

Ordinal numerals in dialects of Dutch

Ruby Sleeman



Leiden University, Faculty of Humanities

Research Master Thesis in Linguistics

Supervisor: prof. dr. L.C.J. Barbiers

Second reader: prof. dr. J.S. Doetjes

2017

Acknowledgements

First of all I would like to thank my supervisor, Sjef Barbiers, for interesting conversations, valuable insights, encouragement, compassion, and all the time he has invested in me. I am indebted to Ben Hermans and Marc van Oostendorp for helping me with phonological issues. I wish to express my gratitude to Geert Booij for his valued input.

My thanks go out to Jelke Bloem who is always ready to listen and give advice. I am indebted to Peter Alexander Kerkhof and Rogier J. M. te Paske for invaluable feedback. I wish to thank my colleagues at the Meertens Institute, where I spent many writing days, and my colleagues at LUCL for moral support and fellow thesis-complaining. Thanks also to my dear fellow students at Leiden University who inspired me by finishing their theses and starting PhD's abroad. Thanks to Lisa E. H. Steenkamer for her help with my Limburgian data.

I am extremely grateful to everyone who helped me find survey respondents and who filled out the survey. And finally, I wish to thank my (extended) family and my closest friends who never stopped believing in me - especially Bram, who has been with me the whole journey.

Contents

1. Introduction	6
2. Background	7
2.1 The cardinals and ordinals of Dutch.....	8
2.2 Conclusion	13
3. Patterns of -ste ordinals in dialects: novel, empirical data	14
3.1 Methods	15
3.2 Results	19
3.3 Conclusion	22
4. Analysis of suffix distribution in the southern dialects of Dutch	23
4.1 Extralinguistic motivation: geographical distribution	24
4.2 Intralinguistic motivation	33
4.2.1 Two ordinal suffixes: suppletive allomorphy?	33
4.2.2 Final consonant of the stem	40
4.2.3 Syllable weight	44
4.2.4 Parsing syllables into feet	54
4.3 Conclusion	60
5. Summary, conclusions and recommendations	61
References	65
Appendix A: Diachronic overview of Dutch cardinals and ordinals	69
Appendix B: Extended survey methods and methodology	82
Appendix C: Additional findings and unresolved issues	100

'Numerals belong in a series, which results in the fact that they influence each other. For this reason the explanation of their form is extremely complicated'

(Beekes & De Vaan, 2011: 237)

1. Introduction

This thesis is a synchronic study of the derivation of ordinal numerals from cardinal numerals in several different dialects of Dutch. Ordinals are formed with one of two ordinal suffixes, *-de* or *-ste*, both of which occur in all Dutch varieties. However, there is variation among several dialects in Flanders (Belgium): they exhibit stem-suffix combinations which differ from those in Standard Dutch (SD). This thesis investigates the patterns of this variation and offers motivations for those patterns.

The main research questions of this thesis are:

- (i) What are the suffix choices in the ordinal number lines of the different varieties of Dutch (including the standard variety)?
- (ii) How can these patterns be motivated?
- (iii) Do the suffixes share the same underlying representation?

After providing an introduction to the ordinal system of SD in chapter 2, I will answer the first research question in chapter 3, where I describe the variation in ordinal formation in Flanders (3.2). The variation falls into three categories: a Flemish pattern, a Standard Dutch pattern, and an intermediate pattern or Transitional pattern. I will also explain the setup of the survey with which these novel empirical results were obtained (3.1).

Chapter 4 provides a tentative answer to the third research question as well as three answers to the second question: an extralinguistic motivation (geographical distribution), an intralinguistic motivation (syllable weight) and a negative effect of the final consonant on the choice of suffix. Firstly, I interpret the survey results as the outcome of geographical factors, thus providing an extralinguistic motivation for the ordinal suffix patterns (4.1). I then proceed to finding intralinguistic motivations. In order to research the intralinguistic factors at play I look at research question (3) first, and examine whether the suffixes are allophonic alternants of the same underlying representation (4.2.1). I conclude that they most likely are suppletive allomorphs, and I describe the consequences this has on how to proceed with the research. In the next section I investigate the possibility of phonological conditioning of the suffix choice by looking at the effect of the final stem consonant on the suffix choice, and I conclude that there is no such effect (4.2.2). I find a positive effect of syllable weight on suffix choice in one of the three patterns (4.2.3): for the transitional area, a presuffixal extra light syllable dictates the

selection of *-ste*. Finally, I look at the foot structure (a direct result of syllable weight in Dutch) and extrametrical syllables in the next section (4.2.4). We will see that this gives an undesired grouping of the ordinals. Chapter 5 summarizes the main findings, discusses which questions were left unanswered and concludes the thesis.

As I will argue in section 4.2.4, the number of underlying syllables in the ordinal's stem form plays a role in the suffix selection, and because the synchronic Dutch number words are the result of historical processes, I have summarized the relevant historical changes in appendix A. Appendix B elaborates on the survey methods and appendix C summarizes some survey results that were not relevant to the main text, but interesting food for thought and could be taken up for future work.

2. Background

Ordinals, or ordinal numerals, are defined by Veselinova (1998: 2) following Hurford (1987) as 'lexico-grammatical expressions which denote position in an ordered sequence of objects.' Veselinova (1998) conducted a typological study over 47 maximally unrelated languages to research the relative frequencies of several strategies for forming ordinals, many of which involved in some way a derivation from the corresponding cardinals in the language:

'With regards to their formation, ordinals represent a (theoretically) indefinite set of lexical items which is derived from the cardinal numerals of the specific language. The lower members of the set, terms for 'first' and 'second' tend to appear as exceptions to this tendency.'

(Veselinova, 1998: 3)

Turning now to the cardinals and ordinals of SD we find that they conform in part to these findings: although *twee-de* 'second' is formed regularly after the cardinal, the term for 'first', *eer-ste*, indeed is *not* derived from the cardinal *één* '1'. Van Loey (1970: 155) says it is a superlative formation from the stem *eer*, cognate with Gothic *áír* and English *ear-* in *ear-ly*.¹

¹ *-ste* is one of the two Dutch ordinal suffixes, but there is a homonymous superlative suffix *-ste*. It is not unthinkable that the ordinal *-ste* suffix was an innovation, a newcomer added to the ordinal paradigm, borrowed from the superlative paradigm. *-de* is very old and can be found as far back as Old Dutch:

Thus, the synchronic ordinal *eer-ste* could be best translated as originally meaning ‘earliest’, referring to the object ranked first in a temporal sense (Buck, 1949 as cited in Veselinova, 1998: 14).

Barbiers (2007) shows with three tests why *eerste* should be considered a true superlative and different from all the other ordinals, regardless of whether those other ordinals are formed with *-de* or *-ste*: (i) *eerste* can modify plural nouns (*de eerste boeken*, ‘the first books’) just as other superlatives can; ordinals cannot. (ii) When in predicative position, *eerste* can be reduced to *eerst* just as superlatives can also drop their final schwa. The ordinal *achtste* ‘8th’ cannot, nor can any other ordinal. (iii) The intensifier *aller-* can be added to the left of *eerste* as well as to superlatives, but not to any other ordinals. QED: *eerste* ‘1st’ is not an ordinal, it is a superlative; the other *-ste* ordinals are ordinals, they are *not* superlatives. For this reason, this thesis will from this point onwards not be concerned with the suffix choice of *-ste* in *eerste* because it is not the ordinal suffix, but rather the homonymous superlative suffix.

2.1 The cardinals and ordinals of Dutch

I have listed the cardinals and ordinals in table 2.1 and highlighted the ordinals formed with the suffix *-ste*. I will now introduce the reader to the system of Dutch ordinals by first mentioning some quotes from the literature and then discussing the individual Standard Dutch number words.

The literature does not say much about the ordinal numerals of Dutch. Van Bree (1987) briefly mentions their historical origins in his historical grammar of Dutch (a sidenote after the cardinal numerals are discussed).² Booij (2010), with the subtitle ‘an analysis of Dutch numerals’, is also focused on the cardinals and gives a brief description of the SD system:

‘Ordinal numerals are created in a regular fashion by adding the suffix *-ste* or the suffix *-de*. The suffix *-ste* [stə] is added after the ordinal allomorph for *een* ‘one’, *eer-*; after *acht* ‘eight’, after the suffix *-tig* (*twintig-ste*, *dertig-ste*, etc.), and after the numerals *honderd*,

‘sivondo’, oldest attestation 1151-1200 (ONW). See appendix A for an overview of the historic stages of the ordinals. I will briefly discuss the historical origins of both suffixes in section 4.2.1.

² See appendix A for information on the historical origins of the cardinals and ordinals.

duizend, *miljoen*, and *miljard*. In all other cases (after 2–7, 9–10 and numerals ending in these numerals), the suffix *-de* [də] is used.’

Booij (2010: 94)

Table 2.1: The cardinals and ordinals of Standard Dutch

	Cardinal	Ordinal
1	één	eer-ste
2	twee	twee-de
3	drie	der-de
4	vier	vier-de
5	vijf	vijf-de
6	zes	zes-de
7	zeven	zeven-de
8	acht	acht-ste
9	negen	negen-de
10	tien	tien-de
11	elf	elf-de
12	twaalf	twaalf-de
13-19	dertien	der-tien-de
20	twintig	twintig-ste
100	honderd	honderd-ste
1,000	duizend	duizend-ste
1,000,000	miljoen	miljoen-ste
1,000,000,000	miljard	miljard-ste

Zonneveld (2007) makes a few cautious statements about the distribution of the suffixes, namely that *-de* seems to be the less productive one of the two.

‘The suffix [-*de*] occurs after numbers below ‘20(th)’, after which *-ste* takes over completely (‘1st’ is *eer-ste*; ‘8th’ is *acht-ste*, possibly because the number ends in a plosive). These limited cases might be taken to indicate that this suffix is of very low

productivity, and the marked case when compared to its sister *-ste* and to fully productive past tense *-de*. However, although this is not commonly recognized, numeral *-de* is very productive, too. It is used in the ‘to-the-power-of’ construction with more than just numbers’.³

Zonneveld (2007: 20)

The ‘to-the-power-of’ construction that Zonneveld refers to is revisited in section 4.2.2 below. Zonneveld’s remark about *achtste* is, I think, going in the right direction, but not sufficiently worked out. Barbiers (2007) is the only work that I have come across to comment on formulating a motivation for the way in which the two ordinal suffixes are distributed along the ordinal number line. The article is mainly concerned with the motivation behind suppletion for 1st. In a footnote he comments on the entire ordinal system:

‘I leave the question as to whether the two ordinal suffixes *-de* and *-ste* are phonological alternants for future research. One could argue that *acht-ste* ‘eight-th’, *honderd-ste* ‘hundred-th’ and *duizend-ste* ‘thousand-th’ are the result of two derivational steps: (i) addition of the suffix *-de*; (ii) insertion of /s/, triggered by the adjacency of two coronal stops that only differ in voicing, the final /t/ of the numeral (after final devoicing in the case of *honderd* and *duizend*) and the initial /d/ of the ordinal suffix, possibly a violation of the Obligatory Contour Principle (Leben, 1973). However, such an analysis would not carry over to forms like *vijf-tig-ste* ‘fifty-th’, etc. or *miljoen-ste* ‘million-th’. A different formulation of the system may be that the ordinal suffix is *-de* for the numbers from two to ten and *-st* for the tens, hundreds, etc. with *achtste* ‘eighth’ the result of a dissimilation process triggered by the OCP.’

Barbiers (2007: 861)

Barbiers makes three interesting remarks here. Firstly, he notes that *-de* and *-ste* may be phonological alternants of the same underlying representation. I will investigate this hypothesis in section 4.2.1 below. Secondly, he notes that an attempt to explain 8th, 100th and 1,000th all with the same rule - namely an underlying ordinal suffix *-de*, and a phonological s-insertion because of t-d clash - does not hold for some of the other numbers along the line. Thirdly, he proposes an alternative analysis, in which 8th is the only one resulting from an

³ Because this includes only one single construction, specifically in the realm of mathematics, I don’t think it can counter the observation that *-ste* is much more productive overall - and I am not at all sure that Zonneveld meant to counter it.

underlying *-de* and the ordinals over 10 all have an underlying *-ste* suffix. Note that the teens are compound cardinals, formed, as in English, by addition of the lower cardinal to *tien* ‘10’ on the right hand side; this makes the latter the head of the compound, following the right hand head rule (Williams, 1981). Anything that applies for 10 therefore by extension also applies to the teens in the Dutch number system. What Barbiers overlooks in this footnote, however, is that 11 and 12 do not follow this formation, but are instead (synchronically) simplex forms.⁴ They, too, use *-de* in SD; so the statement ‘the ordinal suffix is *-de* for the numbers from two to ten’ should in my view be altered slightly to ‘the ordinal suffix is *-de* for the numbers from two to *twelve*’.⁵

Building on Barbiers (2007), I propose the following analysis for the Dutch ordinal system:

- (1) (a) The default morpheme for ordinal formation in Standard Dutch is *-ste*. This morpheme is applied to **all cardinals** except the lower cardinals 2-12. Some varieties in Flanders extend the use of the *ste*-morpheme all the way down to 7.
- (b) In Standard Dutch, the lower cardinals 2-19 get a different morpheme: *-de*. In the other dialects, it is applied only to the cardinals 2-6. *Achtste* ‘8th’ is formed with *-de* but the surface form has an inserted *-s-*; see below.
- (c) The ordinal for 1 is unique compared to all other ordinals: its *-ste* is a superlative suffix, which is different from the larger ordinals.

