Polish numerals and quantifiers: A syntactic analysis of subject-verb agreement mismatches

Heidi Klockmann

3275752

Master's Thesis

Linguistics: The Study of the Language Faculty

Utrecht University

August 2012

Supervisors: Dr. Marjo van Koppen

Prof. dr. Norbert Corver

ACKNOWLEDGMENTS

The writing of this thesis was a long and arduous process which could not have been accomplished without the help and support of my supervisors, friends, and family. I would first like to thank Marjo van Koppen for the many meetings and discussions we had together. They were always enlightening and truly the source for this thesis. Secondly, I would like to thank Norbert Corver – his comments and suggestions were very helpful and much appreciated. Importantly, I would also like to thank my informants: my boyfriend, Łukasz Głąb, for always answering my questions, no matter the hour, Joanna Borkowska for her helpful and insightful intuitions and comments concerning the data, and the friends of Łukasz, who were kind enough to take the time to fill out my questionnaire. Without their willingness to help, this thesis could not be what it is. I'd also like to thank my parents for constantly supporting me in this and for always showing interest in what I do, even if they didn't really know what it is that I do. A special thanks goes to my boyfriend, Łukasz, whose help and support was invaluable during this thesis; he was my encouragement and has been throughout this program. Lastly, I'd like to thank my friends here in Utrecht and at home – you helped to keep me sane when I was most stressed, and brought me some very fun and relaxing evenings.

Table of Contents

Summary

Polish numerals and quantifiers present a complex pattern of agreement and case assignment that defy the current instantiation of Agree as given in Chomsky (2000, 2001). In Polish, the presence of numerals and quantifiers leads to so-called "agreement mismatches" – the phi-features on the probe and the goal differ unexpectedly, i.e. there is a mismatch in features. Additionally, they lead to interesting patterns of case assignment, such as a shared genitive on the numeral and noun, and a type of case-agreeing, case-assigning alternation dependent on structural position. In this thesis, I explore these agreement mismatches and case phenomena, providing an analysis for the behavior of numerals and quantifiers.

I argue that Polish numerals are semi-lexical nouns (Corver & Riemsdijk, 2001) and from this I derive much of their behavior. The functional aspect of these numerals is conceived of as the absence of certain phi-features, or for certain numerals, those phi-features being uninterpretable and unvalued. Due to this, subject-verb agreement fails (as the verb cannot agree with the underspecified/under-valued numeral) leading to default agreement (Preminger, 2011) and default case assignment. This defectiveness of the numeral is also inherited by the numeral's genitive case assigner, leading to a phenomenon of "case leaking", i.e. the occurrence of the same genitive on both numeral and noun. This is argued to be an instance of cyclic agreement, in which after searching downwards, a probe also extends its search space upwards (Rezac, 2003), producing two instances of genitive case marking. Furthermore, due to the functional nature of the numeral, it is unable to carry a theta-role, producing the case-assigning, case-agreeing alternation. The differing behaviors of numerals are viewed as an effect of their semi-lexical nominal status, and the mechanisms involved in this analysis (default agreement, cyclic agreement, case stacking) have consequences for theories of agreement and case assignment.

List of Abbreviations

1	1st person
2	2nd person
3	3rd person
ACC	accusative
C(P)	complementizer (phrase)
DAT	dative
DIST	distributive marker *po*
D(P)	determiner (phrase)
DU	dual
F	feminine
GEN	genitive
INSTR	instrumental
LOC	locative
M	masculine
MP	masculine personal
N	neuter
NOM	nominative
NON-VIR	non-virile
N(P)	noun (phrase)
OBL	oblique
PL	plural
Q(P)	quantifier (phrase)
REFL	reflexive
SG	singular
TO	predicative marker *to*
T(P)	tense (phrase)
VIR	virile
V(P)	verb (phrase)

1 Introduction

Agreement is defined as the relationship between two elements in a sentence, where one of these elements determines the form of the second (controller and target, respectively, in Corbett (2006)'s system); a core example is subject-verb agreement, in which the verb manifests the features of the subject. Chomsky (2000, 2001) defines this relationship as an operation Agree; an active probe (by virtue of its unvalued features) seeks out an active goal within its c-command domain. The probe takes on the features of the goal and in exchange, assigns it case, after which both become inactive for further agreement relationships. If any elements are left with unvalued features at the end of the derivation, the derivation crashes.

In this thesis, I will explore cases in which Agree does not function as defined above, producing what I call "agreement mismatches", to be demonstrated shortly. Such mismatches are abundant in the numeral systems of Polish and various other languages. The purpose of this thesis is to explain why they might occur, and what changes are necessary to increase the empirical coverage of Agree, specifically with regard to agreement mismatches.

Polish numerals in subject position present a clear counterexample to the definition of Agree given above. By Agree, the same features which are present on the subject should appear in some form on the verb. For numerals 2,3,4 in Polish, this is the case (1a); with numerals greater than 5 (but less than 1000), however, it turns out that there is a mismatch between the features of the subject and the features of the verb, yet it does *not* lead to a derivation crash (1b):

(1) a. *Dwa* *ptaki* *spały*
Two.M.NOM bird.M.NOM.PL slept.NON-VIR.PL[1]
"Two birds slept"

[1] Examples with numerals occur in 3rd person. I leave out the person feature in the glosses, as it makes no difference for the data or analyses to come.

b. *Pięć* *ptaków* *spało*
Five bird.M.GEN.PL slept.N.SG
"Five birds slept"

The noun in (1a) is valued for masculine gender and plural number. This is reflected on the verb through non-virile plural agreement;[2] thus, agreement has occurred successfully in (1a), as expected by Agree. In example (1b), the same noun and verb is used, the only difference being the value of the numeral (which here produces a genitive case marking on the noun). By analogy to (1a), we would expect non-virile plural agreement on the verb. Instead however, neuter singular marking appears. This constitutes an agreement mismatch – the verb has the ability to manifest matching agreement for the noun in question, but it does not, instead surfacing with a different set of features from the noun. This is what we see in example (1b).

A similar phenomenon occurs with quantified masculine personal nouns (masculine gendered nouns referring to humans) in subject position. This is found even with the numerals 2,3,4 which previously showed agreement. Compare (1a) with (2a) below – with (1a) there was no agreement mismatch, but with (2a) there is:

(2) a. *Dwóch* *chłopców* *spało*
Two.GEN/ACC boy.MP.GEN/ACC.PL slept.N.SG
"Two boys slept"
b. *Chłopcy* *spali*
Boy.MP.NOM.PL slept.VIR.PL
"The boys slept"

The difference between (1a) and (2a) is only the choice of noun, *boy* (masculine and human) instead of *bird* (just masculine). The noun is gendered and numbered as masculine personal plural, yet the verb only shows neuter singular marking, instead of the expected virile agreement (2b). Thus, here, the gender of

[2] In the plural, the only relevant gender distinction in the verbal domain is that between masculine personal (masculine gender + human) and non-masculine personal (everything else), which is marked on the verb as virile and non-virile, respectively. In (1), the subject is masculine, but non-human, so we expect non-virile agreement.

the noun plays a role in whether there is an agreement mismatch or not. Agreement mismatches like these are the driving force of this thesis.

Note that it is not the numeral that the verb is agreeing with. Firstly, if the numeral carried neuter singular features, it would be unexpected for the gender of the following noun to play any role in agreement mismatches, as it did between (1a) and (2a); secondly, it would make little sense as to why the verb agrees with the noun for one gender (1a), but the numeral for another (2a). Coordination data also provides evidence against such an account – the coordination of two neuter singular nouns produces non-virile verb agreement (3a), while the coordination of two quantified subjects still results in neuter singular marking (3b):

(3) a. *Krzesło* *i* *biurko* ***zapadły*** *się*
Chair.NOM.SG and desk.NOM.SG collapsed.NON-VIR.PL REFL
"A chair and a desk collapsed"

b. *Pięć* *krzeseł* *i* *pięć* *biurek* ***zapadło*** *się*
Five chair.GEN.PL and five desk.GEN.PL collapsed.N.SG REFL
"Five chairs and five desks collapsed"

If the numeral were the head of the construction, carrying neuter singular features for the verb to agree with, then like the coordination of two neuter singular nouns in (3a), the coordination of two neuter singular numerals (3b) would be expected to produce non-virile plural agreement. However, this does not occur, as we still find neuter singular marking on the verb, suggesting that it is not the numeral with which the verb agrees. Rather, this suggests that the neuter singular features on the verb are not the result of the operation Agree at all, as neither the noun nor the numeral could be the source of those features.

The question then becomes, why do these agreement mismatches occur? Under the instantiation of Agree described above, the verbal probe should find the presumably active subject and agree with it, as it appears to do in (1a). Instead, however, we find agreement mismatches, accompanied by non-nominative case marking (on the noun for 5+ numerals, on the noun and numeral for masculine personal nouns). The issue cannot be solved by appealing to the role of case marking either – presumably, if the noun (and numeral) have

already been cased for whatever reason, then they are inactive for agreement; with no active goal for the verb, the verb cannot Agree and should finish the derivation with remaining unvalued features. However, by Agree, if there are any remaining unvalued uninterpretable features at the end of the derivation, that derivation should crash. Clearly, that does not happen, as these agreement mismatches are perfectly grammatical. The noun cannot be the source for the verb's features and neither can the numeral – thus, there are no sources for the neuter singular features on the noun, and the case marking of the noun cannot explain the issue either, as it predicts that these sentences are ungrammatical. Therefore, we have a theoretical problem for Agree.

These problems are not restricted to Polish numerals. Inari Sami presents similar issues: with numerals in subject position, the nouns are obligatorily singular, while verbs still show dual or plural marking:

(4) *Meecist* *kaččáin* *kyehti* *almaa*
forest.LOC ran.**3DU** two man.**SG**.GEN
"Two men ran in the forest"

(Toivonen, 2007)

Thus, here again, we see an agreement mismatch – singular marking on the subject noun, yet dual marking on the verb.

These strange patterns of agreement, which seem particular to numerals and some quantifiers, occur in numerous languages. All Slavic languages, except Bulgarian and Macedonian, have this split between numerals 2,3,4 and 5+, where 2,3,4 show true verbal agreement, while 5+ have this odd agreement mismatch (Corbett, 2000). Similarly, the Inari Sami pattern of plural verb with singular subject noun is also a cross-linguistic phenomenon found in Skolt Sami (Feist, 2010), Kurdish (Ortmann, 2000), and East Circassian, a Kabardian dialect (Ortmann, 2000; Colarusso, 1992). Thus, this phenomenon of agreement mismatches is more common in the world's languages than a look at just Polish might suggest. As a whole, it presents a challenge to the theory of agreement.

For many of the above-named languages, this behavior is particular to quantifiers and numerals, with their absence leading to full agreement. This can be seen in the two examples below, (5) for Polish and (6) for Inari Sami.

(5) a. *Pięć* *ptaków* *spało*
Five bird.M.GEN.PL slept.N.SG
"Five birds slept"
b. *Ptaki* *spały*
Bird.M.NOM.PL slept.NON-VIR.PL
"(The) birds slept"

(6) a. *Kyehti* *almaa* *láin* *meecist*
two man.**SG**.GEN were.**3DU** forest.LOC
"Two men were in the forest"
b. *Almah* *kuá'lásteh* *onne*
man.NOM.**PL** fish.**3PL** today
"The men are fishing today"

(Toivonen, 2007)

In the (a) examples, we see agreement with a numeral in the subject, which leads to an agreement mismatch for both languages; the (b) examples demonstrate what occurs *without* a numeral – for Polish, the verb agrees in both gender and number and for Inari Sami, both verb and subject are marked as plural. Altogether, this behavior makes the class of quantifiers and numerals rather interesting from a cross-linguistic perspective, as they present a class of items which are more prone to agreement mismatches. They are also interesting from the perspective of their syntactic behavior within the noun phrase, which has led to considerable disagreement about their nature (heads or phrases) and their syntactic category (noun, adjective, functional element, etc.) (see Cardinaletti and Giusti, 2006 and Corbett, 1978 for discussions of these issues). A careful study of the juxtaposition of numerals and verb agreement provides a rich opportunity to further our understanding of the process of agreement and the nature of numerals in general.

In this thesis, I focus particularly on Polish numerals, with an emphasis on their behavior as regards agreement. I have two main goals here. The first is to explore the behavior of Polish numerals and identify their source. I will build

an account which draws on the concepts of default agreement (Preminger, 2011), multiple agree (Radford, 2006; Chomsky, 2005), and cyclic agree (Rezac, 2003). I will argue that Polish numerals are actually feature-defective, case-assigning nouns. As defective nouns, verbs cannot agree with them, leading to a default agreement on the verb. This argues against the idea of derivational crashes if uninterpretable features remain unvalued and undeleted at the end of a derivation. In the case of masculine personal nouns, I will argue that the masculine personal gender of the noun will lead to "case leaking", resulting in a genitive case on both the noun and numeral, which as I will argue blocks verb agreement. The phenomenon of case leaking is a combination of agreement with two goals (multiple agree) and agreement upwards (cyclic agree), which together lead to this dual appearance of genitive. My second goal is to discuss how the use of these concepts in my analysis affects our current conceptualization of Agree and what consequences this has for the theory.

Before turning to the data, I present a short typology on the agreement behavior of verbs with numerals in the subject.

1.1 A Small Typology of Agreement with Numerals

From a broader perspective, verb agreement with subject-numerals can be classified into two main types – mismatching and matching. As we saw above, Polish and Inari Sami numerals evoke mismatching agreement. The agreement patterns of these two languages, however, suggest that among the mismatching agreement numerals, there is a further division. With Polish numerals, subjects were plural, yet verbs were singular with a different gender. In Inari Sami, on the other hand, subjects were singular, yet verbs were plural. Thus, with Polish numerals, verbs were unable to take on the features of the subject; with Inari Sami, however, verbs seemed to be taking on the features that one would *expect* the subject to have, were there no numeral to force singular marking on the noun. This is a critical difference, as Polish verbs seem to lack agreement, while Inari Sami verbs agree, but unexpectedly, given the features of the noun. For simplicity, I will refer to these languages as *Plural-Singular* (Polish) and *Singular-Plural* (Inari Sami).

With the matching agreement numerals, there is also an important division to be made. With numerals, nouns can either be plural or singular, and for matching agreement numerals, this means that the verb will also be singular or plural. Thus, there are two more types of languages to add to our small typology – *Singular-Singular* languages and *Plural-Plural* languages.

Plural-Plural languages are perhaps the most familiar. This includes languages like Dutch, English, and Spanish. In these languages, subjects are plural with or without a numeral, and verbs agree either way. The example below is taken from Dutch.

(7) a. *De* *mannen* *hebben* *geslapen*
The man.PL have.3PL slept
"The men slept"
b. *De* *twee* *mannen* *hebben* *geslapen*
The two man.PL have.3PL slept
"The two men slept"

Singular-Singular languages are also fairly common. They include Turkish, Hungarian, and Finnish. In these languages, the presence of a numeral forces the noun to be singular, and the verb agrees accordingly. Without a numeral, however, both the subject and the verb would then be plural. The example below is taken from Hungarian.

(8) a. *Mókus-ok* *szalad-nak*
Squirrel-PL run-3PL
"Squirrels are running"
b. *Tíz* *mókus* *szalad*
Ten squirrel.SG run.3SG
"Ten squirrels are running"

(Corbett, 2000: 211)

There are two factors which lead to this four-way division of agreement with numerals in subject position. The first concerns the number marking on the noun following the numeral – this can be either singular or plural. The second concerns verb agreement, which is either matching or mismatching. The table below captures these divisions.

Table 1: Number marking with plural simplex numerals

Verb Subject-Noun	Singular	Plural
Singular	Hungarian, Turkish, Finnish	Inari Sami, Skolt Sami, East Circassian, Russian (2-4)
Plural	Polish (5+), Russian (5+), (other Slavic languages 5+)	English, Dutch, Spanish, Polish (2-4)

In this thesis, I focus on the second factor, the question of verb agreement and the lack thereof in the Plural-Singular corner, basing my analysis on Polish. For a potential analysis of the first factor, the number on the noun, I direct the reader to Ionin and Matushansky (2004).

1.2 Outline

This thesis consists of four chapters in addition to this introduction. In chapter 2, I present an overview of the data. There are three main phenomena with regards to Polish: verb agreement, masculine personal agreement, and case assignment. Both verb agreement and masculine personal agreement have been introduced already, but I will discuss their properties in more detail, especially concerning their behavior with differently valued numerals. The case assignment section addresses a peculiar case assigning versus case agreeing dichotomy of Polish numerals. The rest of the chapter discusses a classification for the numerals. I present a short section on the behavior of adjectives and nouns with regards to the three phenomena mentioned above; this allows a comparison between the behavior of numerals and that of their cousins, adjectives and nouns. Based on the behavior of the numerals, I will propose a four-way division, consisting of (1) numeral 1, adjective, (2) numerals 2,3,4, adjective-like, (3) numerals 5+ (5-999), noun-like, and (4) numerals 1000+, nouns.

In chapter 3, I address the previous solutions to the phenomena discussed in chapter 2. This includes theories concerning the case of the numeral (whether accusative, nominative, or nominative and genitive), as well as the syntactic category of numerals. Based on this discussion, I will promote the view that

numerals are semi-lexical nouns, which can be either nominative or genitive in subject position. The last section of this chapter turns to the case assignment dichotomy and the running theories of that phenomenon.

In chapter 4, I present my analysis, building off the data discussed in chapter 2 and the theories of chapter 3. I will begin by discussing the similarity of the quantitative construction (prevalent in chapter 2) to the partitive construction (section 4.1), with regard to numeral behavior, and the insights this comparison brings. Following this, I turn to my own analysis. I discuss my theory of the neuter singular verbal marking as default agreement (4.2) and how cyclic agree can explain the behavior of numerals with masculine personal nouns (4.3). In section 4.4, I provide an explanation for the case assigning – case agreeing alternation. Sections 4.5 and 4.6 are devoted to syntactic structure; I address quantitatives in the first section and partitives in the second. Finally, in section 4.7, I suggest a tentative explanation for defectiveness, i.e. what it means for a numeral to be defective.

Chapter 5 is the discussion and conclusion. After a quick summary, I address the implications of my analysis and findings for theories of agreement and case, the syntax of numerals, and categories. I present avenues for future research at the end.

2 Polish Numerals and Quantifiers: The Data

As it turns out, the behavior I will describe below is found with both numerals and quantifiers; so, for ease of exposition, I will refer to them together as "numerals".

There are three main domains concerning Polish numerals: verb agreement, masculine personal agreement, and case assignment. I devote a section to each of these domains. Based on the data, I will claim that there are four main classes of numerals. The first is the numeral 1, which behaves identically to adjectives. The second are numerals 2,3,4; these will be found to share properties with both numeral 1 and numerals 5+. However, they differ too much from either group to be treated the same, and are treated as their own class. They are similar to adjectives, but not identical, thus, adjective-like. The third group is the 5+ numerals; these actually include numerals 5-999, although I focus here only on the simplex numerals. This group of numerals behaves similarly to nouns in terms of case assignment, but not so in terms of verb agreement and masculine personal agreement; this leads me to call them noun-like. The last and final group are the 1000+ numerals; these numerals act so much like nouns, that I will claim that that is what they are.

Section 2.2 addresses verbal agreement, section 2.3 masculine personal agreement and section 2.4 case assignment. In section 2.5 I give a short description of adjectives and nouns with regards to the previous discussions. Section 2.6 defends the four-way division among the numerals, and section 2.7 quickly introduces quantifiers.

To facilitate discussion concerning these numerals in the next sections, I present a series of agreement paradigms for the numerals 2 and 5, contrasting case value and gender of the following noun

Table 2: Numeral 2 (Swan, 2002: 190)

	Fem	Neut	Masc	MascPers (Masc + Human)
Nominative Accusative	dwie	dwa		dwaj, dwóch dwóch
Instrumental	dwiema		dwoma	
Genitive Locative		dwóch, dwu		
Dative		dwom, dwóm, dwu		

Table 3: Numeral 5 (Swan, 2002: 191)

	MascPers (Masc + Human)	Non-MascPers (Masc, Fem, Neut)
Nominative	pięciu	pięć
Accusative	pięciu	pięć
Instrumental	pięcioma	
Genitive		
Dative	pięciu	
Locative		

2.1 Verb Agreement

With numeral 1, agreement is singular; with 2,3,4 it is plural, and with 5+ it is neuter singular. This is shown in example (9) below.

(9) a. *Jeden* *ptak* *spał*
One.M.NOM.SG bird.M.NOM.SG slept.M.SG
"One bird slept"

b. *Dwa* *ptaki* *spały*
Two.M.NOM bird.M.NOM.PL slept.NON-VIR.PL
"Two birds slept"

c. *Pięć* *ptaków* *spało*
Five bird.M.GEN.PL slept.N.SG
"Five birds slept"

Agreement appears as expected for numerals 1,2,3,4. For numerals 5+, however, agreement is always neuter singular, regardless of the gender of the subject. This is where the agreement mismatch occurs.

For the 1000+ numerals, agreement occurs with the numeral itself (10).[3] As 1000+ numerals are morphologically masculine, they take masculine adjectival and verbal agreement:

(10) *Cały* *tysiąc* *krów* *spał*
entire.M.NOM.SG thousand.M.NOM.SG cow.F.GEN.PL slept.M.SG
"An entire one thousand cows slept"

Numerals 1000+ also appear to have a number feature, as they can appear in the singular (as above) or the plural (see below):

(11) *Tysiące* *ludzi*
thousand.M.NOM.PL people.M.GEN.PL
"Thousands of people"

In terms of gender features, while 1000 seems to have masculine gender, it is unclear whether numerals 2,3,4 and 5+ have an interpretable gender feature. Numerals 2,3,4 have separate forms for different genders (12a,b), while numerals 5+ for the most part do not (12c):

(12) a. *dwa* *ptaki /* *krzesła*
two.M/N bird.M chair.N
"Two birds / chairs"

---

[3] Rutkowski (2005) suggests that the numerals 1000+ are in a process of grammaticalization, becoming more similar to numerals 5+. For example, while agreement can and does occur as shown above, if the adjective is omitted, agreement no longer occurs, e.g.

i. *Tysiąc* *krów* *spało*
Thousand.M.NOM.SG cow.F.GEN.PL slept.N.SG
N A thousand cows slept"

The lack of adjective leads to neuter singular verbal agreement as we see with 5+ numerals. Note that it is not the case that the adjective is actually some kind of noun, as when the numeral is replaced by a noun of a different gender, the adjective and verb both show the gender of the noun:

ii. *Cała* *głowa* *bolała*
entire.F.NOM.SG head.F.NOM.SG ache.F.SG
"(her/his) entire head ached"

Thus, without the adjective, an agreement mismatch occurs. I will not have the space to address this behavior in this thesis here, and so it will remain an open question; for simplicity, I will ignore it in the rest of the thesis.

b. *dwie* *dziewczyny*
two.F girl.F
"two girls"
c. *pięć* *ptaków* */ krzeseł* / *dziewczyn*
five bird.M chair.N girl.F
"five birds / chairs / girls"

The examples suggest that numerals 2,3,4 at least have an uninterpretable gender feature, which agrees with the following noun. For the numerals 5+, however, the consistent appearance of an agreement mismatch obscures any understanding of what the gender might be, if there even is one. Numeral 1 behaves similarly to numerals 2,3,4, agreeing in gender with the following noun:

(13) a. *jeden* *ptak*
one.M bird.M
"one bird"
b. *jedno* *krzesło*
one.N chair.N
"one chair"
c. *jedna* *dziewczyna*
one.F girl.F
"one girl"

This suggests that 1 also has some uninterpretable gender feature which is filled through agreement with the noun.

As for number, it is also unclear if numerals 2,3,4 and 5+ have such a feature, and if they do, whether it can be valued for anything other than plural. For one, these numerals can only ever be used with plural nouns, becoming ungrammatical if followed by singular nouns:

(14) a. **Dwa* *ptak* *spał / spały*
Two.M.NOM bird.M.NOM.SG slept.M.SG / slept.NON-VIR.PL
b. **Pięć* *ptaka* *spało*
Five bird.M.GEN.SG slept.N.SG

When followed by mass nouns, the noun is still required to appear in the plural:

(15) a. *dwie* *kawy / *kawa*
two.F coffie.F.NOM.PL / *coffie.F.NOM.SG
"two coffies / *coffee"
b. *pięć* *kaw / *kawy*
five coffee.F.GEN.PL / *coffee.F.GEN.SG
"five coffies / *coffee"

This behavior contrasts with the numeral 1, which has no such restriction.[4] With *pluralia tantum* nouns, like 'door', the numeral appears in the plural, as does the verb.

(16) *Jedne* *drzwi* *zapadły* *się*
One.NOM.PL door.NOM.PL collapsed.NON-VIR.PL REFL
"One door collapsed/fell"

As it turns out, even if the noun is not *pluralia tantum*, but just in the plural, 1 can still agree in number, taking on the meaning of existential 'some'.

(17) *Jedne* *dziewczyny* *spały*
One.F.NOM.PL girl.F.NOM.PL slept.NON-VIR.PL
"Some girls slept"

Thus, it appears that numeral 1 is free to modify any count noun, in any number or gender, and the verb will agree accordingly. In this sense, it differs from 2,3,4 and 5+ numerals, which are restricted to plural nouns and agreeing or non-agreeing verbs (depending on the numeral). This may be evidence for an inherent plural feature with 2,3,4 and 5+ numerals. Additionally, it also suggests that numeral 1 and numerals 2,3,4 and 5+ differ in terms of the number feature, where numeral 1 allows its number feature to be valued by the noun after it, while 2,3,4 and 5+ do not.

