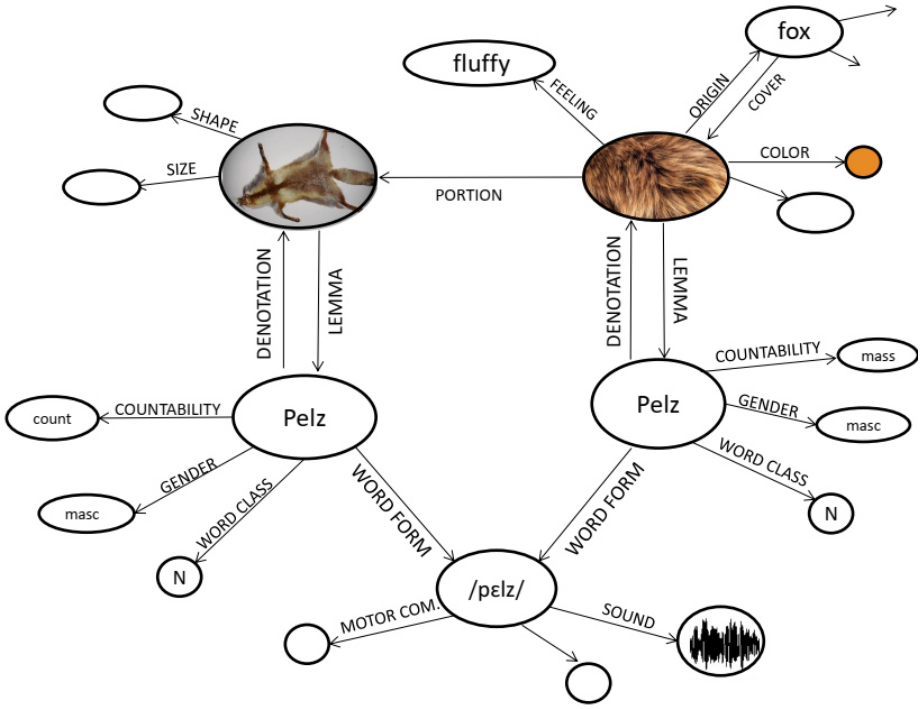


Figure 27: Lexical entry frame of the dual noun *Pelz, pelt/fur*
Note: Arrows represent attributes, round nodes represent values, empty attributes and values represent larger complexity of the frame representation in the mind than can be depicted here
Abbreviations: DET = determiner, N= noun, masc = masculine, motor com = motor command

During the first discussion of frame representation in this dissertation (Subsection 2.4.2) as well as in Beckmann, Indefrey & Petersen (2018) it was mentioned that reference shifts can only occur with conceptual value nodes that are connected via a bi-directional link. As Figure 27 shows, this is not the case for the conceptual nodes representing the bounded entity (connected to the lemma specified for count) and the conceptual node representing the unbounded entity (connected to the lemma specified for mass). This is because we are not dealing with a shift of reference here, as in the case of a countability shift (such as *ein Wasser, a water*, in Subsection 2.4.2.), but with two distinct lemmata, each connecting to a different conceptual node. Here, there is a conceptual node "fox" that is bi-directionally linked to the node representing the unbounded entity of the Pelz-concept. This fox-node on the conceptual level could be the reference node of a fox-frame, there connected to the lemma and word form frames for the noun *fox*. For that frame one could assume an ability for a metonymic shift of reference similar to the case of countability shifts so that the fox-lemma and fox-word form connect to the conceptual aspect of the unbounded Pelz-node to enable interpretation of sentences like "Her coat is made of fox." where *fox* does not refer to the animal in its entirety, but only to its fur which the node for the unbounded entity represents.[46]

The assumption of two distinct lemma-nodes for dual nouns is further supported by considerations about the processes involved in phrase comprehension to which I will turn in the next subsection.

In sum, the finding of a significant congruency effect (in both lexical decision experiments) clearly speaks for the hypothesis that countability is stored in the lemma of the lexical entry of both noun and article. As countability was determined by the frequency of co-occurrence of articles and

[46] Note that the described reinterpretation also works the other way around. There are cases in German where a compound noun including *Pelz, fur,* is used to denote the whole (living) animal, such as *Graupelz, graypelt,* for wolf.

nouns and the significant difference was found based on this classification it must be that a representation of this classification is stored and also utilised at some point during processing. The nature of this utilisation is discussed in the next subsection.

4.1.2. Countability during phrase processing

At which stage of processing the countability value information is utilised is made evident by the absence of the congruency effect in the Phoneme Monitoring Experiment. As discussed at length in Section 3.4., it is generally accepted that an effect that occurs with lexical decision but not with phoneme monitoring has a post-lexical locus because the phoneme monitoring task – unlike lexical decision – is not sensitive to processes of post-lexical integration and well-formedness checking mechanisms. A lack of an effect during a phoneme monitoring task places the influence of the factor in question at that stage rather than at the stage of lexical access, which both tasks are sensitive to.

I assume the processing of noun phrases to occur in accordance with the considerations about lexical entries outlined above. The article is processed, that is at least the word form frame and the lemma frame (probably also the conceptual frame, via spreading activation) of its lexical entry is lexically selected. When the noun entry has been selected, an attempt is made to integrate the two elements into a noun phrase. Friederici & Jacobsen (1999), drawing on the model by Levelt, Roelofs & Meyer (1999) described above, suggest that phrase integration includes a grammatical checking mechanism that activates an expected grammatical property (gender, in their case) for the noun based on information gained from the previously processed article. This pre-activation of the expected value does not limit search space of lexical candidates, but is upheld as a prediction until the relevant information of the noun is actually perceived. Then – that is after lexical access of the noun – the expected and the actual value are compared. Friederici & Jacobsen (1999) highlight the postlexical locus of this process. They argue that problems during integration of this sort (i.e. wrong grammatical gender agreement) play a role during the second stage of processing described in Friederici's (1995) model of comprehension; a phase where lexical information of a retrieved element is analysed. That element definitively includes grammatical information "such as number, tense or mode of a bound morpheme" (p. 477) and can thus be equated with the lemma[47]. Hindered processing like in the case of a mismatch between predicted information from the article and actually perceived information of the noun is said to be reflected by ERP effects in the 300ms-600ms time window. Note that while all these considerations were made to discuss grammatical gender and not countability, I believe them to be applicable in the current discussion. In the previous subsection, I have laid out arguments why countability (at least in the context of these experimental results) should be treated as a grammatical property stored within the lemma. Further, the ERP result on countability reported by Steinhauer et al. (2001) described in Subsection 2.5.2 was found in the same time window, pointing towards a similar temporal locus of the effect as the one discussed by Friederici must have.

[47] Friederici (1995) also argues for simultaneous activation of semantic information which in the present account would not be stored in the lemma but rather in the conceptual representation or by means of the connection between the two levels, that is the DENOTATION-attribute. However, under Friederici's account, these properties are also stored separately and interact at post-lexical stages of comprehension.