(1a) is in line with Zonneveld’s (2007: 20) quote above: *-ste* is more productive. Some scholars consider words denoting the last or middle entity in a row to also be an ordinal (e.g. *middel-ste* ‘middle-*ste*’, *laat-ste* ‘late-*ste*’/‘last’). See for example Petra Sleeman (2010: 1), who mentions ‘the ordinals *first* or *last*.’ One could argue that these words are superlatives, but that would not do justice to the fact that *middelste* and *laatste* do what other ordinals do: they denote the position in an ordered sequence. If Barbiers’ (2007) tests show that these words pattern with *eerste*, and thus have the superlative suffix *-ste*, not the ordinal suffix *-ste*, well, then a good analysis of the differences and similarities between superlatives and ordinals is needed. Other examples of ordinal-like words that are formed from non-numerals with *-ste* include the

⁴ *Elf* ‘11’ and *twaalf* ‘12’ are non-controversially described in the literature as being formed from ‘1 left’ and ‘2 left’ (relative to the base, ten). See Van Bree (1987: 259) and see the appendix A for a more detailed history.

⁵ German has a *-ste*-morpheme for larger ordinals as well; their *-te* morpheme for lower ordinals also reaches from 2-12, also with exception of 1. 8 behaves ‘normally’ in the sense that it takes the *-te* morpheme (Render, 1805: 109-110).

question word *hoeveelste* ‘how-many-th’ and colloquial non-standard formations such as the predicates *enig-ste*⁶ ‘the only one’, and *tweedst* ‘second’.⁷

Turning again to table 2.1, there are a few other individual SD ordinals that I would like to discuss. Firstly, the stem form of the ordinal *derde* ‘3rd’ has undergone metathesis (Van Bree, 1987: 259). I don’t consider the metathesis in this stem to be relevant to the current study, because its main focus lies with variation among the dialects (microvariation) and I am not aware of any synchronic variation in suffix choice for 3rd.

Secondly, I would like to elaborate on *acht-ste*, which is said above to be formed with *-de* but with an inserted *-s-* in the surface form. After valuable input from Peter Alexander Kerkhof (personal communication, April 2016) I claim in this thesis that the formation of *achtste* and the choice of *-ste* have a completely different backstory than the other ordinals. This is corroborated by the fact that the survey data showed no dialectal variation for the suffix choice: ‘8th’ is formed with *-ste* across the language area.

According to MNW⁸, *achte* existed alongside *acht*, the former existing both as a cardinal as well as as an ordinal. The ordinal was supposedly formed through reduction from *achtede* (Vr), related to Got. *ahtu-da* and OHG *ahto-do*. Other ordinal forms in MDu include: *achtende* (with inserted *-n-*, probably analogous to the adjacent forms ‘7th’ (ODu *sivondo*) and ‘9th’ (cf. Got. *niunda*) (Vr); *achtenste* (formed after *zevenste* ‘7th’ and *negenste* ‘9th’, which occurred alongside their *-de* counterparts in MDu (in Flanders) and may have been formed after *eerste* and *twintigste* (Vr)). Thus, in MDu, we find **four** competing ordinal forms so far: *achte*, *achtede*, *achtende* and *achtenste*⁹; of these, *achtede*, if it is indeed from *ahtuda*, it is the oldest; *achte* is a reduction thereof; *achtende* was formed after the older forms *zevende* and *negende* and *achtenste* is the newest, formed after *zevenste* and *negenste* which were newer than their *-de* counterparts.

⁶ See http://taaladvies.net/taal/advies/vraag/262/enigste_enige/.

⁷ Among the results of the survey I did for this study, there were quite a number of answers in the form *ik ben tweedes/twee-d-st* ‘I am second’. This kind of predicative use of ordinal numbers with its own formation rules is a topic which in my view deserves to be examined, as it seems to make use of a morpheme *-st-* (but notably not *-ste*) which looks quite similar to *-ste*. When discussing this construction with speakers of standard colloquial Dutch, they often refer to it as children’s language in gameplay (to call the order of turns: I am first, you are second), but the survey results revealed that it is used by adults as well in some regional varieties, for example when calling one’s position in a queue. Researching this construction may shed light on the origins of *-ste* as an ordinal suffix.

⁸ In this diachronic interlude, I use abbreviations as is common practice in diachronic studies: Got: Gothic; MDu: Middle Dutch; ModDu: Modern Dutch; ODu: Old Dutch; OHG: Old High German; Vr: de Vries (1971 as cited in Sijs, 2010). MNW, WNT refer to Middelnerlands woordenboek and Woordenboek Nederlandse Taal, searchable online in the *Geïntegreerde taalbank* at gtb.inl.nl.

⁹ In the VMNW we find two additional forms, which are said most likely to have been errors in the texts: *achtechste* and *achstende*.

In the (late) MDu era, a fifth form must have arisen: ModDu *achtste*. WNT argues that it occurred in MDu, but was less frequently used than *achte* and *achtende*; according to Vr, it did not occur until the ModDu era.¹⁰ Although I cannot find a source that argues for its origins, I propose (with Peter Alexander Kerkhof, p.c.) that *achtste* evolved from *achtede* due to a sound change known as the Middle Dutch schwa-syncope¹¹: *achtede* > *achtde*; an *-s-* was inserted to relieve the phonotactically unpleasant combination of two consonants pronounced in the same place of articulation, also known in the field of phonology as a violation of the Obligatory Contour Principle. The final step is to assimilate *d* > *t* in voice: *achtsde* > *achtste*. This Middle Dutch sound law provides a neat confirmation from a diachronic perspective of what Barbiers (2007) proposes. From a synchronic perspective, one could argue that *acht-ste* undergoes morphological reanalysis as though *-ste* is the suffix. I have neither the means nor the desire to decide whether the underlying synchronic suffix is *-de* or *-ste*; all that is relevant for the current purpose is to state that there is no synchronic linguistic motivation for 8 to be an exception to the group of ‘lower cardinals’ 2-12.

2.2 Conclusion

In this chapter I have introduced the reader to the system of cardinal and ordinal numerals in Dutch, with a focus on the standard variety. The most important conclusions can be found in (1) above. I repeat them here.

(i) There are two ordinal suffixes, *-de* and *-ste*. *-ste* is the most productive of the two and thus I dub it the default morpheme. (ii) In SD, all ordinals are formed with *-ste* except 2-12, which are formed with *-de*. 13-19 follow the formation of 10, as 10 is the dominant half of the compound numerals 13-19. (iii) 8th belongs to the *-de* group, but due to diachronic sound changes the surface form now looks as if the suffix is *-ste* (and one could argue for morphological reanalysis, but there is no synchronic motivation for 8 to be an exception in the 2-12 group). (iv) Some varieties in Flanders form 2-6 with *-de*, and start using *-ste* at 7 (more details in chapter 3 below). (v) *Eerste* ‘1st’ is not an ordinal, but a superlative; and it is formed not from a cardinal but from a suppletive stem *eer-*. Literally translated it means ‘earliest’.

In chapter 3 below I describe the novel results found by conducting an online survey.

¹⁰ WNT says nothing about the origins of *achtste* except that it **differed from the origin of *achte/achtende***. I do not hold this to be true.

¹¹ This sound change is described in Bloemhoff & Streekstra (2015: 147-8); it occurred mainly in the domain of verbs, resulting in such changes as *du wonedes* > *du woendes* ‘you dwell’ and *ghi makedet* > *ghi maectet* ‘you make’ in the weak declension paradigms of the preterite; and in participials, for example *ghewonet* > *ghewoent* ‘dwelled’.

3. Patterns of *-ste* ordinals in dialects: novel, empirical data

In this chapter I answer the question: What are the suffix choices in the ordinal number lines of the different varieties of Dutch (including the standard variety)? During preliminary research I studied the diachronic stages of the Dutch ordinals in the Integrated Language Bank (Geïntegreerde Taalbank, GTB) and found mentions of variation in use of the suffixes, concentrated in a certain area of Flanders (Brabant).¹² Dialect grammars from most regions of the Netherlands and Flanders either did not mention the numbers at all (suggesting that they may be the same as for the SD system), or showed that the varieties described followed the SD system. Variation was located mainly in East and West Flanders and Flemish Brabant, and thus I designed a questionnaire and focused the respondent search on these areas. Of course the survey attracted respondents from outside this area as well and they confirmed the areal premise by exhibiting little to no variation with respect to the standard ordinals.

In table 3.1 I give a preview of the forms for the three biggest patterns. As I mentioned in the introduction, I label them the Flemish pattern, the Transitional pattern and the Standard Dutch (SD) pattern. Note that the dialectal forms have been abstracted to resemble the SD forms; the suffix remains recognizable throughout my survey data.

The table illustrates that the ‘break’ between *-de* and *-ste* at 20 in SD is not the same in all dialects: there are dialects in Flanders which have the break at 7 instead, shown here as the Flemish and Transitional patterns. They have *-de* under 7, and *-ste* for 7 and up. The Transitional pattern is called transitional because as opposed to the Flemish pattern, there is an exception ordinal: 10, which is formed not with *-ste* but with *-de*. (And, as explained above, by extension 13-19 are also exceptions. The survey data indeed show that 13-19 pattern with 10 in almost all speakers. More details are given in chapter 3 below.)

In this chapter I discuss the empirical data I collected and the methodology I used to do so. In section 3.1 I briefly discuss the methods; in section 3.2 I present and discuss the results; section 3.3 concludes the chapter.

¹² The GTB is accessible online at <http://gtb.inl.nl/>.

Table 3.1: Ordinals of the three main patterns

	Flemish pattern (West and East Flanders)	Transitional pattern (Antwerp, Flemish Brabant)	Standard Dutch pattern (rest of Flanders and Netherlands)
1	eer-ste	eer-ste	eer-ste
2	twee-de	twee-de	twee-de
3	der-de	der-de	der-de
4	vier-de	vier-de	vier-de
5	vijf-de	vijf-de	vijf-de
6	zes-de	zes-de	zes-de
7	zeven-ste	zeven-ste	zeven-de
8	acht-ste	acht-ste	acht-ste
9	negen-ste	negen-ste	negen-de
10, 13-19	tien-ste	tien-de	tien-de
11	elf-ste	elf-ste	elf-de
12	twalf-ste	twalf-ste	twalf-de
20-99	twintig-ste	twintig-ste	twintig-ste
100	honderd-ste	honderd-ste	honderd-ste
1,000	duizend-ste	duizend-ste	duizend-ste
1,000,000	miljoen-ste	miljoen-ste	miljoen-ste
1,000,000,000	miljard-ste	miljard-ste	miljard-ste

3.1 Methods

To collect dialectal data on cardinals and ordinals, I designed and conducted an online questionnaire. The motivation behind this form of data collection lies in the nature of the research goal. Bowerman (2008: 80) argues the following for questionnaires: ‘[T]his method of data collection is very good if you need a standard data set over multiple respondents, for example in examining potential variation.’ Numerals, and ordinal numerals to an even greater extent, are a quite specific type of words that I wanted to acquire data on; I wanted to gather every cardinal and every ordinal in the system of many different regional varieties of Dutch. Alternative research tools such as corpora would not be able to meet these research needs,

simply because there are no Dutch dialectal corpora that I know of where I could extract every single cardinal and ordinal in the system.

The remainder of this section describes the contents and design of the questionnaire, including some methodological decisions, and a description of the areas targeted. For details, and screenshots of the questionnaire as well as the promotional texts used for distribution via Facebook, I refer the reader to appendix B.