To summarize: numeral 1 agrees with the noun, regardless of its number feature, while numerals 2,3,4 and 5+ require a plural. Numeral 1 leads to

[4] In this sense, Polish 1 also differs from English 1 – whereas Polish 1 can modify plural nouns, English 1 cannot:

i. **one scissors*
ii. *one pair of scissors*

singular (or plural) agreement, numerals 2,3,4 to plural agreement, numerals 5+ to neuter singular and numerals 1000+ to either neuter singular or masculine singular. This is evidence for a four-way distinction among the numerals.

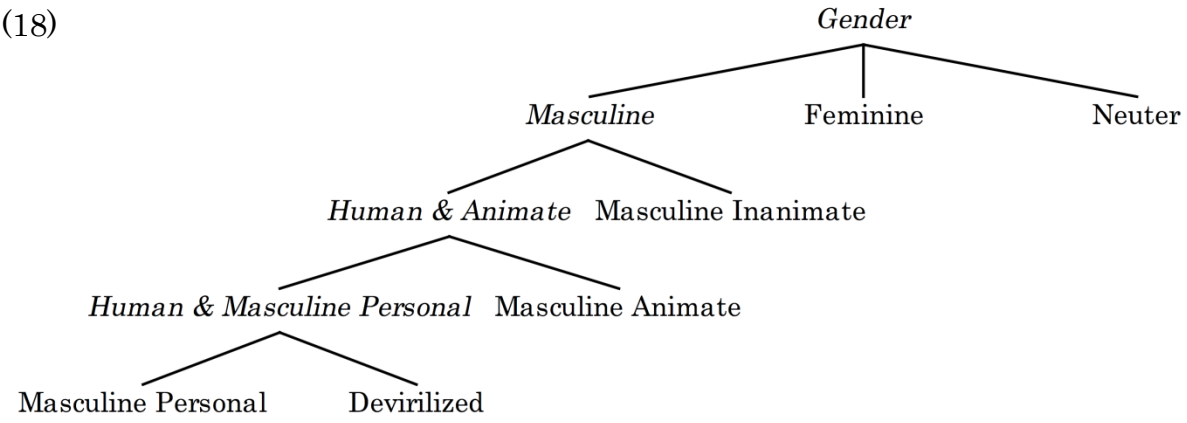
2.2 Masculine Personal Agreement

This section concerns agreement when the noun has masculine personal gender, i.e. grammatically masculine nouns which refer to a human. To address the data properly, it is necessary to first discuss the gender system in Polish.

2.2.1 Gender in Polish

Polish has three main grammatical genders: masculine, feminine, and neuter. Within the masculine gender, there are four subdivisions, making a total of six genders (Corbett, 1983). These subdivisions are: masculine inanimate, masculine animate, masculine personal, and devirilized. This can be visualized as the following tree (redrawn from Corbett, 1983: 87).

(18)



Masculine personal gender consists of nouns which are masculine, and refer to a human (*man* 'mężczyzna', *uncle* 'wujek', etc.). Devirilized nouns (*dwarf* 'karzeł', *slob* 'brudas', etc.) also refer to masculine humans, but they fail to get virile agreement outside of the NP, e.g. with adjectives or verbs. Masculine animate nouns refer to animals (*dog* 'pies', *cat*, 'kot', etc.), although there are various non-animate exceptions (*banana* 'banan', *checkmate* 'mat'). The set of masculine inanimate nouns, the largest masculine set, are the residue; basically

any noun which is not masculine personal (human) or masculine animate (alive) is masculine inanimate (*table* 'stół', *house* 'dom', *computer* 'komputer', etc.).

The masculine genders have a subset relationship – masculine inanimate, masculine animate, and masculine personal are all subsets of masculine proper. This subset relationship is reflected in the morphology. In the genitive, nominative, and accusative cases, there are visible distinctions for the different subgenders, but in the dative, instrumental, and locative, those distinctions are all collapsed and the same morphology is used. This is schematized in table 4.

Table 4: Case and gender morphology for masculine nouns[5]

	Nom	Acc	Gen	Dat	Instr	Loc
Masc Inan sg	0	0	-u	-owi	-em	-'e / u
Masc Anim sg	0	-a	-a	-owi	-em	-'e / u
Masc Pers sg	0	-a	-a	-owi	-em	-'e / u
Masc Inan pl	-e / y	-e / y	-ów / y	-om	-ami	-ach
Masc Anim pl	-e / y	-e / y	-ów / y	-om	-ami	-ach
Masc Pers pl	-e / 'y	-ów / y	-ów / y	-om	-ami	-ach

Note: Singular endings are presented before plural endings
Symbols: / (slash) indicates different morphological possibilities, conditioned phologically. ' (apostrophe) indicates a phonological phenomenon known as "softening". This is usually done through a palatalization of the final consonant before the morpheme is added. 0 (zero) indicates a null ending.

In the dative, instrumental, and locative columns, the morphology for each subgender is identical, for a given number. The nominative, accusative, and genitive columns, however, differ depending on the subgender. Thus, we do see evidence for subgenders, and importantly, evidence that these subgenders belong to a single masculine gender.

Notice that the morphologically relevant subgender varies between the singular and plural. In the singular, it is the animate/personal – inanimate distinction which plays a role (in the accusative and genitive cases only), but in the plural, it is the inanimate/animate – personal distinction which is important (in the nominative and accusative genders only). These same distinctions are also found in adjectives and similar ones in verbs. In the verbal domain, the

[5] I omit the devirilized gender in the table here, as its case morphology is identical to the masculine personal gender. It is only in the adjectival and verbal domains that a difference between the two is seen – masculine devirilized nouns pattern with masculine animate nouns for adjectival and verbal agreement.

singular does not care about any of the subgenders, unlike the nominal domain. All masculine nouns are marked with the masculine morpheme, regardless of subgender. The plural, however, makes the same distinctions as the nominal domain – masculine personal subjects are marked specially to the exclusion of non-masculine personal subjects. Yet, whereas the nominal domain also makes distinctions between feminine and neuter, in the verbal domain all non-masculine personal subjects are collapsed. Verbs take virile agreement (-i) if there is at least one masculine personal subject; otherwise, they take non-virile agreement (-y). For numerals, this same distinction between masculine personal and non-masculine personal is vital to their behavior, and the discussion below.

The importance of this section is this: masculine nouns differ in their behavior depending on whether they are animate or human. This suggests that in addition to the usual number, gender, and person phi-features, Polish also has "human" and "animate" encoded in some way in the feature system. I claim that the feminine gender (at least) also has these sub-divisions, although they are not realized morphologically. My evidence for this comes from the following paradigm:

(19) a. *Kobieta i dziewczyna* ***spały*** */ *spali*
Woman.F.SG and girl.F.SG slept.NON-VIR.PL / slept.VIR.PL
"A woman and a girl slept"

b. *Kot i pies* ***spały*** */ *spali*
Cat.M.SG and dog.M.SG slept.NON-VIR.PL / slept.VIR.PL
"A cat and a dog slept"

c. *Kobieta i pies *spały /* ***spali***
Woman.F.SG and dog.M.SG slept.NON-VIR.PL / slept.VIR.PL
"A woman and a dog slept"

d. *Kot i dziewczyna *spały /* ***spali***
Cat.M.SG and girl.F.SG slept.NON-VIR.PL / slept.VIR.PL
"A cat and a girl slept"

In example (19a), the coordination of two feminine (human) nouns produces non-virile (or non-masculine personal) verb agreement. Similarly, in example (19b), two masculine animate nouns are coordinated – again, the result is non-virile agreement. However, if one of the feminine human nouns is coordinated with one

of the masculine animate nouns (19c, 19d), then virile agreement occurs. Thus, "human" is presumably encoded in the feminine conjunct and "masculine" in the masculine animate conjunct. The combination of the two produces virile agreement. Importantly, this suggests that human and masculine are independently encoded in the feature system.[6] Additionally, it suggests that feminine nouns also encode human. These examples argue for a feature which can capture human separate from a gender feature.

The feature system I will adopt for Polish gender is based on Brown (1998). Brown distinguished two features, namely GENDER and SUBGENDER. The values of GENDER can be *masculine, feminine, neuter,* or *masculine personal*, and the values of SUBGENDER can be *inanimate, animate*, or *person*. Masculine personal nouns would be a combination of *masculine personal* gender and *person* subgender, and masculine animate nouns a combination of *masculine* and *animate* for example. This system can derive Corbett's six Polish genders. They are illustrated in the diagram below, with the masculine genders bolded.

(20)

	Inanimate	Animate	Person
Masculine	**Masc.Inanim.**	**Masc.Anim.**	**Devirilized**
Feminine	Feminine	Feminine	Feminine
Neuter	Neuter	Neuter	Neuter
Masc.Pers.	X	X	**MasculinePersonal**

I turn now to the behavior of masculine personal nouns with numerals.

2.2.2 *Masculine Personal Nouns with Numerals*

With the masculine personal nouns, the numeral 1 acts as any adjective would, agreeing in case, number, and gender with the following noun. In the singular, it makes only the animate/person-inanimate distinction (21a), and in the plural, the masculine personal – non-masculine personal distinction (21b).

[6] These examples also have interesting ramifications for agreement: either (1) both features are somehow realized on the &P through feature resolution or (2) there are separate probes for each feature. As interesting as this is, I will not have the chance to address it here, and leave it for future research.

(21) a. *Jeden* *młody* *chłopiec* *spał*
One.MP.NOM.SG young.MP.NOM.SG boy.MP.NOM.SG slept.M.SG
"One young boy slept"
b. *Jedni* *młodzie* *chłopcy* *spali*
One.MP.NOM.PL young.MP.NOM.PL boy.MP.NOM.SG slept.VIR.PL
"Some young boys slept"

This is an important piece of evidence for classifying numeral 1 as an adjective.

In contrast to the numeral 1, numerals 5+ produce additional morphology with masculine personal nouns. Numerals 5+ have only three morphological forms: the nominative/accusative *pięć*, the instrumental *pięcioma*, and the oblique *pięciu* (see table 3). The oblique form is found with the genitive, dative, and locative cases. Additionally, this is the form which is found with masculine personal nouns in structural case positions. Compare (22a) with (22b). The use of a masculine personal noun changes the case morphology on the numeral to oblique:

(22) a. *Pięć* *ptaków* *spało*
Five.NOM/ACC bird.M.GEN.PL slept.N.SG
"Five birds slept"
b. *Pięciu* *chłopców* *spało*
Five.OBL boy.MP.GEN.PL slept.N.SG
"Five boys slept"

The oblique form of the numeral occurs with masculine personal nouns, but not with regular masculine nouns. From this example, it is unclear what case appears on the numeral, due to the numerous syncretisms. I will address this question in detail later.

The behavior of numerals 2,3,4 falls in between that of numeral 1 and numerals 5+. In subject position, numerals 2,3,4 have two strategies for marking agreement with masculine personal nouns. The first is a special nominative form, available only with masculine personal nouns in the nominative, which allows agreement throughout the numeral-noun construction and with the verb.

(23) *Dwaj* *chłopcy* *spali*
Two.MP.NOM boy.MP.NOM.PL slept.VIR.PL
"Two boys slept"

The second strategy involves a non-nominative case, either genitive or accusative (again, obscured by the syncretisms). This case appears throughout the numeral-noun construction, and similarly to 5+ numerals, blocks verb agreement.

(24) *Dwóch* *chłopców* *spało*
Two.MP.GEN/ACC boy.MP.GEN/ACC.PL slept.N.SG
"Two boys slept"

Thus, of the two strategies with numerals 2,3,4, one leads to full agreement, and the other to an agreement mismatch. According to Swan (2002), these variants differ only stylistically, the first being reserved more for written language than spoken language.

In object position, there is only one strategy possible, that being the genitive-accusative syncretism:

(25) *Wiedziałam* *dwóch* *chłopców*
I[F].saw two.MP.GEN/ACC boy.MP.GEN/ACC.SG
"I saw two boys"

Lastly, the 1000+ numerals appear in the nominative case in subject position, with the masculine personal noun in the genitive (26).

(26) *Tysiąc* *chłopców* *spało*
Thousand.M.NOM.SG boy.MP.GEN.PL slept.N.SG
"A thousand boys slept"

The behavior of 1000+ numerals does not vary with the gender of the following noun.

To summarize: numeral 1 agrees with masculine personal nouns and there is full verb agreement. There are no alternative strategies with this numeral, and it behaves the same regardless of gender. The same can be said for 1000+ numerals, which still appear in the nominative, with the noun in the

genitive. Numerals 5+ exhibit additional morphology, which is morphologically related to the oblique case; verbs still do not agree. Numerals 2,3,4 have two strategies in subject position – the first patterns with numeral 1 and leads to full agreement; the second patterns with numerals 5+ and produces non-nominative case marking throughout the numeral-noun construction, leading to an agreement mismatch.

2.3 Case Assignment

Numerals 5+ seem to assign a genitive case to the following nominal, as do numerals 1000+.

(27) a. *Pięć* *ptaków* *spało*
Five.NOM/ACC bird.M.**GEN**.PL slept.N.SG
"Five birds slept"
b. *Tysiąc* *ptaków* *spało*
Thousand.M.NOM.SG bird.M.**GEN**.PL slept.N.SG
"A thousand birds slept"

This behavior does not occur with the numeral 1 or numerals 2,3,4. These numerals show homogeneous case throughout the numeral-noun construction:

(28) a. *Jeden* *ptak* *spał*
One.M.NOM.SG bird.M.NOM.SG slept.M.SG
"One bird slept"
b. *Dwa* *ptaki* *spały*
Two.M.NOM bird.M.NOM.PL slept.NON-VIR.PL
"Two birds slept"

The case assigning property of 5+ numerals is limited to structural case positions. In oblique case positions, the genitive case disappears, and instead, the oblique case appears on both the numeral and the noun; this does not change for masculine personal gender. This pattern of homogeneous case assignment in oblique positions is also found with numerals 1,2,3,4. In contrast to numerals 5+, numerals 1000+ do not lose their case assigning abilities in oblique positions.

(29) a. *Spałam z jednym kotem*
I[F].slept with[INSTR] one.M.INSTR.SG cat.M.INSTR.SG
"I slept with one cat"

b. *Spałam z dwoma kotami*
I[F].slept with[INSTR] two.M.INSTR cat.M.INSTR.PL
"I slept with two cats"

c. *Spałam z pięcioma kotami*
I[F].slept with[INSTR] five.M.INSTR cat.M.INSTR.PL
"I slept with five cats"

d. *Spałam z tysiącem kotów*
I[F].slept with[INSTR] thousand.M.INSTR.SG cat.M.GEN.SG
"I slept with a thousand cats"

To summarize: numerals 1,2,3,4 always act as case-agreers, in both structural and oblique case positions; they consistently produce homogeneous case throughout the numeral-noun construction. Numerals 1000+, on the other hand, always function as case assigners, regardless of their position. In contrast, numerals 5+ alternate between acting as case-assigners and case-agreers. In structural positions, they assign a genitive case to the following noun, producing a heterogeneous case environment; in oblique positions, they seem to agree in case, producing a homogeneous case environment. This is summarized in the table below:

Table 5: Numeral case assignment versus case agreement

	Structural position	Oblique position
1	Agree	Agree
2,3,4	Agree	Agree
5+	Assign	Agree
1000+	Assign	Assign

2.4 Adjectives and Nouns

I will only consider adjectives and nouns in the context of the previous sections. I begin with adjectives, then move on to nouns. To maximize the comparison between numerals, and adjectives and nouns, all adjectives are given as modifiers of nouns, and all nouns in noun-complement structures.

2.4.1 Adjectives

With regards to agreement with non-masculine personal nouns, adjectives agree in case, gender, and number with the noun following them. The verb in turn agrees with both the adjective and noun in number and gender. Adjectives and nouns have nominative case in subject position.

(30) a. *Ładna* *dziewczyna* *spała*
pretty.F.NOM.SG girl.F.NOM.SG slept.F.SG
"The pretty girl slept"
b. *Ładne* *dziewczyny* *spały*
pretty.F.NOM.PL girl.F.NOM.PL slept.NON-VIR.PL
"The pretty girls slept"

There are no agreement mismatches with non-masculine personal gendered nouns modified by adjectives.

With regards to masculine personal gender, again adjectives show full agreement, as do verbs. Both adjective and noun appear as nominative.

(31) *Młodzi* *chłopcy* *spali*
Young.MP.NOM.PL boy.MP.NOM.PL slept.VIR.PL
"The young boys slept."

Again, in these examples, adjectives do not lead to any agreement mismatches.

Finally, regarding case – adjectives do not act as case assigners; they always agree, regardless of their position, whether it be structural (32a) or oblique (32b).

(32) a. *Czerwony* *kot* *spał*
Red.M.NOM.SG cat.M.NOM.SG slept.M.SG
"The red cat slept"
b. *Spałam z* *czerwonym* *kotem*
I[F].slept with[INSTR] red.M.INSTR.SG cat.M.INSTR.SG
"I slept with the red cat"

2.4.2 Nouns

Polish nouns do not present any surprises either. In noun-complement constructions, the verb always agrees with the head noun; it makes no difference whether the lower noun is masculine personal (33b) or not (33a). The head noun always takes nominative case in subject position.

(33) a. *Student* *fizyki* *spał*
student.MP.NOM.SG physics.F.GEN.SG slept.M.SG
"The student of physics slept"
b. *Studenki* *bratów* *spały*
student.F.NOM.PL brother.MP.GEN.PL slept.NON-VIR.PL
"(My) brothers' students slept"

In terms of case assignment, nouns assign genitive to the second noun; there is no number requirement with noun complementation, like there is for numerals. This can be seen in the above example, where it is grammatical to have a singular (33a) or plural (33b) complement. Genitive case is still assigned in oblique positions:

(34) *Spałam* *z* *kotem* *brata*
I[F].slept with[INSTR] cat.M.INSTR.SG brother.M.GEN.SG
"I slept with (my) brother's cat"

This concludes the section on nouns and adjectives. In the next section, I address the categorization of numerals into four classes.

2.5 Classifying the Numerals

The four classes of numerals I promote here are numeral 1, numerals 2,3,4, numerals 5+, and numerals 1000+.

Numeral 1 agrees in case, number, and gender with the following noun, regardless of its position or the gender or number of the noun. Verb agreement is always successful with this numeral. Notice that this behavior is identical to the behavior I just described for adjectives. For this reason, I consider the numeral 1 to be an adjective.

Numerals 5+ clearly differ from the other numerals. They assign genitive to the following noun in structural positions (unlike 1,2,3,4, but like 1000+), while agreeing in oblique positions (like 1,2,3,4, but unlike 1000+); additionally they always block verb agreement (unlike 1,2,3,4, and 1000+). Lastly, they have special morphology with masculine nouns (like 2,3,4, but unlike 1 or 1000+).

Numerals 2,3,4 are like a hybrid of the numeral 1 and numerals 5+. Similarly to 1, they agree in case, gender (to some extent), and often lead to successful verb agreement. However, there are a few key points in which they differ, patterning with 5+ numerals instead. First of all, they require the noun to be plural (unlike 1), and secondly, they have an additional strategy with masculine personal nouns that leads to an agreement mismatch.

Numerals 1000+ are very similar to nouns – they lack the special morphology with masculine personal nouns, consistently have a genitive noun following them, and do show verbal agreement. For these reasons, I am most inclined to regard them as nouns and will treat them as such here.

In order to push this difference between numeral 1 and numerals 2,3,4, which has not yet been recognized in the literature, I present some further evidence. There is another domain in which numerals 2,3,4 pattern with 5+ instead of 1. Polish has a distributive marker *po* which assigns locative case to following nouns. With numerals 2+ however, the case is not locative, but accusative.[7] Note that this is specifically related to the fact that these are numerals, and not that they are plural (Przepiórkowski, 2006).

[7] Case here depends on the position – in nominative position, it would be nominative (Przepiórkowski, 2010)

(35) *Dałam* *każdej* *dziewczynie …*
I[F].gave every.DAT girl.DAT …
"I gave each girl …"

a. *po* *cukierku*
DIST candy.M.LOC.SG
"a candy"

b. *po* *jednym* *cukierku*
DIST one.M.LOC.SG candy.M.LOC.SG
"one candy"

c. *po* *dwa* *cukierki*
DIST two.M.NOM/ACC candy.M.ACC.PL
"two candies"

d. *po* *pięć* *cukierków*
DIST five.NOM/ACC candy.M.GEN.PL
"five candies"

e. *po* *tysiąc* / *tysiącu* *cukierków*
DIST thousand.NOM/ACC / thousand.LOC candy.M.GEN.PL
"thousand candies"

With a regular noun (no numeral), locative case appears on the noun following *po* (35a). This same locative case appears with the numeral 1 (35b). With numerals 2,3,4 (35c) and 5+ (35d), however, the case on the numeral is accusative, and cannot be locative. Numerals 1000+ allow both the locative and accusative (35e). Thus, for the distributive marker *po*, 2,3,4 and 5+ numerals pattern together (along with 1000+ numerals somewhat). This supports the distinction between 1 and 2,3,4.

When looking at conjunction, we see even more evidence for this division between 1 and 2,3,4 with 5+, providing support also for the four-way distinction. The following examples demonstrate this:

(36) a. *Pięć kobiet i jedna dziewczyna siedziały / ?siedziało*
Five women.GEN and one.NOM girl.NOM sat.NON-VIR.PL / sat.N.SG
"Five women and one girl sat"

b. *Sześć kobiet i dwie dziewczyny nosiły / nosiło brązowe sukienki*
Six women.GEN and two girls.NOM wore.NON-VIR.PL / wore.N.SG brown.ACC dresses.ACC
"Six women and two girls wore brown dresses"

(37) a. *Jedna kobieta i pięć dziewczyn siedziały / siedziało*
One.NOM woman.NOM and five girls.GEN sat.NON-VIR.PL / sat.N.SG
"One woman and five girls sat"

b. *Cztery kobiety i osiem dziewczyn *siedziały / siedziało*
Four.NOM woman.NOM and eight girls.GEN sat.NON-VIR.PL / sat.N.SG
"Four women and eight girls sat"

In the examples in (36), a 5+ numeral is conjoined with a 1,2,3,4 numeral. In (36a), it is conjoined with 1, and in (36b), with the numeral 2. In the examples in (37), the same occurs, but the order of conjunction is different, with the 1,2,3,4 numerals in the first conjunct, and the 5+ numeral in the second. These four sentences were given to ten Polish native speakers who were asked to fill in the verb form.[8] The pattern that was found consisted of the following: with the order (5+ and 1,2,3,4), plural agreement was found 90% of the time with numeral 1 and 60% of the time with numeral 2. With the order (1,2,3,4 and 5+), plural agreement was found 55%[9] of the time with numeral 1 and 0% of the time with numeral 4. Thus, this constitutes an important difference between numeral 1 and numerals 2,3,4 – numeral 1 can resolve agreement with a 5+ numeral in both first and last conjunct position, while numerals 2,3,4 can only do so in last conjunct position. The conjunction of two 5+ numerals consistently failed (95% failure), further promoting the distinctions between these numerals. The point is that numerals 1 and 2,3,4 differ with regards to feature resolution in coordination.

To summarize these findings more succinctly, I present Table 6 below. This highlights the similarities and differences between the four types of numerals, comparing them to nouns and adjectives.

[8] Questionnaire participants aged from 23 to 53 (age 23 (1), age 27 (5), age 28 (2), age 29 (1), age 53 (1)), were primarily male (8 male, 2 female), and lived currently in Kraków (originating from cities in the three most southern states in Poland, e.g. Małopolska (cities: Kraków (2), Gorlice (4), Proszowice (1)), Podkarpackie (cities: Jarosław (1) and Przeworsk (1)), and Śląskie (city: Bielsko-Biała (1)))

[9] One response was excluded as it came from a previous version of the questionnaire in which the order of subject and verb was switched. 5/9 informants choose plural with numeral 1.

Table 6: Numeral classification

	Adj	1	2,3,4	5+	1000+	Noun
Verb Agreement (non-masculine personal)						
Full	x	x	x		x	x
Mismatch				x	x†	
Verb Agreement (masculine personal)						
Full	x	x	x		x	x
Mismatch			x	x		
Case of numeral (non-masculine personal; subj)						
Oblique*						
Nom/Acc	x	x	x	x	x	x
Case of numeral (masculine personal; subj)						
Oblique*			x	x		
Nom/Acc	x	x	x		x	x
Case Assignment (structural position)						
Agree	x	x	x			
Assign				x	x	x
Case Assignment (oblique position)						
Agree	x	x	x	x		
Assign					x	x
Case after *po*						
Locative	x	x			x	x
Accusative			x	x	x	

** For numerals 1,2,3,4, and 1000+ oblique here refers to the genitive/accusative form; for numerals 5+ this is the oblique form, a syncretism of the genitive, accusative, dative, and locative*
† c.f footnote 3 – I do not address this issue in this paper.

The table above highlights a number of interesting things. First of all, numeral 1 patterns with adjectives across the board. Secondly, numerals 1000+ pattern for the most part with nouns, with only a few exceptions – however, they *can* pattern completely with nouns. Lastly, there are phenomena which are completely unique to numerals, and these are the agreement mismatches (for all genders), the oblique case marking with masculine personal gender, and the accusative marking with *po*.

2.6 Quantifiers

Up until this point, I have failed to say anything specific about quantifiers. Essentially, there are two broad classes of quantifiers which pattern identically to numerals.

The first group of quantifiers patterns with numeral 1. This includes quantifiers *niektóry* (some), *żaden* (none), and *każdy* (every), among others. These quantifiers agree in case, number, and gender with the following nominal (regardless of the gender), and always lead to successful verb agreement.

(38) a. *Niektóra dziewczyna spała*
Some.F.NOM.SG girl.F.NOM.SG slept.F.SG
"Some girl slept"

b. *Niektórzy chłopcy spali*
Some.MP.NOM.PL boy.MP.NOM.PL slept.VIR.PL
"Some boys slept"

c. *Spałam z niektórymi kotami*
I[F].slept with[INSTR] some.M.INSTR.PL cat.M.INSTR.PL
"I slept with some cats"

Example (38a) demonstrates agreement with a feminine singular noun – the verb agrees as expected. With a plural masculine personal noun (38b), there is no mysterious case marking, as there is with numerals 2,3,4; additionally the verb agrees. Lastly, the quantifier agrees also in oblique case positions (38c). Thus, these quantifiers pattern with numeral 1.