Applying this view on the processes of phrase comprehension, processing of the current experimental stimuli may have proceeded as follows: The lemma-node of the article is activated after the article is lexically accessed. Arguably, the lemma-node of the article is specified for countability, since all articles and quantifiers regarded here (see Subsection 3.1.1. for a list) combine only with either count nouns or mass nouns (or their lemmata, to include dual nouns). A countability value as part of the article's lemma would be a way to encode the combinatorial restrictions between articles and nouns and enable the process responsible for phrase integration to make its prediction of the upcoming noun's countability value. The model by Gauvin and Hartsuiker (2020) describes predictive processes during comprehension and though their example concerns semantic cloze probability, similar mechanisms could apply to predictions based on grammatical cues and thus the processes envisioned here.

> "During listening, the perceived words become active first in the perception system, and via the interconnections between modalities the representations in the production system also become active. Based on past experience, the production system activates related nodes, which in turn become active in the perception system, thereby creating a prediction of the upcoming percept. […] [T]hereby preparing the perception system for these items, thus creating expectancies. When the predicted utterance is met, the active nodes in the perception system increase in activation, until after an amount of time the word is selected for comprehension. When perceived words match the predicted percept, speech perception thus becomes a low-effort process, as the perception of that word is prepared." (Gauvin & Hartsuiker, 2020, p. 17)

When the noun has been lexically accessed, its lemma frame is activated via the word form frame. Then, during phrase integration, the predicted, pre-activated countability value is compared with the actual value that is part of the noun's lemma frame. In the case of congruent count and mass noun trials, the prediction based on the article's countability value is verified by the perceived input and a grammatical noun phrase is formed. Contrary, in the case of incongruent trials, the predicted countability value is different from the one encoded in the noun's lemma frame and phrase integration is hindered. Note that, as pointed out by Gauvin and Hartsuiker, no special machinery or module is required for the detection of a mismatch between predicted and perceived input. Rather the checking mechanism "exploits information that is computed during language production and perception" (p. 17) i.e. the activation of the given nodes in the specific layers of representation. If one imagines the countability value node in the lemma-frame as being able to receive activation on its own, one could see the hindrance in integration as a conflict arising from competition between two countability value nodes; one receiving activation because it was predicted due to the determiner, the other receiving activation via the selected noun. While the cited passage of Gauvin and Hartsuiker's work suggests that the predicted value acts upon processes during lexical selection ("When the predicted utterance is met, the active nodes in the perception system increase in activation, until after an amount of time the word is selected for comprehension." Gauvin & Hartsuiker, 2020, p. 17), this is not assumed here as an influence of countability congruency on processing stages of lexical selection has been disproven empirically by the current experiments. Rather, a comparison between the predicted and the perceived countability value of the noun is part of the phrase integration process.

It is only after the described processes have occurred that the lexical decision required by the experimental task can be started. If that were not the case, there would have been no effect of

congruency found during the first and second lexical decision experiment. Bölte & Connine (2004) state that the delay in lexical decision of incongruent trials could be due to a lack of familiarity of the incongruent phrases. They cite the work of Balota and Chumbley (1984, 1990) who developed a model for single word lexical decision that assumes a familiarity/meaningfulness dimension. Where a given stimulus is found on that dimension determines how fast a lexical decision on it can be performed, with (frequent) real words having a very high familiarity/meaningfulness value and are thus quickly identified as a word, and nonwords having a very low value (as the subjects have never encountered anything like them before) and are thus also responded to rather quickly. Pseudowords on the other hand, require more extensive analysis, as it takes longer to identify them as not being real words. Bölte & Connine (2004) describe a comparative account for congruent and incongruent noun phrase (here, again, discussing gender congruency). They argue that because incongruent phrases are less familiar (and less meaningful) to subjects, lexical decision of the noun is delayed. How exactly the hindrance of the integration process by the failure of the well-formedness check leads to an influence on the lexical decision for the noun is not explained in detail, either in Bölte & Connine (2004) or in Friederici & Jacobsen (1999). Conflict-resolution processes described in Gauvin & Hartsuiker (2020) unfortunately concentrate on erroneous lexical selection and resolution thereof and do thus not apply to the processes discussed here, namely integration and grammatical well-formedness in phrases.

As discussed earlier, familiarity of a specific determiner-noun combination should have resulted in an advantage of that particular combination (e.g. *ein Bus, a bus*) as compared to another (*viel Bus, much bus*). However, since in the lexical decision experiments countability classes were used that were based on a noun's co-occurrence with multiple determiners (e.g. *ein/zwei/einige/viele Bus(se), a/two/several/many bus(ses)*), this explanation is not the preferred interpretation of the current results. Rather, I suggest that the detection of two countability value nodes with similar activation level (i.e. incongruency) initiates a checking mechanism that starts an attempt of reanalysis, trying to resolve the conflict and establishing the correct countability value. The first countability value receives activation because it is predicted for the noun based on the information received from the determiner, that is if one hears for example the indefinite determiner, one predicts the upcoming noun to be marked for count. The other countability value receives activation because it is the countability value the perceived noun is actually marked for, in case of incongruent trials this would be mass. This conflict and the attempt made to resolve it put extra demands on the processing operator. Since the lexical decision draws on the same operator, namely analysing the noun's lexical entry, decision time is longer when that operator is engaged in additional processing.

For dual nouns, however, the conflict caused by incongruency does not arise and therefore no delay is predicted. In cases where a lemma with the correct countability value (i.e. the one fitting the value of the article) can be found connected to the word form, be it one out of two, integration and lexical decision can both proceed without hindrance. This is exactly what we find in the data of Lexical Decision Experiment 1: The congruency effect was found to be modulated by where a certain noun was found on the scale of countability, with the congruency effect being significant for pure count nouns and mass nouns, but not for dual-count and dual-mass nouns, which combine with both sets of articles to a relevant degree. The slight (non-significant) difference between congruent and incongruent trials with dual nouns found in Lexical Decision Experiment 1 can be attributed to the competition between the two lemmata of the dual noun entry. Considering Gauvin & Hartsuiker's (2020) proposal of competition for the node with the highest activation, it could be the case that during incongruent dual noun trials (e.g. *viel Tuch, much cloth/fabric*), the count-specific lemma and the mass-specific lemma of the dual noun experience similar activation; one due to its higher frequency ($Tuch_{\text{COUNT}}$ is more frequent as *Tuch* is a dual-count noun) and the other due to expectations formed because of cues from the determiner

(*viel, much*, leads to the expectation of a mass noun). This momentary conflict may delay the response to the lexical decision enough to affect the reaction times, even if not significantly so.

It has been suggested, that facilitation effects point towards influences during the build-up of a noun phrase, whereas inhibition effects are better interpreted as reflecting an effect locus during a well-formedness check (see Friederici & Jacobsen, 1999, Bölte & Connine, 2004, and Brenner, 2016). This suggestion in Bölte & Connine (2004) is again based on works by Balota & Chumbley (1984, 1990) regarding word frequency effects. The authors state that – were a similar logic applied to gender effects – a preceding article congruent in gender would increase "the initial familiarity of a word and subsequently facilitate [lexical decision] responses" whereas neutral or incongruent articles would "provide no such increase in familiarity" (Bölte & Connine, 2004, p. 1028). This familiarity assessment is thought of as occurring "independently of the time course of word recognition" and Balota and Chumbley assume lexical frequency to be "an information source that is utilised after a word has been recognised" (Bölte & Connine, 2004, p. 1027).