The questionnaire was created using Google Forms and reachable through a link. Participants were not required to log in before they could participate. I did not ask the respondents to write out their answers like they pronounced it, but all participants did this (some to a larger extent than others) even without being instructed to do so. The questionnaire is made up of 4 subparts. The first subpart consists of 5 background questions that ask about the respondent's identity; specifically about their dialect. Subsection 2 asks the participant to give 25 cardinal numeral forms from their dialect; each cardinal is depicted in Arabic numerals, as 1, 2, 3...21, 100, 1,000, 1,000,000 and 1,000,000,000. Section 3 looks very much like section 2, but it asks about the ordinal forms instead. Section 4 is a small closing section with a non-required, open question: does the respondent have anything they wish to say? Each subsection has its own title followed by a short introductory text to explain to the respondent what is asked of them.

The questionnaire asks every participant every number only once. What this entails for the data is that whenever something rare occurs or I come across a typo, there is not really any way to check it. However, adding something like example sentences to the questionnaire would have made it longer thus making the threshold for potential respondents higher: the smaller an effort it promises to be, the less of an impact people will assess it to have on their time, the more likely they are to be willing to participate - the higher the response rate. My questionnaire consists of 2x25 number items, plus 5 opening questions and 2 closing questions; that makes 57 questions, which in total take no more than 10 minutes to complete. Any ambiguous data was considered noise in the analysis phase.

The questionnaire specifically and explicitly targets dialect speakers and asks for dialect answers; however, it is written itself in Standard Dutch because SD is the common language in the entire Dutch speaking area, and most if not all dialect speakers also speak SD. This comes with a disadvantage: asking for the forms in SD may prime the respondents into giving the SD forms for the ordinals; this may result for example in an answer that says there is variation between let's say *7-ste* and *7-de*, when in fact the dialect they learned does not have the SD *-de* form, only the *-ste* form, but the speaker also speaks SD and therefore thinks that the dialect allows both options; whereas should I really go do fieldwork among the older speakers of the

dialect maybe I would find that there is only one form. However, we cannot ignore the fact that there is language contact between the standard language and the dialects; thus it might not be meaningful to speak of a ‘pure’ form if it doesn’t really exist in any speaker.

Disadvantages aside, writing the questionnaire in SD also has some obvious benefits: it saves time not to have to translate the questionnaire in different dialects; some dialects lack a standard spelling, making it difficult to gauge how many versions of the questionnaire should be made; all participants get the same questionnaire and this controls for the factor ‘means of data collection’ when considering differences in the data, factoring out such complications as translation mistakes or translation gaps between the dialect and SD.

To factor out the disadvantages as well as possible, I stated explicitly what I am looking for in the questionnaire. I asked the respondents to keep the forms of their own dialect in mind and also to report it if they had different forms for the same thing (such as *7-ste* and *7-de*); AND if they did have variation, to also please mention whether or not one of both forms is preferred over the other.

The questionnaire was distributed through Facebook, a social media platform at <http://www.facebook.com>.¹³ It targeted speakers from specific parts of the Dutch language area. For those not familiar with the latter I provide a map below (fig. 3.1) with the names of those provinces in the Dutch language area that were the subject of this research.¹⁴ The current study focuses on roughly the same language area as a recent paper on several syntactic phenomena by Sjef Barbiers, Marjo van Koppen, Hans Bennis and Norbert Corver (2016); the dialect descriptions that follow below are based on their descriptions (*ibid.*, 2016: 11). The current study focuses on Flemish (spoken in Belgium, in the provinces of West and East Flanders), Zeeuws (spoken in the Dutch province of Zeeland and the southern-most part of South Holland), North-Brabantish (spoken in the Dutch province of North-Brabant), South-Brabantish (spoken in the Belgian provinces of Antwerp and Brabant), Flemish Limburgish

¹³ I chose to use Facebook as the platform for distributing my questionnaire because I believed it would very quickly lead to a large number of responses, an expectation which was indeed confirmed within days. Facebook has many users and a lot of these Facebook users have a large network of contacts. A downside is that it may very well be the case that some respondents from the same area are related to each other and may even come from the very same background, same community, same street or even same household, thus potentially skewing the results: we should be aware that it is possible that this number is not representative for the entire region because there could have simply coincidentally been a few members of the same family responding with their identical answer sets, even if their pattern is not representative for their entire town or region. This is due to the type of sampling I used to find respondents: snowball sampling, a type of sampling where your contacts lead to other contacts ‘within the same network’. (Dollinger, 2015: 273)

¹⁴ This map was created by the author on the basis of the blank map including colored dialect areas following Daan & Blok (1969), provided at <http://www.meertens.knaw.nl/sand/zoeken/kaart/frames.php>.

(spoken in the Belgian province of Limburg) and Standard Dutch (the standard language spoken in the Netherlands).¹⁵ The colors indicate the dialect areas as established by Jo Daan and D.P. Blok (1969), which roughly but not entirely correspond to the geographical province borders.



Figure 3.1: The relevant dialect areas in the Netherlands and Belgium

After having discussed the methods and methodology of the data acquisition, let us now turn to the results of the questionnaire.

¹⁵ See the appendix for more detailed information about these dialect groups.

3.2 Results

The questionnaire results consist of datasets - one full set per respondent - of forms for the cardinals 1-21, hundred, thousand, million and milliard (billion), and their corresponding ordinals. What follows is a description of these results in terms of patterns; I have compared datasets of individual speakers and from these comparisons I have been able to draw generalizations over the dataset as a whole.

Let me first describe what I mean by ‘pattern’. Every respondent submitted an answer set containing their cardinal and ordinal forms. In the ordinal paradigms, I looked at the use of *-de* and *-ste* suffixes. The generalizations I made by comparing different ordinal paradigms found in the data, then, are concerned with the distribution of these suffixes. Thus, if one answer set contained only *-ste* forms from 7 upwards in their ordinal number line, this was considered to be one ‘pattern’; this is a different suffix distribution then, for example, that of the Standard Dutch ordinal paradigm.

In the initial data organization stage every pattern was considered a new pattern on the basis of one single ordinal diverging from the existing patterns. This led to a large amount of patterns: 63 on a total number of 240 speakers; and many of these patterns were unique, found only for one speaker. See appendix C for this kind of rare findings, which are possibly interesting, possibly noise.

From the 63 patterns, some tendencies became immediately visible: (i) three patterns were much more frequent than the other 60; and (ii) certain ordinals on the number line would frequently take the same suffix within the same speaker. For example, whenever 11 combined with *-ste* in a certain speaker, that same speaker would often also form 12 with *-ste*. Whenever a speaker would form one ordinal within the range 13-19 with *-ste*, there was almost always at least one other ordinal in the same range to take that suffix. Following these tendencies, I set out to cluster the patterns together on the basis of which ordinals tended to behave similarly.

For each speaker, I checked the forms occurring in their dataset against the following criteria. Note that some speakers gave two forms for an ordinal - if an ordinal could be derived with both *-ste* and *-de* I treated it as checking the relevant criterion.

- (2) a. At least one ordinal in the range [7-9] derived with *-ste*¹⁶
 - b. 10 derived with *-ste*
 - c. At least one ordinal in [11-12] derived with *-ste*
 - d. At least one ordinal in [13-19] derived with *-ste*

¹⁶ There was no variation concerning the ordinal 8-*ste*, so criteria 2a means a speaker has either 7-*ste*, or 9-*ste*, or both.

These criteria were empirically found. I find it vital to let the data speak, because we cannot derive the criteria from presupposed assumptions. One important example of this is that we expect all respondents to have the same suffix for all numbers in the range 13-19, and that suffix should be the same as for 10. This is *very often* the case in the data, but *not always*. Seeing that it is *very often* the case was the motivation for setting up criteria b and d; seeing that it is *not always* the case was the motivation for letting b and d be separate criteria. In the same way 7 and 9 often pattern together, as do 11 and 12.

Checking each dataset against these criteria resulted in a reduction of 63 different patterns to 12 patterns, listed in table 3.2. The ‘numeric scope’ in the first column refers to which ordinals on the ordinal number line are formed with the *-ste* suffix. The scope should be read as follows: in the Flemish pattern, all ordinals from 7 upwards are formed with *-ste*; all ordinals below 7 are thus formed with *-de*. In the Transitional pattern, the numeric scope is ‘7-9, 11-12, ≥ 20 ’, meaning that speakers with this pattern exhibited the forms *zeven-ste* ‘seventh’, *acht-ste* ‘eighth’, *negeen-ste* ‘ninth’ (but not *tien-ste*; *tien-de* instead); *elf-ste* ‘eleventh’, *twalf-ste* ‘twelfth’; and all ordinals above and including *twintig-ste* ‘twentieth’ are also formed using *-ste*. This includes all of the compound numerals between twentieth and a hundredth, as well as hundredth itself, thousandth, millionth and milliardth. All ordinals not mentioned in the scope (2-6, 10) are formed by use of the other suffix *-de*.

By checking every dataset against the criteria listed above, I have taken steps to make a motivated generalization among the 63 different patterns found in the earliest stage of data processing. The logically possible combinations of the criteria are listed in table 3.2, where we see that certain combinations of criteria occur more frequently - these are the so called biggest three patterns - while other combinations don’t occur at all. Recall from chapter 2 that 1st and 8th are not relevant at this point; their formations can be linguistically motivated to be different from the general system. All I will add to that in this chapter is that the data corroborates their special status: no variation (apart from small numbers, probably noise) was found. See appendix C for more details. 20 and higher are also free from variation and formed with *-ste* by all respondents.

I have highlighted the most common criteria combinations, or ‘patterns’. They will be discussed below. It is first interesting to note that of the logical combinations of criteria a, b, c and d, some are not found. These not-attested combinations are: (b), (a, b), (b, c) and (a, b, c). Now, we must keep in mind that some of the attested patterns are *very* infrequently found, meaning that in a more extensive study, the missing combinations may well be discovered; but, looking at the data at hand, the missing combinations are peculiarly *non-random*. One could tentatively draw the conclusion that any dataset never checks criterion (b) without also

checking criterion (d); or, translated to the data in the datasets: within one speaker, ordinal [10] is not derived with *-ste* unless at least one ordinal in [13-19] is.

Table 3.2: All logically possible combinations of the four criteria a, b, c, and d

Numeric scope of <i>-ste</i>	Checks which criteria?	Number of speakers
7-9, ≥ 20	a	13
10, ≥ 20	b	N/A
11-12, ≥ 20	c	9
≥ 13	d	3
7-10, ≥ 20	a, b	N/A
7-9, 11-12, ≥ 20	a, c	55
7-9, ≥ 13	a, d	4
10-12, ≥ 20	b, c	N/A
10, ≥ 13	b, d	1
≥ 11	c, d	4
7-12, ≥ 20	a, b, c	N/A
7-9, 10, ≥ 13	a, b, d	1
7-9, ≥ 11	a, c, d	13
≥ 10	b, c, d	6
≥ 7	a, b, c, d	32
≥ 20	-	87
-	rest category ¹⁷	12

Table 3.2 may look a little chaotic, but there is a system underlying it. Although many of the combinations are very infrequently found, three combinations are *not* infrequent. They are the dominant three, and I list them separately in table 3.3. We will see that they have a very neat areal distribution in section 4.1 below. The labels I give them correspond to the regions where they are found.

¹⁷ Some datasets included rare *-de* formations, which are beyond the scope of the current study. Formations that rarely ever occurred were *1-de*, *8-de* and *-de* ordinals anywhere on the number line above and including '20'. This may very well be noise, but worth further investigation with more respondents. Some notes on these findings can be found in appendix C.

Table 3.3: The three biggest patterns

Numeric scope of -<i>ste</i>	Number of respondents	Pattern label
≥ 20 (and 8)	87	Standard-Dutch pattern
7-9, 11-12, ≥ 20	55	Transitional pattern
≥ 7	32	Flemish pattern

The other patterns in table 3.2 can be considered transitional systems between the larger systems: the scope of *-ste* gets gradually more and more restricted as we move from the Flemish pattern via the Transitional pattern to the SD pattern. The steps in which the scope is diminished - for example, 11 and 12 forming one step together - are not random. These will be discussed more in section 4.1 below, where the relation between the patterns is also examined more elaborately.

The unfrequent patterns of table 3.2 show no clear areal distribution and I will not use them in my analysis. If in future research the ordinal paradigms of the southern part of the Dutch-speaking area were to be documented in a more fine-grained way, the less structural patterns may become visible as transition zones between the larger central areas of 1, 2 and 3. I will elaborate on this idea below, and I will discuss some literature on transition zones: we need this in order to establish some of the criteria for labeling a specific area as such, as well as to decide what we can expect from such an area in terms of linguistic features of the dialects inside the transitional area and the dialects surrounding it.