The second group of quantifiers patterns with numerals 5+. Like 5+ numerals, they assign a genitive case in structural positions (39a,b) and agree in case in oblique positions (39c). With masculine personal nouns, additional morphology appears on the numeral (39b). Additionally, verb agreement always fails with these quantifiers (36a,b). This group includes quantifiers such as *wiele* (many), *kilka* (a few), and *ile* (how many). Like 5+ numerals, each of the quantifiers has only three forms, the nominative-accusative (*wiele, kilka, ile*), the instrumental (*wieloma, kilkoma, iloma*), and the oblique (*wielu, kilku, ilu*). These quantifiers are indistinguishable from 5+ numerals in behavior.

(39) a. *Wiele* *dziewczyn* *spało*
Many.NOM girl.F.GEN.PL slept.N.SG
"Many girls slept"
b. *Wielu* *chłopców* *spało*
Many.OBL boy.MP.GEN.PL slept.N.SG
"Many boys slept"
c. *Spałam* *z* *wieloma* *kotami*
I[F].slept with[INSTR] many.INSTR cat.M.INSTR.PL
"I slept with many cats"

There are no quantifiers which pattern with the numerals 2,3,4 or 1000+.

2.7 Summary

What this section has shown is that there are four main classes of numerals, numeral 1, numerals 2,3,4, numerals 5+, and numerals 1000+. There are no quantifiers that pattern with numerals 2,3,4 or 1000+, although there are quantifiers which pattern with 1 and 5+. The class of the numeral, the gender of the following noun (whether masculine personal or not), and the case position are all relevant factors in determining the form of the numeral and noun, as well as whether agreement will be successful or not.

When it comes to agreement, there are two configurations that lead to an agreement mismatch: (1) anything with 5+ numerals and (2) any masculine personal nouns with 2+ numerals. This suggests that 5+ numerals independently lead to agreement breakdowns as do masculine personal nouns.

This leads to a number of important questions which I will address in the following chapters:

(a) Why do 5+ numerals result in agreement mismatches?

(b) Why do masculine personal nouns result in agreement mismatches?

(c) What is the source of the case on numerals with masculine personal nouns?

(d) How can we account for the case assigning-case agreeing alternation?

3 Previous Solutions to the Polish Numeral Puzzle

In this chapter, I look at previous solutions to the Polish data; many of the solutions I will discuss focus on the 5+ numerals, ignoring the 1,2,3,4 and 1000+ numerals. Previous research has made use of the numeral's case feature and syntactic category to explain the lack of verbal agreement, attributing the oblique morphology on the numeral with masculine personal nouns to gender agreement and the lack of verb agreement to an accusative case on the numeral.

There are three main sections here. In section 3.1, I address three theories about the case feature of the numeral; I adopt the third in this thesis, the nominative-genitive hypothesis which states that numerals are genitive with masculine personal nouns, but nominative with non-masculine personal nouns. Section 3.2 looks at the syntactic category; I will argue that numerals are semi-lexical nouns, after discussing alternative hypotheses that numerals are QPs, NPs, or adjectives with a silent noun. Lastly, section 3.3 addresses the case-assigning, case-agreeing alternation.

3.1 Case Feature on the Numeral

One of the prevalent hypotheses concerning Polish numerals is that 5+ numerals sit in the accusative case (Franks, 1994, 2002; Przepiórkowski, 2004; Rutkowski, 2002). Under the assumption that the numeral is the head of the construction, this results in a simple explanation for the lack of verbal agreement – the case feature in the subject is already valued for accusative (and the noun for genitive), so there is no active goal for the verb to agree with. As a result, verbal agreement fails, but instead of the derivation crashing, the phi-features of the verb are given a default value, namely neuter singular. The Accusative Hypothesis is only relevant for explaining the pattern of verbal agreement in Polish (compared to Serbo-Croatian and Russian – Franks, 1994).

The Accusative Hypothesis contrasts with two other hypotheses in the literature (Przepiórkowski, 2004): the Nominative Hypothesis and the Nominative-Genitive Hypothesis. Under the Nominative Hypothesis, the claim is

that all numerals are in the nominative case. The Nominative-Genitive Hypothesis takes a different approach; it takes into account the agreement differences between masculine personal nouns and non-masculine personal nouns, claiming that numerals are nominative with non-masculine personal nouns and genitive with masculine personal nouns. This is the position I will take in this paper.

3.1.1 The Accusative Hypothesis

As Franks (1994) points out, morphologically, there is nothing to rule out the Accusative Hypothesis. With masculine personal nouns, there is a genitive-accusative syncretism and with non-masculine personal nouns, a nominative-accusative syncretism. By this, we would expect to find masculine personal numerals in a form that resembles the genitive and non-masculine personal numerals in a form that resembles the nominative, despite the fact that both are accusative. This is indeed found. As mentioned in section 2.2.2 above, 5+ numerals have three forms, the nominative/accusative, the instrumental, and the oblique (see table 3). With non-masculine personal nouns, the nominative/accusative form is used and with masculine personal nouns, the oblique form. Example (22) is repeated below.

(40) a. *Pięć* *ptaków* *spało*
Five.NOM/ACC bird.M.GEN.PL slept.N.SG
"Five birds slept"
b. *Pięciu* *chłopców* *spało*
Five.OBL boy.MP.GEN.PL slept.N.SG
"Five boys slept"

By this hypothesis, the case on the numerals in the example above would be accusative.

A second argument in favor of the Accusative Hypothesis (and contra the Nominative Hypothesis) concerns demonstratives (Przepiórkowski, 2004). Demonstratives can appear in one of two cases – the genitive (apparently agreeing with the lower noun) or the nominative-accusative (apparently agreeing

with the numeral). Crucially, however, with masculine personal-nouns they cannot appear in the nominative, but only the accusative-genitive.

(41) a. *Te / tych pięć dziewczyn spało*
These.NOM/ACC these.GEN five.NOM/ACC girl.F.GEN.PL slept.N.SG
"These five girls slept"

b. **Ci / tych pięciu chłopców spało*
These.MP.NOM these.GEN/ACC five.OBL boy.MP.GEN.PL slept.N.SG
"These five boys slept"

According to the proponents of this hypothesis, the case of the demonstrative should not vary based on gender. Since the nominative form is impossible with masculine personal nouns, it should be impossible with all genders. With non-masculine personal nouns, there is an accusative-nominative syncretism on the numeral – thus, if it cannot be nominative, then it must be accusative. This is one of the arguments for the Accusative Hypothesis and against the other hypotheses. It allows the hypothesis to provide a unified account of the case present on numerals – all numerals have accusative case in subject position.

The last and most important argument concerning this hypothesis is that it explains why verbs do not agree – with the numerals only appearing in the accusative, the verb is blocked from agreeing with them, hence, the lack of agreement.

The most obvious issue with the Accusative Hypothesis concerns this accusative case marking – in subject position, where does the accusative case come from? And why is it assigned before the nominative? For the most part, these questions have remained unanswered with regards to Polish numerals. Rutkowski (2004) suggests in a footnote that numerals simply lack a nominative case slot in their lexical entry. But as these numerals must take *some* case, they take the other remaining structural case, namely the accusative. Franks (2002) also addresses this issue in a footnote, suggesting there is some covert prepositional head immediately above the numeral which assigns the accusative case, thereby blocking nominative case from the verb; alternatively, he suggests that this accusative case might not even need licensing. Despite this agreement in the literature as to the case of numerals being accusative, there is little

evidence to explain *why* this might be the case and only minimal reference to possible solutions.

An additional argument against the Accusative Hypothesis concerns the lower numerals 2,3,4. Recall that with masculine personal nouns, these numerals are optionally able to take the genitive-accusative in subject position, with the result being that verb agreement is no longer possible. If we extend the analysis of 5+ numerals to 2,3,4 numerals, it is unclear why this accusative case would appear only with masculine personal nouns, as if it is conditioned by the gender of the noun. As mentioned above, one of the advantages of the Accusative Hypothesis is that it removes any stipulations based on gender and unifies the different forms of the numerals as being a single case. Thus, it makes little sense that with the lower numerals, this case marking would be conditioned by gender, whereas it is not with the higher numerals. If we turn to Rutkowski's claim, the reason the accusative appears is because the numerals lack a nominative form; however, numerals 2,3,4 *do* have nominative forms and *can* have verbal agreement. Additionally, the suggestion of Franks' does not get us any further. Why should it be the case that this prepositional head only appears when the noun is masculine personal for numerals 2,3,4 but with every gender for numerals 5+? Rutkowski's suggestion simply does not work for the lower numerals and Franks' loses the unification of the Accusative Hypothesis, and introduces the unwarranted stipulation that case is conditioned by gender, but only for certain numerals.

One alternative might be to assume that the numerals 2,3,4 and 5+ are entirely unrelated in behavior. However, this too is problematic. First of all, we lose the connection between numerals 2,3,4 and 5+ when it comes to masculine personal nouns. Secondly, with the loss of this connection, there is no longer an explanation (or even an inkling of an explanation) as to why there is this special behavior for masculine personal nouns with numerals 2,3,4, as it is supposedly entirely unrelated to the behavior of masculine personal nouns with numerals 5+. This is clearly disadvantageous. On the other hand, assuming the two phenomena are related allows us to treat them in a similar way, but under the Accusative Hypothesis, it remains to be explained why the accusative is possible

only with masculine personal nouns for numerals 2,3,4, but with all genders for numerals 5+.

3.1.2 The Nominative Hypothesis

Under the Nominative Hypothesis, all numerals in subject position are in the nominative case, regardless of gender. The problem with this hypothesis, however, is that if the numeral is valued for the nominative case, it is unclear why the verb cannot simply show agreement with the numeral. Additionally, there is the question of why the demonstrative cannot appear in the nominative case with masculine personal nouns. Example (41b) is repeated below:

(42)	**Ci*	*pięciu*	*chłopców*
	These.MP.NOM.PL	five.NOM	boy.MP.GEN.PL

If the numeral is in the nominative, the demonstrative should also be able to appear in the nominative. The fact that it cannot suggests that this hypothesis is incorrect.

3.1.3 The Nominative-Genitive Hypothesis

Under the Nominative-Genitive Hypothesis, the numeral is usually in the nominative case; however, if the numeral modifies a masculine personal noun, then it is in the genitive case.[10] Importantly, this hypothesis is able to explain the differences between masculine personal and non-masculine personal nouns for numerals 2,3,4 and numerals 5+. If the noun is masculine personal, then the numeral sits in the genitive case. For numerals 2,3,4, this predicts that there will be no agreement with the verb and there will appear genitive case morphology; likewise, for numerals 5+, this predicts that the oblique form of the numeral should be used, as it is. Furthermore, this hypothesis also provides an explanation for the lack of the nominative demonstrative with masculine personal nouns – since the numeral is in the genitive case, there is no nominative noun or numeral for the demonstrative to agree with; it is expected

[10] This hypothesis cannot apply as is to numerals 1 or 1000+ as they do not show additional morphology for masculine personal nouns. Presumably, they fall under the nominative hypothesis, although it is unclear if this is entirely true for 1000+ numerals.

that the nominative form cannot occur. Thus, the Nominative-Genitive Hypothesis is better able to explain the issues surrounding masculine personal gender.

There are, however, some remaining issues with this hypothesis. Przepiórkowski (2004) points out two relevant issues. First of all, if the 5+ numerals are in the nominative case, it is unclear why the verb cannot agree with them; this same argument applied to the Nominative Hypothesis. Secondly, there is the question of *why* there should be a different case with masculine personal and non-masculine personal nouns. However, as numerals 2,3,4 show, it *is* the case that masculine personal nouns can elicit a different case than non-masculine personal nouns. Presumably, the same is possible with 5+ numerals. There are no morphological reasons for not assuming this hypothesis, as the numeral forms are compatible – thus, it is only the theoretical issues that work against it: why should case be determined by gender and why can agreement with a nominative subject not occur? For now, I will not address either of these issues; I push them aside temporarily until Chapter 4.

In this paper, I follow the Nominative-Genitive Hypothesis and claim that numerals are nominative with non-masculine personal nouns and genitive with masculine personal nouns. For lower numerals, 2,3,4, there is an optionality in whether the numeral-noun construction appears in the nominative case (full verb agreement) or the genitive case (no verb agreement), whereas with the higher numerals, 5+, masculine personal nouns can *only* be in the genitive. Numerals 1 and 1000+ are treated here as adjective and noun, respectively, and the nominative-genitive hypothesis cannot apply; instead, I assume the Nominative Hypothesis for numerals 1000+ and an agree relationship between the numeral and noun for numeral 1.

3.2 The Categorial Status of Numerals

The cross-linguistic categorial status of numerals has been subject to much debate. According to Corbett (1978), lower numerals tend to be adjective-like and higher numerals noun-like; in-between numerals seem to have properties of both. For Polish, such a categorization is rather apt. Numeral 1 acts most like an

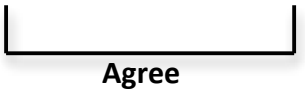
adjective (and arguably is an adjective). Likewise, numerals greater than 1000 act most like nouns, assigning genitive in both structural and oblique cases, and often leading to verb agreement with the numeral (Rutkowski, 2005). Numerals 5+ have the genitive assigning property of nouns, but lack verb agreement, and numerals 2,3,4 share properties with numeral 1 and numerals 5+. Thus, there seems to be this continuum between adjective and noun, where different numerals sit on different parts of the continuum.

In the literature, there is little discussion on the category of numerals 1,2,3,4 or 1000+. The lowest numerals (1,2,3,4) are often assumed to be adjectives and the highest numerals (1000+) nouns. For numeral 1 and numerals 1000+, I will follow these same assumptions. The numerals 5+ are those which lead to the most controversy. For the numerals 5+, there seem to be two main camps – either these numerals are QPs (Franks, 1994, 2002; Rutkowski, 2002; Rutkowsi & Szczegot, 2001), or they are NPs (Matushansky & Ionin, 2004). An alternative position, taken by Kayne (2005; 2007) for English *few* and extended by Zweig (2005) to numerals, argues that numerals are adjectives modifying a silent noun. I will explore each of these accounts in the following sections.

3.2.1 5+ Numerals are QPs

Franks (1994) makes a distinction between numerals that project to a QP and numerals that project to a full DP, assuming that in both cases, they are still QPs. He takes this to be one of the parameters governing the behavior of numeral phrases in Polish, Russian, and Serbo-Croatian. In Russian, verb agreement with 5+ numerals is either neuter singular or plural; this correlates with the position of the subject numeral phrase, whether it precedes the verb (agreement occurs) or follows the verb (agreement fails). This can be visualized in the diagrams below.

(43) a. [[Num [N]] V] *precedes = Agree*



b. [V [Num [N]]] *follows = no Agree*

To explain this correlation, Franks suggests that when numeral phrases project to QPs, they remain in a VP internal position as they do not require case and hence, do not need to raise to subject-position; thus, verbs cannot agree with them. In contrast, when they project to full DPs, they must raise to subject position for case, and hence, allow verb agreement. Numeral phrases always project to full DPs in oblique case contexts. The alternation is given below.

(44) a. $[_{TP}\ [_{\mathbf{DP}}$ Num-Noun$]_i\ [_{T'}$ T+V$_j\ [_{VP}\ t_i\ [_{V'}\ t_j$ … *precedes = DP + Agree*

b. $[_{TP}\ [_{T'}$ T+V$_i\ [_{VP}\ [_{\mathbf{QP}}$ Num-Noun$]\ [_{V'}\ t_i$ … follows = QP + no Agree

For Serbo-Croatian, the claim is that all numeral phrases project to full DPs, since the noun gets genitive case from the numeral regardless of structural position; however, the genitive case, which is inherent, somehow blocks the percolation of features up in Serbo-Croatian, thus blocking verb agreement. Finally, Franks extends his tests to Polish – by his account, all numeral phrases in Polish project to full DPs, and hence, they should agree. The lack of verb agreement, however, is due to an unrelated factor, namely the accusative case on the numeral (see section 3.1.1 for a discussion of the Accusative Hypothesis); as Franks (2002) later argues, QPs can only be licensed in accusative DPs. Thus, under this account of Franks (1994, 2002), numeral phrases in Polish are actually QPs, but in addition, they sit inside of a DP which also contains the noun, e.g.

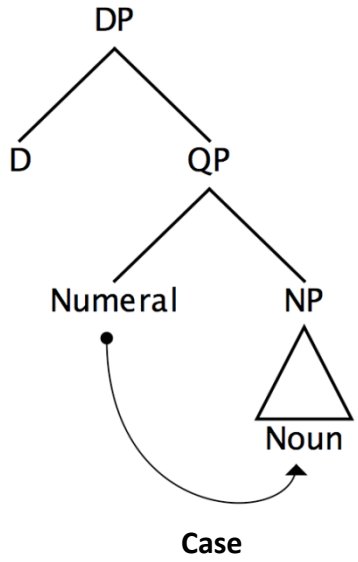
(45) $[_{DP}\ [_{QP}$ Numeral $[_{NP}$ Noun]]]

Note that this suggests that in non-accusative DPs (oblique environments), numerals do not occur as QPs, which forces them to be case-agreeing adjectives.

Rutkowski & Szczegot (2001) and Rutkowski (2002) also promote an analysis in which numerals are QPs. They attempt to support this idea by claiming that numerals are functionals, i.e. they belong to the class of functional elements. As functionals, they carry only minimal semantic meaning (cardinality) and they represent a closed class of items. This is essentially the basis of their argument – if numerals are functional, it is more plausible that they are not some lexical category, but instead a functional category, like QP.

Rutkowski (2002) also argues in favor of a three-layered DP, in which there is a DP at the top, the NP at the bottom, and the functional structure, the QP, in between. Numerals are to be found within the QP and nouns in the NP. Case is assigned via a head-complement relationship between the quantifier (numeral) and its complement. This is diagramed below.

(46)



As Rutkowski and Szczegot (2001) and Rutkowski (2002) point out, numerals do indeed behave like functional elements. This aspect of numeral behavior is well captured in these analyses. However, a disadvantage is that these analyses are forced to posit a new case, often referred to as quantitative or Gen-Q, which is only assigned by numerals to nouns. This case looks remarkably similar to the genitive (and as far as I know is identical to it in Polish), and

additionally, it constitutes case assignment *within* a DP, a domain in which case assignment does not usually occur. This makes the mechanism of case assignment and the new case rather suspect. Furthermore, this approach fails to capture the resemblance of this case assignment to the case assignment which occurs between two nouns. For these reasons, I will not adopt the QP analysis in this paper; however I will make use of the ideas concerning the functional status of the numeral.

3.2.2 5+ Numerals are NPs

The alternative approach states that numerals are actually NPs. This is the position held by Ionin & Matushansky (2004), and also promoted by Corver and Zwarts (2006) for Dutch. Similarly to the QP stance, Ionin & Matushanksy (2004) claim that the numeral is the head of the construction, and assigns case through the head-complement relationship. Essentially numerals can take other numerals in the complement position, as well as nouns. This allows for the formation of complex numerals, such as multiplicatives (e.g. *three hundred*, which is *three* x *hundred*) and additives (e.g. *hundred and five*, which is *hundred + five*). The structure of multiplicatives mimics that of simplex numerals with nouns:

(47) a. [[*three* NP] *men* NP]
b. [[[*three* NP] *hundred* NP] *men* NP]

Basically, since numerals are nouns, then just like regular nouns they can take other nouns as complements and assign a genitive case.[11]

Rutkowski (2005), however, argues that Polish numerals cannot be nouns. An important question he brings up is why it is that, with regards to case, 5+ numerals behave differently from 1000+ numerals and regular nouns. This

[11] Kayne (2005) objects to this structure, however, on the grounds that [[three hundred] men] seems to be a more natural structure than [three [hundred [men]]]. This idea is supported by topicalization and ellipsis (examples from Kayne 2005, examples 116-117 and 119-120):

i. a. *??Linguistics books they have three hundred*
b. **Hundred linguistics books they have three*
ii. a. *They have three hundred linguistics books and we have three hundred too*
b. **They have three hundred linguistics books and we have three, too.*

problem falls to the distinction between structural and oblique cases. Recall that whereas with 5+ numerals, the numeral assigns genitive case in structural environments and agrees in case in oblique environments, nouns and 1000+ numerals always assign genitive case, regardless of structural position.

(48) a. *Z* *tysiąc* *kotów*
With[INSTR] thousand.INSTR cat.GEN
"With a thousand cats"
b. *Z* *studentem* *fizyki*
With[INSTR] student.INSTR physics.GEN
"With a student of physics"

A second argument against the numerals as nouns approach, noted by Rutkowski (2005), is that Polish already has a set of noun-like numerals, which unlike 5+ numerals, do have verb agreement (49a), lack the strange agreement patterns with masculine personal nouns (49a), and do not enter into the case-assigning, agreeing alternation (49b) (see Rutkowski, 2005 for additional arguments that this set of numerals are in fact nouns).

(49) a. *Piątka* *Koreańczyków* *spała* *tutaj* *wczoraj*
Five.F.NOM.SG Korean.MP.GEN.PL slept.F.SG here yesterday
"(A) five (of) Koreans slept here yesterday"
b. *Ufam* *piątce* *Koreańczyków*
I[F].trust five.F.DAT.SG Korean.M.GEN.PL
"I trust (a) five (of) Koreans"

(Rutkowski, 2005: 108, 109)

According to Rutkowski, it is rather redundant for a language to have two sets of nominal numerals (especially considering these differ so much in behavior), and thus, numerals cannot be nouns.

What these two sections show is that while Polish numerals seem to be like nouns, they are not exactly like nouns. Because of this, researchers such as Franks, Rutkowski, and Szczegot have rejected the numerals-as-nouns approach, instead adopting a numerals-as-quantifiers approach. In this way, they capture the functional aspect of numerals. However, researchers such as Matushansky,

Ionin, Corver, and Zwarts are correct in noting that numerals *do* share a number of properties with nouns, most notable the genitive case assigning ability. When presented with the problem of numerals, researchers can choose to take one road or the other – numerals are lexical nouns or functional quantifiers, and attempt to wade their way through any remaining discrepancies.

In this paper, rather than choosing quantifiers or nouns, I will take a middle stance: I recognize that numerals share various properties with nouns, such as genitive case assignment (hence, I do not argue for a quantitative case or Gen-Q as under the QP stance), but at the same time, they differ from nouns, for many of the reasons that have been demonstrated in Chapter 2. I claim here, instead, that numerals are a functional noun, or rather, a semi-lexical category in the terms of Cover and Riemsdijk (2001). Thus, the lexical component carries the semantic meaning of cardinality (and allows them to assign genitive in the same way as nouns), while their functional component produces the phenomena that make them unique from nouns. For my account here, 5+ numerals are semi-lexical nouns.

3.2.3 Numerals are adjectives followed by a silent noun NUMBER

A final account I would like to discuss here is one by Zweig (2005), adapted from Kayne (2007). As I will discuss towards the end, this account is not applicable to the numerals 5+, although it is a possibility for numerals 2,3,4. I begin by discussing Kayne's data and its application to numerals.

Kayne (2007) proposes for lexical item *few* that there is a silent noun NUMBER which occurs immediately after it. Under this account, *few* is actually an adjective (supported by its ability to take degree morphology, *fewer, fewest*), and NUMBER is a singular noun. This accounts for the so-called *every*-construction, where *every* takes a singular noun (50a) except when *few* intervenes (50b), as well as the ability of *few* to sit in argument positions (50c):

(50) a. *every day / *days*
b. *every few days / *day*
c. *Few are intelligent these days*

This analysis of *few* gives it a morphology of adjectives, while allowing it to have the distribution of nouns, thus, solving the issue of its dual-natured behavior as both adjective and noun.

Zweig (2005) takes this hypothesis of Kayne for *few* and applies it to numerals, arguing that numerals are adjectives modifying a silent noun NUMBER. This is taken to provide an explanation for the simultaneously noun-like and adjective-like behavior of numerals. For example, similar to *few*, numerals also participate in the *every*-construction:

(51) a. *every two days / *day*
b. *co* *dwa* *dni* */ *dzień*
every two.M.NOM day.M.NOM.PL / day.M.NOM.SG
"every two days /*day"

Additionally, like adjectives, numerals 2,3,4 show gender agreement:

(52) a. *dwaj* *mężczyzni*
two.MP.NOM men.MP.NOM.PL
"two men" (masculine personal)
b. *dwie* *dziewczyny*
two.F.NOM girl.F.NOM.PL
"two girls" (feminine)
c. *dwa* *krzesła*
two.N.NOM chair.N.NOM.PL
"two chairs" (neuter)

This account posits a structure for numerals in which there are two noun phrases within the numeral-noun construction, one containing the adjectival numeral and silent noun and the other, the overt noun:

(53) $[_{NP}\ [_{NP}$ numeral [NUMBER]] noun]

This account does well to explain the agreeing behavior of numerals 2,3,4. However, it is clear that this same account cannot be applied to both numerals 2,3,4 and 5+. Numerals 5+ always assign a genitive case, while numerals 2,3,4 do so only in special cases (involving masculine personal nouns); furthermore, these numerals differ with regards to verb agreement, where numerals 2,3,4

show agreement and numerals 5+ do not. Thus, if we were to assume the silent noun NUMBER with all Polish numerals, we would be unable to explain the differences between these types of numerals, under the assumption that there is one and only one silent noun NUMBER. Numerals 2,3,4 are the most likely candidates for having this silent noun, as they do show adjectival agreement to some extent; numerals 5+, however, fail in this regard. Thus, I reject this account as being unlikely for the 5+ numerals, while I do maintain that it is a plausible explanation for numerals 2,3,4. Importantly, this account acknowledges the adjective-like behavior of the numerals 2,3,4.