As discussed earlier, familiarity-related explanations of the congruency effect are not favoured for the current experimental results. Further, the explanation of how the familiarity of a determiner-noun combination influences two different post-lexical operations differently given in Bölte & Connine (2004) is rather difficult to grasp. In Brenner (2016) no more detailed descriptions of the integration and the checking mechanisms after lexical selection could be found that would help to retrace the argumentation that facilitatory effects as compared to inhibitory effects are a sign for problems of integration rather than well-formedness.

Friederici & Jacobsen (1999) also state that inhibitory effects should be interpreted to reflect processes during a well-formedness check, since the check of a condition without prime information (i.e. a neutral condition) should require the same amount of time as the check of a congruent condition, but a detection of incongruency would require more time. Here, however, the well-formedness mechanism is not contrasted to a post-lexical integration process, but to an initial phrase-building process that takes place before lexical selection is completed. Therefore, it is unclear whether the authors would analyse an inhibitory effect as signalling an influence of a post-lexical well-formedness check as opposed to a post-lexical integration process, or merely as signalling an influence of a post-lexical well-formedness check as opposed to a pre-lexical processing stage.

Compared to the neutral condition, the reaction time patterns of the congruent and incongruent trials did not yield a clear picture that would allow a differentiation between facilitation and inhibition of the current effect with absolute certainty. In Lexical Decision Experiment 2, the neutral condition yielded reaction times that were similar to those yielded by the incongruent condition but significantly higher than those of the congruent condition. Numerically, this pattern was observable for all the countability classes (except for pure count nouns, where the incongruent condition yielded the highest reaction times). Thus, here the results point towards a facilitatory effect of congruency on processing (see Subsection 3.3.2). However, in Lexical Decision Experiment 1, trials of the neutral condition yielded reaction times that were significantly higher than those of trials both of the congruent and incongruent condition for count nouns and dual-count nouns, but for dual-mass nouns and mass nouns reaction times from the neutral condition did not differ significantly from those of either the congruent or the incongruent condition.

Because of the interaction between Congruency and Countability Class in Lexical Decision Experiment 1 the current results have to be interpreted with caution with regard to this issue. Additionally, since it is not completely understood why and how the effect should be placed on one specific post-lexical operation, it seems wise to leave the issue unresolved for now.

In conclusion, the influence countability information has on the processing of nouns as heads of determiner-noun NPs can be pinpointed to post-lexical stages with some certainty. Specifically, an analysis is suggested where a prediction regarding the countability value of the noun is formed based on the countability value of the previously processed article, which is compared to the actual value of the noun after it is lexically accessed. This comparison may involve the comparison of activation levels of two countability values in the case of incongruency.

4.2 What's left to count or Limitations of the current research

The studies reported in this dissertation are not able to shed light on the processing of any conceptual difference between count noun and mass noun referents. There was no significant effect of semantic factors in the first lexical decision experiment, neither of Concreteness which would have reflected an influence of physical properties of the noun's referent, nor of Shiftability which would have suggested a difference between successful versus unsuccessful (re-)interpretation processes. As pointed out earlier, the task utilised here does not target the conceptual level during phrase processing in any meaningful way. Reanalysis processes as occurring during countability shifts probably take place after the stage of phrase integration described above, thus semantic information may not be utilised before the lexical decision is made.

The model by Friederici (1995) assumes that during the critical time window of 350ms-600ms semantic information is activated alongside information which I have associated with the lemma frame here. The well-known N400 ERP component of semantic integration problems (first found by Kutas & Hillyard, 1980) supports that assumption. So, if countability is processed in the same time window, why were there no semantic effects found alongside the syntactic effects?

As pointed out earlier, this could be due to the specific requirements of the lexical decision task. With this task, the only semantic effects found are suggested to be the result of automatic spreading activation (Bates et al., 1996) while actual shift or reanalysis operations require more complex, controlled processing. Before these processes can be initiated, processing is aborted because the lexical decision has been made and its response is initiated.

Note that a slight, but non-significant difference between the error rates for trials with shiftable and unshiftable nouns in the incongruent condition has been observed in Lexical Decision Experiment 1, with the unshiftable nouns in incongruent phrases eliciting more errors than shiftable nouns in incongruent phrases. While this result should be interpreted tentatively because of the lack of significance, it could represent a reflection of spreading activation towards the conceptual level indicating that a reference shift is possible in the case of shiftable nouns. However, it could also represent properties of a syntactic nature. It's conceivable that because a shift is possible, the participants tested in the experiment have

indeed encountered the noun in the incongruent phrase context before and are thus not as completely unfamiliar with the construction as with an unshiftable noun in an incongruent phrase. Consequently, one had to assume the presence of a second lemma specified for the other (shift-congruent) countability class, similar to dual nouns. This second lemma could receive less activation due to its infrequency and thus influencing the lexical decision only slightly (that is insignificantly, in the same way as for the dual nouns, but even less so). A corpus, such as the one on the basis of which the noun stimuli were classified as pure count nouns and pure mass nouns, can never adequately represent the entirety of a language, let alone accurately predict the language experience of specific participants. It only ever gives an approximation of frequent constructions at a specific time.

Considering all the arguments against an influence of conceptual features during the currently employed tasks, it seems better to attribute this result to the imperfections in categorising the stimulus nouns than to assume an influence of conceptual features that was somehow limited to this specific difference in error rates. Interpreting this difference at all is dangerous, since it very well could reflect a random property of the collected data with no meaningful origin with regard to the current research questions. It seems prudent, therefore, to state the speculative nature of this line of thought explicitly and leave the burden of answering the question of a semantic influence on the comprehension of count language to further research.

The current results provide no insights about how a shift or reinterpretation is performed by the language processing mind. It's possible to speculate, however, that the here observed detection of countability incongruency is what triggers a reinterpretation attempt in the first place. This seems logical since without the detection of a mismatch there would be no reason to start additional cognitive processes during noun phrase interpretation.

A theoretical approach to how countability shifts may take place is given in Beckmann, Indefrey & Petersen (2018). That approach, however, has not been empirically tested to this day. In the case that the hypothesised ungrammaticality detection during phrase integration or by the grammaticality well-formedness check would set in motion a reinterpretation process, a shift in reference could proceed as described in this earlier work and in Subsection 2.4.2 here, namely by a reference shift from the default value of the lemma's DENOTATION-attribute to another node in the same conceptual frame.

More research with a different experimental paradigm is needed to test this theory. I propose that if the stimuli from the first lexical decision experiment, that included the factor Shiftability, were to be tested in an EEG study, the resulting ERP components yielded by shiftable and non-shiftable nouns in the incongruent condition would shed light on the reinterpretation process. Specifically, because for shiftable nouns a reinterpretation would take place and that process would be reflected in the EEG signal, possibly yielding a similar response as cases where a less-frequent polyseme must be selected instead of the previously selected high-frequent version (eg. *Brazil* interpreted as the government of the country rather than the country itself) or cases where interpretation of a noun must change from literal meaning to metaphorical meaning. For non-shiftable nouns, however, this reinterpretation process would not be possible and the ERP response to those nouns in incongruent environment should resemble one yielded by grammatical errors, such as number violations or gender incongruency, namely a P600 component.