3.3 Conclusion

Now that the data have been organized into a small set of patterns and varieties of these patterns, a few questions arise. What do the numbers 7-9 have in common with each other? What do 11-12 have in common? Do these two subsets of the ordinals have something in common with each other, so 7, 8, 9, and 11 and 12? If so, what is the nature of this common property? Do they share form-characteristics? Syllable structure, (residual but no longer transparent) morpheme complexity? And if so, how do they relate to the higher numbers on the number line, which also get *-ste*?

Between the Flemish pattern and the more restricted Transitional pattern, the big difference is the number 10. If these are included within a dialect's grammar, is this perhaps

due to 10 having a bimorphemic status in those dialects? Can it be shown for example that these dialects pronounce it disyllabically, and if so, does that indicate bimorphemity?

I will work out these questions in section 4 below, where I analyze how the Flemish pattern differs from the SD pattern, and what motivates these differences; and how the intermediate pattern is different and what motivates those differences.

4. Analysis of suffix distribution in the southern dialects of Dutch

As was stated in the introduction chapter, this thesis is focused around three research questions:

- (i) What are the suffix choices in the ordinal number lines of the different varieties of Dutch (including the standard variety)?
- (ii) How can these patterns be motivated?
- (iii) Do the suffixes share the same underlying representation?

In chapter 3 I described the variation in ordinal forms in the Dutch language area, thus answering research question (i). We saw that the variation falls into three categories: a Flemish pattern, a Standard Dutch pattern, and an intermediate pattern or Transitional pattern. Chapter 4 answers research questions (ii) and (iii). Question (iii) is answered in (4.2.1) where I examine whether the suffixes are allophonic alternants of the same underlying representation and conclude that they are most likely suppletive allomorhps. Question (ii) will be answered in threefold: firstly, by way of an extralinguistic motivation, namely geographical distribution explains a change in grammatical system (4.1); secondly, with a negative result, namely that there is no effect of the final stem consonant on the suffix choice (4.2.2); and thirdly, with an intralinguistic motivation, namely for the transitional area, a presuffixal extra light syllable dictates the selection of *-ste* (4.2.3). This is further elaborated by looking at the syllabification (4.2.4), where I must conclude that there are also phonologically disyllabic ordinal stems *without* presuffixal schwa syllable (4, 5, 10) which are problematic for an analysis of the suffix choice based on stress patterns.

4.1 Extralinguistic motivation: geographical distribution

Recall the blank map with Daan & Blok's (1969) dialect areas from figure 3.1. If we draw the three most frequent patterns onto that map we get figure 4.1, and it becomes visible that the patterns have a clear geographical distribution. The pattern with *-ste* ordinals from 7 upwards (depicted by blue squares) occurs exclusively in West and East Flanders, and is therefore labeled the 'Flemish pattern'. The pattern depicted by white circles is the SD pattern. The third of the three big patterns lies in between the SD pattern and the Flemish one - both with regard to their numeric scopes as well as their locations - and is therefore labeled 'Transitional pattern'. We can see that the yellow triangles of this pattern lie in the dark orange area that corresponds to Daan & Blok's (1969) South-Brabantish dialect area (the Belgian provinces of Antwerp and Brabant). A final important observation is that the Flemish and the transition pattern occur only in the Belgian part of the language area: the Netherlandic part contains only the SD pattern.¹⁸ Table 3.3 with the three largest patterns including their numeric scopes and frequencies is repeated below for the reader's convenience as table 4.1. Note that the numbers of occurrences given between parentheses in the upper left corner of the map do not match those in the table. This is because the map shows only one symbol per pattern per location, while the table lists all data points, thus including all respondents for locations with more than one respondent per pattern. In other words, the legend in the upper left corner of the map denotes not the number of speakers per pattern, but the number of locations illustrated on the map for each symbol.

¹⁸ Note that this means that the **Flemish pattern and the Transitional pattern** are **restricted** to certain parts of Flanders only. The Standard-Dutch pattern occurs more widely than is depicted in figure 4.1: not only in the locations marked by white circles, but all over the map. For the sake of clarity, these other locations have been omitted from figure 4.1. The SD pattern most likely occurs in the entire Dutch language area, but because the area of research was limited to the regions discussed in section 3.1 above, naturally the white circles on this map are also limited to these regions. Additionally, in many locations in the Belgian provinces of West and East Flanders, Antwerp and Flemish Brabant - those provinces where we find the Transitional pattern and the Flemish pattern - other participants reported SD paradigms, showing that the standard language is spoken throughout the Dutch language area.

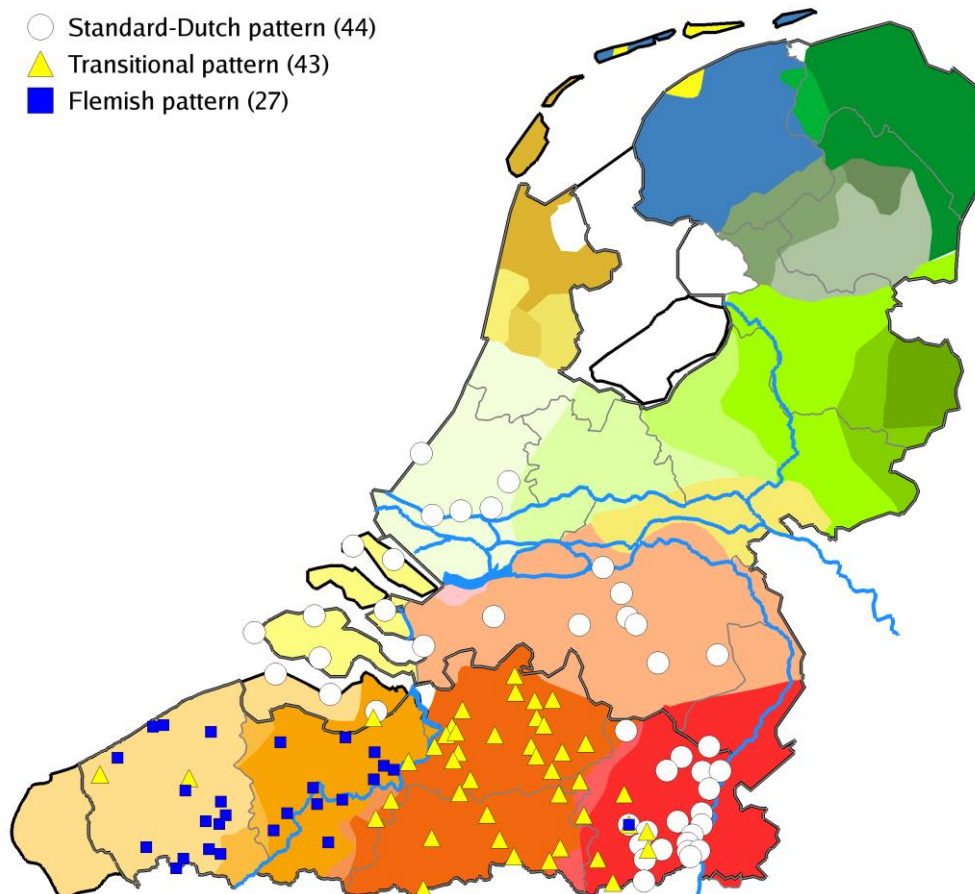


Figure 4.1: Distribution of ordinal patterns in the southern provinces of the Dutch language area

Table 4.1: The three biggest patterns

Numeric scope of - <i>ste</i>	Number of respondents	Pattern label
≥ 20 (and 8)	87	Standard-Dutch pattern
7-9, 11-12, ≥ 20	55	Transitional pattern
≥ 7	32	Flemish pattern

This is a simplified map.¹⁹ Had I drawn a map with the smaller patterns listed in table 2.1, it would become visible that the Netherlandic part of the language area also contains a few instances of non-Standard-Dutch ordinal datasets; however, these instances are not many in number. Future research could shed light upon the subject; for now, I think it can be safely

¹⁹ The map was created by the author with the help of the Kloeketabel (see Van den Berg, 2003) on the basis of the blank map including colored dialect areas following Daan & Blok (1969), provided at <http://www.meertens.knaw.nl/sand/zoeken/kaart/frames.php>.

concluded that non-Standard-Dutch ordinal paradigms are not common in the Netherlands, while there are many Flemish dialects (and many speakers of those dialects) that have an ordinal paradigm that is clearly distinguishable from the SD ordinal set.

The novel findings laid out in the sections above seem to pattern neatly into three areas: (i) a large part of Flemish Belgium, on the West side of Flanders, roughly coinciding with the two provinces known as ‘West Flanders’ and ‘East Flanders’; (ii) an area in the middle of Flemish Belgium, roughly coinciding with the provinces of Antwerp and Belgian Brabant; and (iii) the Limburg province of Belgium, and the southern parts of the Netherlands (in the provinces of Dutch Limburg, Dutch Brabant, and Zeeland). That the variation in ordinal suffix distribution within a speaker’s ordinal number line can be mapped onto three different regions is an interesting finding; but what makes the finding really interesting is that there are previous publications that the current, novel finding perfectly agrees with. They describe the exact same geographical regions that I mentioned above, showing interregional differences and intraregional similarities in terms of the linguistic features that the authors researched. Two examples of such literature are De Vogelaer (2008) and Barbiers, van Koppen, Bennis and Corver (2016). I will discuss their findings below. What is interesting for the current dataset is that this previous literature describes the three areas as being a ‘core’ area and a ‘peripheral’ area with the third one a transitional area or *transition zone* between the first two. By comparing their motivations for doing so with my own dataset, and by consulting other literature on transition zones in general, I will show how and for what reasons the geographical distribution of my data can also be interpreted as being situated in a ‘core’ and a ‘peripheral’ region and one transition zone between them.

De Vogelaer (2008) discusses subject doubling, a syntactic phenomenon displaying a great deal of variation among the dialects of Dutch.²⁰ As the name suggests, subject doubling occurs whenever there are two instances of the same subject in one sentence - for example, two subject pronouns, or one pronoun and one clitic; but, crucially, both instances refer to the same entity. See the example in (3), which is (1c) in De Vogelaer (2008: 230):

- (3) Gij gaat gij naar Brussel.
 you_{strong} go.2SG you_{strong} to Brussels
 ‘You are going to Brussels.’

²⁰ It was noted by Sjeff Barbiers (p.c.) that the doubling status of first person doubling in Brabantish has been disputed by several studies. However, the current purpose is simply to sketch the features and characteristics of a transition zone displayed by the relevant areas, and therefore I choose to show De Vogelaer’s map nevertheless.

De Vogelaer writes this paper based on data from the Syntactic Atlas of Dutch Dialects (or 'SAND', Barbiers et al., 2005). Have a look at his 'map 1', which is shown in figure 4.2. He denotes the black area as the 'core' area, the light grey area as 'peripheral' and the grey zone in between as transition zone:

'The most productive area, or the 'core', is the western, Flemish dialect area, i.e. French-Flanders, and East and West-Flanders, where doubling is found for all grammatical persons. In between the core area and the peripheral area, a transition zone is found, where both second person pronouns and first person singular pronouns can be doubled'.

(De Vogelaer, 2008: 236)

The light grey, 'peripheral', area shows doubling only on 2SG/2PL pronouns. Thus, De Vogelaer presents the data as if on a cline from least restricted (all pronouns are affected) to most restricted (only 2nd person pronouns), with the transition zone in the middle (something halfway between the other two zones). This is exactly how Kortmann (2010: 853-854) describes the phenomenon 'transition zone': 'the typical situation for syntactic variation, especially on a supra-regional scale, is that there are no sharp isoglosses (...) but that the situation can rather be likened to a cline or slope.' It is important to stress here that a transition zone is 'in the middle' in two ways: the data displayed in a transition zone is a sort of compromise between the patterns in region A and B; and the zone is physically between the regions: it is *geographically located between them*.

The resemblance between the maps in figures 4.1 and 4.2 is striking: the blue squares in 4.1 coincide roughly with the black 'core' area in 4.2 and the yellow triangles fall inside the grey transition zone. The white circles of 4.1 should be expected, then, to be found in the 'peripheral area', or most restricted area; and from the perspective that we have taken so far, this is exactly what we find. The *-ste* suffix is combined with 7 upwards in the 'core' area, while it is restricted to 20 upwards in the 'peripheral' area; the transition zone in Antwerp and Brabant shows different kinds of mixes between these two extremes.