In this paper, however, I prefer to develop an analysis of these two classes of numerals in which they are of the same category and have identical structures, while also deriving the variation found between them through lexical factors, e.g. specific properties of the numerals. With the silent noun NUMBER, differences in structure and category are immediately introduced – 2,3,4 numerals are adjectives and 5+ numeral semi-lexical nouns, and there is necessarily additional structure for 2,3,4 numerals; thus an identical account cannot be developed for these numerals, if NUMBER is assumed. For this reason, I do not adopt this approach, despite its availability for numerals 2,3,4.

To summarize this section, there are two major accounts in the literature for the treatment of Polish numerals in terms of categories – these are the numerals-as-nouns approach, and the numerals-as-QPs approach. Outside of the Slavic domain, the numerals-as-adjectives-modifying-a-silent-noun approach is also relevant, although it does not apply to the numerals 5+. In my discussion, I have declined from fully adopting any of these approaches, instead taking the middle route in which 5+ numerals are semi-lexical nouns, with both lexical and functional features, as are the numerals 2,3,4. With the goal of presenting a unified account for the 2,3,4 and 5+ numerals, I have set aside the numerals-as-adjectives approach. In section 4.7, I will give a tentative discussion on how the adjectival agreeing behavior of the 2,3,4 numerals can be derived even if they are treated as semi-lexical nouns.

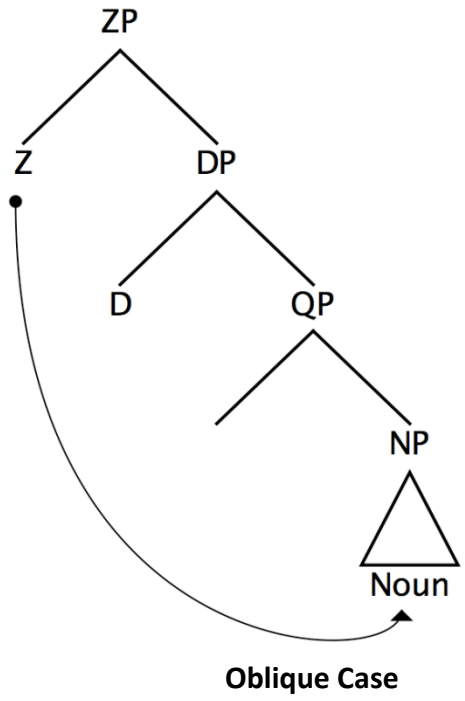
I turn now to the final sub-section of this chapter concerning the case agreement alternation.

3.3 The Case-Agreeing, Case-Assigning Alternation

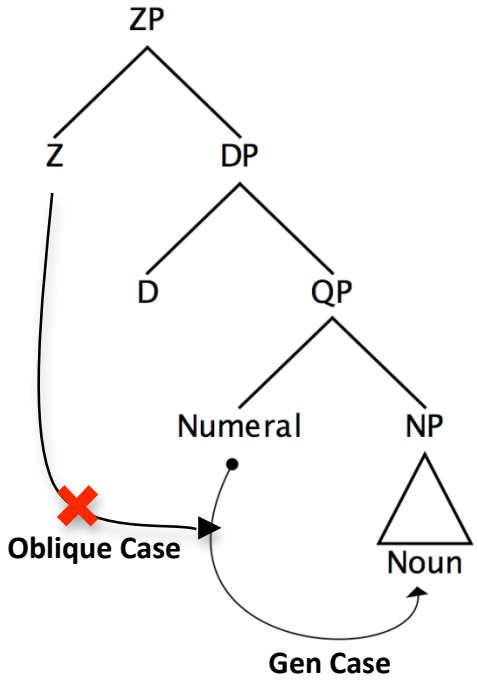
There have been numerous attempts to explain the case-agreeing, case-assigning alternation of Polish numerals. An early account is found in Franks (1994); Franks makes a comparison between Russian and Serbo-Croatian, applying those findings to Polish. Russian has the same case-agreeing, case-assigning alternation as Polish, while Serbo-Croatian does not; in Serbo-Croatian, the numeral *always* assigns genitive, regardless of structural position.

Franks argues that in Russian the genitive assigned by the numeral is a structural case, whereas in Serbo-Croatian it is inherent. Drawing on the differences between D-structure and S-structure (under the G&B framework), he argues that inherent cases are assigned at D-structure and structural cases at S-structure. Thus, for Russian, in oblique contexts, the inherent case is assigned *before* the numeral has a chance to assign its case, effectively blocking the assignment of genitive at S-structure; in structural positions, no other case is assigned to the numeral and genitive appears without issue. This produces the case-assigning, case-agreeing alternation. In Serbo-Croatian, on the other hand, both the genitive of the numeral and the inherent case of the oblique would be assigned in D-structure; the genitive, having a more local case assigner, wins over the oblique case, and thus surfaces. By this, genitive will always surface in Serbo-Croatian. The mechanism of this is illustrated below, with Z filling in for some inherent case case-assigner.

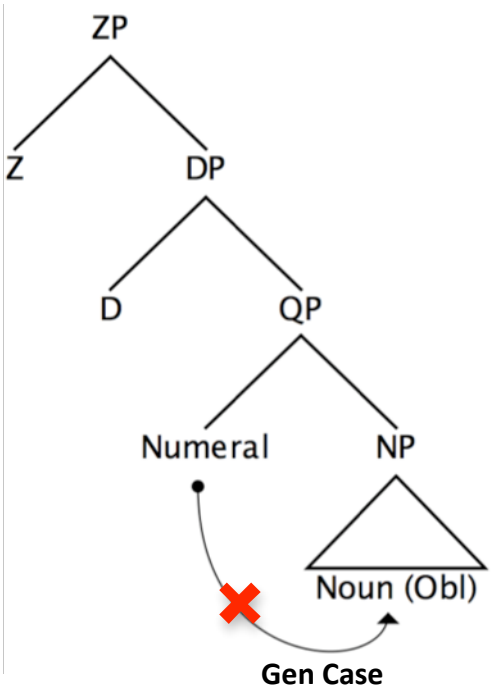
(54) D-Structure (Russian)



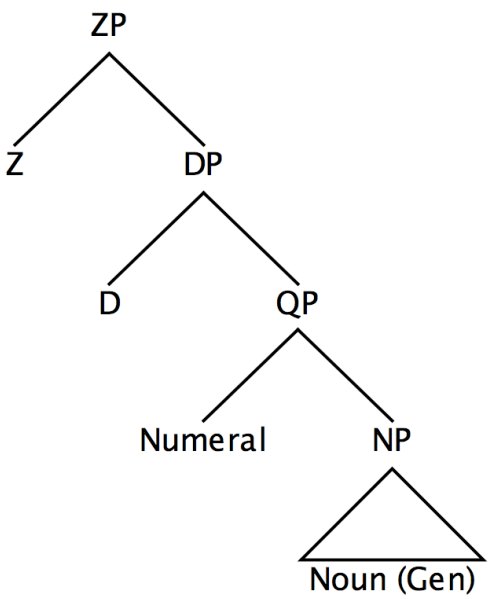
D-Structure (Serbo-Croatian)



S-Structure (Russian)



S-Structure (Serbo-Croatian)



Although Polish patterns like Russian in terms of the case-alternation, Franks argues it cannot work the same. His diagnostic for whether the case is inherent or not – if it gets overridden by inherent case assigners (then case is

structural) or not (then case is inherent) – indicates that the genitive is a structural case, since it is overridden by external oblique case assigners. However, Polish also provides evidence that the case is inherent, in that it allows genitive modifiers to appear before the numeral, as in the example below:

(55) *Tych* *pięć* *ptaków* *spało*
These.M.GEN.PL five bird.M.GEN.PL slept.N.SG
"These five birds slept"

Presumably the genitive on the modifier is assigned prior to movement; by this logic, the genitive must be assigned at D-structure for the modifier to have that case by the time it moves, making it an inherent case. Using this argument, Franks claims that the genitive is an inherent case. The fact that it also has a case alternation, he explains in a different manner, drawing on the Accusative Hypothesis. Franks (1994) claims that the genitive assigned by the numeral (the case actually being Gen-Q which is syncretic with the nominal genitive) can only be assigned in accusative DPs, although he cannot explain why; this restricts Gen-Q assignment to nominative and accusative positions. In oblique positions, an accusative DP is impossible (since the DP must be in the oblique case), which effectively disables the QP from assigning Gen-Q. Thus, the reason for the case alternation is because the Gen-Q case can*not* be assigned in oblique positions, and instead the oblique case appears on both numeral and noun. This is diagramed below.

(56) *Structural Position:* (*Accusative DP*)
[**AccDP** *d* [QP Numeral [NP Noun]]]
Gen Case

Oblique Position: (*Not Accusative DP*)
[**DP** *d* [QP Numeral [NP Noun]]]
Gen Case

Rutkowski (2005) and Rutkowski and Szczegot (2001) make a similar account relying on D-structure. However, they follow Franks' first piece of

evidence, namely the case assigning alternation, and decide that the genitive assigned by the numeral is in fact a structural case. They claim that numerals are functionals, and as functional items, they are inserted late in the derivation, immediately before spell-out (i.e. S-structure). Thus, this gives them the pattern of Russian described above: oblique case assigners assign their case at D-structure; by the time the numerals are inserted at S-structure, the noun is already cased and so the numeral cannot give its case. In structural positions, there is nothing to interfere with the case assignment on the noun, and thus, case is assigned. Refer to the diagrams in (54) for the mechanism.

In this thesis, I do not assume a D-structure or S-structure, but follow the basic tenants of Minimalism. Thus, neither of these accounts can be adopted here. However, they do provide valuable insight into the nature of case assignment, such as the generalization that oblique case is more privileged than structural case. It still seems to remain an open question whether the genitive case that appears on the numeral is actually structural or inherent. However, because D-structure is no longer a factor, Franks' second argument that the case must be inherent for the modifier to be genitive no longer holds. Presumably, genitive case modifiers can move after they have been assigned genitive case, while nominative/accusative modifiers can be base generated in a numeral preceding position. This leaves only the argument that the case is structural (by analogy with the Russian pattern), and I will indeed assume that this is the case. However, a detailed explanation as to how to explain any other differences between Russian and Polish remains beyond the scope of this paper.

A final account worth considering here, which is more modern in that it follows Minimalism, is one by Rappoport (2003). To derive the case-assigning alternation, Rappaport proposes a mechanism for inherent case assignment, in which a lexical item is only visible to Merge "if associated in the lexicon with the case feature required by a governing lexical item, in association with its semantic role assigning properties." (Rappaport, 2003: 129). This basically means that a lexical item will be *invisible* to Merge with an inherent-case assigner if not already in the required oblique case. Due to this, he also claims that numerals and nouns have an optionally valued case feature, meaning that

they can enter into the derivation pre-cased[12]. With these two claims in hand, he explains the case-assigning alternation. It goes as follows. In structural case positions, nouns enter into the derivation uncased, while numerals are cased. In Rappaport's account, numerals do not have a nominative case form and cannot be valued as nominative, so an uncased numeral in structural environments would be ungrammatical. Instead, numerals have a quantitative case associated with them, which is copied onto the noun through an Agree operation; this results in the same case on both numeral and noun.

(57) a. Numeral [case:quantitative] noun [case: —]

b. Numeral [case:quantitative] noun [case:quantitative]

However, note that the case *appears* to be different on the numeral and noun (recall that the case on the noun is genitive and the numeral a nominative/accusative). Here, Rappaport claims that the reason for this comes from spell-out – for nouns, the quantitative case is syncretic with the genitive and is thus spelled-out as genitive; numerals, on the other hand, have a spell out which is conditioned by gender: genitive with masculine personal nouns and null with non-masculine personal nouns.[13]

(58) *Spell-out (noun):*
[case: quantitative] → [case:gen]

Spell-out (numeral):
(i) [case:quantitative] → [case:gen] / __ [gender: masculine personal]
(ii) Elsewhere, [case: quantitative] → 0

[12] Note, I do not intend to imply that there has been any previous case assignment. By Rappaport's account, when there is an inherent case, there are items in the lexicon which come already marked with those inherent cases. This presupposes that for each noun, there are potentially five instantiations of it in the lexicon – the uncased form, instrumental, dative, locative, and genitive.

[13] Through this spell-out rule, he derives the morphological differences between masculine personal and non-masculine personal nouns with 5+ numerals.

In this way, he derives the surface facts of structural positions – pre-cased numerals Agree with the lower noun, passing on their case feature, which through spell-out rules is realized as genitive on the noun.

With the oblique case, the opposite occurs – nouns come pre-cased, whereas numerals do not. In this way, numerals act as pure modifiers, same as attributive adjectives, and simply agree with the case on the following noun. Because the case position is oblique, or inherent if you will, the noun *must* be pre-cased in the proper oblique case to be visible for Merge. This forces the oblique case assigner to select a complement which occurs in its oblique case, since all other options would be invisible to it. Because the numeral has no case value of its own, it agrees with the noun following it, borrowing the case feature.

(59) a. Numeral [case: —] noun [case: oblique]

b. Numeral [case:oblique] noun [case:oblique]

Thus, we see a reverse of what occurred in structural environments. No syncretism rules are necessary here.

The actual mechanism of lexical and inherent case assignment is not very clear in the literature. Rappaport makes an attempt to define how inherent case actually works, drawing on his concept of visibility to Merge. This produces a clever explanation of the case alternation found in Polish, which once inherent Merge is defined, is dependent on the idea of valued or unvalued case features on nouns and numerals. Importantly, this approach highlights a need to treat inherent case assignment differently from structural case assignment. Although I will not adopt Rappaport's exact implementation in my analysis, his work rightly points in the needed direction: the idea that structural cases and oblique cases are inherently different, most likely due to the intimate relationship between theta marking and inherent case. I will return to these issues in section 4.4, where I discuss an analysis of the case-agreeing, case-assigning alternation.

3.4 Summary

In section 3.1, I introduced three hypotheses concerning the case of the numeral, the Accusative Hypothesis, the Nominative Hypothesis, and the Nominative-Genitive Hypothesis. The Accusative Hypothesis has been used as an explanation for why verbal agreement does not occur with 5+ numerals. Despite the arguments for this hypothesis, I have rejected it in favor of the Nominative-Genitive Hypothesis, which makes a case distinction based on gender. This hypothesis brings up two important points, which I will address in the next chapter – why verbs do not agree with nominative 5+ numerals, and how gender can dictate case assignment. Section 3.2 discussed the syntactic category of the numeral. Here, I opted for an approach in which numerals are a combination of functional and lexical, i.e. semi-lexical nouns. This captures the functional nature of numerals, ensconced in the numerals-as-QPs approach, as well as the lexical nature of numerals, found in the numerals-as-NPs approach. However, it is still necessary to explain the adjective-like nature of numerals 2,3,4 under such an approach. Finally, in section 3.3, I discussed the case assigning alternation. This established that numerals assign a structural case, and that structural case assignment and oblique case assignment are different. I will attempt to provide an account of numerals in the next chapter which is consistent across position for all numerals 2,3,4 and 5+.

4 Solving the Puzzle of Polish Numerals

In this chapter, I propose an analysis to the data I have presented in Chapter 2. The purpose of this chapter is to account for the syntax of numerals, using this to explain the behavior of numerals in subject position, especially with regard to verbal agreement. Chapter 2 presented three main phenomena which require explanation: (1) verb agreement mismatches with 5+ numerals, (2) verb agreement mismatches and the accompanying oblique case marking with masculine personal nouns, and lastly, (3) the case assignment dichotomy – genitive case in structural positions and oblique case in oblique positions. I will address each of these issues in turn.

This chapter is divided into eight sections. The first section delves into a description of the behavior of partitives. Partitives involve both numerals and quantifiers and appear to respect the same rules as the quantitative constructions addressed in chapter 2; thus, they are subject to the same puzzling phenomena as quantitatives. I will address this connection and discuss the insights it provides on the nature of the quantitative construction. In section 4.2, I turn to the first issue of chapter 2: the lack of subject-verb agreement. I propose that numerals are case-assigning, semi-lexical nouns with defective features. This is what blocks verb agreement, resulting in the appearance of default agreement. In section 4.3, I extend this analysis to masculine personal nouns. Drawing on the work of Rezac (2003), I argue that the case on the numeral is genitive, which is the result of a type of "case leaking" – the case assigner is unable to fully agree with the noun and instead, extends its domain upwards, towards the numeral; thus, both receive case, and the verb cannot agree with an already cased, defectively featured noun. Finally in section 4.4 I address the case assignment dichotomy. I claim that numerals are unable to function as arguments; because of this, when inherent case is assigned to the numeral-noun projection, it spreads down to the noun, which can take the theta role; assuming case stacking is possible, it is the last assigned case which is pronounced, namely the oblique. Section 4.5 is devoted to the structure of quantitatives and section

4.6 the structure of partitives. Lastly, in section 4.7, I provide a tentative explanation as to what it means for Polish numerals to be defective. I suggest that the loss of certain features may be at fault, describing the differences between numerals in terms of feature sets. Section 4.8 concludes.

4.1 The Connection to Partitives

Partitives present a construction which looks very similar to numeral-noun constructions. In languages like English and Dutch, partitive constructions lead to a genitive preposition between the numeral and noun:

(60) *Three **of** the girls were sleeping*

(61) *Drie* ***van*** *de* *meisjes* *hebben* *geslapen*
Three of the girls have.3PL slept
"Three of the girls slept"

As we have seen many times before, numerals 5+ assign a genitive case to the noun following them; the main surface difference between Polish numeral-noun constructions and Dutch and English partitives would then be that the genitive assigner is overt in Dutch and English, but covert in Polish.

The actual partitive construction in Polish involves an overt genitive assigning preposition-like element.

(62) *Trzy* *z* *dziewczyn* *spały*
Three from[GEN] girl.F.GEN.PL slept.NON-VIR.PL
"Three of the girls slept"

In Polish partitives, the numeral acts as the head of the partitive, and determines agreement. Thus, if the numeral is 1, agreement is singular, and if the numeral is 5+, there is an agreement mismatch. As shown in the example above, if the numeral is of the 2,3,4 numerals, plural agreement occurs.

(63) a. *Jedna* *z* *dziewczyn* *spała*
One.F.NOM.SG from girl.F.GEN.PL slept.F.SG
"One of the girls slept"

b. *Pięć* *z* *dziewczyn* *spało*
Five from girl.F.GEN.PL slept.N.SG
"Five of the girls slept"

Additionally, if the noun is masculine personal, the same agreement patterns as with the quantitative (i.e. numeral-noun) construction are found on the numeral. The noun itself, however, remains in the genitive.

(64) a. *Jeden* *z* *chłopców* *spał*
One.M.NOM.SG from boy.MP.GEN.PL slept.M.SG
"One of the boys slept"
b. *Dwaj* *z* *chłopców* *spali*
Two.MP.NOM from boy.MP.GEN.PL slept.VIR.PL
"Two of the boys slept"
c. *Dwóch* *z* *chłopców* *spało*
Two.GEN/ACC from boy.MP.GEN.PL slept.N.SG
"Two of the boys slept"
d. *Pięciu* *z* *chłopców* *spało*
Five.OBL from boy.MP.GEN.PL slept.N.SG
"Five of the boys slept"

With the numeral 1, nothing special occurs; with numerals 2,3,4 there are two strategies, one which involves full agreement, and the other which blocks verbs agreement. With numerals 5+, the oblique form of the numeral is used. This is the same pattern found with numerals in quantitative constructions.

One important difference from the quantitative constructions concerns oblique case environments. In oblique positions, the genitive of the noun is not overridden by the oblique case, as the case in the quantitative construction is. The numeral, however, is marked with the oblique case.

(65) a. *Spałam* *z* *jednym* *z* *kotów*
I[F].slept with[INSTR] one.M.INSTR.SG from[GEN] cat.M.GEN.PL
"I slept with one of the cats"
b. *Spałam* *z* *dwoma* *z* *kotów*
I[F].slept with[INSTR] two.M.INSTR from[GEN] cat.M.GEN.PL
"I slept with two of the cats"
c. *Spałam* *z* *pięcioma* *z* *kotów*
I[F].slept with[INSTR] five.INSTR from[GEN] cat.M.GEN.PL
"I slept with five of the cats"

This data provides an important insight – it is possible for the genitive case of the quantitative construction to be overwritten, but not the genitive of the partitive. Presumably, this has to do with the overt presence of the genitive assigner *z* in partitives; alternatively, it may suggest the structures between the two are different – oblique case can access the noun in quantitatives, while it cannot in partitives. Importantly, whatever explanations are given for the behaviors of quantitives with regards to masculine personal gender and verbal agreement must be applicable to partitives as well.

I turn now to my analysis of the quantitative construction.

4.2 Verbal Agreement as Agreement Failure

As was shown in chapter 2, verb agreement with 5+ numerals consistently produces an agreement mismatch – nouns follow the numeral and are marked as plural with genitive case, and regardless of the gender of that noun, verbs appear marked as neuter singular. This is an agreement mismatch, a discrepancy between the number and gender markings on the subject and verb. The example below demonstrates this once again, with masculine gender and plural number on the noun, but neuter gender and singular number on the verb:

(66) *Pięć ptaków spało*
Five bird.M.GEN.PL slept.N.SG
"Five birds slept"

Recall that this is not the result of agreement with the numeral, since the coordination of two numerals does not produce a plural on the verb, unlike the coordination of two neuter singular nouns. Example (3) from chapter 1 is repeated below:

(67) a. *Krzesło i biurko zapadły się*
Chair.NOM.SG and desk.NOM.SG collapsed.NON-VIR.PL REFL
"A chair and a desk collapsed"
b. *Pięć krzeseł i pięć biurek zapadło się*
Five chair.GEN.PL and five desk.GEN.PL collapsed.N.SG REFL
"Five chairs and five desks collapsed"

This phenomenon does not follow from Chomsky's (2000, 2001) instantiation of Agree. If Agree were responsible for the number and gender marking on the verb, then there is only one logical possibility here – there should exist some neuter singular element in the subject for the verb to agree with. However, the noun is clearly not neuter singular in the example above (and is never singular when quantified by 5+ numerals), and the numeral is not necessarily neuter singular either, since its coordination does not produce a plural. Essentially, there is no source for the neuter singular features. Thus, by Agree, it is predicted that the sentence should be ungrammatical, which it is not. Hence, Agree cannot explain the marking on the verb with 5+ numerals.

The solution to this issue follows contra Agree, yet is very simple. The appearance of neuter singular features is the result of *default agreement*, which is produced by a failure to agree (Preminger, 2011). This solution, though not stated explicitly, was hinted at in the Accusative Hypothesis (Franks, 1994, 2002; Rutkowski, 2004; Przepiórkowski, 2004). The Accusative Hypothesis made the assumption that the numeral is not available for agreement, since it has already been cased for accusative; the noun isn't either, as it has been cased for genitive. With nothing to agree with, the verb is assigned the set of default features. Thus, unlike Agree, which assumes that this situation should lead to a derivation crash, these authors assumed that a default agreement occurred, produced by a failure to find an adequate goal to agree with.

This idea of default agreement is an interesting problem for Agree, as it suggests that not all derivations fail if a proper goal cannot be found. English also has cases which support the idea of default agreement. Consider the examples below.

(68) a. *<u>That she was happy</u> is obvious to me.*
b. *<u>For him to leave me alone</u> is all I want.*
c. *<u>To struggle with this book</u> suggests that you may need help reading.*

Presumably, none of the underlined subjects in these sentences have a set of interpretable phi-features since they are not nouns. The subjects in (68a) and (68b) are both complementizer phrases (CPs), and in (68c), the subject is a tense

phrase (TP). In each case the verb shows third person singular marking. The default nature of the verb agreement in these sentences becomes even more obvious once coordination is involved – for each of the subjects given above, agreement is still third person singular when coordinated with a similar subject, as shown below.

(69) a. *That she was happy and that Bill was unhappy makes / *make us happy.*
b. *For him to leave me alone and for her to go home surprises / *surprise me.*
c. *To struggle with this book and to ask me for words suggests / *suggest that you may need help reading.*

Theoretically, without any phi-features on these subjects, agreement should fail, making the sentences in (68) and (69) ungrammatical. However, CP and TP subjects are perfectly acceptable, even when coordinated. This suggests that the third person singular verbal marking we see in these examples is actually a default marking.

Polish has similar cases in which so-called default marking appears on the verb. This occurs with infinitive subjects in predicational sentences, and with dative subjects of impersonal verbs (Swan, 2002). It is also found with weather verbs.

(70) *Wyłysieć* *to* ***byłoby*** *dla* *mnie* *tragedia*
to.grow.bald TO would.be.N.SG for me tragedy
"To grow bald would be for me a tragedy"
(Swan, 2002: 391)

(71) ***Nudziło*** *mi* *się*
was.bored.N.SG me.DAT REFL
"I was bored"

(72) ***Padało***
Rained.N.SG
"It was raining"

In example (70), an infinitival verb sits in subject position; the main verb is marked as neuter singular. Similarly, in example (71), with the subject in the

dative case, the verb takes neuter singular marking; finally, in example (72), the verb appears to lack a subject and shows neuter singular instead.[14] These cases are reminiscent of what occurs with 5+ numerals. In (70), the subject is an infinitival verb, and thus lacks features for the main verb to agree with; likewise, in example (71) the subject has already been marked as dative and is therefore not active for agreement with the verb. Lastly, in example (72), there seems to be no subject for the verb to agree with. What is common to all of these cases is that Agree cannot be achieved, but the sentences are still grammatical. With an appeal to default agreement, we can succinctly explain the neuter singular marking on the verb in these cases, as well as in the case of 5+ numerals: agreement fails and so default features are inserted on the verb.