4.3 Conclusion

The research reported here answers the questions whether countability is a lexical property of nouns and when it is utilised during the grammatical processing of noun phrases. It has been shown that the mental representation of the noun in the mental lexicon must include information about the degree to which nouns combine with specific articles, i.e. syntactic countability, be it as a statistic information of co-occurrences or a categorical feature. Further, it has been shown that the locus of the influence of this property must be at a post-lexical stage of processing rather than during the lexical access of the noun. These results are in line with results regarding other noun properties, especially gender and concept type.

Additionally, a formal approach has been proposed that pinpoints countability to the lemma level of a lexical entry and describes how it could be utilised during the post-lexical processes of phrase processing.

While the literature reviewed in this thesis clearly proofs the existence of a conceptual difference between bounded and unbounded entities in the mind, it could not be determined how exactly this property is mapped onto (grammatical) properties of nouns.

It was however theorised that many concepts denoted by nouns are actually complex webs of conceptual aspects with rather fuzzy boundaries. Denotation, that is the connection between a noun and a conceptual aspect, is strongly influenced by the context the noun appears in and the precise denotation target (or value, in frame terminology) varies for many nouns, not just classic examples of polysemy.

Dual nouns have been analysed as having two lemma-frames as part of their lexical entry, one denoting the bounded conceptual aspect of a concept-frame and the other denoting the unbounded aspect. Similarly, countability shifts are analysed as a shift of the denotation from one conceptual aspect to the other. In both cases, the switch or shift, respectively, is triggered by the syntactic environment of the noun, that is the (absence of a) determiner it combines with in the specific instance.

Further research involving the recording of ERPs during phrase processing was proposed to test some of the aspects of these theoretical considerations.

In sum, it can be said that the current research shows that it is impossible to narrow down a phenomenon as complex as countability at the syntax-semantic interface to one specific feature of either language or the things it describes.

The German language, like any language, is not a stable system of unchangeable rules and invariable parts. Language – as any linguist knows – is something that happens rather than something that has been invented once and has existed unchanged and unaffected ever since. Because language is used by people and people differ from each other, change over time and develop new ways of perceiving the world around them, language as a means of describing that perception, too, is and must be an ever-changing and highly flexible tool. Neither properties of the world, the human perception of it or

the humans themselves on its own is sufficient to explain all the phenomena of language, even if one focusses on a single topic such as countability. The investigation of this complex interplay between the world, the mind and a specific language is what makes psycholinguistics such an amazingly interesting subject of study. Only by taking into account all the pieces of the puzzle and not losing sight of the interdependencies and reciprocal effects is it possible to find satisfactory answers.

Regarding the current topic, noun countability showed itself to touch many mental capacities; from learning to perceive multitudes and storing number words, over making out groups and their relevance compared to an individual member, to combining grammatical elements skilfully enough to change fine-grained conceptual aspects in the mind of our speech's recipient. The current investigation showed that all these aspects of the topic need to be considered in order to explain otherwise seemingly contradictory observations and that much work is still needed to explore all the processes involved in sufficient detail. The work described in this dissertation presents one piece of the puzzle that is noun countability. Hopefully, future work will uncover the remaining pieces so that one day we may know for sure what actually counts.

References

Alkuhlani, S., & Habash, N. (2011). A corpus for modeling morpho-syntactic agreement in Arabic: gender, number and rationality. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 357-362).

Allan, K. (1980). Nouns and countability. *Language*, 541-567.

Bale, A. C., & Barner, D. (2009). The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics*, *26*(3), 217-252.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human perception and performance*, *10*(3).

Balota, D. A., & Chumbley, J. I. (1990). Where are the effects of frequency in visual word recognition tasks? Right where we said they werep Comment on Monsell, Doyle, and Haggard (1989).

Barber, H., & Carreiras, M. (2003). Integrating gender and number information in Spanish word pairs: An ERP study. *Cortex*, *39*(3), 465-482.

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience*, *17*(1), 137-153.

Barner, D. & Snedeker, J. (2005). Quantity judgments and individuation: Evidence that mass nouns count. *Cognition* 97(1), 41-66.

Barner, D., & Snedeker, J. (2006). Children's early understanding of mass-count syntax: individuation, lexical content, and the number asymmetry hypothesis. *Language Learning and Development*, *2*(3), 163-194.

Barner, D., Thalwitz, D., Wood, J., Yang, S. J., & Carey, S. (2007). On the relation between the acquisition of singular–plural morpho-syntax and the conceptual distinction between one and more than one. *Developmental Science*, *10*(3), 365-373.

Barner, D., Wagner, L., & Snedeker, J. (2008). Events and the ontology of individuals: Verbs as a source of individuating mass and count nouns. *Cognition*, *106*(2), 805-832.

Barner, D., Wood, J., Hauser, M., & Carey, S. (2008). Evidence for a non-linguistic distinction between singular and plural sets in rhesus monkeys. *Cognition*, *107*(2), 603-622.

Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Perception & Psychophysics*, *58*(7), 992-1004.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Beckmann, N. S., Indefrey, P., & Petersen, W. (2018). Words count, but thoughts shift: A frame-based account to conceptual shifts in noun countability. *Voprosy Kognitivnoy Lingvistiki (Issues of Cognitive Linguistics)*, *2*, 79-89.

Bisiacchi, P., Mondini, S., Angrilli, A., Marinelli, K., & Semenza, C. (2005). Mass and count nouns show distinct EEG cortical processes during an explicit semantic task. *Brain and Language*, *95*(1), 98-99.

Bock, K., & Levelt, W. J. (1994). *Language production: Grammatical encoding* (pp. 945-984). Academic Press.

Bölte, J., & Connine, C. M. (2004). Grammatical gender in spoken word recognition in German. *Perception & Psychophysics*, *66*(6), 1018-1032.

Brenner, D. (2016). *Why his Mother is Better Than a Mother. Psycholinguistic Investigation of Concept Types & Concept Type Shifts* (Doctoral dissertation).

Brown, C. H. (1999). *Lexical acculturation in Native American languages.* Oxford University Press.

Carey, S. (2009). Where our number concepts come from. *The Journal of Philosophy*, *106*(4), 220.

Carey, S., & Xu, F. (2001). Infants' knowledge of objects: Beyond object files and object tracking. *Cognition*, *80*(1-2), 179-213.

Champollion, L. (2017). *Parts of a whole: Distributivity as a bridge between aspect and measurement* (Vol. 66). Oxford University Press.

Cheng, L. L. S. (2012). Counting and classifiers. *Count and mass across Languages*, 199-219.

Chierchia, G. (1998b). Reference to kinds across language. *Natural Language Semantics*, *6*(4), 339-405.

Chierchia, G. (2010). Mass nouns, vagueness and semantic variation. *Synthese* 174(1), 99-149.

Chierchia, G. (2015). How universal is the mass/count distinction? Three grammars of counting. *Chinese syntax: A cross-linguistic perspective*, 147-177.