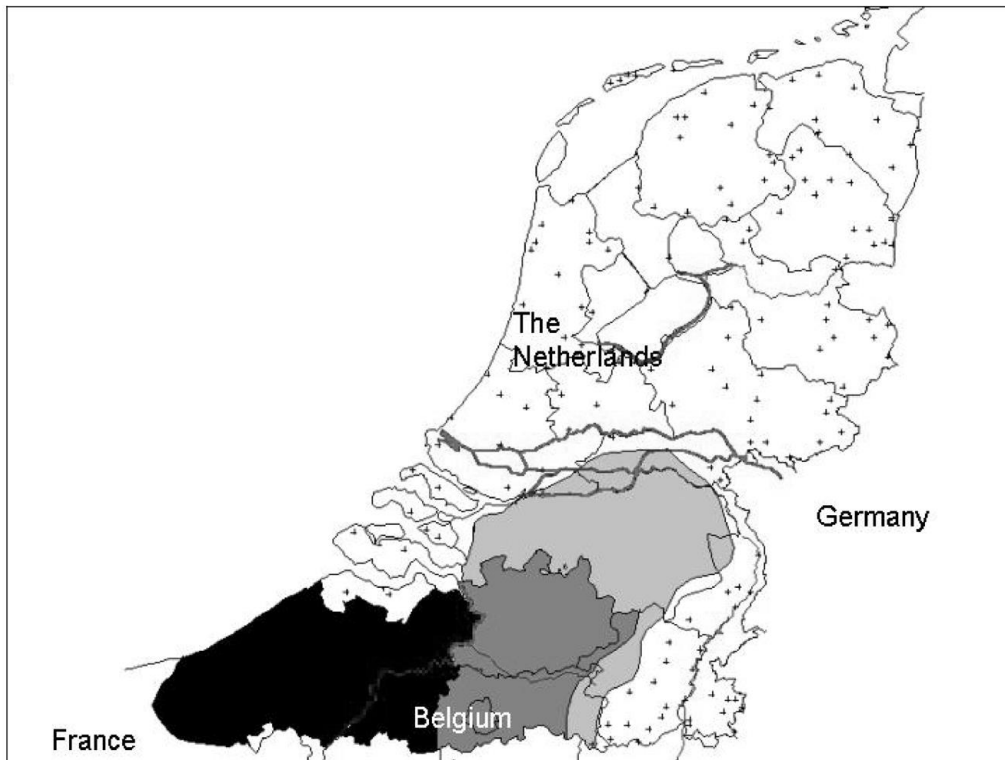


Figure 4.2: Map 1 (De Vogelaer, 2008: 235)

Barbiers, van Koppen, Bennis and Corver (2016: 6) also use data from SAND (Barbiers et al., 2005). They compare four syntactic phenomena and the geographical distribution of variation for each phenomenon. Just like De Vogelaer (2008) they research the dialectal variation in subject doubling; additionally, they look at demonstrative doubling, fronting in imperatives and complementizer agreement. Comparing the data, they discover correlations between the variation that they find in certain areas. Driven by the empirical findings, they group the datasets together into four dialect areas: Dutch, Flemish, South-Brabantish and North-Brabantish. They describe the relation between these areas as follows:

‘South-Brabantish (i.e. the dialects spoken in the Belgian provinces of Antwerp and Belgian Brabant) can be characterized as a transitional zone between the dialects of Flemish and those of North-Brabant. In turn, the whole Brabantish area can be characterized as a transition zone between Flemish and northern Dutch.’

(Barbiers et al., 2016: 6)

This is made visual in figure 4.3, which is a copy of table (36) in Barbiers et al. (2016: 23).

(36)	North-Brabantish	South-Brabantish	Flemish
Subject doubling	Restricted	Restricted	Non-restricted
Comp agreement	-	-	+
D-fronting imperatives	+	-	-

Figure 4.3: Table (36) (Barbiers et al., 2016: 23)

What this chart states is that the dialects of North Brabantish, South Brabantish and Flemish all display the phenomenon of subject doubling. South Brabantish has in common with North Brabantish that it doesn't allow complementizer agreement, while it has in common with Flemish that it doesn't allow D-fronting in imperatives. Therefore, it can be categorized as being 'in between' both dialects: it shares features with both dialects, but also differs from both dialects.

In the same way, North- and South-Brabantish together differ from Standard Dutch and South Hollandic in that the Brabantish dialects allow doubling whereas the Hollandic dialects don't; this is the same observation that we saw from De Vogelaer (2008) above. Flemish has the full-fledged, non-restricted doubling system; the Hollandic dialects have no subject doubling whatsoever, and the Brabantish dialects are in the middle of this cline: they display doubling, but in a much more restricted way.

From the two studies discussed in this paragraph, it seems that there is some precedent to corroborate the grouping of the data from the present study into three zones: a core Flemish zone, a peripheral zone in Belgian Limburg and the Netherlands, and a transition zone in Antwerp and Belgian Brabant. But this leads to the question: what exactly is a transition zone? Let us have a look at some of the literature on transition zones and see how they can help interpret the novel data.

Kortmann (2010), who bases his work on Chambers and Trudgill (1998), says the following about transition zones:

‘[T]he variants of a given syntactic phenomenon which are exclusively found or strongly preferred in the clearly identifiable syntactic areas (e.g. variant A in zone 1 and variant B in zone 2) are found side-by-side in the transition zone (i.e. both A and B) and may even combine (AB).’

(Kortmann 2010: 853-854)

In the influential work of Chambers and Trudgill (1998) for dialectal variation between two phonological systems, this is referred to as a ‘mixed lect’: forms from both systems occur in the transition zone.²¹ Taeldeman (2010: 361) argues, when discussing the differences between central areas and transition zones, that ‘areas of the second type are characterized by an intensive spectrum of local variation (in the structural sense)’. Guido Seiler (2004: 381-384) makes a three-way distinction between diatopic variation, free grammar-internal variation, and conditioned grammar-internal variation. All three can be argued to be present in our data.

We find at least three types of grammar-internal variation - in the core area as well as in the transition zone. Less so in the peripheral zone, and this is to be expected, because the peripheral zone displays the number line of the standard language, which is always more stable: the dialects converge towards the standard, but the standard is less likely to shift towards any dialects: ‘[T]he converged-to variety holds a higher status in social space than the converging variety. In many cases the converged-to variety is an overarching spoken standard variety’ (Røyneland, 2010: 259).

Firstly, there is a form of variation which we should analyze as *free* grammar-internal variation: many speakers have two forms for the same ordinal; some speakers have a preference for one of both, while other speakers have no preference.²²

Secondly, apart from speakers with numberlines as filed under the Flemish pattern, the Standard-Dutch pattern, and the Transitional pattern, there was a number of less frequent patterns, most of which were found in the core area and transition zone. These ‘special patterns’ seem to follow different restrictions from the general patterns and could therefore possibly fall under conditioned grammar-internal variation - what needs to be done, then, is to

²¹ This is in opposition with a ‘fudged lect’, where the forms from system A and system B are mashed together and form a new form which does not occur in either A or B; in the present study this would translate to such forms occurring as *elf-de-ste* ‘uth’. Through further elicitation with speakers, these forms may well be uncovered, but while some forms in the raw data may adhere to this type of finding, the type of data collection for the current study (online survey) leaves too much room for error of interpretation to seriously investigate whether we can find Chambers & Trudgill’s fudged lects in the ordinal paradigms of the Dutch language area. Alternatively, one could consider the paradigms of the transitional dialects as fudged, because they combine properties of both neighboring areas.

²² It should be noted that more research is needed to get a clearer picture of this free grammar-internal variation.

find the conditions. Thirdly, not all speakers adhere perfectly to the ordinal suffix distribution that rules the area (which seems to be a mix between the suffix distributions in the core and peripheral area): while all datasets under a certain pattern check all the required boxes, checking a box does not entail that ALL numbers within the relevant range are formed with *-ste*: as I described in section 2.3 above, I generalized over the data in a controlled way by filing a dataset under ‘Transitional pattern’ only if that dataset was able to check the required criteria. What this means is that there are instances of Transitional pattern speakers in my dataset who did not in fact have a perfect ordinal numberline with exactly those numbers that are mentioned in the numeric scope ‘7-9, 11-12, ≥ 20 ’ as found in table 2.2 above: while most of the speakers in this group *did* have all of these ordinal numbers combining with *-ste*, there are some who have almost the same number line, but they have an exception for [12], for example. Their number line then is ‘7-9, 11, ≥ 20 ’. This could mean that these speakers have some additional conditions in their grammar, resulting in a slightly different numberline.

One may object to this methodology. However, I would like to point out that if no generalization is made, what we see is a tentative division of the speakers into three groups, and a lot of speakers that don’t fall into these categories but just have some kind of undefined mix of *-de* and *-ste* ordinals. But this would fail to capture *the diatopic variation*: it is the kind of variation that Kortmann (2010) describes. The transition zone shows signs that its speakers lack a clean cut-off point like the Flemish pattern has at 7, or the Standard-Dutch pattern at 20. It is almost as if the speakers in the transition zone are realizing forms that are neither grammatical in the Flemish zone, nor in the Dutch zone, which would suggest that these transitional systems are unstable. The forms are [-grammatical] but they are [+realized] nonetheless - the only combination you would not expect to be possible.²³ As Røyneland (2010: 260) puts it: ‘Vertical convergence of linguistic varieties will normally reduce the degree of *inter-systemic* variation. *Intra-systemically*, however, **the degree of variation may increase since speakers can choose linguistic variants from a larger repertoire.**’

Apart from the grammar-internal variation that we find in high amounts in the transition zone (but also in the core area, which may lead us to conclude that that area is actually also starting to transition under influence from the standard language), Seiler’s definition of diatopic variation seems equally relevant to the situation in the southern part of the Dutch language area. Diatopic variation, according to Guido Seiler (2004), is ‘all kinds of geographical contrasts between grammars’, including the possibility that ‘A and B [are] different solely in their preference for one or another option, but not in their inventories of devices.’ This is the

²³ [+grammatical], [-realized] are forms that are not disallowed by the grammar, but for some reason just don’t occur in a language (see for example Barbiers, 2014: 198).

situation for ordinal formation in the Dutch language area: in the entire area, we find variation in vowels and consonants, so that the number words sound slightly different - but we can always recognize the same two suffixes: *-de* and *-ste*. Thus, the ‘inventories of devices’ are always the same. What varies per region - apart from sound changes - is the ‘preference for one or another option’.

Let us now take a closer look at the descriptions of these three systems. Is there a way to connect them? As described above, both the SD pattern and the Flemish pattern seem to have a cutoff point on the ordinal numberline: the former at 20, the latter at 7. Under this point, ordinals are formed with *-de*, above it with *-ste*. It is desirable to have one rule that is able to generate two such very similar systems. An informal notation of such a rule might be the following:

(4a) Standard-Dutch pattern ordinal formation rule:

Take a cardinal number as the input form.

If the value of the cardinal number <20: use *-de* suffix;

Otherwise: use *-ste* suffix.

(4b) Flemish pattern ordinal formation rule:

Take a cardinal number as the input form.

If the value of the cardinal number <7: use *-de* suffix;

Otherwise: use *-ste* suffix.²⁴

The in-between pattern, then, does *not* have a clear cutoff point. Instead, it mixes both systems: it has 7 as a cutoff point like the Flemish pattern, but for 10 and the compound numbers 13-19 (with 10 as their second component) the forms on *-de* are ‘borrowed’ from the SD system, even though they are ungrammatical in the Flemish system. This is typically what happens in a transition zone between dialects: there is a clash between two systems, causing

²⁴ We might even add another if-clause at the point where the cardinal number is selected as input and extend it to other ordinal formations in Dutch, like this:

(4c) If the input form is a cardinal number:

If the value of the cardinal number <20: use *-de* suffix;

Otherwise: use *-ste* suffix.

Otherwise: use *-ste* suffix.

This way, the rule could potentially be extended to non-cardinal-based ordinal formations like *umpteenth* and *how-manyeth*; see section 2.1 for a motivation for regarding *-ste* as the default ordinal suffix.

forms which would be ungrammatical in system A to occur anyway, due to influence from system B. In our case, this means that in the Flemish system, *10-de* is ungrammatical, but in the transition pattern, there is influence from the SD pattern, causing *10-de* to be imposed upon the clean ordinal rule that dictates all ordinals above and including 7 to be formed with *-ste*. Why, then, are specifically these numbers targeted for a cross-over with the SD pattern? In 4.2, I examine the possibility of an intralinguistic motivation for the suffix choices in the different patterns.

4.2 Intralinguistic motivation

We saw in section 4.1 that a possible explanation for the differences between the three ordinal suffix patterns is their clear geographical distribution. However, the geographical distribution can only account for the empirical fact that there are three different systems, not for the specific differences. What is so special about 10 that it makes the difference between the Transitional pattern and the Flemish one? For answers to this kind of questions, we must look for intralinguistic motivations.²⁵

4.2.1 Two ordinal suffixes: suppletive allomorphy?

Before we can say anything about the intralinguistic factors at play during suffix selection, it is important to decide whether the suffixes in question are suppletive allomorphs or derivable from the same underlying form, and to decide this, it is useful to discuss what allomorphy is.