Preminger (2011) makes a case in favor of default agreement. Based on data from Kichean, Basque, Icelandic, and French, Preminger argues that the derivational time bomb approach ensconced in Agree is inadequate in a larger empirical domain. Similar to what I have demonstrated above, he discusses cases in which there is no source for the features that appear on the verb, claiming that agreement has failed and default features have appeared instead. Preminger argues that although Agree is obligatory, its success is not. Thus, Agree must be attempted, but if it fails, the sentence does not become ungrammatical, as instead default features can be inserted. This effectively rules out mismatching agreement and default agreement in cases where full agreement is possible (73), while still allowing for the mismatching agreement patterns we have seen above, in those cases where agreement fails (74).

(73) *The women *am / *is / are happy*

(74) *Pięć ptaków spało / *spały*
Five bird.M.GEN.PL slept.N.SG / slept.NON-VIR.PL
"Five birds slept"

[14] Weather verbs in Polish also allow overt subjects in which agreement occurs. Compare example (72) above, with the example below:

i. *Deszcz padał*
Rain.M.NOM.SG rained.M.SG
"It was raining"

In (73), the subject *the women* is active and available for agreement. Since Agree is obligatory and can be successful here, it is only grammatical when it has been successful and there are no agreement mismatches. In (74), the claim is that the verb cannot successfully agree with the subject, *pięć ptaków*. Thus, only default agreement is grammatical; verb marking that matches the noun in the subject is ungrammatical, since agreement fails with quantified nouns.

Having established that default agreement is at play here, the next question is why does agreement fail with 5+ numerals? In chapter 3, I discussed three hypotheses concerning the case value of the numeral and three hypotheses concerning its categorial status. From that discussion, I rejected both the Accusative Hypothesis and the Nominative Hypothesis, opting for the Nominative-Genitive Hypothesis. Additionally, I argued that Polish numerals were semi-lexical nouns. Here, I expand on these conclusions, to demonstrate how they produce the agreement failure that leads to default marking on the verb.

My analysis is as follows. As nouns, Polish numerals are defective; this is a part of their having semi-lexical status. Because they are nouns, they assign genitive case to their complements; but since they are defective, they cannot value all the features of the verb during agreement. This is what produces the agreement failure. In section 4.7, I will speculate more on what it means to be defective, but for my purposes here, I will simply assume that this defectiveness results from missing features, or rather, underspecification – as a defective noun, the numeral does not have the full set of phi-features which are required by the verb. As a result, not all uninterpretable features on the verb can be checked through agreement with the numeral. Furthermore, the verb cannot agree with the noun either, as it has already been cased for genitive by the numeral. Because not all uninterpretable features on the verb can be valued through agreement with the numeral, agreement will fail. This will lead to default agreement.

I provide a derivation below to illustrate how this works; for simplicity, I only mention the relevant projections. The noun and numeral are merged first, after which genitive is assigned to the noun. In this paper, I will assume there is

an intermediate functional head, *g*, which assigns the genitive case; the reason for this will become clearer in sections 4.3 and 4.5 (see section 4.5 for the final structure). We can assume that the numeral has at least a number feature, say plural for 2+, but unvalued for 1.

(75) [Numeral *g* [N]]
[*uCase*] [*GenCase*]
[*iNum*] [*iPhi*]

The assignment of genitive case (through an Agree operation) checks the case feature of the noun. Thus, it is inactive for further agreement.

The merged numeral and noun then merge with the verb.[15] They sit in the specifier position, accessible for the next higher phase (Chomsky, 2001).

(76) [[Numeral *g* [N]] V]
[*uCase*] [*GenCase*]
[*iNum*] [*iPhi*]

This is then merged with T, which has uninterpretable phi-features, as well as interpretable tense.

(77) [T [[Numeral *g* [N]] V]]
[*uPhi*] [*uCase*] [*GenCase*]
[*iTense*] [*iNum*] [*iPhi*]

The uninterpretable phi-features make tense an active goal. Similarly, by virtue of the uninterpretable case feature on the numeral, it too is active. As a probe, T searches downward in its c-commanding domain and encounters the active numeral. However, because the numeral is defective and lacking in interpretable features, it cannot deactivate the verb. Thus, after encountering the numeral, the verb remains active, as does the as-of-yet uncased numeral.[16]

[15] I abstract away from little *v* here.
[16] Polish T seems to have an optional EPP feature, as subjects can appear both pre and post verbally. Presumably, when the EPP feature is present, in addition to the valuation of number on the verb and case on the numeral, the numeral-noun construction would move to the specifier position of TP. This step is not represented in the derivation above.

(78) [T [[Numeral *g* [N]] V]]
[*Sg/Pl-Num*] [*uCase*] [*GenCase*]
[*iTense*] [*iNum*] [*iPhi*]
[*uPhi*]

As the probe still has unvalued features, it can continue its search downwards. However, the only other element downwards is the noun, which is inactive and invisible to the probe. At this point, there is nothing more the probe can do – it has attempted agreement, and failed. Under the analysis we are assuming here, this is enough. Default features can be inserted on the verb, deactivating it.

There is a remaining question here – how does the numeral get case? Just as there is a mechanism for default agreement, I claim that there is also a mechanism for default case assignment. It is this default case which appears on the numeral. Thus, under my analysis, neither the verb nor the numeral are deactivated through the conventional Agree, but through a process of default agreement and default case assignment.

In Polish, the default case is nominative. This contrasts with English where the default case is accusative. This can be seen, for example, when picking out people from a crowd – in English, one would say "you, me, and him", using the accusative forms, whereas in Polish it would be the nominative forms "ty, ja, i on". Similarly, in questions referring to one's self, accusative is found in English, but nominative in Polish:

(79) a. *Me? I don't feel like going today.*
b. *Ja? Nie chce mi się iść dzisiaj*
I.NOM? No want I.DAT REFL to.go today
"Me (lit. I)? I don't feel like going today"

Additionally, in English one can coordinate subjects in which the second conjunct is in the accusative, whereas in Polish, this is entirely unacceptable – it must always be nominative:

(80) a. *John and me went to the cinema*

b. *Jan* *i* *ja/*mnie* *poszliśmy* *do* *kina*
John.NOM and I.NOM/*ACC went to cinema
"John and me went to the cinema"

Thus, the default case in Polish is nominative. By the end of derivation, both the numeral and verb will be deactivated and marked with default features – neuter singular for verbs and default nominative for numerals.

So far, I have shown what must occur with the 5+ numerals. I now turn to the other numerals, beginning with 2,3,4. I claim that numerals 2,3,4 are also defective nouns like numerals 5+, although they have a different sort of defectiveness (I will expound on this in section 4.7). In addition to being a phi-incomplete goal, this defectiveness also deprives them of their ability to assign genitive case (usually), and it is this lack of case assignment which produces the agreement differences between numerals 2,3,4 and numerals 5+. Essentially, since there is no genitive case assignment with numerals 2,3,4, the lower noun remains active after a merger with the numeral. When it comes time for T to probe, it finds both the numeral and noun as active — whereas with 5+ numerals, T was unable to agree with the lower noun (which was inactive), with the numerals 2,3,4, this is entirely possible (since it is active). Thus, agreement is found between the noun and the verb. Presumably, because the numeral is defective, then like the numerals 5+ it cannot be cased through agreement with the verb and instead it receives the default nominative case. Under this account, the differences in agreement between the two types of numerals is due precisely to the lack of genitive case assignment with the numerals 2,3,4.

The numeral 1 presents a much simpler case. As established in chapter 2, numeral 1 is an adjective. Thus, we expected it to behave like an adjective for agreement. As an adjective, it will not be a proper goal for T (thus, T will never agree with it) and it will never assign any case to the following noun (thus, will not "bleed" T of its goal). Instead, T agrees with the noun, assigning it nominative case, and in turn the numeral agrees with the noun in case, gender, and number. For the lexical numerals 1000+, agreement should proceed as with any noun – genitive is assigned to the lower noun making it inactive; the

numeral, however, is fully active, and in this case phi-complete. Thus, the verb can and does agree with the numeral. The behavior of numeral 1 and numerals 1000+ with regards to agreement is a consequence of their categories – noun and adjective.

Before concluding this section, there is one further question I would like to address: what are the advantages of this approach over the Accusative Hypothesis in dealing with the 5+ numerals? Recall that with the Accusative Hypothesis, the numeral was assumed to be accusative as an explanation for the default verb agreement. This represents an important ingredient for any analysis of Polish numerals, the need to derive an agreement failure. The nominative hypotheses were criticized for the fact that despite being in the nominative, the verb still could not agree with them; these hypotheses would presumably not derive an agreement failure. This is an issue that I hope I have dispelled in my above discussion – I have shown how it is possible to have the numeral in the (default) nominative case, yet still have an agreement failure. In this sense, my analysis is superior – it does not involve any stipulations about some mysterious accusative case. Instead, Agree works as would be expected were the numeral a regular noun, with the catch that defective features prevent agreement, and it is these defective features, a property of more functional items (as compared to numerals 1 or 1000), which cause agreement to fail. This sort of approach to the category of the numeral is in line with the observations of Rutkowski (2002) that 5+ numerals are similar to functional elements and potentially QPs. It has the added advantage that we do not expect these numerals to act identically to nouns, another objection against the numerals as nouns approach. Thus, by treating numerals as semi-lexical nouns, i.e. nouns with defective features, we avoid the problems of the Accusative Hypothesis, while still recognizing the similarities between numerals and nouns.

There are two main points to this section: numerals are defective nouns and verbs are marked with default agreement when agreement is unsuccessful. Agreement will always apply (Preminger, 2011), but since numerals are defective nouns, verbs cannot value all their features with the numeral, nor can they with the noun, as it is already genitive, and hence, not an active goal. This leads to

default agreement, which accompanied by default case on the numeral, derives the patterns which are found with 5+ numerals, and the "nominative" portion of the nominative-genitive hypothesis.

4.3 "Case Leaking" and Cyclic Agree

In the previous section, I addressed the agreement mismatches of 5+ numerals. In this section, I extend that analysis to cover the behavior of numerals with masculine personal nouns, i.e. nouns of masculine gender that refer to a human. There are two major questions concerning masculine personal nouns: what and why. What is the source of the agreement mismatch and why does it occur?

4.3.1 The source of the agreement mismatch

The first question is the easiest to answer and follows directly from the previous section. Recall that with masculine personal nouns, the numeral appears to take on an oblique form:

(81) a. *Pięć* *ptaków* *spało*
Five bird.M.GEN.PL slept.N.SG
"Five birds slept"
b. *Pięciu* *chłopców* *spało*
Five.OBL boy.MP.GEN.PL slept.N.SG
"Five boys slept"

According to the Accusative Hypothesis, the case of the numeral in both sentences above would be accusative. However, due to the discrepancies under this hypothesis between lower and higher numerals with masculine personal nouns, I have rejected this hypothesis in favor of the Nominative-Genitive Hypothesis (c.f. section 3.1). The Nominative-Genitive Hypothesis claims that numerals are marked as nominative with non-masculine personal nouns and as genitive with masculine personal nouns. By assuming this hypothesis, we have an answer for the first question: what is the source of the agreement mismatch? Because the numeral is already cased as genitive, it is inactive for agreement.

Thus, when the verb acts as an active probe, both the numeral and the noun are invisible to it; with no available goal, agreement fails. This leads to default marking.

With 5+ numerals, I am claiming that the reason for the agreement mismatch is different between masculine personal and non-masculine personal numeral constructions, despite the obvious similarity between the two. With non-masculine-personal nouns, agreement fails because the numeral is defective in phi-features. With masculine-personal nouns, the numeral is still defective, but in addition, it is already pre-cased as genitive. This is an additional factor, which although not obvious with 5+ numerals, is clearly visible with the 2,3,4 numerals, which alternate between a nominative form with full agreement (for masculine personal (82b) and non-masculine personal (82a)) and a non-nominative form with an agreement mismatch (for masculine personal only (82c)):

(82) a. *Dwa* *ptaki* *spały*
Two.M.NOM bird.M.NOM.PL slept.NON-VIR.PL
"Two birds slept"

b. *Dwaj* *chłopcy* *spali*
Two.MP.NOM boy.MP.NOM.PL slept.VIR.PL
"Two boys slept"

c. *Dwóch* *chłopców* *spało*
Two.GEN boy.MP.GEN.PL slept.N.SG
"Two boys slept"

Because there is no genitive to block agreement with the noun with 2,3,4 numerals, we are able to see agreement between the noun and the verb. However, with masculine personal gender, there is an optional agreement mismatch, which, when it occurs, is due to the genitive case on the numeral and noun, which makes both unavailable for agreement. With the first factor of defectiveness not playing a blocking role here, it becomes more obvious that there is a second factor involved, namely, the genitive case marking on the numeral. Despite its invisibility with 5+ numerals, I claim that the phenomenon which occurs with 2,3,4 masculine personal numerals also occurs with 5+ masculine personal numerals. The appearance of default agreement with 5+

masculine personal and non-masculine personal numerals, although seemingly identical, is derived in different ways.

There is some evidence to suggest that this distinction I propose for 5+ numerals does actually exist. In data gathered from one of my native speaker informants, coordination of a 5+ non-masculine personal numeral and a 5+ masculine personal numeral with a 2,3,4 numeral (of non-masculine personal gender) was compared. The following paradigm was found:

(83) a. *Pięciu mężczyzn i dwie kobiety*
jadło / *jadły / ?jedli *razem*
Five.GEN man.MP.GEN.PL and two women.F.NOM.PL
ate.N.SG / *ate.NON-VIR.PL / ?ate.VIR.PL together
"Five women and two men ate together"

b. *Pięć białych krzeseł i trzy czarne*
krzesła ****zapadło / zapadły*** *się*
Five white.GEN chair.N.GEN.PL and three black.NOM
chair.N.NOM.PL *collapsed.N.SG / collapsed.NON-VIR.PL REFL
"Five white chairs and three black chairs collapsed"

In (83a), a 5+ masculine personal numeral is coordinated with a non-masculine personal 2,3,4 numeral (e.g. 5mp + 2f); in (83b), both numerals modify non-masculine personal nouns (e.g. 5n + 3n). For both sentences the first numeral is a 5+ numeral which is either masculine personal or non-masculine personal and the second a 2,3,4 numeral, which is non-masculine personal – syntactically the sentences differ only in terms of the gender of the first numeral-noun conjunct. In (83a) where the numeral modifies a masculine personal noun, default agreement is the unmarked choice (*jadło*), while actual gender and number resolution (**jedli*) is marginal. In contrast, when no masculine personal gender is involved (83b), the opposite is found – resolution is fully possible (*zapadły*), while default agreement is ungrammatical (**zapadło*).[17] Thus, gender resolution is

[17] Data from a larger group of native speakers (10), showed that in fact, both resolution and default agreement were possible with 5+ and 2,3,4 non-masculine-personal nouns, with around a 60% resolution rate. The ability to resolve features here is likely subject to interspeaker variation and factors of animacy (as the example above has non-animate nouns, while the examples tested with the group of 10 all involved animate nouns). The point of this example is to demonstrate that within a single speaker, this pattern can be found. Future research is still necessary to determine how widespread and productive the dissociation between 5+ masculine personal and non-masculine personal numerals is.

possible (and obligatory for this speaker) with non-masculine personal nouns, but not so with masculine personal nouns. This shows that despite the fact that masculine personal and non-masculine personal genders both result in the same failure in verb agreement with 5+ numerals in non-coordinate structures, they are syntactically different, and this becomes visible during coordination.

Although I do not intend to provide an account of the resolution phenomena here, it suffixes to say that the pattern in this example demonstrates a real difference between the behavior of 5+ numerals with masculine personal and non-masculine personal nouns. This provides support for my hypothesis that numerals are nominative with non-masculine personal nouns (thus, the only factor being their defectiveness) and genitive with masculine personal nouns (there being two factors, defectiveness and genitive case). It is likely that it is this additional factor which makes resolution more difficult with masculine personal nouns.

To summarize, I claim that numerals modifying masculine personal nouns are marked as genitive. This occurs with both 2,3,4 and 5+ numerals. With 2,3,4 numerals, this is more obvious, as it means there is an agreement mismatch for masculine personal nouns, but not for non-masculine personal nouns. With 5+ numerals, this difference is subtler, but traces of it can still be found in the resolution behavior of coordinate structures. Importantly, it is this genitive marking which leads to the agreement mismatch. Because the numeral has been pre-cased, it is invisible to the verbal probe. Agreement will be attempted by the probe, but since neither the numeral nor the noun is available for agreement, agreement will fail, and default marking will surface.

The Nominative-Genitive Hypothesis solves the first question: what is the source of the agreement mismatch? The source is the genitive case on the numeral. Now the more difficult question comes: why is the numeral even marked as genitive?

4.3.2 The reason for the agreement mismatch

Determining that the numeral is marked as genitive is not so difficult, as the morphological marking of these numerals is very suggestive of it being

genitive. However, explaining why this behavior is gender specific is the more troublesome aspect. This question was in fact one of the major objections against the Nominative-Genitive Hypothesis. In this section, I intend to show that these objections are unwarranted – under the system I propose, it will be natural for the numeral to be cased as genitive.

The first clue is the genitive case – genitive appears on both the numeral and the noun. Using this important clue, I claim that the genitive marking on the noun is no mere coincidence. Rather, the genitive on the numeral and noun are related, by virtue of being assigned by the same case-assigner. This is a phenomenon that I will descriptively refer to as "case leaking" – the case on the noun leaks out and spreads to the preceding numeral. In this way, the genitive on the noun and the numeral has the same source.

The next question is: just how might this occur? To explain where case leaking comes from, I will build on work by Rezac (2003, 2004), Chomsky (2005) as discussed in Radford (2006), and Hiraiwa (2005).

Let us begin with Chomsky (2005), as discussed in Radford (2006), and Hiraiwa (2005), focusing on a phenomenon known as *multiple agree*. In sentences with *there*-expletives, verbs appear to be agreeing with two separate DPs, the expletive *there* in subject position and a second DP in a post-verbal position; number agreement occurs with the second DP:

(84) a. *There were many children in the school.*
b. *There was only one girl in the school.*

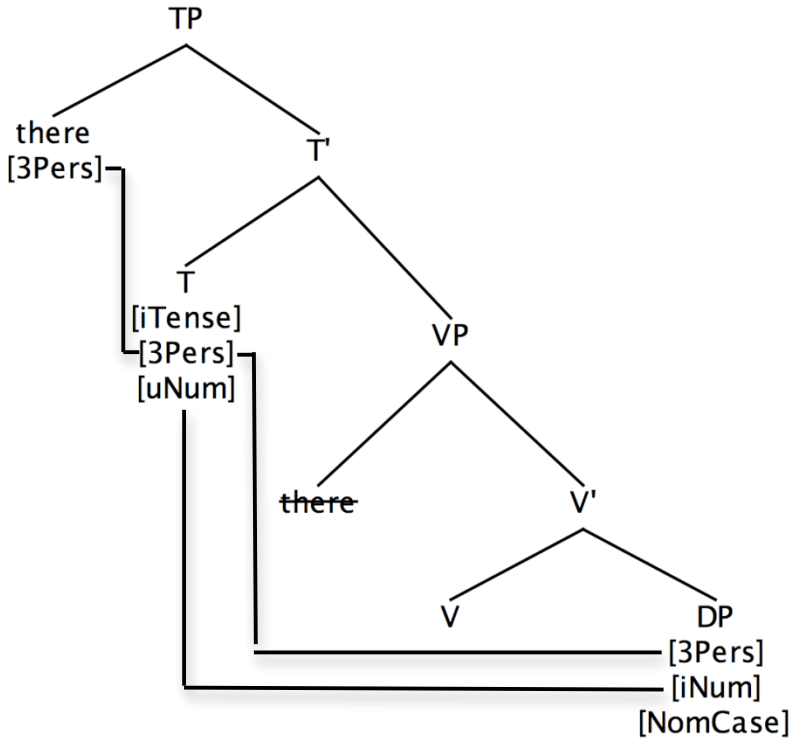
In the first sentence, *were* appears to be agreeing with the second DP *many children*, which is plural; similarly, in the second sentence, *was* agrees with the second DP, which in this case is singular. Presumably, the verb can satisfy all of its features through agreement with the second DP – so why does the *there*-expletive appear? The idea is that *there* is merged higher in the structure than the second DP, for example, in the specifier position of a VP with the DP in the complement position; furthermore, it is supposedly active due to an

uninterpretable, but valued person feature.[18] This makes an intervention effect in that the verb is forced to Agree with *there* before it can Agree with the lower DP (abstracting away from the simultaneity condition for the moment).

Because *there* is merged higher in the structure and is also active (due to its uninterpretable person feature), then when T searches for an active goal, it will encounter *there* first. *There*, however, does not have all the features needed by the T probe, as it lacks a number feature. T agrees with *there*, pulls it up to specifier position by its EPP feature, and deactivates it. However, T is not finished, as it is still active, by virtue of its unvalued number feature; it continues the search downwards, encountering the second DP, which is also active, for a lack of case. Agreement with this DP values the probe's number feature and the goal's case feature. Instances of Multiple Agree are assumed to respect the Simultaneity Condition, meaning that agreement with multiple probes occurs simultaneously. Thus, this agreement with both *there* and the DP should occur simultaneously. I illustrate the agreement relationships in the diagram below.

[18] Because it is a valued uninterpretable feature, the expletive *there* only needs its feature to be deleted. This is a point at which the use of Multiple Agree with expletive *there* and with numerals will differ – *there* needs its feature deleted, whereas numerals are assumed to have an unvalued case feature, just like nouns; despite this, both are active. Thus, although the trigger for Multiple Agree will differ, the general idea of this mechanism for *there* and numerals is the same.

(85)

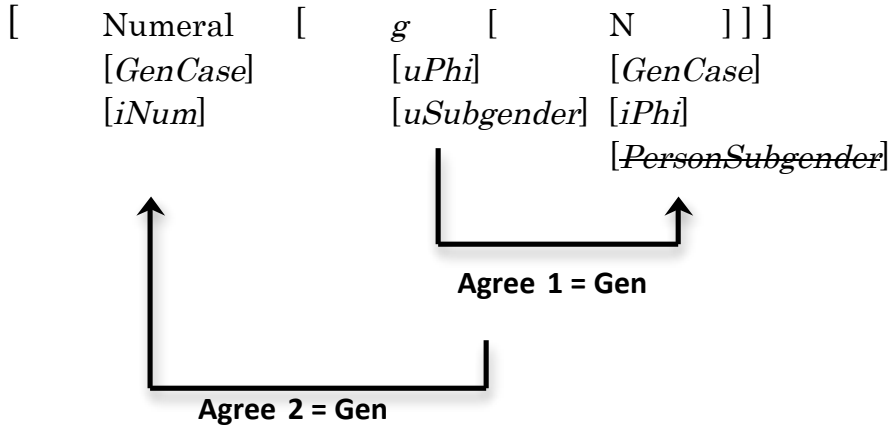


The diagram in (85) illustrates where T gets its features from. *There* has moved from Spec, VP to Spec, TP due to the EPP feature. Third person is valued through agreement with *there*; this feature matches to the person feature in the lower DP, and so T is able to value its number feature through the lower DP. The lines demonstrate the valuation of phi-features. This is multiple agree. Hiraiwa (2005) mentions cases in which Multiple Agree also results in multiple case assignment – thus, if *there* were to also have a case feature, then hypothetically, both *there* and the DP would have nominative case; this is the way in which Multiple Agree will be used here – multiple agreement with multiple case assignment.

This mechanism of multiple agree is useful for our purposes with masculine personal nouns modified by numerals. As stated earlier, I am claiming that the genitive which appears on the numeral and noun is assigned by the same head. Assuming that there must be some intermediate functional head to assign case to the noun, it encounters a problem when trying to agree with a masculine personal noun. This problem is what leads to an instance of multiple agree – assuming the case assigner cannot acquire all the necessary features

through agreement with the noun, it continues its search upwards (Baker, 2008; Rezac, 2002, 2003). As the numeral is active (but defective), the case assigner can also agree with the numeral, assigning it genitive case in the process.[19] When it is time for the verb to agree with the subject, both numeral and noun have already been cased as genitive, and are unavailable for agreement. This is the basic idea for the analysis of masculine personal nouns, illustrated in (86) below. For purposes of illustration, let us assume it is the person value on the subgender feature which is problematic.

(86)



The first instance of agree values the case as genitive on the noun; the second instance of agree, due to the problematic person value in the illustration above, values the numeral as genitive. Letter *g* simply refers to the case assigning head.

There are two theoretical issues here that require further development: upward agreement, and this "problem" with masculine personal nouns. Empirically, it is also necessary to say a few words concerning the differences between noun-noun constructions and numeral-noun constructions (excluding 1 and 1000+), since noun-noun constructions do not have this case leaking phenomenon. I will address each of these three issues in the next section.

[19] Recall that in the previous section, T could not agree with the numeral and instead it received default nominative case. Here, I am claiming that that same defective numeral can be agreed with and assigned case by the genitive case assigner. This is not a discrepancy as numerals can and do receive cases other than the default nominative – this suggests that they can be agreed with or at least the very least, can be assigned case. However, the question remains what the relationship is between case assignment and phi-features, especially defective phi-features. For my purposes here, it is enough to say that the genitive case assigner must be in some way different from the nominative case assigner with the result that it can assign genitive case to the numeral. This may be the result of the T and *g* probes having different phi-requirements, such that the numeral goal is defective for one probe, but not for another. This is an area that requires further development, consideration, and research.

4.3.2.1 Theoretical issues: Cyclic Agree and masculine personal gender

According to Chomsky's Agree, a probe can only search downwards in its c-command domain. What I am proposing instead is that in addition to the probe searching downwards, it can also search upwards, if it is still has unvalued features after a downward search. This is the mechanism for the genitive case assignment of the numeral.