Chomsky, N. (1996). *Powers and Prospects Reflections on Human Nature and the Social Order*. Pluto Press.

Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge University Press.

Cole, P., & Segui, J. (1994). Grammatical incongruency and vocabulary types. *Memory & Cognition*, *22*(4), 387-394.

Connine, C. M., Titone, D., Deelman, T., & Blasko, D. (1997). Similarity mapping in spoken word recognition. *Journal of memory and language*, *37*(4), 463-480.

Constantinescu, C. (2011). *Gradability in the nominal domain*. Netherlands Graduate School of Linguistics.

Copestake, A., & Briscoe, T. (1995). Semi-productive polysemy and sense extension. *Journal of semantics*, *12*(1), 15-67.

Corbett, G. (2000). *Number*. Cambridge University Press.

De Lima, S. O. (2015). Quantity judgments in bilingual speakers (Yudja/Brazilian Portuguese). *Letras de Hoje*, *50*(1), 84-90.

De Long, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature neuroscience*, *8*(8), 1117-1121.

De Vincenzi, M. (1999). Differences between the morphology of gender and number: Evidence from establishing coreferences. *Journal of Psycholinguistic Research*, *28*(5), 537-553.

De Vincenzi, M., & Di Domenico, E. (1999). A distinction among features: The role of gender and number in the retrieval of pronoun antecendents. *Italian Journal of Linguistics*, *11*(1), 41-74.

Del Pinal, G. (2018). Meaning, modulation, and context: a multidimensional semantics for truth-conditional pragmatics. *Linguistics and Philosophy*, *41*(2), 165-207.

Di Pace, E., Marangolo, P., Pizzamiglio, L., & Burr, D. (1994). Colour-priming does not depend on spatial coincidence or conscious encoding strategies. *Manuscript submitted for publication*.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, *67*(3), 547-619.

Dromi, E. (1987). Early lexical development. London. Cambridge University Press.

DWDS, Digitales Wörterbuch der deutschen Sprache, https://www.dwds.de/, Text Corpus of Berliner Zeitung, 1994 – 2005. Data retrieved between March and September 2016.

El Yagoubi, R., Mondini, S., Bisiacchi, P., Chiarelli, V., Angrilli, A., & Semenza, C. (2006). The electrophysiological basis of mass and count nouns processing. *Brain and Language*, *99*(1-2), 199-200.

Erbach, K. (2021). *Object Mass Nouns: A Frame Based Analysis*. Heinrich-Heine-Universität, Düsseldorf. Phd Thesis.

Falkum, I. L. (2011). *The semantics and pragmatics of polysemy: A relevance-theoretic account* (Doctoral dissertation, UCL (University College London)).

Fardin, M. A. (2014). On the rheology of cats. *Rheology Bulletin*, *83*(2), 16-17.

Faussart, C., Jakubowicz, C., & Costes, M. (1999). Gender and number processing in spoken French and Spanish. *Rivista di Linguistica*, *11*(1), 75-101.

Feigenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of more: Object files versus analog magnitudes. *Psychological science*, *13*(2), 150-156.

Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition*, *97*(3), 295-313.

Fieder, N., Nickels, L., & Biedermann, B. (2014). Representation and processing of mass and count nouns: A review. *Frontiers in Psychology*, *5*, 589.

Filip, H., & Sutton, P. (2017, December). Singular count NPs in measure constructions. In *Semantics and Linguistic Theory* (Vol. 27, pp. 340-357).

Foraker, S., & Murphy, G. L. (2012). Polysemy in sentence comprehension: Effects of meaning dominance. *Journal of Memory and Language*, *67*(4), 407-425.

Frazier, L., & Rayner, K. (1990). Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory and Language*, *29*(2), 181-200.

Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and language*, *50*(3), 259-281.

Friederici, A. D., & Jacobsen, T. (1999). Processing grammatical gender during language comprehension. *Journal of psycholinguistic Research*, *28*(5), 467-484.

Frisson, S. (2009). Semantic underspecification in language processing. *Language and Linguistics Compass*, *3*(1), 111-127.

Frisson, S., & Frazier, L. (2005). Carving up word meaning: Portioning and grinding. *Journal of Memory and Language*, *53*(2), 277-291.

Gamerschlag, T., Gerland, D., Osswald, R., & Petersen, W. (Eds.). (2013). *Frames and concept types: applications in language and philosophy* (Vol. 94). Springer Science & Business Media.

Gauvin, H. S., & Hartsuiker, R. J. (2020). Towards a new model of verbal monitoring. *Journal of Cognition*, *3*(1).

Gillon, B. S. (1999). The lexical semantics of English count and mass nouns. In *Breadth and Depth of Semantic Lexicons* (pp. 19-37). Springer, Dordrecht.

Gillon, B., Kehayia, E., & Taler, V. (1999). The mass/count distinction: Evidence from on-line psycholinguistic performance. *Brain and Language*, *68*(1-2), 205-211.

Glaser, W. R. (1992). Picture naming. *Cognition*, *42*(1-3), 61-105.

Grice, H. P. (1975). Logic and conversation. In *Speech acts* (pp. 41-58). Brill.

Grimm, S. (2018). Grammatical number and the scale of individuation. *Language 94*(3), 527-574.

Grimm, S. & Levin B. (unpublished manuscript). *Furniture and other artifactual aggregates.*

Hespos, S. J., & Baillargeon, R. (2001). Reasoning about containment events in very young infants. *Cognition*, *78*(3), 207-245.

Heyman, T., Bruninx, A., Hutchison, K. A., & Storms, G. (2018). The (un)reliability of item-level semantic priming effects. *Behavior Research Methods*, *50*(6), 2173-2183.

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118-135.

Imai, M., & Gentner, D. (1997). A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition*, *62*(2), 169-200.

Imai, M., & Mazuka, R. (2003). Re-evaluating linguistic relativity: Language-specific categories and the role of universal ontological knowledge in the construal of individuation. *Language in Mind: Advances in the Study of Language and Thought*, 429-464.

Imai, M., & Mazuka, R. (2007). Language-Relative Construal of Individuation Constrained by Universal Ontology: Revisiting Language Universals and Linguistic Relativity. *Cognitive Science*, *31*(3), 385-413.

Inagaki, S. (2014). Syntax-semantics mappings as a source of difficulty in Japanese speakers' acquisition of the mass-count distinction in English. *Bilingualism*, *17*(3), 464.

Jackendoff, R. S. (1990). Semantic Structures Current Studies in Linguistics Series 18.

Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: object specific integration of information. *Cognitive Psychology*, *24*, 175-219.

Kammann, R., & Streeter, L. (1971). Two meanings of word abstractness. *Journal of Verbal Learning and Verbal Behavior*, *10*(3), 303-306.

Klein, D. E., & Murphy, G. L. (2002). Paper has been my ruin: Conceptual relations of polysemous senses. *Journal of Memory and Language*, *47*(4), 548-570.

Klepousniotou, E., Titone, D., & Romero, C. (2008). Making sense of word senses: the comprehension of polysemy depends on sense overlap. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1534.

Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics. *Semantics and Contextual Expression*, *75*, 115.