There are a number of different linguistic situations in which the term *allomorphy* may be used. Generally speaking, allomorphy is ‘the phenomenon that a morpheme has more than one phonetic form’ (Booij, 2016). This general phenomenon can be further divided into different categories.

Firstly, we can distinguish between allomorphs that share a common underlying form and those that do not:

²⁵ It is not unthinkable that there is another kind of extralinguistic factor at play: a mathematical (conceptual) factor. I have left this factor outside of the scope for this thesis, but considering the fact that we are dealing with numbers, it may be worth considering.

‘The classical generative approach to allomorphy is that all allomorphs of a morpheme are derived from a common underlying form by means of a set of (possibly ordered) phonological rules. The choice of a specific allomorph is governed by the phonological shape of the stem.’

(Booij, 2012: 9)

Within this category, the phonological rules mentioned above can either be rules which are generally applied in the language (final devoicing, for example) and thus automatically also apply to the morpheme in question; or they are non-general rules, also known as morpholexical rules, applying only in ‘the presence of a specific lexical or morphological feature’ (Booij, 2012: 11). This is what happens when a set of allomorphs are clearly derivable from one underlying representation (or UR, as Smith, 2015 calls it), but the rules that derive the different allomorphs do not appear elsewhere in the language. Booij (2012: 11) gives the example of Dutch diminutives: the underlying coronal stop /t/ of the morpheme /-tje/ place-assimilates to the final sound of the stem to which it attaches. This rule does not apply elsewhere in the Dutch language and therefore is a subtype of classic rule-based allomorphy.

In testing the derivability of allomorphs from a potentially shared underlying form, if allomorphs can be traced back to different historical origins, this can be a good indication that they lack a common UR (see Booij, 1998: 145 and Kager, 1996: 3). However, one should bear in mind that morphemes may always be reanalyzed as time passes, thus two initially unrelated morphemes which come to fulfill the same grammatical function may end up being so phonologically similar that it can appear as though they do share one UR - even if the language does not have an independent phonological rule to derive them, because we can still analyze those data as the result of a morpholexical rule in the domain of that specific suffix. Vice versa, the phonological rules that derived a morpheme from another morpheme in an historical language stage may become unproductive, resulting in two synchronically non-derivable allomorphs (Booij 2012: 17 and Kager, 1996: 3, footnotes 5 and 6).

So, do *-ste* and *-de* have different historical origins? Yes, they do. According to van Loey (1970: 155), in Middle Dutch, forms on *-ste* were formed analogously to *eerste* and *twintigste*: *sevenste*, *achtste*, *tienste*. This is supported by Ross & Berns (1992: 622), who say that the ordinals of the decads are formed with ‘a suffix identical with one of the two superlatives of adjectives (*-ôsto*)’ and that ‘from the decads, this suffix often spreads to the lower ordinals’. The *-de* suffix (as well as the Germanic cognate forms English *-th* and German *-te*, among others) was developed from the Proto Indo-European forms *-tē/-tō* (van Bree, 1987: 259; van Loey (1970: 155). Thus, *-ste* and *-de* have different historical origins. Following the reasoning

above, there are still two options open: either they are *still* derivable, even with different origins, because of historical or contemporary phonological rules; or they simply do not share one UR.

The two suffixes are *-ste* and *-de*, both ending in schwa, making the first possible alternation *-st-* versus *-d-*. There is, to my knowledge, no phonological rule in Dutch that produces such alternations neither currently productive nor was there one at any earlier stage. I have not come across any literature that said *-ste* and *-de* are to be derived from the same underlying morpheme, save one exception: Van Beers (1854), who claims that the two ordinal suffixes must originally have been one, and to support his claim, sums up a few Dutch minimal word pairs with *-d-* and *-st-*, and two similar meanings in each pair.²⁶ I subjected these minimal pairs to scrutiny and found no evidence for his claims in contemporary etymological dictionaries.

Thus, I must conclude that *-st-* and *-d-* cannot be phonological alternations in Dutch. However, this is not the only way to derive *-ste* and *-de* from the same UR: /t/ and /d/ differ only in the voicing dimension (not in manner or place of articulation) and /s/ shares the same place of articulation as well, making it not unlikely that *-ste* and *-de* are related through a voice alternation of /t/ and /d/ accompanied by either a deletion or an insertion of /s/. This could go in two directions: either /ste/ > [de] or, more likely in my opinion, an underlying voiced /d/ resulting in /de/ > /sde/ > [ste]. Of course the insertion of /s/ needs then to be motivated. For example, for 11, it is not at all clear why not simply use *-de*; and if it would devoice, *elfte* should also work fine in Dutch phonotactics. It is not apparent upon first sight what motivates the insertion of *-s-* into a hypothetical formation *elfde* > *elf-s-de* > *elfste*. Whether this is a possible analysis for the suffix distribution as we find it in the data, will be discussed in section 4.2.2 below, where I compare the final sounds in the stem forms.

Even if the two Dutch ordinal suffixes do not stem from the same underlying form, their distribution can still be determined by phonology. This is known in the literature as Phonologically Conditioned Allomorphy (PCA: Carstairs 1988 and subsequent work). It is sometimes termed Phonologically Conditioned **Suppletive** Allomorphy (PCSA), to denote (one of) the affixes as being suppletive rather than coming from the same UR, as ‘suppletion’ is a term commonly used for the appearance of a non-derivable form among the regular forms in

²⁶ His examples: *moest* - *moetede* (sg. past tense of the modal verb *moeten*, ‘must’); *begost* - *begonde* (sg. past tense of the verb *beginnen*, ‘begin’); *dorst* - *dorde* (archaic sg. past tense form of *durven* ‘dare’); *stuwen* - *duwen* (infinitives, meaning: ‘to push (forward)’); *stom* - *dom* (adjectives ‘stupid’ or ‘dumb’); *dezelfde* - *dezelfste* (according to Van Beers, *dezelf-ste* is the superlative of *dezelve*, a demonstrative pronoun; *dezelf-de* would simply be an alternant of this superlative). For *duwen/stuwen*, the *Nederlands Etymologisch Woordenboek* (Dutch etymological dictionary) by De Vries (1971) gives two different origins: *duwen* PIE **teuk^w*- ‘squeeze/press/push’; *stuwen*: PIE **steh₂u-*, a root-extension of **steh₂-* ‘stand’.

a paradigm, either from another paradigm or from an unknown source.²⁷ Here is an example of PCA from Smith (2015: 1), concerning the subject marker in Korean: ‘It has two allomorphs: *-i* occurs after consonants, and *-ga* occurs after vowels. Though the distribution is phonologically conditioned, the two allomorphs cannot be derived from a single UR.’

How this phonological conditioning takes place has been the subject of much debate. I will give a brief overview, based on Smith’s (2015) excellent introduction to his dissertation. We can distinguish two general approaches: the grammatical approach on the one hand, and the lexical approach on the other. We find the grammatical approach in constraint-based frameworks, mostly Optimality Theory (OT: Prince and Smolensky, 1993, 2004; McCarthy & Prince, 1993). Works such as Mester (1994), Kager (1996), Mascaró (1996), and Wolf (2008) all present data of affix competition, and explain those data by means of OT tableaux, representing a ranking of several phonological and morphological constraints at work in the relevant language.²⁸ In these accounts, the most optimal suffix is selected by the phonological grammar: suffix selection occurs during the phonology.

These phonological constraints can be of two general directions: one direction resembles the ‘classic’ allomorphy type, in the sense that the stem form, the input, determines which affix is selected, as in the case for the Korean subject marker above; the only difference with the classic allomorphy type, then, is the absence of a shared underlying form for the allomorphs in the case of PCA.

Alternatively, and more commonly in this type of constraint-based PCA literature, it is not the *input* that determines the allomorph selection, but it is constraints on the *output* that do so. Booij (1998) gives a few examples in Dutch, where the choice of a plural suffix *-en* or *-s* is determined by a prosodic condition, namely the preference of Dutch to form disyllabic trochees; the difference between the two suffixes being that *-en* adds a syllable, while *-s* does not. Constraint rankings can show which affix must be the preferred one in a language; Kager (1996: 155) calls it output optimization, and provides examples from the Djabugay genitive and the Estonian genitive and partitive.

Alternative to a language choosing the most optimal suffix, suffix distribution may also be governed by the desire to *avoid* certain sequences, thus looking not to select the most optimal

²⁷ Booij (1998: 145) prefers for the sake of clarity to only ‘use the term ‘allomorph’ if the competing affixes show a clear phonological similarity’. Instead, morphemes that ‘have nothing in common phonologically, but have the same grammatical function’ may be more appropriately labeled as ‘competing affixes that compete for expressing the same information’ (Booij, 2016). However, I think - at least in the context of the current thesis - this is a bit of a trivial matter and I choose to use the term allomorph to refer to the competing suffixes throughout this thesis, regardless of whether or not they share an UR.

²⁸ The works mentioned in this paragraph are as cited by Smith (2015).

candidate but rather to avoid the least optimal one. Booij (1998: 156-157) gives the example of the Dutch comparative suffix *-ər*: to avoid sequences like [rər], the allomorph *-dər* can be chosen instead.

There have been accounts presenting data that do not conform to the relevant language's phonological rules. Instead, a non-optimal form, or even what the ranking predicts would be the *least* optimal form, is selected. Paster (2009) calls this a 'perverse' allomorph distribution. The accounts presenting this kind of evidence (see for example Bye, 2008; Embick, 2010; and Paster, 2005, 2007, 2009, 2015 as cited by Smith, 2015) contra the constraint-based grammatical approach often fall under or are akin to the framework of Distributed Morphology (DM: Halle and Marantz, 1993).

Smith (2015) refers to these works as undertaking a lexical approach, because they propose that affix selection is determined in the lexicon: they argue that some affix selection phenomena simply cannot be adequately explained by use of phonological and morphological constraint rankings and cannot always be predicted. Instead, each suffix in the lexicon comes with its own subcategorization, or subcat, frame with phonological requirements. This can consist of a set of phonological and/or morphological requirements on what kind of stem the suffix can attach to (lateral subcategorization, the most common type), or a set of requirements for the output - referred to as *vertical* subcategorization by Bennett et al. (2015: 2-3).

In the midst of this dichotomy between literature from the grammatical approach and from the lexical approach, there are works that argue for a combination of the two: Dolbey (1997); Lapointe and Sells (1997); Booij (1998); Bonet, Lloret, and Mascaró (2003); Nevins (2011); and Smith (2015)²⁹ defend the grammatical approach by showing that certain generalizations can be overlooked if every affix competition in every language would be ascribed to subcategorization frames with requirements specified in the lexicon for every allomorph; they also show that for some cases, the rules are not sufficient to explain all the facts and then the lexical approach can help account for those facts. Recall for example the plural suffix choice mentioned above. It is governed according to Booij (1998) by phonological rules focused on creating the most optimal output in terms of stress pattern. Adding *-en* adds a syllable to the word, while *-s* does not; thus, depending on the amount of syllables present in the stem, choosing one over the other will result in a good amount of syllables for creating disyllabic trochees, while the other one will not have the same effect. However, when forming the plural of certain diminutive forms, both suffixes yield the same disyllabic trochee: *toe-tje-s* versus *toe-*

²⁹ All works mentioned are cited by Smith (2015) and were not consulted by me separately.

tje-n: clearly there is no differentiating effect left between the two. Still, Dutch diminutive words always get the *-s* suffix, there is no free variation. Thus, phonology alone is not enough to decide between both forms, reasons Booij (1998: 148), and he concludes that the diminutive suffix subcategorizes for a specific plural suffix, namely *-s*; and that this subcategorization steps in to aid the suffix selection where the phonologically optimal ordering of constraints fails to make the distinction.

To distinguish whether or not allomorphy facts can be accounted for by means of a constraint ranking, one important criterion is that the ranking must not be alien to the language in question. A similar situation is encountered by Booij (1998), who comes across two affixes of which the allomorphy facts cannot be accounted for with one single constraint ranking:

‘[D]ifferences in behaviour [between two affixes] could be accounted for in terms of different constraint rankings. However, I do not want to assume affix-specific rankings of constraints, because this implies that the language learner would have to acquire a number of affix-specific grammars for his native language, not a very attractive position from the learnability point of view.’

Booij (1998: 153)

In this case, he then proceeds to assume subcategorization frames for all different affix allomorphs involved, in order to account for their idiosyncratic behavior.