A similar system is implemented by Rezac (2003), and utilized in Baker (2008). Rezac draws heavily on the Earliness Principle of Pesetsky and Torrego (2001) to explain how a probe can extend its search space upwards. The Earliness Principle states that "an uninterpretable feature must be marked for deletion as early in the derivation as possible" (Pesetsky & Torrego, 2001: 400). Rezac argues that adopting this principle has effects on the possible search space of a probe. Consider a head *h*, complement *c*, and specifier *s*. The first application of Merge results in a head complement structure, [*h* [*c*]]. By the Earliness Principle, if the head has any uninterpretable features, it will need to immediately probe for matching features. Once the complement is merged to the head, a search can be conducted. For many languages, a probe needs to search no farther – complements often have the set of matching features necessary to value the uninterpretable features of the probe; this is the canonical case of agreement. However, for some languages, such as Basque and Georgian, there are classes of complements which are underspecified for certain features which are needed by the verb. In those instances, the probing head does not become deactivated after finding a goal in the complement. A second application of Merge joins the specifier to the head complement structure: [*s* [*h* [c]]]. By the Earliness Principle, those active features on the head still need to be deleted – at this point, the head extends it search space to include the newly merged specifier; if the search for a goal is successful, the head will be deactivated, with no need to search farther (if it is unsuccessful, but cannot search farther, then likely default features will be assigned). Essentially, with each Merge, the possible search space of a probe is extended. Hypothetically, even higher levels past the specifier might be searchable; this, however, is restricted by phases, as the search cannot

extend beyond the limits of a single phase (by the Phase Impenetrability Condition: see Chomsky, 2001).

According to Rezac, it is very natural for the search space to extend beyond the head-complement structure, with the addition of new material by Merge. He calls it a "questionable stipulation" to restrict the search space to the complement, and by the Earliness Principle, it is unnecessary to do so. Empirically, the predictions are correct, even for the more familiar canonical cases of agreement, in which goals only seem to sit within a probe's c-command domain. In those cases where there is a phi-complete goal in the complement, the search space will not extend upwards and the specifier will remain untouched. When a new, active head is merged above that, making our specifier-head-complement structure into a complement, the initial downward search will lead the head to find the as-of-yet untouched specifier. In this way, upward probing *only* occurs after downward probing has been attempted.

To account for the patterns of Basque and Georgian, Rezac also proposes slight, but significant changes to the agree machinery. Building on work by Bejar (2000), Rezac promotes a view in which individual features are probes. This approach, taken from Bejar (2000), treats features as occupying separate projections in the syntactic structure, rather than as occurring as a feature bundle on the probe and goal. Due to this treatment of features, the concepts of deactivation and case require reformulation. Rezac discusses a new case system which views case as a realization of the features on the probe, similar to the view taken by Pesetsky and Torrego (2001) that interpretable features on the probe appear as uninterpretable features (i.e. case) on the goal. Deactivation of a goal then becomes an intervention effect.

In this paper, while I do adopt Rezac's analysis of cyclic agreement, I do not adopt his case system or feature system. For the data at hand, such a system is not necessary and only complicates the analysis.[20] Instead, I take the view

[20] Note however that if I were considering those cases in which the probe uses one feature from one goal and another feature from a second goal (c.f. example 19 in section 2.2.1), such an approach might be necessary. It is cases like these in Basque and Georgian that drives the features-as-probes approach taken by Rezac (2003) and Bejar (2000). However, this approach would also require some reformulation, as it is not compatible with the idea of default agreement

that cyclic agreement is an instance of multiple agree as described earlier: if the feature bundle on the goal does not fully value the feature bundle on the probe, then and only then can the probe continue its search upwards. Agree with a second goal can only occur if the features on that second goal are identical to the corresponding features on the probe. Note that this predicts that the gender, subgender, number, and person features on the noun and numeral must be identical, a prediction that seems correct.

A key component of Rezac's system is underspecification – a probe will only extend its search space upwards if it reaches an underspecified goal, i.e. a goal which lacks certain features needed by the probe. My system differs on this point. While I do agree that underspecification would produce upward agreement, with Polish masculine personal nouns, it cannot be the source. If it were, then upward probing would occur in all agreement domains, such as with tense and adjectives, which does not seem to be the case. Let us examine the gender system of Polish more closely to determine what factor is in play here.

Recall in section 2.2.1, I presented accounts by Corbett (1983) and Brown (1998) which suggested that there are six genders in Polish (feminine, neuter, masculine inanimate, masculine animate, masculine personal, and devirilized) and two features in the gender system (GENDER and SUBGENDER), with varying values – masculine, feminine, neuter, and masculine personal for GENDER and inanimate, animate, and person for SUBGENDER. Let us assume that this is indeed the case. Under this system, underspecification of the value *masculine personal* is not an accurate solution, as then tense and adjectives should also show agreement upwards, which they do not. Because of this, it cannot be the case that the issue lies with the goal – for other probes, the goal has no trouble valuing uninterpretable features. This means that it is the probe itself that is problematic.

With this knowledge in hand, we can better understand what the trigger of this upward agreement might be. There are various ways to go on this issue. One route is to say that the probe is unable to "see" certain features. However,

(e.g. how does the system know when to assign default features if individual features can act as probes? What happens if only one feature fails to agree? Etc.).

then we have the issue that if it were the actual feature that the probe has trouble with, it would have that trouble for all genders and subgenders, regardless of their values. But the fact of the matter is that it is specifically the masculine personal gender which is problematic. This suggests that it is not the feature itself that the probe has trouble seeing, but the *value* of that feature. This leads us to the conclusion that for this particular probe, the masculine personal gender is unreadable.[21]

Two questions arise from this conclusion: (1) why can the probe not see the masculine personal value? and (2) why does this particular probe have issues, but no other probes do?

My answer to these questions perhaps verges on stipulation. To answer the first question, I point to the fact that masculine personal gender is in fact a constructed gender. It is constructed from a masculine gender feature and a person subgender feature. If this were not true, it would not be possible to have cases in which the gender feature from one noun and the subgender feature of another can combine to produce masculine personal verbal agreement, as in example (87) below, repeated from example (19c) in section 2.2.1.

(87) *Kobieta i pies* ****spały / spali***
Woman.F.SG and dog.M.SG slept.NON-VIR.PL / slept.VIR.PL
"A woman and a dog slept"

As a constructed gender, it is possible that the system at some point reads that the gender is "masculine", the subgender is "person" and then reassigns the gender feature to "masculine personal". This reassignment of gender may leave some sort of mark on the gender feature which makes it illegible to this particular probe.

As to why this probe and only this probe is problematic, I tentatively suggest that it may have to do with the defective nature of the numeral. As it is the numeral which is selecting the noun, the case assigner which appears

[21] Alternatively, one could claim that verbs are only interested in the gender feature and not the subgender feature. In this way, there could be underspecification of the subgender feature. However, adjectives need the subgender feature for agreement in the singular, making this an unlikely solution.

between the numeral and noun is in some way related to the numeral, perhaps even selected by it. The case assigner may in some sense inherent this defective nature from the numeral, meaning both numeral and case assigning head are in some way defective. This would explain why it is only numerals which lead to case leaking – the defective qualities of the numeral appear on the case assigner, which makes it defective and unable to read constructed genders; fully-fledged nouns, however, would be expected to have fully-fledged case assigners as well, thus suggesting a reason for the lack of case leaking with regular nouns.

Despite these speculations, the point remains that there is something special about masculine personal gender, which causes trouble for the numeral's probe. Whether it is a problem of the gender being constructed or not is not so very important; what remains important is that this difficulty with masculine personal gender serves as a trigger for upward agreement, leading to the genitive that we see on the numeral.

A final remaining question concerns the numerals 2,3,4. In the previous discussion, there has been an underlying assumption that we are dealing with a genitive case assigner and that this genitive case consistently appears on the lower noun. For numerals 5+, this assumption is perfectly fine, as indeed, the noun following 5+ numerals does appear in the genitive case. However, with 2,3,4 numerals, we also see the case-leaking phenomenon, but in contrast to 5+ numerals, they do not usually assign a genitive case. How can this fact be reconciled given the current account?

Given that this occurs only with three very specific lexical items, it is tempting to claim that this is a purely lexical phenomenon, particular to numerals 2,3,4. However, such an answer is not very satisfying; let us see how far we can get, before such a claim becomes necessary.

I will assume here that the numerals 2,3,4 and 5+ are not so very different; they are identical structurally and also in terms of their behavior with masculine personal nouns. The areas where they differ concern verbal agreement and genitive case assignment. However, the difference in terms of verbal agreement can be explained with regards to their lack of genitive case

assignment – since they do not assign genitive case with non-masculine personal nouns, the verb can still agree with the active noun, allowing for verbal agreement. Thus, there is only one relevant difference here, and it is the lack of genitive case assignment with non-masculine personal nouns. If they are so similar, why do numerals 2,3,4 not assign genitive case with non-masculine personal nouns? The difference I claim lies in the feature make up of numerals 2,3,4 versus numerals 5+. In a sense, numerals 2,3,4 are even more defective than numerals 5+.[22] Thus, even though there is a case assigner present, due to its relation to the numerals 2,3,4, it lacks the ability to assign its genitive case during agreement. When it is time for the verb to agree, the noun is still active, and the verb can agree with it.

The trick comes with the masculine personal nouns, and it is here that I must resign myself to simply treating this as an idiosyncracy, perhaps built on analogy with the 5+ numerals. Essentially, with masculine personal nouns, the case assigner recovers its ability to assign genitive case, and case leaking ensues. It might even be the case that masculine personal nouns, in some sense, can select for a numeral with an active case assigner, whereas non-masculine personal nouns cannot; whatever is exactly happening, it is related to the gender features of the noun. Note that this is an optional property, as the case assigner does not necessarily need to recover its case assigning ability with masculine personal nouns. This may even be evidence that numerals 2,3,4 are becoming more like numerals 5+, this being the first step. For now, it seems safe to say that the case assigner is only able to assign case with masculine personal nouns, and that this is a choice for the speaker. However, once this choice has been made, the same process as occurs for 5+ numerals will occur for 2,3,4 numerals.

To summarize, there are two mechanisms made use of in this section: multiple agree and cyclic, or rather, upward agree. When masculine personal gender is encountered by the defective case assigners of numerals 2,3,4 and 5+, it is unable to fully satisfy its features through agreement; presumably the probe

[22] Alternatively, I could say that they are even more *adjective-like* than numerals 5+, by virtue of the fact that they seem to agree in gender to some extent. A more adjective-like feature make-up may rob these numerals of their case assigning abilities.

has issues in reading the masculine personal gender. As a result, it also searches upwards, encountering the active numeral. This it too values as genitive. Thus, both numeral and noun are marked as genitive by the time they are merged onto the verbal structure. Because of this, T cannot agree with either, and default agreement surfaces instead.

4.4 *The Case of the Overwritten Genitive*

In this section, I address the case-assigning, case-agreeing alternation of Polish numerals. Recall that in structural case positions, 5+ numerals assign a genitive to the following noun. In oblique positions, however, both numeral and noun appear in the oblique case. This is shown again below.

(88) a. *Pięć* *ptaków* *spało*
Five bird.M.GEN.PL slept.N.SG
"Five birds slept"
b. *Spałam* *z* *pięcioma* *ptakami*
I[F].slept with[INSTR] five.INSTR bird.M.INSTR.PL
"I slept with five birds"

In the previous sections, it was established that there is a genitive case assigner between the numeral and the noun. It is this genitive case assigner that is responsible for the genitive on the noun. Based on the data above, the question then becomes, why does the genitive case disappear to be replaced by the oblique in oblique positions?

Assuming numerals are nouns, it is expected that the oblique will appear on the numeral, but unexpected that it also shows up on the noun. Compare the example above with regular noun-noun construction:

(89) *Rozmawiałam* *ze* *studentem* *brata*
I[F].talked with[INSTR] student.MP.INSTR.SG brother.MP.GEN.SG
"I talked to (my) brother's student"

When it is not a numeral but a regular noun that appears initially, the oblique case does not spread to the second noun.

In section 3.3, it was argued based on the works of Franks (1994), Rutkowski (2005), and Rutkowski and Szczegot (2001) that the case assigned by the numeral is structural. As we consider numerals to be nouns in this account, it is predicted that the case of regular nouns is also structural.[23] Thus, this difference between the behavior of numeral-noun constructions and noun-noun constructions with regard to oblique case must be explained. There are two issues at present: (1) what happens to the genitive case assigned by the numeral? And (2) how does the oblique case spread to the noun?

My solution to these two issues also comes in two parts. In order to avoid claiming that the numeral is somehow able to lose its case assigning status and become similar to an adjective (as is done by Rappaport, 2003 and Franks (1994; 2002)), I will claim that both the genitive and oblique case are assigned to the noun, although only one appears, that being the oblique case. This in line with work on case stacking and the "overwriting" of case by Matushanksy (2008, 2010), Richards (2007), Pesetsky (2009), and Yoon (2004). Secondly, to explain how the oblique case arrives at this lower noun, I will rely on the functional nature of the numeral and the idea of case percolation.

I will discuss the plausibility of "overwriting" the genitive case first, followed by a discussion of how the oblique case reaches the lower numeral.

---

[23] There are also cases in which the noun appears to assign an inherent case. For example:

(i) *zagrożenie* *wyginięciem*
threat.N.NOM.SG extinction.M.INSTR.SG
"the threat of extinction"

Inherent cases in the nominal domain differ from structural cases in that the structural genitive must be adjacent to the assigning noun (ii), whereas inherently cased nouns can be scrambled (iii). These examples are based on Zlatic (1997, chapter 4).

(ii) a. *darowizna* *pieniędzy* *dla* *szpitali*
donation.NOM money.GEN for hospitals.DAT
"donation of money to hospitals"
b. **darowizna* *dla* *szpitali* *pieniędzy*

(iii) a. *zagrożenie* *wyginięciem* *przez* *ludzkość*
threat.NOM extinction.INSTR by mankind.ACC
"the threat of extinction by mankind"
b. *zagrożenie* *przez* *ludzkość* *wyginięciem*

This suggests that with the genitive, we are indeed dealing with a structural case.

4.4.1 Overwriting the genitive

The idea of case-stacking, which is a precursor to this idea of overwriting genitive case, is found in Richards (2007) for the language of Lardil and Yoon (2004) for Korean. It is argued that cases can be stacked, one after another. For example, in Lardil, nouns can be marked as both genitive and instrumental:

(90) *Ngada* *latha* *karnjin-i* *marun-ngan-ku* *maarn-ku*
I spear wallaby-ACC boy-GEN-INSTR spear-INSTR
"I speared the wallaby with the boy's spear"
(Richards, 2007: ex. 3)

The possessor is initially marked as genitive to indicate its possessive relationship. Later, instrumental case is assigned to the entire DP which spreads to both the noun and its possessor. As Lardil allows multiple case marking, both the genitive and instrumental can appear overtly on the possessor.

In contrast, in Polish, only one case is ever allowed overtly on a single noun phrase. Despite this, there is a strong similarity between the example above and the behavior of Polish numerals in oblique positions – in both cases do we find the oblique case on the upper and lower noun, but whereas in Lardil both the assigned cases can appear on the noun, in Polish, only one case is allowed, that being the oblique case.

Accounts by Pesetsky (2009) and Matushansky (2008, 2010) draw on this idea of case stacking and apply it to Russian. They both claim that Russian is a case-stacking language, taking the numeral-noun construction as direct evidence for it. I adopt the same approach here. Pesetsky argues that the last case assigned is the one that appears overtly. He verbalizes this idea as the "One-Suffix Rule", which states that "only the final overt inflectional suffix on a noun is pronounced" (Pesetsky, 2009: 2). Matushansky (2008) makes a similar claim to that effect, stating that as oblique cases are more marked than structural cases, they are ordered before the structural cases, and during lexical insertion rules, this causes only the oblique case to be spelled out.

The behavior of inherent case with respect to numerals and negated sentences suggests that Matushansky's proposal on case ordering is more

accurate for Polish. As we have seen before, when there is an oblique case assigner, it overrides the genitive case assigned by the numeral. In negated sentences, genitive case usually overrides accusative objects (91a); if there is an oblique case assigner, however, the object remains in the oblique case (91b). This is similar to what occurs with numerals:

(91) a. *Nie* *widziałam* *książki* */ *książkę*
Not saw[F] book.GEN / *book.ACC
"I did not see the book"
b. *Nie* *handlowałam* *książkami* */ *książek / *książki*
Not deal.in[F] books.INSTR / *books.GEN / *books.ACC
"I do not deal in books"

The proof for Matushanksy's approach over Pesetsky's can be found in the structures. In numeral-noun constructions, the genitive case would be closer and so, assigned first, followed by the oblique case. In negated sentences, however, the opposite occurs – the oblique case is assigned first, followed by the genitive:

(92) *Numeral-Noun:* [o [Numeral [g [Noun]]]]
Negated: [g [$\text{Verb}_{\text{Oblique}}$ [Noun]]]

'o' stands for oblique case assigner, and *'g'* for genitive; the structure assumes a silent case assigner.

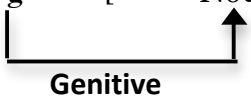
Thus, it does seem to be the case that oblique cases are ordered before structural cases when it comes to spell-out. Note that Pesetsky's rule is likely relevant in the domain of structural cases – in negated sentences, technically the accusative of the verb is assigned before the genitive of the negation, but it is the last-assigned genitive which appears. Thus, both rules are applicable.

However, to avoid positing that with oblique cases the first case is pronounced, but with structural cases it is the last, I claim that when the ordering of oblique and structural cases occurs, oblique cases are placed *at the end*. Then, when the One-Suffix Rule applies, it can apply across the board to all instances of case-stacking. This will lead to consistent pronunciation of the final case, after the reordering of obliques and structurals.

Now, assuming Polish can case-stack, then the noun in the numeral-noun construction will be marked with *both* the genitive and oblique case. As the oblique is the last assigned, no reordering is necessary; by the One-Suffix Rule, only the oblique case will be pronounced. I demonstrate how this would work.

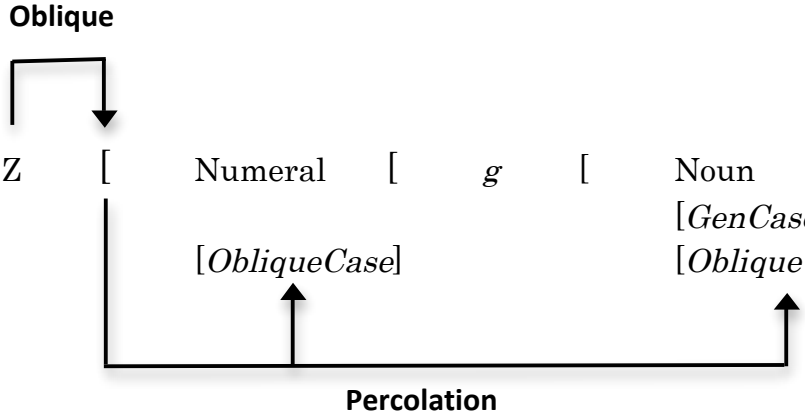
Assuming there is an intervening case assigning head between the numeral and the noun, the first Merge would consist of this head merged with the noun. This head would then Agree with the noun, assigning it case:

(93) [*g* [Noun]]



This is followed by Merges with the numeral and the oblique head. The oblique head has its oblique case to assign. Let us assume this case is assigned to the entire numeral-noun projection, at which point it percolates down. It will percolate down to the numeral, as well as the noun, assigning both the oblique case. For the numeral, this will be the first case it has received (unless it already has genitive due to the presence of a masculine personal noun), whereas for the noun, it will be the second case, making a concatenation of genitive + oblique.

(94) [Z [Numeral [*g* [Noun]]]]
[*GenCase*]
[*ObliqueCase*] [*ObliqueCase*]



By Pesetsky's One-Suffix Rule, only the last case assigned will be pronounced. Thus, for both the numeral and the noun, this will be the oblique case.

This mechanism is not case overwriting *per se*, but it is a mechanism in which although two cases are assigned, only one will be pronounced. Under this account, it is unnecessary to assume that the genitive case assigner somehow loses its case assigning ability in oblique positions; by appealing to case stacking,

a uniform featural and structural analysis can be maintained for numerals in all structural positions.

4.4.2 Case Percolation

The next question to address is why the percolation of the oblique case down to the noun is possible in numeral-noun constructions, but not in noun-noun constructions. This is the point at which it becomes important to discuss barriers to the percolation of case and Theta Theory.

In the systems of Matushansky (2008, 2010) and Pesetsky (2009), both structural and inherent case is assigned from a head to its complement, which is then allowed to percolate downwards, blocked only by case barriers. Pesetsky argues that phases are boundaries for this percolation, as once a phase has been sent to spell out, no more case markers can be concatenated; this possibility is acknowledged by Matushansky, but for the purposes of her work, she states that it is non-verbal heads which block case percolation. At any rate, as true as these might be, they cannot explain this difference between numerals and nouns. As numerals and nouns are the same category, they are expected to have the same phase barriers – if there were a phase barrier to block the percolation of case with the noun-noun construction, it would do so as well in the numeral-noun construction.

To solve this, I turn to Theta Theory. It has been acknowledged that inherent case is tightly linked to the assignment of theta roles in the verbal domain; Franks (2002) mentions the idea that inherent case is necessary to make the noun visible for theta-role assignment. I assume here that something of the sort is going on and there is a tight link between theta role assignment and inherent case assignment.[24] Turning to the Theta Criterion, it reads as follows: "each argument bears one and only one theta-role, and each theta-role is assigned to one and only one argument" (Chomsky, 1986: 36, as quoted in Radford, 2006). This means that when the oblique case assigner is assigning its

[24] This cannot be the only possible link though, since nominative subjects and accusative subjects are not inherently marked even though they also receive theta roles.

case which allows it to assign its theta-role, it must find something which can serve as an argument and that argument must not yet have a theta-role.

Taking this into account, I claim that the problem is simply that numerals cannot function as arguments by themselves. This is due to the fact that in addition to being lexical, they are also functional, i.e. semi-lexical. As a result, they are too functional to play the role of the argument as needed by the oblique case assigner. Conceptualizing this in terms of theta-role "slots" where the theta-function can be encoded, numerals do not have such a slot, whereas nouns do. As a result, when inherent case is assigned at the phrase level, it percolates downwards, in search of an empty theta slot; it will saturate all elements with that inherent case along the way, stopping only once it has found an item which can hold a theta role slot, that being the noun.[25] This will put inherent case on both the numeral and noun. In noun-noun constructions, the first noun will already have a theta-slot, in that way blocking further percolation of that case. Essentially, it is the fact that the numeral cannot serve as an argument (due to its functional nature) that leads to the appearance of inherent case on both the numeral and the noun.

This explanation can apply in other domains as well, also involving semi-lexical elements. For example, indefinite pronouns in Polish also show the case assigning-case agreeing alternation that has been the subject of this discussion (examples adapted from Rutkowski, 2001):

(95) a. *Widziałam* *coś* *miłego*
I[F].saw something nice.GEN
"I saw something nice"

b. *Spałam* *z* *czymś* *miłym*
I[F].slept with[INSTR] something.INSTR nice.INSTR
"I slept with something nice"

[25] Phase boundaries will likely also play a role, depending on how much structure there is in the noun phrase.

If these elements, being semi-lexical, also cannot serve as arguments, lacking the slot for theta-roles, then it will again be this percolation of inherent case which leads to instrumental on both indefinite pronoun and adjective.[26]

Cross-linguistically, the case assigning – case agreeing alternation occurs with numerals in multiple languages, including Inari Sami, Skolt Sami, Finnish, Estonian, and other Slavic languages (Nelson and Toivonen, 2000; Feist, 2010; Rutkowski and Szczegot, 2001; Karlsson, 1999; Franks, 1994). It is in these languages specifically that it is most tempting to consider the numeral the syntactic head of the construction, as it assigns a structural case to nouns in structural environments. Thus, it may be the case that like in Polish, these numerals are also semi-lexical and cannot serve as arguments; a stronger stance might claim that *all* numerals in *all* languages are semi-lexical and unable to serve as arguments. Under such a stance, case alternations would even be expected in languages where numerals assign case.[27]

This explanation provides us with a further insight. There has been much confusion associated with numerals due precisely to the fact that the semantic and syntactic heads of the construction appear to differ. Thus, the numeral seems to act as the syntactic head, dictating agreement and case assignment, while the noun acts as the semantic head, i.e. the true holder of meaning. This produces a misalignment in terms of semantic and syntactic head which is not so common. However, by this analysis, this is entirely expected. Syntactically, the numeral *is* the syntactic head, but semantically, it cannot hold the theta role and this falls to the noun instead, thus, making the noun the "semantic head". The confusion is a result of this misalignment.

---

[26] These examples lead to another important question, which I will not discuss here – if the percolation of case does *not* find a theta slot – should the derivation crash? Presumably, the adjective cannot hold the theta role either, so there are no theta slots within the projection in example (95). This may suggest that the theta role can appear on the highest projection (which perhaps always comes equipped with a theta slot), but in addition, it also seeps downwards to find the element which is the actual theta role holder; this would prevent violations in the Theta Criterion, like seem to occur in (95). This idea remains to be developed.

[27] Serbo-Croatian is a case-assigning language which does *not* enter into the alternation. However, the fact that it does not agree in inherent contexts might be explained independently; for example, Serbo-Croatian might not allow case-stacking, or instead, it might be the case that the first assigned case is the one that is pronounced instead of the last.

To summarize, there are two main claims here: Polish is a case-stacking language, and numerals cannot serve as arguments. Using these two assumptions, it is possible to derive the case alternation found in Polish. Because the numeral cannot serve as an argument (it has no theta slot), inherent case will need to extend down to the lower noun, assigning it a theta role. Since case-stacking is possible, the oblique case will be assigned after the genitive, but due to the One-Suffix Rule and the rule concerning ordering, only the last assigned case, the oblique, can be pronounced. This account has the advantage that numerals are treated uniformly in both structural and oblique positions; furthermore, it explains how the numeral can act like a syntactic head, while failing to act as the semantic head.