Krifka, M. (1995). Common nouns: A contrastive analysis of English and Chinese. *The Generic Book*, 398-411.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203-205.

Landman, F. (2011). Count nouns, mass nouns, neat nouns, mess nouns. *Baltic International Yearbook of Cognition, Logic and Communication 6*(1), 1-67.

Landman, F. (2016). Iceberg semantics for count nouns and mass nouns: The evidence from portions. *The Baltic International Yearbook of Cognition Logic and Communication 11*, 1-48.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences*, *22*, 1-38.

Levin, B., L. Glass, & Jurafsky, D. (2016). Systematicity in the semantics of noun compounds: The role of artifacts vs. natural kinds. Stanford University

Li, P., Dunham, Y., & Carey, S. (2009). Of substance: The nature of language effects on entity construal. *Cognitive Psychology*, *58*(4), 487-524.

Lima, S. (2014). *The Grammar of Individuation and Counting* (Doctoral dissertation, University of Massachusetts, Amherst).

Lin, J., Hacohen, A., & Schaeffer, J. (2018). The interpretation of the mass-count distinction across languages and populations: Introduction. *Glossa: A Journal of General Linguistics*, *3*(1).

Link, G. (1983). The logical analysis of plurals and mass terms. *Formal semantics: The essential readings*, *127*, 147.

Löbner, S. (2011). Concept types and determination. *Journal of Semantics*, *28*(3), 279-333.

Lukatela, G., Kostić, A., Todorović, D., Carello, C., & Turvey, M. T. (1987). Type and number of violations and the grammatical congruency effect in lexical decision. *Psychological Research*, *49*(1), 37-43.

MacDonald, D., & Carroll, S. E. (2018). Second-language processing of English mass-count nouns by native-speakers of Korean.

Mani, N., & Plunkett, K. (2011). Phonological priming and cohort effects in toddlers. *Cognition*, *121*(2), 196-206.

Marangolo, P., Di Pace, E., & Pizzamiglio, L. (1993). Priming effect in a color discrimination task. *Perceptual and Motor Skills*, *77*(1), 259-269.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1-2), 71-102.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1-86.

McQueen, J. M., Dahan, D., & Cutler, A. (2003). Continuity and gradedness in speech processing. *Phonetics and phonology in language comprehension and production: Differences and similarities*, 39-78.

Middleton, E. L., Wisniewski, E. J., Trindel, K. A., & Imai, M. (2004). Separating the chaff from the oats: Evidence for a conceptual distinction between count noun and mass noun aggregates. *Journal of Memory and Language*, *50*(4), 371-394.

Mondini, S., Angrilli, A., Bisiacchi, P., Spironelli, C., Marinelli, K., & Semenza, C. (2008). Mass and count nouns activate different brain regions: an ERP study on early components. *Neuroscience Letters*, *430*(1), 48-53.

Moravcsik, J. M. (1975). Aitia as generative factor in Aristotle's philosophy. *Dialogue: Canadian Philosophical Review/Revue canadienne de philosophie*, *14*(4), 622-638.

Morton, J. (1964). The effects of context on the visual duration threshold for words. *British Journal of Psychology*, *55*(2), 165-180.

Morton, J. (1968). A RETEST OF THE RESPONSE-BIAS EXPLANATION OF THE WORD-FREQUENCY EFFECT. *British Journal of Mathematical and Statistical Psychology*, *21*(1), 21-33.

Morton, J. (1969). Interaction of information in word recognition. *Psychological review*, *76*(2), 165.

Morton, J. (1979). Facilitation in word recognition: Experiments causing change in the logogen model. In *Processing of visible language* (pp. 259-268). Boston, MA: Springer US.

Naumann, R., & Petersen, W. (2019). Bridging inferences in a dynamic frame theory. In *Language, Logic, and Computation: 12th International Tbilisi Symposium, TbiLLC 2017, Lagodekhi, Georgia, September 18-22, 2017, Revised Selected Papers 12* (pp. 228-252). Springer Berlin Heidelberg.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theory. In *Basic processes in reading: Visual word recognition* (pp. 264-336).

Nichols, L. (2008). Lexical semantic constraints on noun roots and noun borrowability. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"*, *32*(3), 683-700.

Noble, C. E. (1952). An analysis of meaning. *Psychological review*, *59*(6), 421.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*(3), 189-234.

Ortega-Andrés, M., & Vicente, A. (2019). Polysemy and co-predication. *Glossa: A Journal of General Linguistics*, *4*(1).

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, *34*(6), 739-773.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of experimental psychology*, *76*(1p2), 1.

Palmer, F. R. (1981). *Semantics*. Cambridge University Press.

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

Pelletier, F. J. (2012). Lexical nouns are both+ MASS and+ COUNT, but they are neither+ MASS nor+ COUNT. In *Count and Mass Across Languages.* Oxford University Press.

Petersen, W. (2007). Representation of Concepts as Frames. *The Baltic International Yearbook of Cognition, Logic and Communication, 2,* 151-170.

Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.

Pinheiro, J. C., & Bates, D. M. (2000). Theory and computational methods for linear mixed-effects models. *Mixed-effects models in S and S-PLUS*, 57-96.

Pisoni, D., Bernacki, R., Nusbaum, H. C., & Yuchtman, M. (1985, April). Some acoustic-phonetic correlates of speech produced in noise. In *ICASSP'85. IEEE International Conference on Acoustics, Speech, and Signal Processing* (Vol. 10, pp. 1581-1584). IEEE.

Prasada, S., Ferenz, K., & Haskell, T. (2002). Conceiving of entities as objects and as stuff. *Cognition*, *83*(2), 141-165.

Pustejovsky, J. (1995). *The Generative Lexicon*. MIT Press. Cambridge, Massachusetts.

Pylkkänen, L., Llinás, R., & Murphy, G. L. (2006). The representation of polysemy: MEG evidence. *Journal of Cognitive Neuroscience*, *18*(1), 97-109.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects and situated vision. *Cognition, 80*, 127-158.

Quine, W. V. (1960). Word and Object, New Edition, 2013.

Quine, W. V. O. (1969). Ontological relativity. *Ontological Relativity and other Essays*, *26*, 68.

Rijkhoff, J. (1990). Explaining word order in the noun phrase. 5-42.

Ritter, E. (1988). A head-movement approach to construct-state noun phrases. *Linguistics*, *26*(6), 909-930.

Ritter, E. (1993). Where's gender?. *Linguistic inquiry*, *24*(4), 795-803.

Roelofs, A., & Ferreira, V. S. (2019). The architecture of speaking. *Human language: From genes and brains to behavior*, 35-50.

Rothstein, S. (2010). Counting and the mass/count distinction. *Journal of Semantics 27(*3), 343-397.

Schlenker, P. (1999). La flexion de l'adjectif en allemand: La morphologie de haut en bas. *Recherches linguistiques de Vincennes*, (28), 115-132.

Schriefers, H., Meyer, A. S., & Levelt, W. J. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, *29*(1), 86-102.

Schulzek, D. (2014). A frame approach to metonymical processes in some common types of German word formation. In *Frames and concept types* (pp. 221-242). Springer, Cham.