We have seen classic allomorphy between derivable allomorphs, phonologically conditioned suppletive allomorphy based on the ranking of phonological and morphological constraints or based only on subcategorization frames, and finally PCA in which not only phonology plays a role, but which features morphological requirements listed in the subcat frame of the relevant affix, without any apparent motivation. To conclude section 4.2.1, I present an overview of the different types of allomorphy in table 4.2. Note that the upper right cell is empty: two allomorphs that derive from the same underlying form are always distributed according to phonological constraints.

Table 4.2: Overview of the different types of allomorphy

	Phonologically governed distribution	Non-phonologically governed distribution (often combined with phonological factors)
Same underlying form	<p>‘Classic’ allomorphy</p> <ul style="list-style-type: none"> Formed by general phonological rule of the language; Or formed by domain-specific ‘morpholexical’ phonological rule 	N/A
Different underlying forms	<p>Phonologically Conditioned Suppletive Allomorphy</p> <p><i>Grammatical approach:</i></p> <ul style="list-style-type: none"> Either the stem form governs the affix choice by a language-specific ranking of phonological and morphological constraints; Or the output is bound by (language-specific) phonological constraints; Both options can be optimization or avoidance of the least optimal option <p><i>Lexical approach:</i></p> <ul style="list-style-type: none"> The affix subcategorizes for specific phonological and/or morphological requirements on the stem (lateral subcategorization); Or it subcategorizes for phon. / morph. requirements on the output (vertical subcategorization). 	<p>Phonology plays no role or only takes care of part of the facts; the rest must be specified in the lexicon, otherwise the language learner has to learn a different ranking for each affix.</p> <ul style="list-style-type: none"> The affix subcategorizes for specific morphological requirements on the stem (lateral subcategorization); Or it subcategorizes for morphological requirements on the output (vertical subcategorization).

In the remainder of this chapter, I will look for a system that can derive the three different suffix distributions in Flemish, Standard Dutch and in the transition zone. In 4.2.2 I consider for the sake of completeness the final consonant of the ordinal stems; I will test specifically on voicing assimilation and I will show that there is no way to derive the patterns found in chapter 3 by looking at stem-final sounds alone. In 4.2.3 I measure and discuss the possible effect of the weight of syllables of the stem and of the output on the suffix choice. In 4.2.4 I take into account the literature on heavy syllables and their underlying disyllabicity when it comes to Dutch stress, and proceed to syllabify the cardinals and ordinals and give an account of their foot structures. Section 4.3 compares and evaluates the different hypotheses, and provides a conclusion to this chapter.

4.2.2 Final consonant of the stem

When two affixes are in competition, one way to distribute them is according to the stem's final sound. We see this for example with plural formation in English (the plural is pronounced voiced after a voiced consonant in *hugs*, voiceless after a voiceless consonant in *hacks*: [hʌg] - [hʌgz], [hæk] - [hæks]) and in the Dutch verbal paradigm with the past tense: the past tense suffix *-de* is pronounced [-te] after a voiceless consonant:

- | | | | |
|--------|--------------|---|------------------|
| (5) a. | landen | - | land- de |
| | land.INF | - | land-PAST |
| | 'to land' | - | 'landed' |
| b. | planten | - | plant- te |
| | plant.INF | - | plant-PAST |
| | 'to plant' | - | 'planted' |
| (6) a. | krabben | - | krab- de |
| | scratch.INF | - | scratch-PAST |
| | 'to scratch' | - | 'scratched' |
| b. | kappen | - | kap- te |
| | chop.INF | - | chop-PAST |
| | 'to chop' | - | 'chopped' |

There are several reasons to immediately assume at this point that such a selection mechanism cannot work for the ordinal suffix. Firstly, both examples given above are generally considered not to involve a selection of any kind: the allomorphs stem from a shared underlying representation (UR) and thus we are dealing with progressive assimilation. It is not apparent that *-ste* and *-de* share an UR, as I discuss in 4.2.1 above, and so we must conclude that comparing the ordinal suffix selection process to the examples of assimilation above is like comparing apples and oranges.

Secondly, Zonneveld (2007: 20) explicitly mentions exactly the ordinal suffix as well as the past tense suffix and argues that they behave differently. He states, contrary to other claims in the literature, that the past tense suffix is the *only* obstruent-initial suffix that undergoes progressive assimilation, and that the ordinal suffix *-de* triggers regressive assimilation on the final stem obstruent instead.

Zonneveld illustrates that ‘regressive assimilation occurs across the board’ by showing the productivity of the mathematical ‘x to the power of x’ construction: non-numeral words, mostly pronunciations of letters of the Roman or Greek alphabet all show regressive voice assimilation when they are paired with the ordinal *-de* which is used in this construction.

- (7) a. 8^f: acht tot-de-e[v]-de (*f* is pronounced [ɛf] in isolation)
 b. 10^s: tien tot-de-e[z]-de (*s* is pronounced [ɛs] in isolation)

He also gives all the Dutch ordinal forms and shows that they can be divided into stems with sonorant endings and stems with fricative endings.³⁰ I repeat here his list of forms with sonorant endings, with my own added IPA transcriptions in table 4.3.

³⁰ The only exception to this is *acht*, which he then speculates probably gets *-ste* because it ends in a plosive. He does not comment on the endings of *honderd* ‘100’ or *miljard* ‘1,000,000,000’; he does *show* the form for ‘100’ but he leaves it undiscussed, and makes no remarks about the final stem consonant, probably because his main focus is on *-de* and not on *-ste*. It strikes me as odd that he mentions ‘100’ in the list of ‘fricative endings’ but no remark is made with respect to this.

Table 4.3: Sonorant cardinal endings

	Cardinal	Ordinal	Transcription of ordinal stem
2	twee	twee-de	/tʷe:/
3	drie	der-de	/dɛr/
4	vier	vier-de	/vi:r/
7	zeven ³¹	zeven-de	/'ze:vən/
9	negen	negen-de	/'ne:ɣən/
10	tien	tien-de	/ti:n/

For the fricative endings, Zonneveld shows that some of them have been devoiced according to the general Final Devoicing rule which occurs widely throughout the Dutch phonology (a famous phonological rule also attested in German, Polish and Catalan, among other languages). He shows the underlying sounds by giving the plural forms of the cardinals. He claims that all of the fricative endings, even *zes*- '6' which is underlyingly voiceless, undergo regressive voicing assimilation to the voiced *-d-* of the ordinal suffix. See table 4.4, which includes my IPA transcriptions.

Table 4.4: Fricative cardinal endings

	Cardinal	Underlying final sound: plural cardinal	Ordinal	Transcription of ordinal stem
5	vijf	vijv-en	vij[v]-de	/vɛiv/ - [vɛiv]
6	zes	zess-en	ze[z]-de	/zɛs/ - [zɛz]
11	elf	elv-en	el[v]-de	/ɛlv/ - [ɛlv]
12	twaalf	twaalv-en	twaal[v]-de	/tʷa:lv/ - [tʷa:lv]
20	twin-tig	twintig-en	twintig-ste	/'tʷɪn.təɣ/ - ['tʷɪn.təχ]

³¹ For 7 and 9, the /n/ is not always pronounced in colloquial Dutch. Depending on the regional variant it can be deleted more often or less often, the details of which are not currently relevant; because once it is deleted, the final sound will be schwa - thus making the alternation between final [n] and final [ə] trivial for the matter of sonority.

In my view, an empirical production experiment is necessary to analyze the degree of voicing in various speakers for *-de* in *zes-de* ‘6th’, to test Zonneveld’s claim that the suffix is not sensitive to any progressive assimilation to the stem-final sound. If the suffix in ‘6th’ is phonetically analyzed not to be voiced, one could argue that *zesde* is written with *-de* due to paradigm analogy, while the actual suffix in use by speakers is *zesste*. This would be hard to distinguish from *zesde* with the naked ear if Zonneveld’s [zɛzdə] is an incorrect representation of the actual pronunciation made by speakers. Furthermore, Zonneveld (2007: 6) claims that ‘20’ has an underlying voiced coda consonant, and that it is pronounced voiceless when the suffix *-ste* is added. I am not sure whether this is true.

Ignoring for the time being Zonneveld’s analysis I will continue to work out the hypothesis that **the stem-final sound determines the suffix choice** to further show that it does not hold. If we assume that not the suffix alters the stem-final sound, but the stem-final sound determines the suffix shape, then we need to base the test on underlying forms, not the surface forms as occurring in the ordinals: we need to use the stem form as it occurs in the plural (see table 4.4). See table 4.5 for a list of all Dutch ordinal forms (as elsewhere in this thesis, excluding *eerste* ‘1st’ because it is not formed from a cardinal).

In this table it becomes evident that there are **only two ordinals** with an underlying voiceless stem-final sound: 6 and 8. The conclusions we can draw from this are the following:

- The underlying voicing features of the ordinal stem forms cannot cater for the *-ste* forms found in all dialects in ordinals ≤ 20 ;
- It’s possible that *acht* ‘8’ get *-ste* instead of *-de* because its final sound is a voiceless obstruent; however, then it needs to be explained why 6 gets *-de* and not *-ste*.

Table 4.5: Dutch ordinals with underlying stem-final sound and IPA transcription

	Cardinal	Underlying final sound: plural cardinal	Transcription of underlying ordinal stem
2	twee	twee-en	/tʷe:/
3	drie	drie-en	/dɛr/
4	vier	vier-en	/vi:r/
5	vijf	vijf-en	/vɛiv/
6	zes	zess-en	/zɛs/
7	zeven	zeven-en	/'ze:vən/
8	acht	acht-en	/ɑχt/
9	negen	negen-en	/'ne:γən/
10	tien	tien-en	/ti:n/
11	elf	elv-en	/ɛlv/
12	twaalf	twaalf-en	/tʷa:lv/
13-19	dertien	der-tien-en	/ti:n/
20	twintig	twintig-en	/'tʷɪn.təγ/
100	honderd	honderd-en	/'hɔn.dɛrd/
1,000	duizend	duizend-en	/'dœy.zənd/
1,000,000	miljoen	miljoen-en	/mɪl'ju:n/
1,000,000,000	miljard	miljard-en	/mɪl'jard/

4.2.3 Syllable weight

In the following two sections, I compare both the ordinal stem forms as well as their cardinal input forms in terms of their syllable weight (4.2.3) and the resulting stress patterns of the output forms (4.2.4); stress is determined in Dutch for a great deal by syllable weight. The literature mainly shows cases in which the *output* form is the relevant domain whenever syllable weight plays a role in determining the choice of a suffix from a set of alternative suffixes; Booij (1998) and Kager (1996) give examples for this. I am yet to find the literature in which through some kind of constraint, the final syllable of the input alone can determine which suffix must follow, without any interference of the output form; I will nevertheless, for the sake of completeness and clarity, begin by looking at exactly that - the weight of the final syllable of the input.

This immediately brings up the question: what should we regard as the input for a Dutch ordinal? For Standard Dutch (SD) ordinals, if we take off the suffix, we are left with a form which is identical to the corresponding cardinal. Thus, what I call the ‘ordinal stem’, *vier-* in (8b), is identical to the cardinal in (8a):

- (8) a. vier
four
 b. vier-de
four-ORD

For SD ordinals, the only exception to this is *drie – der-de*, in which the stem has undergone metathesis. See appendix A for the etymology of *derde*; from a synchronic perspective, it shows us that the ‘ordinal stem’ does *not* necessarily have to be identical to the cardinal input in all cases. And, as we will see below, in some Flemish and Zeeuws dialects, the cardinals all end in a schwa which is not present in the ordinal form.

Thus, to be complete in our endeavor to answer the research question ‘**What governs the distribution of ordinal suffixes in the three different patterns of Standard Dutch, Flemish, and the transitional Brabantish zone?**’ we will test the following two hypotheses in this section:

- (9) i. The weight of the rightmost syllable of the **ordinal stem form** determines the choice of suffix.
 ii. The weight of the rightmost syllable of the **cardinal input form** determines the choice of suffix.

The focus will be on hypothesis (i), after which I will briefly digress toward the cardinals to cover the cases where they differ from the ordinals, but I will not dwell on this for long, because they do not differ to a great extent. Before I get into either hypothesis, however, it is necessary to start by determining a definition of syllable weight.

In Dutch, syllable weight plays an important role in the distribution of word stress (Van der Hulst, 1984; Booij, 1998). The clearest cases of this phenomenon in Dutch phonology are visible in 1), the fact that syllables with a schwa in their nucleus can never be stressed; and 2), the fact that superheavy syllables are very strong attractors of word stress (Van der Hulst, 1984; Van Oostendorp, 2012: 353). Thus, after I look at the input forms strictly for syllable weight in 4.2.3,

I will consider the resulting word stress both in input forms as well as output in the following section, 4.2.4.