4.5 Syntactic Structures

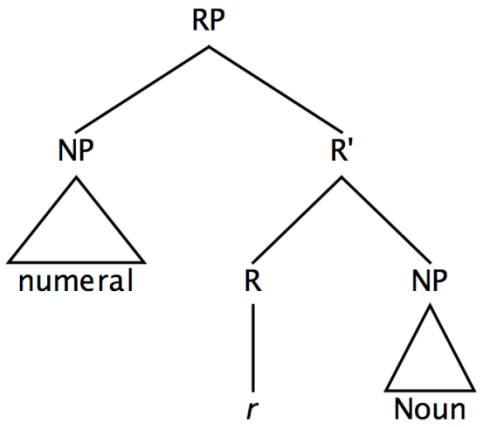
In this section, I provide a detailed description of the structure of the numeral-noun construction. Up to this point, I have only alluded to the potential structure. Based on my analysis so far, there are a few points which must be incorporated into any structure. Firstly, the case assigner must be a head separate from the numeral. This is necessary to derive the cyclicity effects found with masculine personal gender. Additionally, the numeral and case assigner must sit in a configuration in which it is possible for the numeral to have an influence on the case assigner; this is necessary to ensure that the case assigner in some way "inherits" the defectiveness of the numeral. Lastly, and more obviously, the case assigner must sit structurally in a position between the numeral and the noun.

Treating the genitive case assigner as a separate head has its advantages theoretically. From the data, it is clear that there must be genitive case assignment, whether it is from the numeral itself or some functional head in between. Theoretically, however, it is more sound to treat the genitive case as coming from a separate head. Consider how things might look if the case actually came from the numeral itself. As a genitive-assigning defective noun, this would mean that it could also act as a probe. This presupposes that the defective noun also has some uninterpretable feature besides case which is in need of a value.

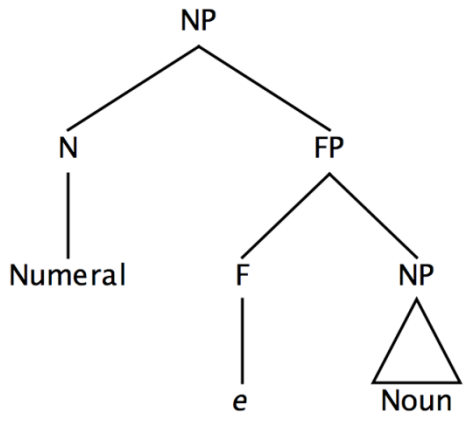
Furthermore, this uninterpretable feature can only be valued by another noun. Unfortunately, this makes little sense – the set of uninterpretable and interpretable features should be the same for different nouns (putting aside those defective ones for the moment), so no nouns should ever be able to value the uninterpretable features on another noun. If we claim the numeral is the case assigner, it amounts to saying that all nouns are case assigners (since they should have identical structures and be nearly identical with regard to case assignment) and furthermore, that nouns can satisfy their features through agreement with another noun. This is a very illogical statement, but is the result of assuming that numerals, and nouns in general, are the actual case assigners. Because of this issue, I reject any structure which treats the numeral as the case assigner.

Thus, we know that there must be a separate case assigner. There are two relevant structural possibilities here, which are presented below. Remember that this structure will also be the basic structure for noun-noun complementation.[28]

(96) a.



b.



The first structure is borrowed from Den Dikken's (2006) Relator. Den Dikken promotes an analysis in which predicational relationships are mediated by a functional element R, a Relator, where the subject and predicate sit in specifier or complement position. Borrowing Den Dikken's structure (but not assuming

[28] Potentially, not all noun-noun constructions have the same structure. Den Dikken (2006) discusses varying syntactic structures for different types of noun-noun constructions in Dutch and English. The structures posited here will likely be relevant for only some constructions.

the predicational relationship here), the numeral would sit in the specifier position and the noun in the complement position, both being mediated by the relator *r*, which in this case is our genitive case assigner.[29]

The alternative structure involves treating the case assigner as occupying its own projection. This is the structure posited by Ionin and Matushanksy (2006), with an additional case assigner assumed. The numeral is a head and selects the case assigner (or rather, selects the noun, and the case assigner comes as a consequence of that selection).

Complex numerals provide a clue as to which structure is best. In Polish complex numerals, it is the last numeral in the complex that dictates the agreement and case assignments. Thus, for example, the numeral 22 would behave differently from the numeral 25.

(97) a. *Dwadzieścia-dwa* *ptaki* *spały*
Twenty-two.M.NOM bird.M.NOM.PL slept.NON-VIR.PL
"Twenty-two birds slept"
b. *Dwadzieścia-pięć* *ptaków* *spało*
Twenty-five.NOM bird.M.GEN.PL slept.N.SG
"Twenty-five birds slept"

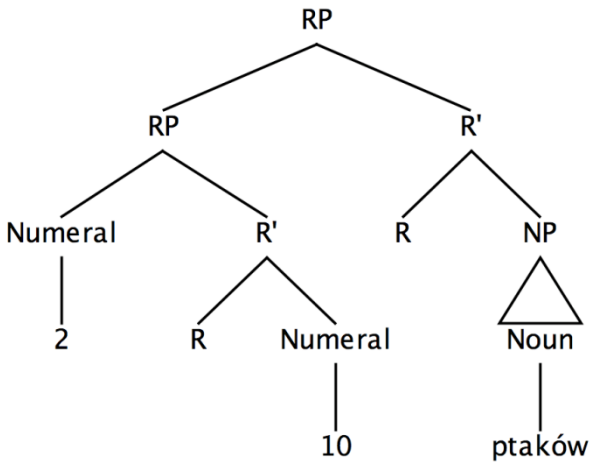
With regard to case assignment (or lack thereof) and verb agreement, the numerals 22 and 25 behave identically to 2 and 5, respectively. Numeral 22 agrees in gender with the following noun,[30] appears throughout the subject as nominative, and has successful verb agreement; numeral 25, on the other hand, assigns genitive and lacks verbal agreement. This suggests that whatever the structure is, it is the last numeral which determines agreement and case.

[29] Corver (2001) suggests the alternative ordering in which the numeral sits in complement position and the noun in specifier position. This structure requires predicate inversion, in which the head *r* raises to a higher functional head as does the complement numeral to the specifier position of that higher phrase. I do not adopt this analysis for Polish, however, due to the need for locality between the case assigner and the numeral – with the numeral in complement position, it presumably is unable to control the properties of the genitive case assigner and hence, leaves us without an explanation for the different behaviors of the case assigner with different numerals.

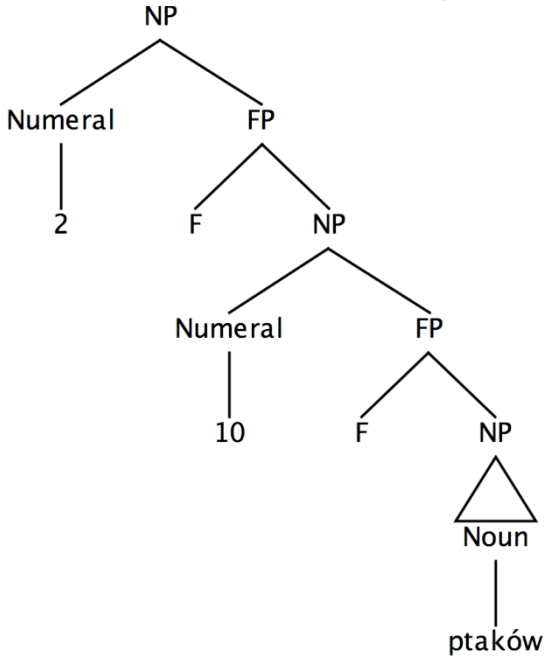
[30] It in fact agrees twice, in each instantiation of *dwa* ('two'); compare *dwiedzieścia-dwie dziewczyny* (feminine, '22 girls'), *dwadzieścia-dwa ptaki* (masculine, '22 birds'), and *dwudziestu-dwóch chłopców* (masculine personal, genitive throughout, '22 boys'). The nominative form with masculine personals is not possible in complex numerals, **dwajdzieścia-dwaj chłopcy* ('22 boys').

Ionin and Matushansky (2004, 2006) present analyses of complex numerals. They treat multiplicatives (2x10 = 20) as cases of complementation and additives (20+2) as cases of coordination, with right node raising. Based on the analysis, the numeral 22 would involve both complementation (of 10 as a complement of 2 to get 20) and coordination (of 20 coordinated with 2 to get 22). To simplify matters, I will diagram the two tree structures for a multiplicative example, the numeral 20 (which behaves like 5+ numerals).

(98)



(99)



The trees above show what the multiplicative structure would be for the numeral 20 (2 x 10). The first tree considers the structure in which the noun and numeral are moderated by a linking element R. When an additional layer is added, it is

added in the specifier. The second tree considers the structure in which both the numeral and case assigner are separate heads in separate projections. Here, additional structure is added at the top, through more and more layers. This is the structure advocated by Ionin and Matushansky for complex numerals.

The structure I support in this paper is the second one, similar to the analysis of Ionin and Matushansky. The reason for it is this. In complex numeral constructions, it is the last element which dictates the agreement. Similarly this last element is what chooses the case assigner, whether it is a case assigner that can actually assign case (5+) or a case assigner that cannot (2,3,4). Recall that one of the requirements is that the numeral and case assigner be in such a configuration that the numeral can have an effect on the case assigner. The tree in (86) does not seem to allow this. The specifier does not hold the numeral which controls R, but rather, another RP. If we take this structure, then it is necessary to somehow cause the defectiveness of the numeral in the complement of the RP specifier to spread to the make the big RP defective, which then makes the R of the big RP defective. This does not seem very likely, and is rather complicated, as the numeral is deeply buried deep; with more complex numerals, it only sits deeper. Because of the inability of this structure to deal well with complex numerals, I reject it. Thus, I promote the second structure, in which the numeral is its own head, as is the case assigner.

4.6 Back to Partitives

In this section, I return to the partitives I presentend in section 4.1. Recall that for the most part, partitives behave identically to quantitatives: depending on the numeral, they either block or allow verb agreement and masculine personal nouns are or are not a problem. The point at which they differ concerns oblique case assignment – whereas oblique assignment overwrites the genitive with quantitatives, it does not do so with partitives. In this section, I explore this difference and propose an analysis of partitives that is consistent with the analysis I have developed throughout this chapter.

An important question concerning partitives is whether they contain one noun or two nouns. Certain researchers like Martí Girbau (2003, 2010) argue

against the two noun approach and claim that there is only a single noun in the structure; others, like Kranendonk (2008) assume that there are two nouns, one immediately below the numeral which is silent (*pro* in Kranendonk's analysis), and a second one, which is overt and follows the preposition-like element. For both analyses, the intermediating preposition is not considered a preposition *per se*, but rather a special partitive head.

Structurally, the two analyses do not differ very much. Both involve a use of Den Dikken's (2006) Linker and Relator heads, along with inversion. In Den Dikken's work, the Relator head is posited to mediate between a subject and a predicate, where either can sit in the specifier or complement position. When there is a need for predicate inversion for whatever reason, the Linker head merges with the RP. The R head then moves to the L head; this move makes it such that the specifier and complement of the RP are now *equidistant*, and the complement can move to the specifier position of the LP. This is demonstrated below.

(100) $[_{LP}$ [R-Comp]$_i$ *l* + r_j $[_{RP}$ [R-Spec] t_j t_i]]

Martí Girbau (2010) uses this same mechanism to derive the partitive structure, although she does not argue that this is a subject-predicate relationship, claiming instead that it is a quantifying relationship and the inversion occurs for reasons of case. Martí Girbau argues that the numeral is generated in a position below the partitive head and noun, i.e. in the complement position of the RP. She renames LP and RP as KP and FP, respectively. The numeral begins in the complement of FP, then after movement of the F to K, the numeral moves to the specifier of KP. As F+K has a case to assign, she argues that by Burzio's generalization, it needs there to be something in its specifier in order for it to assign case; this is the numeral, which has moved to the specifier position. Case is assigned and through merges of a Num head (producing NumP) and D head (producing DP), the DP is completed.

In Kranendonk's work, the same mechanism is used: the numeral begins in the complement position of the lower phrase, then after movement of the

head, moves to the specifier position of the higher phrase. However, Kranendonk differs in that the numeral is not just the numeral, as in Martí Girbau's analysis, but is a numeral + *pro*. It is this complex which moves up to the specifier position.

In this thesis, I will also adopt the predicate inversion analysis of these two authors. However, there remains one issue – is there one noun or two? I will argue here that there must be two nouns in the Polish partitive.

My first reason for claiming there are two nouns falls from the case assignment, agreement alternation – partitives do not enter this alternation, while quantitatives do. Recall that the explanation for this was because semi-lexical numerals cannot act as arguments – they are missing a theta-slot. Because of this, when inherent case is assigned to the numeral-noun projection, it must percolate down until it finds an empty theta slot, that being on the noun. In the partitive construction, the genitive noun does not receive the oblique case. This must mean that there is some element which can block the percolation of inherent case, most likely by virtue of having a filled theta-role. This element, I argue, is the covert noun, which appears immediately after the numeral. If this *pro* takes the theta role, it will stop the further percolation of case. Under the single noun approach, there is nothing to block the percolation of case, and it would be expected to appear throughout the partitive.

A second argument comes from the case assignment found with masculine personal nouns. Recall that when a masculine personal noun is present, in both partitives and quantitatives, the numeral is also marked as genitive. Under the single noun approach, this would mean that the partitive head was forced into upward agreement. In contrast, the dual noun approach assumes that it is the silent case assigner between *pro* and the numeral which is forced into upward agreement. This is an important difference between the single and dual noun approaches. In the quantitative construction, the probe searches upward because it cannot deal with the masculine personal gender value; the deficiency in the probe is assumed to be related to the deficiency in the numeral. The single noun approach would force us to apply that same reasoning to the partitive head, where depending on the numeral the partitive head does or does not agree

upwards. This is exactly where the problem lies with the single noun approach. First of all, the numeral is initially merged below the partitive head and noun – thus, it is not sufficiently local to determine the behavior of the case assigner. Secondly, assuming this first issue is overcome somehow, if the defectiveness of the numeral were at play, there would be no genitive with numeral 1, and numerals 2,3,4 would only assign genitive with masculine personal nouns; this shows that the defectiveness can*not* be inherited from the numeral, and it is unclear *why* the partitive head would need to probe upwards. It is impossible to maintain the account for masculine personal nouns under the single noun approach. With the dual noun approach, however, the phenomena with masculine personal nouns can occur exactly as it would for a quantitative, restricted to the numeral, silent genitive case assigning head, and *pro* noun.

A third reason concerns the verb agreement. In partitives, verb agreement occurs just as it would were it the quantitative construction, with the success of agreement depending on the value of the numeral. Considering that the overt noun in the partitive construction is already marked as genitive, under the single noun approach there is nothing for the verb to agree with when the numeral is 1 or 2,3,4. Under the dual noun approach, however, the verb can be assumed to agree with the *pro* noun with numerals 1 and 2,3,4, with the numeral for numerals 1000+, and with the defective numeral that leads to agreement failure with the 5+ numerals – exactly as would occur in the quantitative construction. Thus, the dual noun approach is better able to handle the agreement facts of Polish numerals.

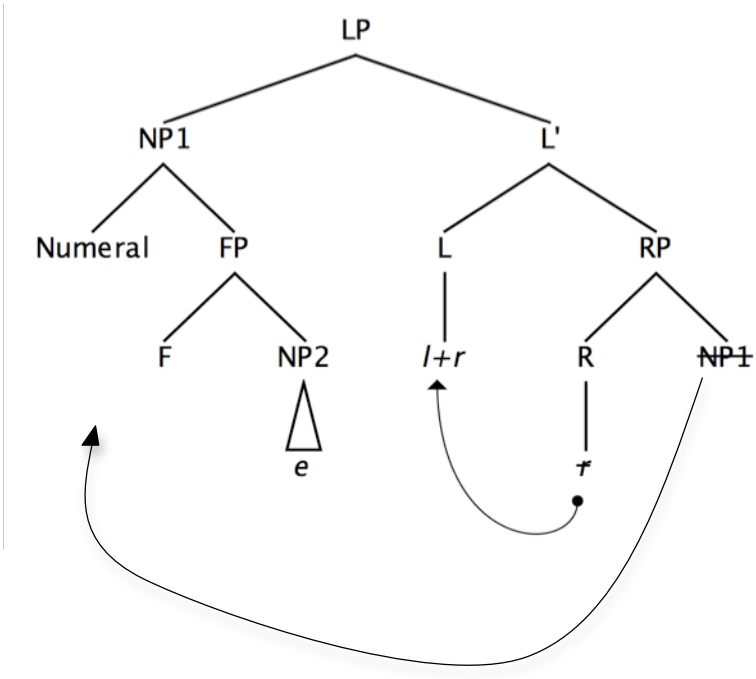
A last and final consideration for the dual noun approach concerns case assignment. Because the overt noun is always assigned genitive by the preposition-like element *z*, there will be, in many instances, two separate cases within the partitive, one on the overt noun and another on the numeral. This is exactly what occurs in oblique positions. For some of the numerals, under the single noun approach, this is fine, as these numerals are also assumed to be nouns and can function as a goal for the case assigner. The problem, however, comes from the numeral 1. As discussed previously, 1 is identical to an adjective. This means that even in the partitive construction, it should still be an adjective

(granted that there must be some restriction to prevent all adjectives from entering into the partitive construction). Because of this, it cannot serve as a proper goal for a verbal case assigner. The curious question then is, under the single noun approach, how does the adjectival numeral 1 get case? Under the dual noun approach, this is accounted for – the verb agrees with the *pro* noun, and the adjective in turn agrees with it in case, gender, number, and subgender.

Thus, to summarize these arguments, the dual noun approach is more compatible with the analysis presented in this thesis and with the data found with Polish partitives. Verbal agreement remains unexplained under the single noun approach, as does the mechanism of case assignment with the numeral 1. The single noun approach is incompatible with the analysis of masculine personal nouns I have promoted here, and similarly, is difficult to consolidate with the account of the case agreement dichotomy. For these reasons and for the fact that these problems do not occur with the dual noun approach, I have adopted the dual noun approach.

To conclude this section, I provide a tree structure. This is essentially the same analysis as Kranendonk, the difference being the structure of the quantitative contained in the numeral portion of the partitive.

(101)



4.7 Interlude: What it Means to be Defective

This section is rather tentative and includes only speculation on what it means to be defective, expressing this in terms of features. With the 5+ numerals I have alluded to the idea that their defectiveness is the result of underspecification – I will expand on this idea here. I have not said anything about the defectiveness of numerals 2,3,4, but I will also approach that in this section.

I begin by exploring the differences between numerals 5+ and numerals 2,3,4. Concerning gender agreement, numerals 5+ seem to lack gender agreement altogether, whereas numerals 2,3,4 show what appears to be agreement. We see this in the example below:

(102) a. *Pieć ptaków / dziewczyn / krzeseł*
Five bird.M girl.F chair.N
"Five birds / girls / chairs"

b. *Dwa ptaki / krzeseł; dwie dziewczyny*
Two.M/N bird.M chair.N two.F girl.F
"two birds / chairs ; two girls"

Numerals 5+ use the same form regardless of the gender of the noun (102a). Numerals 2,3,4, on the other hand, vary the form depending on the gender of the noun (102b). It is this property of numerals 2,3,4 that has led many researchers to group them with the numeral 1, as adjectives. I do not take this approach here and maintain that numerals 2,3,4 are more like numerals 5+ than numeral 1. However, I will use the idea that numerals 2,3,4 are like adjectives in terms of agreement to derive their agreeing behavior.

Baker (2003) outlines an analysis of categories in which a syntactic category is based on the features and properties of that lexical item and on the potential structures it can have. In this system, verbs are the only category which can license specifiers, nouns are the only category with referential indices, and adjectives, as a category, lack both. Additionally, each lexical category can be immediately dominated by a functional head which roughly matches it in features. Since verbs license specifiers, only verbs can have subjects; similarly,

since nouns have referential indices, only they come with a full set of phi features; and lastly, since adjectives have neither, they cannot host subjects nor they do not have their own features. In principle, each lexical category can act as a probe and agree. However, because nouns have their own set of phi-features, and an immediately dominating functional head with those same features, they are blocked from agreeing with anything but themselves (intervention effect of the functional head). Adjectives, on the other hand, have no features, so the functional head is never an intervener and they can always agree. A similar story occurs for verbs.

Certain ideas of this system are useful for my discussion here. In Baker's system, one of the defining differences between an adjective and a noun is this referential index which leads to a full set of phi-features on the noun, but not on an adjective. This leads to the generalization that nouns enter a derivation with a full set of phi-features, and adjectives with an empty set. If nouns and adjectives are determined by whether they have a full set of phi-features or not, it is possible to play with the nature of adjectival and nominal categories, by adjusting the featural set of an element. For example, a "cross" between an adjective and a noun, if such a thing exists, might have half the phi-features inherent and valued (say, through a referential index) and the other half unvalued looking to probe. I am proposing that in this way, it is possible to derive the adjective-like and noun-like numerals featured in the data. Let me discuss how this would work.

In chapter 2, I determined numerals 5+ were noun-like and numerals 2,3,4 adjective-like. What would this mean in terms of features? For 5+ numerals, this would mean that like nouns, they come with phi-features; however, unlike nouns, they do not come with the full set of phi-features. They are *missing* certain features. These numerals do not act like adjectives in any way, so it is unlikely that they would enter the derivation with unvalued uninterpretable features like adjectives would. For the numerals 2,3,4, however, they are adjective-like. They are more similar to this idea of half-noun, half-adjective. Numerals 2,3,4 would presumably come with some valued phi-features, but in addition, some phi-features might be expected to be unvalued

and uninterpretable, able to act as probes just like what is found with adjectives. Thus, in terms of features, both numerals 5+ and 2,3,4 have some phi-features; however, they are either missing certain features (5+) or have unvalued uninterpretable features (2,3,4), which contributes to their semi-lexical status.

The next question is: what sort of features do these numerals come pre-equipped with, and which features are missing or unvalued? The most logical answer to the first question is number. Considering that numerals encode cardinality, it is reasonable to assume that they have an inherent number feature, most likely valued for plural. Because of this feature, numerals 2,3,4 differ from adjectives; the presence of this feature makes them similar to nouns, as valued features are a property of nouns. As for missing or unvalued features, the most likely candidates are the gender and subgender features, due to the fact that numerals 2,3,4 appear to probe for gender. By analogy, it is plausible 5+ numerals are also missing those features. Because of these missing and unvalued features, both numerals 2,3,4 and 5+ will make inadequate goals for any probe – numerals 5+ due to underspecification, and numerals 2,3,4 due to unvalued uninterpretable features.

These unvalued, uninterpretable features of numerals 2,3,4 make these numerals similar to adjectives. Thus, if these numerals do truly have such features, it would mean that they could be probes in the same manner that adjectives are probes and could also agree (with the genitive case assigner still there as a remnant of their once nominal status). The fact that these numerals do agree in gender is evidence for such a view. If the numerals are probes, then they should be able to probe into the noun and value their gender and subgender features, producing the gender marking found on these numerals.[31] Furthermore, because they are like adjectives and not verbs, they do not assign case when they agree, leaving the noun open for agreement with the verb.

[31] An alternative to the analyses presented in the previous sections might be to assume that the numeral also probes for case. In this way, it is never assigned case by the verb or the inherent case assigner, but picks it up through agreement with the noun. This idea is congruent with the data, since the case on the numeral and noun is identical regardless of position or gender.

To summarize, I am tentatively suggesting that the defectiveness in 2,3,4 and 5+ numerals can be found in their feature sets. Both still have interpretable number features, likely inherently value for plural, but in addition, the numerals 2,3,4 have uninterpretable gender and subgender features, while the 5+ numerals are missing these features altogether. Due to the uninterpretable features, numerals 2,3,4 can act as probes and agree in gender and subgender with the lower noun. Because numerals 5+ do not have uninterpretable features, they never act as probes, although they are still defective for verb agreement due to their missing features. To conclude this discussion, I present an overview of what the feature sets of the elements discussed in this paper might look like:

(103) **Adjectives, numeral 1:** all uninterpretable features
Numerals 2,3,4: interpretable number, uninterpretable gender, subgender
Numerals 5+: interpretable number; no gender, subgender
Nouns, numerals 1000+: all interpretable features

4.8 Summary

The mechanisms introduced in this chapter are fairly simple, but through their interaction, produce the complex set of data surrounding Polish numerals. The first idea used here is default agreement. Default agreement is a phenomenon by which agreement is attempted, but fails, resulting in default feature assignment. This is found with 5+ numerals and with masculine personal nouns that trigger cyclic agreement. A second idea is multiple agree. Multiple agree allows the probe to agree with more than one goal, assuming the first goal is underspecified and cannot value all the features of the probe. Multiple agree is found with numerals 2,3,4 in nominative position (deriving nominative on both the numeral and noun), and with masculine personal nouns, during cyclic agree. The next important concept is cyclic agree, the idea that if a probe is not able to deactivate after a search in its c-commanding domain, it can spread its search space upwards. This was argued to be the case with masculine personal nouns for both numerals 2,3,4 and 5+, thus deriving the genitive on the numeral and noun. Finally, the last concept introduced here was case stacking, i.e. the idea that nouns can be assigned multiple cases, but only one of those

cases is pronounced – the oblique if there is an oblique case, and the last assigned structural case if there is not. Combining these various innovations to Agree and Case Theory with the assumptions that numerals are defective nouns which are unable to host a theta role, it is possible to derive the wide array of data surrounding Polish numerals, for the numerals 2,3,4 and 5+.