Schumacher, P. B. (2013). When combinatorial processing results in reconceptualization: toward a new approach of compositionality. *Frontiers in Psychology*, *4*, 677.

Soja, N., S. Carey, & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition* 38(2), 179-211.

Spelke, E. S., Kestenbaum, R., Simons, D. J., & Wein, D. (1995). Spatio-temporal continuity, smoothness of motion and object identity in infancy. *British Journal of Developmental Psychology, 13*, 113-142.

Steinhauer, K., Pancheva, R., Newman, A. J., Gennari, S., & Ullman, M. T. (2001). How the mass counts: An electrophysiological approach to the processing of lexical features. *Neuroreport*, *12*(5), 999-1005.

Subrahmanyam, K., Landau, B., & Gelman, R. (1999). Shape, material, and syntax: Interacting forces in children's learning in novel words for objects and substances. *Language and Cognitive Processes*, *14*(3), 249-281.

Sutton, P. & H. Filip (2016a). Counting in context: count/mass variation and restrictions on coercion in collective artifact nouns. *Semantics and Linguistic Theory 26*(0), 350-370.

Sutton, P. R. & H. Filip (2016c). Mass/count variation: A mereological, two-dimensional semantics. *Baltic International Yearbook of Cognition, Logic and Communication 11*(1), 11.

Sutton, P. R. & H. Filip (2018). Restrictions on subkind coercion in object mass nouns. In *Proceedings of Sinn und Bedeutung 21* (pp. 1195-1213). University of Edinburgh.

Sutton, P. R. & H. Filip (2019). The count/mass distinction for 'granular' nouns. In *Counting and Measuring in Natural Language* (pp. 375-415). Cambridge University Press.

Swinney, D. A. (1979). Lexical access during sentence comprehension:(Re) consideration of context effects. *Journal of verbal learning and verbal behavior*, *18*(6), 645-659.

Tanenhaus, M. K., & Lucas, M. M. (1987). Context effects in lexical processing. *Cognition*, *25*(1-2), 213-234.

Van Valin Jr, R. D. (2001). *An introduction to syntax*. Cambridge University Press.

Van Valin Jr, R. D. (2005). *Exploring the syntax-semantics interface*. Cambridge University Press.

Viebahn, E. (2020). Copredication, polysemy and context-sensitivity. *Inquiry*, 1-17.

Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition*, *61*(3), 261-298.

Weinreich, U., & Webster. (1964). Webster's third: A critique of its semantics. *International Journal of American Linguistics*, *30*(4), 405-409.

Wiese, H., & Maling, J. (2005). Beers, kaffi, and schnaps: Different grammatical options for restaurant talk coercions in three Germanic languages. *Journal of Germanic Linguistics*, *17*(1), 1-38.

Whorf, B. L. (1956). *Language, thought, and reality: selected writings of… Edited by John B. Carroll*.). Technology Press of MIT.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive psychology*, *24*(2), 220-251.

Wynn, K. (1992b). Addition and subtraction by human infants. *Nature*, *358*, 749-750.

Yin, B., & O'Brien, B. A. (2018). Mass-count distinction in Chinese-English bilingual students. *Glossa: A Journal of General Linguistics*, *3*(1).

Appendix

1. Stimulus Set 1 (used in Lexical Decision Experiment 1)

Countability	Shiftability	Determiner	Noun
count	not shiftable	eine	Bürgerin
count	not shiftable	ein	Bus
count	not shiftable	ein	General
count	not shiftable	ein	Kardinal
count	not shiftable	eine	Puppe
count	not shiftable	eine	Politikerin
count	not shiftable	ein	Asiat
count	not shiftable	eine	Runde
count	not shiftable	ein	Bestandteil
count	not shiftable	eine	Gruppe
count	not shiftable	eine	Lektion
count	not shiftable	eine	Nation
count	not shiftable	ein	Paar
count	not shiftable	eine	Situation
count	not shiftable	ein	Versprechen
count	shiftable	ein	Ohr
count	shiftable	eine	Folie
count	shiftable	ein	Signal
count	shiftable	eine	Beilage
count	shiftable	ein	Beweismittel
count	shiftable	eine	Pflanze
count	shiftable	ein	Produkt
count	shiftable	ein	Diebstahl
count	shiftable	eine	Waffe

count	shiftable	eine	Strecke
count	shiftable	eine	Auffälligkeit
count	shiftable	eine	Besonderheit
count	shiftable	eine	Illusion
count	shiftable	ein	Risiko
count	shiftable	eine	Figur
dual-count	NA	ein	Besteck
dual-count	NA	ein	Tuch
dual-count	NA	ein	Krieg
dual-count	NA	ein	Werkzeug
dual-count	NA	ein	Besuch
dual-count	NA	ein	Schatz
dual-count	NA	eine	Irritation
dual-count	NA	ein	Ausfall
dual-count	NA	eine	Nacht
dual-count	NA	ein	Tier
dual-count	NA	ein	Management
dual-count	NA	ein	Angebot
dual-count	NA	eine	Diskussion
dual-count	NA	ein	Verbrechen
dual-count	NA	eine	Gelegenheit
dual-mass	NA	viel	Bier
dual-mass	NA	viel	Feuer
dual-mass	NA	viel	Joghurt
dual-mass	NA	viel	Pelz
dual-mass	NA	viel	Tee
dual-mass	NA	viel	Wein

dual-mass	NA	viel	Reklame
dual-mass	NA	viel	Kontrolle
dual-mass	NA	viel	Urlaub
dual-mass	NA	viel	Hoffnung
dual-mass	NA	viel	Kritik
dual-mass	NA	viel	Wahrheit
dual-mass	NA	viel	Religion
dual-mass	NA	viel	Miete
dual-mass	NA	viel	Angst
mass	not shiftable	viel	Bronze
mass	not shiftable	viel	Dreck
mass	not shiftable	viel	Laub
mass	not shiftable	viel	Sand
mass	not shiftable	viel	Schlamm
mass	not shiftable	viel	Equipment
mass	not shiftable	viel	Gerümpel
mass	not shiftable	viel	Misstrauen
mass	not shiftable	viel	Neid
mass	not shiftable	viel	Personal
mass	not shiftable	viel	Bedarf
mass	not shiftable	viel	Chaos
mass	not shiftable	viel	Einigkeit
mass	not shiftable	viel	Gedrängel
mass	not shiftable	viel	Humor
mass	shiftable	viel	Geschirr
mass	shiftable	viel	Müll
mass	shiftable	viel	Unterwäsche

mass	shiftable	viel	Beute
mass	shiftable	viel	Gold
mass	shiftable	viel	Menthol
mass	shiftable	viel	Gebäck
mass	shiftable	viel	Knoblauch
mass	shiftable	viel	Nachwuchs
mass	shiftable	viel	Obst
mass	shiftable	viel	Karamell
mass	shiftable	viel	Gepäck
mass	shiftable	viel	Volumen
mass	shiftable	viel	Nahrung
mass	shiftable	viel	Ruhm
pseudoword	NA	eine	Pölgerin
pseudoword	NA	ein	Dus
pseudoword	NA	ein	Bümeral
pseudoword	NA	ein	Toldinal
pseudoword	NA	eine	Doppe
pseudoword	NA	eine	Barütikerin
pseudoword	NA	ein	Eschiat
pseudoword	NA	eine	Nemde
pseudoword	NA	ein	Piftandteil
pseudoword	NA	eine	Dloppe
pseudoword	NA	eine	Rübtion
pseudoword	NA	eine	Mätion
pseudoword	NA	ein	Boor
pseudoword	NA	eine	Schüoation
pseudoword	NA	ein	Salsprechen