I summarize here their interactions. For the numerals 2,3,4, the verbal probe attempts to agree with the numeral but fails; instead, it continues the search downwards, finding and agreeing with the noun. The numeral takes default nominative. Thus, both numeral and noun have nominative case marking, and the verb agrees. With masculine personal nouns, however, the case assigner is somehow able to recover its ability to assign genitive case; then, because it cannot read the masculine personal gender value, it extends its search space upwards, also assigning genitive to the numeral. Thus, both appear marked as genitive. Later when Tense tries to agree, both numeral and noun are already marked as genitive, and thus deactivated, and agreement fails, leading to default agreement. In oblique positions, because the numeral cannot hold a theta role, the inherent case percolates down to the level of the noun, with inherent case appearing on both the numeral and noun. This derives the behavior of 2,3,4 numerals.

For the 5+ numerals, the story concerning the masculine personal gender and oblique case is the same. The genitive case assigner (which *is* able to assign its case), ends up assigning genitive to both the numeral and the noun in an attempt to value its features. In oblique positions, the numeral cannot hold a theta role, and so the inherent case percolates down, assigning case to both numeral and noun (note that the case assigner still assigns its genitive here, but because of the One-Suffix Rule, it does not appear overtly). The point in which 2,3,4 and 5+ differ concerns case assignment – numerals 2,3,4 do not usually assign genitive case, while 5+ numerals always do. As a consequence, verbs can agree with the noun with 2,3,4 numerals, but not with 5+ numerals, and since 2,3,4 and 5+ numerals are defective, lacking in (valued) phi-features, verb agreement fails with the 5+ numerals.

Numeral 1 is an adjective; as an adjective it simply agrees in case, gender, and number with the following noun. Cyclic agree, multiple agree, and default agreement are unnecessary and do not occur with this numeral.

Numerals 1000+ are nouns and fully lexical.[32] Because of this, they assign a genitive case consistently, carry their own phi-features, and seem to even have their own theta-slot. Verb agreement does not usually fail with them, oblique cases do not overwrite their complements, and masculine personal gender does not have any effects on them. Presumably, as fully lexical nouns, they do not have a defective case assigner and can carry a theta-role, these elements being the start of the problems with 2,3,4 and 5+ numerals.

[32] Although this is changing in the domain of verbal agreement, c.f. footnote 3. I do not address that behavior in this paper.

5 Discussion and Conclusions

To conclude, once we adopt a system in which cyclic agreement, default agreement, and case stacking are normal and acceptable processes, the patterns of Polish falls out from a simple fact: Polish numerals are semi-lexical nouns which are phi-defective. Because Polish numerals are defective, verbs cannot agree with them; furthermore, the genitive case assigner inherits this defective nature from the numeral and as a result is unable to read masculine personal gender, leading to upward agreement. Additionally, because the numerals are semi-lexical, they cannot carry theta roles; thus, oblique case must percolate down to the noun, creating the case assigning, case agreeing alternation. Numerals 2+ differ from 5+ numerals only in that the case assigner is unable to assign case, except with masculine personal nouns. For this reason, verb agreement occurs with 2+ numerals, but not with 5+ numerals. Thus, as a result of the phi-defectiveness of 2,3,4 and 5+ numerals, we see the patterns that we do in Polish.

In the rest of this chapter, I will consider what the implications of this analysis are for theories of agreement and case theory. Considering that my analysis implements a number of innovations, which although proposed by various authors, are not yet mainstream, it is important to explore what these innovations would mean to the current theory. This I do in the following section.

5.1 Implications

5.1.1 Agree

I proposed two innovations to Agree: default agreement (Preminger, 2011) and cyclic agree (Rezac, 2003; 2004). What do these mean to the theory of Agree and case assignment?

Default agreement has consequences for the time-bomb approach to Agree. In the instantiation of Agree I introduced in chapter 1, if a probe does not value all of its features through agreement, the derivation will crash. With default agreement, derivations do not crash even if agreement fails. In fact, this idea of

default agreement is very important, considering there are numerous subjects allowed in language which would *not* be expected to have phi-features – these include CPs, TPs, even VPs, APs, and PPs if such subjects are possible in a language; this also includes those cases in which there is no subject and no raising to subject position. Because such sorts of sentences are very common in natural language, it is important to introduce a mechanism of dealing with them. Default agreement does this very well and is both necessary and logical.

Thus, introducing default agreement into the theory of Agree presents a rather smart change. In the empirical world, it does not shrink the set of sentences for which Agree is explanatory; rather it increases that set. Note, however, it is important to implement Preminger's (2011) idea that there must be at least an attempt to Agree before default agreement occurs; this is necessary to rule out the ungrammatical agreement mismatches. For grammatical agreement mismatches, however, it is likely that default agreement is at play.

I turn now to cyclic agreement. This was proposed by Rezac (2003; 2004) and promoted also in Bejar (2000) and Baker (2008). What these researchers agree on, is that Agree can target goals which do not sit within a probe's c-commanding domain. They support this heavily throughout their work, with examples from Georgian, Basque, and Icelandic, for example. So how logical is this idea of upward agreement and how does it fit into the theory?

Under Rezac's (2003) implementation of cyclic agreement, Agree turns upward only if it has failed in its downward search. This predicts that canonical agreement will be found within the c-command domain, and that upward agree will only be found in those instances in which it fails. Furthermore, similarly to the case with default agreement, upward agree does not diminish the empirical domain of Agree, but rather increases it. For the Polish numeral system, it was previously necessary to posit some sort of accusative case assigner immediately above the numeral; under the nominative-genitive hypothesis, it might be necessary to claim there is some genitive case assigner immediately above the numeral which is only present with masculine personal nouns. This is messy and theoretically unsound – upward agree allows us to explain the appearance of genitive case in a more elegant way. Rezac's use of the Earliness Principle makes

any stipulations concerning the domain of Agree unnecessary and in fact, because of it, upward agreement follows naturally. Thus, it seems that allowing the probe to extend its search space with the addition of new material is also a rather clever and smart idea.

There are two innovations here, each which are logical and welcome changes to Agree. They increase the explanatory adequacy of the agreement theory and allow it to capture those cases in which agreement seems to look upwards, or agreement is unable to find a goal and receives default features instead. Using these ideas, we can redefine Agree such that the search space can extend if complements are inadequate, and default agreement occurs when no proper goals are found.

5.1.2 Case Theory

There was one addition with regard to Case Theory and that was the idea that cases can stack. Data from Lardil (Richards, 2007) and similar languages have triggered an interest in the idea that even languages which do not show case stacking overtly allow case stacking. This was the approach taken by Matushanksy (2008, 2010) and Pesetsky (2009). In addition, the idea of case stacking has promoted even more widespread changes concerning case. For example, in the work of Matushansky and Pesetsky, case is assumed to be assigned from a sister to a complement, and percolate down as far as possible; the mechanism for inherent and structural case assignment is the same, but barriers and phases serve to limit this spreading of case. Pesetsky (2009) takes his new case idea even farther and argues that case is not really assigned; it is an instantiation of a lexical category, where, for example, nouns come pre-marked with genitive case, and that case spreads to its sister once the phrase is complete. The next case assignment will overwrite that genitive case to, say, accusative or nominative depending on the final position.

In the approach I have taken here, I have not adopted these theories, in part because I have not yet had a chance to fully understand their impact on the language, and also in part because they were unnecessary. However, such theories may end up solving a number of case related theoretical issues, and they

are important on the wider scale. For the data and analysis at hand, however, it was enough to acknowledge that cases can stack.

The move towards case stacking forces a rejection of the Case Filter, which states that a noun can have one and only one case. Clearly the data found in Lardil argues against the Case Filter as is, which might instead be a language specific PF requirement rather than a language universal. Furthermore, dropping the Case Filter and adopting case stacking (with special rules on spell-out) allowed for a more uniform syntactic treatment of certain phenomena and constructions, for example numerals. By adopting case stacking it was possible to argue that numerals *always* assign genitive case. Furthermore, in the negative construction discussed in section 4.4.1, it is possible to claim that verbs still assign their accusative case and negation still assigns its genitive, even when it may not seem so overtly; this prevents the need for analyses in which the case assigning abilities of one item are dependent on the type of case assigned by another. With case stacking, it is possible to present a more uniform approach to these phenomena, and that is definitely an advantage over the Case Filter. Furthermore, this innovation allows for more elegant solutions for the more messy data.

5.1.3 Numerals

The third discussion I have planned for this section concerns numerals. As mentioned before, Corbett (1978) has noted a continuum among numerals, in which lower numerals are more adjective-like and higher numerals more noun-like. This continuum has made the classification of numerals rather difficult. Additionally, numerals have lead to much discussion as to whether they are heads, what their structures might look like, what sorts of categories they might be, and so on. In this paper, I have attempted to provide answers to many of these questions – numerals are heads in Polish, they sit in a head position with a genitive assigning complement, and they are nominal, although semi-lexical.

Potentially, many of these properties are language specific. In other languages, numerals might not actually be nouns; as we have seen a wide array of variation in one language, it is not so strange to assume that such variation

could occur across languages as well. Thus, in some languages, all numerals might be nouns (semi-lexical or full), adjectives, or something else entirely. Furthermore, depending on what category they are, it will affect whether or not they are heads and what the structures would look like. This is clear in Polish for the numeral 1, which must have an adjective structure, and the numerals 1000+ which need a nominal structure. To determine the category of a numeral, its headedness, and the structure, likely careful investigation is needed in each language. However, it is still possible that the range of variation among numerals is limited. In Polish, we saw four different classes of numerals; however, these numerals all sat on a continuum between adjective and noun. Thus, the range of variation in Polish seemed to be restricted to adjective, noun, or something in between. It is possible that there are similar restrictions in other languages, and this might be where cross-linguistic generalizations can be made and universals can be found.

For the Polish data at hand, the ideas concerning the numerals have fit naturally, given the data. Thus, it is also reasonable to assume that there are other languages which pattern in very similar ways. As the previous paragraph hinted, I do not intend to claim that all numerals are the same cross-linguistically – much the opposite, I believe they are subject to variation, but I do believe that that variation might be constrained in some way, although this is a topic left for future research. Essentially, as a class of items, numerals are subject to much variation both within and likely between languages.

5.1.4 Categories

In section 4.7, I discussed the featural make up of Polish numerals, suggesting that numerals 2,3,4 and 5+ are semi-lexical, while 1 and 1000+ are adjective and noun. This produces a continuum between numeral 1 and 1000+ in which items are more or less similar to adjective or noun, depending on where they might lie on that continuum, and what their set of features might be. This idea has consequences for current conceptualizations of "categories", as well as consequences for the lexical functional distinction.

Recall that I claimed that numerals 5+ have an inherent number feature, but are missing the gender and subgender features; numerals 2,3,4, on the other hand, have gender and subgender features (along with the inherent number feature), but those features are unvalued and uninterpretable. Gender and subgender are properties of lexical nouns: gender is found only on the class of lexical nouns and subgender is a semantic based distinction that further divides those lexical nouns. Thus, these are *lexical* features, and numerals lack them. Rutkowski (2006) argues that Polish numerals are in a process of grammaticalization, from lexical to functional, as centuries ago, they were indistinguishable in behavior from nouns. If numerals are in the process of grammaticalization, becoming more functional (Van Gelderen, 1993; Haspelmath, 2994; Roberts, 1993; Roberts, forthcoming; Roberts and Roussou, 1999), then it is possible that they might begin this process by first losing lexical features, either by them becoming unvalued instead of valued, or them disappearing altogether. If this is the case, then being functional, lexical, or semi-lexical is not a matter of being assigned a certain category, but rather a matter of what features that element contains and to what degree those features are lexical or functional and match to a particular category.

Essentially, this suggests that the loss of features or the change from valued to unvalued is a mechanism of change from lexical to functional.[33] Presumably, since this change would not happen all at once, it is predicted that there will be elements which are somewhat in between, partly lexical and partly functional, due to still carrying some lexical features even after the loss of others; these are our semi-lexical items. This approach also suggests that the notion of category is not a stable concept, as elements are predicted to slip from category to category as their feature sets change. Furthermore, it also predicts that certain elements can sit on the boundaries of categories, like the numerals I have been discussing. A category would then be defined as a set of elements which share a similar feature set and are, thus, close to a particular prototypical ideal which represents that category. Thus, adjectives would be adjectives

[33] Perhaps even, the change from interpretable to uninterpretable is but a stage in the process, and is followed by the eventual loss of that feature. Careful analyses of diachronic data is necessary to explore this idea.

because they have uninterpretable phi-features and nouns would be nouns because they have all interpretable phi-features – semi-lexical items would be expected to straddle the boundary of those categories or move to the realm of functional.

This idea is able to explain Corbett's (1978) generalization that lower numerals tend to be more adjective-like and higher numerals more noun-like – as numerals *can* sit on a continuum between adjective and noun (or head towards the functional), then Corbett's generalization is simply a consequence of this. Certain numerals are similar due to their feature sets, but they can also show a gradient behavior as some might differ more from the prototypical lexical category than others. Note that this also provides a mechanism for explaining the fact that even within a class of numerals (e.g. 5+), there might be specific numerals that behave slightly differently.[34] This different behavior can be attributed to a difference in features.

The ideas here suggest a view of categories that is based on features. Large numbers of items with the same sets of features would represent a prototypical category. Being semi-lexical would then be the result of having non-prototypical features, whether those features make the item an intermediate category or a more functional category. This makes the distinction between lexical and functional a gradient distinction.

To summarize this section, the innovations I have implemented in my analysis have implications in the areas of agreement, case theory, numerals, and categories. My ideas concerning agreement suggest that change is necessary to increase the explanatory adequacy of Agree. Furthermore, the case stacking I have implemented forces us to drop the Case Filter and adopt new ideas concerning case assignment. My discussion on numerals suggests that numerals

[34] For example, there are certain numerals which can take a plural, such as *hundred* in English, as in *hundreds of men*; other numerals, like *ten*, usually cannot, e.g. **tens of men*. One could imagine an analysis in which this numeral is just like a noun in that it has valued phi-features, but similarly to nouns (and unlike the numerals here), the numeral does not enter the derivation with an inherent number feature, but gets it through the syntax. Thus, a potential difference between *ten* and *hundred* could be that *ten* has an inherent frozen number feature, but *hundred* does not.

will vary cross-linguistically. In this paper, I have treated them as semi-lexical nouns, but this does not guarantee that they will be the same across all languages, or even within a language as we see in Polish. Finally, if my ideas concerning the feature set of categories is adopted, it may lead to new ideas and insights into the nature of features, categories, and the lexical versus functional distinction.

5.2 Future Research

Based on the data I have presented and discussed in this thesis, there are many rather interesting areas remaining for future research.

One such avenue concerns the historical development of Polish numerals. According to Rutkowski (2005), Polish numerals were once lexical nouns, behaving identically to the nouns we see today. However, over the last few hundred years, this has changed resulting in the current system. Considering my ideas on categories and the change from lexical to functional, the diachronic development of Polish numerals presents a wonderful testing ground for exploring whether such ideas are really true or not.

Another area of interest concerns this idea of features as probes, as advocated in Bejar (2000). With regards to example (19) in section 2.2.1, in which the features of two separate nouns contributed to verbal agreement, it provides an interesting testing ground for the idea that features might probe separately for their goals. Whatever is going on in such cases, it is bound to provide valuable insights into the nature of agreement.

A final avenue of research I will suggest here concerns conjunct resolution. Recall that one of my informants reported a difference in the ability to resolve agreement with a masculine personal noun versus resolution with a non masculine personal noun (c.f. example 77). I have only scratched the surface with this phenomenon. Future research might help us to understand the extent of the phenomenon, as well as what it is that makes resolution with a masculine personal quantified noun more difficult than with a non-masculine personal quantified noun. Such research might also shed light on the question of what

exactly causes difficulty for the numeral's genitive case assigner with regards to masculine personal gender.

Thus, this research is only the beginning and it opens up various other interesting areas of research, highlighting numerous questions that still need answering. How did numerals come to be the way they are? Why is resolution more difficult with masculine personal noun than with non-masculine personal nouns? How do features in separate conjuncts produce masculine personal agreement? What are the limitations on the cross-linguistic variation of numerals? These questions cannot be answered here, but they present various areas that still need careful attention.

6 References

Baker, M. C. (2003). *Lexical Categories: Verbs, Nouns, and Adjectives*. Cambridge: Cambridge University Press.

Baker, M. C. (2008). *The Syntax of Agreement and Concord*. Cambridge: Cambridge University Press.

Bejar, S. (2000). Locality, Cyclicity, and Georgian verbal morphology. Generals Paper, University of Toronto.

Brown, D. (1998). Defining subgender and devirilized nouns in Polish. *Lingua* 104: 187-233.

Cardinaletti, A. & Giusti, G. (2006). The syntax of quantified phrases and quantitative clitics. In *The Blackwell Companion to Syntax, Volume V*, Everaert, M. & van Riemsdijk, H. (eds.), 23-93. Blackwell Publishing Ltd: Oxford.

Chomsky, N. (1986). *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. (2000). Minimalist Inquiries: The Framework. In *Step by Step: Essays on minimalist syntax in honor of Howard Lasnik*, Martin, R., Michaels, D. & Uriagareka, J. (eds.), 89-153. Cambridge, MA: MIT Press.

Chomsky, N. (2001). Derivation by phase. In *Ken Hale: A life in language*, M. Kenstowicz (ed.), 1-52 .Cambridge, MA: MIT Press.

Chomsky, N. (2005). On Phases. Unpublished paper, MIT (to appear in *Foundational Issues in Linguistic Theory*, C. P. Otero et. al. (eds.), MIT Press, Cambridge Mass.).

Colarusso, D. J. (1992). *A grammar of the Karbardian language*. Alberta: University of Calgary Press.

Corbett, G. (1978). Universals in the syntax of cardinal numerals. *Lingua* 46: 355-368.

Corbett, G. (1983). The number of genders in Polish. *Papers and Studies in Contrastive Linguistics*, Volume VXI. 83-89.

Corbett, G. (1991). *Gender*. Cambridge: Cambridge University Press.

Corbett, G. (2000). *Number*. Cambridge: Cambridge University Press.

Corbett, G. C. (2006). *Agreement*. Cambridge: Cambridge University Press.

Corver, N. (2001). On predicate numerals. In *Linguistics in the Netherlands 2001*, T. van der Wouden & H. Broekhuis (eds.). Amsterdam: John Benjamins, 65-67.

Corver, N. & van Riemsdijk, H. (2001). *Semi-lexical categories*. Mouten de Gruyter: Berlin.

Corver, N. & Zwarts, J. (2006). Prepositional numerals. *Lingua* 116: 811-835.

Den Dikken, M. (2006). *Relators and Linkers: The Syntax of Predication, Predicate Inversion, and Copulas*. Cambridge, MA: MIT Press.

Feist, Timothy (2010). *A Grammar of Skolt Saami*. PhD dissertation. University of Manchester.

Franks, S. (1994). Parametric properties of numeral phrases in Slavic. *Natural Language and Linguistic Theory* 12: 597-674.

Franks, S. (2002). A Jakobsonian feature based analysis of the Slavic numeric quantifier genitive. *Journal of Slavic Linguistics* 10: 141-181.

Haspelmath,M. (1994). Functional categories, X-bar-theory, and grammaticalization theory." *Sprachtypologie und Universalienforschung* 47(1): 3-15.

Hiraiwa, K. (2005). *Dimensions of symmetry in syntax: agreement and clausal architecture*. PhD dissertation. MIT

Ionin, T. & Matushansky, O. (2004). A singular plural. In: B. Schmeiser, V. Chand, A. Kelleher, & A. Rodriguez (eds.), *WCCFL 23 Proceedings*, 101-114. Somerville, MA: Cascadilla Press.

Ionin, T. & Matushansky, O. (2006). The composition of complex numerals. *Journal of Semantics* 23: 315-360.

Jackendoff, R. (1977). *X-bar Syntax: A Study of Phrase Structure*. MIT Press. Cambridge, MA.

Karlsson, F. (1999). *Finnish: An Essential Grammar*. London: Routledge.

Kayne, R. (2005). A note on the syntax of numerical bases. Ms. NYU.

Kayne, R. (2007). *Several, few,* and *many*. *Lingua* 117: 832-858.

Kranendonk, H. (2008). *Quantificational Constructions in the Nominal Domain*. PhD Dissertation. University Utrecht.

Martí Girbau, N. (2003). Partitives: one or two nouns? *Proceedings of the XXIX Incontro di Grammatica Generativa.* Urbino.

Martí Girbau, N. (2010). *The Syntax of Partitives*. PhD dissertation. Universitat Autonoma de Barcelona.

Matasovic, R. (2010). *A Short Grammar of East Circassian (Kabardian) [version 9]*. ZAGREB. Retrieved from: [http://mudrac.ffzg.unizg.hr/~rmatasov/KabardianGrammar.pdf]

Matushansky, O. (2008). Predication: A case study. In *Studies in Formal Slavic Linguistics. Contributions from Formal Description of Slavic Languages 6.5*, 213-239. Frankfurt am Main: Peter Lang.

Matushansky, O. (2010). Some cases of Russian. In *Formal Studies in Slavic Linguistics*, G. Zybatow, P. Dudchuk, S. Minor, & E. Pshehotskaya (eds) Proceedings of FDSL 7.5. 117-135. Frankfurt am Main: Peter Lang.

Matushansky, O. (to appear). Gender Confusion. In *Syntactic Diagnostics in Human Language (working title)*, L. Cheng & N. Corver (eds.). Oxford: Oxford University Press.

Merchant, J. (2011). Where's Gender? Evidence from Greek [handout]. Workshop on quirky ellipsis, Rijksuniversiteit Groningen, November 15.

Nelson, D. & Toivonen, I. (2000). Counting and the grammar: case and numerals in Inari Sami. In. D Nelson & P. Foulkes (eds.), *Leeds Working Papers in Linguistics* 8: 179-192.

Ortmann, A. (2000). Where plural refuses to agree: Feature unification and morphological economy. *Acta Linguistica Hungarica* 47: 249-288.

Pesetsky, D. (2009). Russian case morphology and the syntactic categories [handout]. Retrieved from: [http://web.mit.edu/linguistics/people/faculty/pesetsky/Pesetsky_UMass_handout_Russian_case.pdf]

Pesetsky, D. & Torrego, E. (2001). T-to-C Movement: Causes and Consequences. In Ken Hale: *A Life in Language*, ed. D. Michaels, 355-426. Cambridge: MIT Press.

Preminger, O. (2011). *Agreement as a Fallible Operation*. PhD dissertation. MIT.

Przepiorkowski, A. (2004). O wartości przypadka podmiotów liczebnikowych [On the grammatical case of numeral subjects]. Bulletin de la Societe Polonaise de Linguistique, fasc, vol LX, 133-144.

Przepiórkowski, A. (2006). O dystrybutywnym po i liczebnikach jedynkowych. Polonica [On the distributive preposition *po* and the 'one'-numerals], XXVI–XXVII, 171–178.

Przepiórkowski, A. (2010). Towards a construction grammar account of the distributive PO. *Études Cognitives* 10: 163–176.

Radford, A. (2006). *Minimalist Syntax Revisited.* [http://courses.essex.ac.uk/lg/lg514].

Rappaport, G. C. (2003). Case syncretism, features, and the morphosyntax of Polish numeral phrases. *Generative linguistics in Poland* 5, 123-137.

Rezac, M. (2003). The fine structure of cyclic agree. *Syntax* 6: 156-182.

Rezac, M. (2004). *Elements of Cyclic Syntax: Agree and Merge*. PhD dissertation. University of Toronto.

Richards, N. (2007). Lardil "Case Stacking" and the structural/inherent case distinction. Unpublished manuscript. MIT. [http://ling.auf.net/lingBuzz/000405]

Roberts, I. (1993). A formal account of grammaticalization in the history of Romance futures. *Folia Linguistica Historica XIII*. 1-2: 219-258.

Roberts (forthcoming). Grammaticalization, the clausal hierarchy, and semantic bleaching. In G. Trousdale & E. Traugott (eds.), *Proceedings of NRG4*.

Roberts, I. & Roussou, A. (1999).A formal approach to grammaticalization. *Linguistics* 37: 1011-1041.

Rutkowski, P. (2002). The Syntax of Quantifier Phrases and the Inherent vs. Structural Case Distinction. *Linguistic Research* 7(1), 43-74.

Rutkowski, P. (2005). Why Polish numerals should not be analyzed as nouns. Manuscript, Warsaw University.

Rutkowski, P. (2006). Grammaticalization in the nominal domain: the case of Polish cardinals. Paper presented at the 4th Workshop in General Linguistics, University of Wisconsin, Madison, February 17, 2006

Rutkowski, P. and K. Szczegot (2001). On the syntax of functional elements in Polish: Numerals, pronouns and expressions indicating approximation, in: *Generative Linguistics in Poland: Syntax and Morphosyntax*, ed. Adam Przepiórkowski and Piotr Bański, Warszawa: IPIPAN, p. 187-196.

Swan, O. (2002). *A Grammar of Contemporary Polish*. Bloominton, IN: Slavica.

Toivonen, I. (2007). Verbal agreement in Inari Sami. In: I. Toivonen and D. Nelson (eds.), *Saami Linguistics*, CILT 288, 227-258. John Benjamins.

Van Gelderen (1993). *The Rise of Functional Categories*. Amsterdam: John Benjamins.

Yoon, J. H. (2004). Non-nominative (major) subjects and case stacking in Korean. In *Non-nominative subjects: Volume 2*, K. V. Subbarao & P. Bhaskararao (eds.), 265-314. Amsterday: John Benjamin Publishing Company. Retrieved from: [https://netfiles.uiuc.edu/jyoon/www/Papers/yoo.pdf].

Zabbal, Y. (2005). The syntax of numerical expressions. Ms. UMass Amherst.

Zlatic, L. (1997). *The Structure of the Serbian Noun Phrase*. PhD dissertation, University of Texas at Austin.

Zweig, E. (2005). Nouns and adjectives in numeral NPs. *Proceedings of NELS 35*.