pseudoword	NA	ein	Uhl
pseudoword	NA	eine	Surie
pseudoword	NA	ein	Füdnal
pseudoword	NA	eine	Poilage
pseudoword	NA	ein	Pißeismittel
pseudoword	NA	eine	Tanze
pseudoword	NA	ein	Bladukt
pseudoword	NA	ein	Künstahl
pseudoword	NA	eine	Soffe
pseudoword	NA	eine	Slucke
pseudoword	NA	eine	Oßülligkeit
pseudoword	NA	eine	Pifanderheit
pseudoword	NA	eine	Ürrosion
pseudoword	NA	ein	Nüsiko
pseudoword	NA	eine	Sügur
pseudoword	NA	ein	Pöftek
pseudoword	NA	ein	Düch
pseudoword	NA	ein	Glieg
pseudoword	NA	ein	Malkzeug
pseudoword	NA	ein	Düsuch
pseudoword	NA	ein	Jutz
pseudoword	NA	eine	Üllätation
pseudoword	NA	ein	Ousfall
pseudoword	NA	eine	Mucht
pseudoword	NA	ein	Dier
pseudoword	NA	ein	Nerogement
pseudoword	NA	ein	Umgebot

pseudoword	NA	eine	Tüschkussion
pseudoword	NA	ein	Sulbrechen
pseudoword	NA	eine	Kiregenheit
pseudoword	NA	viel	Plonze
pseudoword	NA	viel	Tleck
pseudoword	NA	viel	Rob
pseudoword	NA	viel	Fond
pseudoword	NA	viel	Zramm
pseudoword	NA	viel	Igübment
pseudoword	NA	viel	Kirümpel
pseudoword	NA	viel	Nusstrauen
pseudoword	NA	viel	Jeid
pseudoword	NA	viel	Balsonal
pseudoword	NA	viel	Pidarf
pseudoword	NA	viel	Goos
pseudoword	NA	viel	Umihkeit
pseudoword	NA	viel	Kidrängel
pseudoword	NA	viel	Ramor
pseudoword	NA	viel	Nöschirr
pseudoword	NA	viel	Nill
pseudoword	NA	viel	Omdurwäsche
pseudoword	NA	viel	Paute
pseudoword	NA	viel	Kuld
pseudoword	NA	viel	Nimthol
pseudoword	NA	viel	Kibäck
pseudoword	NA	viel	Gmoplauch
pseudoword	NA	viel	Muchwuchs

pseudoword	NA	viel	Abst
pseudoword	NA	viel	Gulamell
pseudoword	NA	viel	Kipäck
pseudoword	NA	viel	Furumen
pseudoword	NA	viel	Mohrung
pseudoword	NA	viel	Luhm
pseudoword	NA	viel	Püür
pseudoword	NA	viel	Wauer
pseudoword	NA	viel	Zakhurt
pseudoword	NA	viel	Bulz
pseudoword	NA	viel	Dee
pseudoword	NA	viel	Foin
pseudoword	NA	viel	Luglame
pseudoword	NA	viel	Gamtrolle
pseudoword	NA	viel	Ochlaub
pseudoword	NA	viel	Russnung
pseudoword	NA	viel	Glutik
pseudoword	NA	viel	Selheit
pseudoword	NA	viel	Narigion
pseudoword	NA	viel	Nüte
pseudoword	NA	viel	Omgst

2. Stimulus Set 2 (used in Lexical Decision Experiment 2 and Phoneme Monitoring Experiment)

Countability	Determiner	Lemma	Minimal Pseudoword	Maximal Pseudoword
count	eine	Kartoffel	Gartoffel	Nartoffel
count	ein	Vorurteil	Sorurteil	Gorurteil
count	ein	Krokodil	Grokodil	Frokodil
count	ein	Marathon	Barathon	Karathon
count	eine	Illusion	Irusion	Immusion
count	ein	Lexikon	Rexikon	Fexikon
count	ein	Fundament	Sundament	Dundament
count	ein	Referat	Leferat	Weferat
count	eine	Schwierigkeit	Schlierigkeit	Schtierigkeit
count	ein	Sündenbock	Schündenbock	Hündenbock
count	ein	Vorgeschmack	Sorgeschmack	Gorgeschmack
count	ein	Mosaik	Nosaik	Losaik
dual_count	ein	Putzmittel	Tutzmittel	Hutzmittel
dual_count	ein	Waschmittel	Saschmittel	Raschmittel
dual_count	ein	Bauchgefühl	Dauchgefühl	Kauchgefühl
dual_count	eine	Depression	Bepression	Hepression
dual_count	eine	Aggression	Adression	Apression
dual_count	eine	Promotion	Tromotion	Gromotion
dual_count	eine	Belegschaft	Melegschaft	Helegschaft
dual_count	ein	Dialekt	Bialekt	Hialekt
dual_count	eine	Süßigkeit	Schüßigkeiten	Küßigkeit
dual_count	ein	Feuerwerk	Seuerwerk	Keuerwerk
dual_count	eine	Problematik	Troblematik	Groblematik
dual_count	ein	Nachgeschmack	Machgeschmack	Laachgeschmack

dual_mass	viel	Mitgefühl	Bitgefühl	Kitgefühl
dual_mass	viel	Potential	Totential	Gotential
dual_mass	viel	Kapital	Gapital	Napital
dual_mass	viel	Gefallen	Defallen	Pefallen
dual_mass	viel	Perfektion	Terfektion	Gerfektion
dual_mass	viel	Rebellion	Lebellion	Febellion
dual_mass	viel	Einsamkeit	Eimsamkeit	Eichsamkeit
dual_mass	viel	Ähnlichkeit	Ähmlichkeit	Ähklichkeit
dual_mass	viel	Helligkeit	Relligkeit	Delligkeit
dual_mass	viel	Mechanik	Bechanik	Rechanik
dual_mass	viel	Keramik	Teramik	Meramik
dual_mass	viel	Fachwerk	Wachwerk	Nachwerk
mass	viel	Edelstahl	Etelstahl	Erelstahl
mass	viel	Sägemehl	Schägemehl	Kägemehl
mass	viel	Gerümpel	Derümpel	Nerümpel
mass	viel	Volumen	Solumen	Rolumen
mass	viel	Misstrauen	Bisstrauen	Kisstrauen
mass	viel	Präzision	Träzision	Gräzision
mass	viel	Übelkeit	Üpelkeit	Ükelkeit
mass	viel	Equipment	Etwipment	Ezwipment
mass	viel	Dynamit	Gynamit	Hynamit
mass	viel	Romantik	Lomantik	Fomantik
mass	viel	Buschwerk	Puschwerk	Huschwerk
mass	viel	Rockmusik	Sockmusik	Gockmusik