In (10) I present the four-way scale of weight for Dutch syllables as it is commonly found in the literature, taken here from Van Oostendorp (2012: 354):

- (10) extra light: C(C)ə(C)
 light: CV(V)
 heavy: CVC
 superheavy: CVVC, CVCC

Note that the syllable onset makes no difference; the distinguishing factors occur only in the syllable rhyme, or nucleus + coda. About this four-way distinction, Van Oostendorp argues the following: evidence shows that the two categories on the extreme ends, namely the schwa syllable and the superheavy syllable, are reflected in stress distribution mechanisms in Dutch³², but there is significantly less evidence for a distinction between the light and heavy types. He shows experimental data that do not conclusively support the latter distinction and therefore concludes that ‘there is little reason why analysts have been so uniform in their adoption of the assumption that Dutch stress is quantity-sensitive’. He additionally points out that traditional quantity-sensitivity implies counting moras or the number of rhyme positions, but that at close inspection Dutch displays neither type of quantity-sensitivity.

A good example of the insensitivity to the number of rhyme positions is that ‘both open and closed schwa syllables count as extra light’: whether the rhyme contains only a schwa, or that schwa is followed by one or more consonants, both [Cə] and [CəC] syllables are considered by the phonology to be extra light - meaning that they can never bear stress. Thus, Van Oostendorp proposes that Dutch is weight-sensitive, but rather than quantity-sensitive, it is *quality-sensitive*, whereby the following generalization is true of Dutch (Van Oostendorp, 2012: 354):

- (11) Syllables with an empty head should occur in the weak position of a trochee.

‘The intuition behind this is that weak metrical positions should not contain too much phonological material, and inversely. One could debate about the precise definition of ‘emptiness’, but it does not seem unreasonable to say that schwa is heavily

³² ‘Schwa syllables are really weaker and superheavy syllables are really stronger than other syllables’ (Van Oostendorp, 2012: 354).

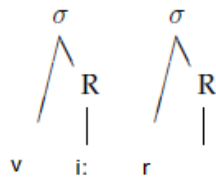
underspecified, at least in Dutch, where it functions as an epenthetic vowel and as the output of vowel reduction (Booij 1995; van Oostendorp 2000). It is thus empty in a relative sense.'

(Van Oostendorp, 2012: 354)

Following this, I label CV, CVV and CVC syllables as 'neutral', so as to distinguish them from extra light and super heavy syllables, without making an internal distinction within this neutral group.

Having discussed the extra light schwa syllables and the neutral syllables, one category remains: the superheavy syllable. Van Oostendorp (2012: 355) cites McCarthy (1979), Langeweg (1988) and Zonneveld (1993) in assuming that superheavy syllables should be treated as **disyllabic complexes**; the first syllable is either a closed syllable with a lax vowel, or an open syllable with a tense vowel. The second syllable consists of only a bare onset vowel, and an empty rhyme. Figure (x) illustrates the syllabification of *vier* 'four', [vi:r].³³

(12)



If we consider the generalization for empty heads in (11), it follows naturally that any superheavy syllable-complex will always attract stress: its second syllable, [r], in the example shown here, occurs in the weak position of a trochee, making [vi:] the strong, stressed head of the trochee, causing the whole complex to attract the stress if it is part of a larger prosodic word.

Traditionally, the superheavy syllable in Dutch is said to consist of a long vowel followed by one consonant, or a short vowel followed by two consonants (as for example in Booij, 1998: 149). However, Van Oostendorp (2000) argues for a tense-lax distinction in Dutch as a more reliable way to analyze the Dutch stress system. Thus, a superheavy syllable is either **a tense vowel followed by one, or a lax vowel followed by two consonants**. Van Oostendorp

³³ The representation in (12) was modeled by the author after figure (16) in Van Oostendorp (2006: 9).

(2000: 32) divides the vowels of Standard Dutch into two groups. The tense vowels of Dutch are a, o, e, ø, i, y, u; the lax vowels ʌ, ɔ, ɪ, ʏ, ɛ.³⁴

The assumptions made above regarding the relevant different kinds of syllable weight in Dutch are summarized in table (4.6):

Table 4.6: Syllable weight in Dutch

Extra light syllable rhyme	'Neutral' syllable rhyme	Superheavy syllable rhyme
schwa, no consonant (-ə)	lax vowel, no consonant (-V)	tense vowel + single consonant (-VVC)
schwa + single consonant (-əC)	tense vowel, no consonant (-VV)	lax vowel + double consonant (-VCC) ³⁵
schwa + double consonant (-əCC)	lax vowel + single consonant (-VC)	tense vowel + double consonant (-VVCC)

In the remainder of this section, the data of three different patterns are to be analyzed: the Flemish pattern, the Transitional pattern found in the Brabant area, and the pattern of Standard Dutch. The suffix distribution in all three patterns is repeated for the reader's convenience in table (4.7). *-ste* endings are in boldface to facilitate a clear picture.

As the more vigilant reader will notice, the ordinal stems in table (4.6) are given in the SD form for all three patterns, instead of a more appropriate dialectal form. The reason is that there is a huge amount of variation in our survey data in terms of stem forms. At this point, I do not have a motivated dialectal ordinal stem form for either non-standard pattern that would accurately represent its entire respective area. To arrive at such a form would require a careful analysis of each dialectal respondent for each different ordinal; such an analysis is too ambitious an aspiration for the present project. Thus, I will start the analysis by comparing the syllable weights of the **standard forms**. As some promising results then call for more fine-

³⁴ He notes hereby that in Standard Dutch, all tense vowels are generally speaking perceived as long, while the lax vowels are short. This is exactly reversed in Antwerp Dutch, where all long vowels are lax and all tense vowels are short (Van Oostendorp, 2000: 122); furthermore, other dialects may have yet different vowel inventories. For example, Hofstade Dutch, as Van Oostendorp (2000: 122) additionally mentions, contains both long and short versions of all the tense as well as all the lax vowels. However, for this thesis, time limits restrict me from looking further into this. I will assume that for the non-standard varieties, the same holds as for SD, namely that tense vowels are heavier than lax vowels; and, with exception of 10, 11 and 12 as I will show below, the pronunciation of the different dialectal stem forms will not be researched.

³⁵ The double consonant at the end may also be a geminate.

grained research, only the accordingly relevant dialectal data is consequently investigated. I realize this is a somewhat crude approach. However, I believe that it is the most efficient way to test the hypotheses on such a varying set of data.

Table 4.7: Suffix distribution in the three patterns

	Flemish pattern	Transitional pattern	Standard Dutch pattern
2	twee-de	twee-de	twee-de
3	der-de	der-de	der-de
4	vier-de	vier-de	vier-de
5	vijf-de	vijf-de	vijf-de
6	zes-de	zes-de	zes-de
7	zeven- ste	zeven- ste	zeven-de
8	acht- ste	acht- ste	acht- ste
9	negen- ste	negen- ste	negen-de
10	tien- ste	tien-de	tien-de
11	elf- ste	elf- ste	elf-de
12	twalf- ste	twalf- ste	twalf-de
13-19	dertien- ste	dertien-de	dertien-de
20-99	twintig- ste	twintig- ste	twintig- ste
100	honderd- ste	honderd- ste	honderd- ste
1,000	duizend- ste	duizend- ste	duizend- ste
1,000,000	miljoen- ste	miljoen- ste	miljoen- ste
1,000,000,000	miljard- ste	miljard- ste	miljard- ste

Having established how to define syllable weight and having determined to regard only the standard forms initially, we can now begin testing the individual ordinal forms in terms of how heavy their presuffixal syllable is. Table (4.8) lists the ordinal stem forms of Standard Dutch. For each form, the rhyme of the rightmost syllable is given a label ‘extra light’ when its head is a schwa, ‘neutral’ when it is -V, -VV or -VC, and superheavy when it is -VVC, -VVC or -VVCC, following the syllable weights from table 4.6.

Table 4.8 Weight of the final syllable of the ordinal stem

	Cardinal	Rhyme of rightmost syllable	Weight
2	twee- /tʷe:/	-VV	neutral
3	der- /dɛr/	-VV	neutral
4	vier- /vi:r/	-VVC	superheavy
5	vijf- /vɛif/	-VVC	superheavy
6	zes- /zɛs/	-VC	neutral
7	zeven- /'ze:.vən/	-əC	extra light
8	acht- /ɑxt/	-VCC	superheavy
9	negen- /'ne:.ɣən/	-əC	extra light
10	tien- /ti:n/	-VVC	superheavy
11	elf- /ɛlf/	-VCC	superheavy
11	variant: /'ɛl.ləf/ ³⁶	-əC	extra light
12	twaalf- /tʷa:lf/	-VVCC	superheavy
12	variant: /'tʷa:.ləf/	-əC	extra light
13-19	dertien- /'dɛr.ti:n/	-VVC	superheavy
20	twintig- /'tʷɪn.tɛɣ/	-əC	extra light
100	honderd- /'hɔn.dərd/	-əCC	extra light
1,000	duizend- /'dœy.zənd/	-əCC	extra light
1,000,000	miljoen- /mil'ju:n/	-VVC	superheavy
1,000,000, 000	miljard- /mil'jard/	-VCC	superheavy

The following generalizations can be made from this table:

- The cardinals with neutral rhyme are: 2, 3, 6;
- The cardinals with superheavy rhyme are: 4, 5, 8, 10, 11 (*elf*), 12 (*twaalf*), 13-19, *miljoen*, *miljard*;
- The cardinals with extra light rhyme are: 7, 9, **disyllabic 11 (*ellef*)**, 12 (*twalef*), 20, 100, **1,000**.

³⁶ In Dutch, if the nucleus consists only of one short vowel, the onset consonant of the following syllable is shared between the two syllables: following Booij (1998: 146) I orthographically indicate ambisyllabic consonants as geminates. The variants *ellef* and *twalef* are very commonly heard in the spoken language (formal as well as informal), but never written.

We can already conclude that for the forms of Standard Dutch, labeling the rhymes of the syllables preceding the suffix does *not* give the desired division. We would expect some grouping of the numbers 2-7 and 9-19 on the one hand, and a grouping of 8 and 20-1,000,000, or perhaps 2-19 and 20-1,000,000, if 8 is formed because of alternative criteria (see 3.4 for an alternative theory for 8). The important observation here is that neither of these two groupings is found for Standard Dutch.

A second observation is that the division we find for the extra light syllables *does* show a remarkable resemblance to the distribution of suffixes in our transitional (Brabantish) zone. I have marked these forms white-on-black in table 4.8. The ordinals with an extra light presuffixal syllable are 7, 9, 11-12 (only the variants with epenthetic schwa), 20, 100, and 1,000. Let us explore the details to such an analysis and the predictions this would make.

I have written out the new hypothesis in (13):

(13) For the transitional area, a presuffixal extra light syllable dictates the selection of *-ste*.

For this hypothesis, the following conditions and predictions apply:

- (i) It is not immediately apparent what could underlie this presuffixal extra light syllable that causes the selection of *-ste*. Thus, pinpointing a presuffixal extra light syllable as the ‘cause’ of selecting *-ste* rather than *-de* does not explain or motivate that selection, it only narrows down the conditions under which *-ste* is selected;
- (ii) As we have seen above, SD ordinals do not fit this grouping. SD 7, 9, 11 and 12 do not combine with *-ste*. Thus, whatever factor it is that causes *-ste* selection in ordinals with a presuffixal extra light syllable, SD must be insensitive to it, and once it is clear what that factor is, this insensitivity must be accounted for;
- (iii) Because the transitional area seems to display a subset of the facts found in the Flemish area, I assume that whatever drives *-ste* selection should be the same factor in both areas;
- (iv) An alternative motivation is needed for transitional and Flemish *acht-ste*, *miljoenste* and *miljard-ste*, since neither 8, nor 1,000,000, nor 1,000,000,000 have a presuffixal extra light syllable;
- (v) The grouping in table 4.8 only applies to disyllabic /'ɛl.ləf/ and /'tva:.ləf/ ordinal stem forms; all Flemish and transitional data should reflect forms with an extra light syllable preceding the suffix;
- (vi) 10 is the differentiating factor between the transitional area and the Flemish pattern: it takes *-ste* in the Flemish region and *-de* in the transition zone. In the

Flemish region, this ordinal stem should be disyllabic with the second syllable a schwa-syllable, something like /ti.ən/, while the other two regions should have superheavy [ti:n].

(i) and (ii) are the most problematic in my opinion, because they pose two very relevant, yet very difficult to answer, questions; two related puzzles which I will not solve within this thesis. (iii) is a relatively straightforward statement. (iv) lies within reach, because 8 is unique in form compared to the other numbers (see 2.1) and the word forms for 1,000,000 and 1,000,000,000 are the youngest number words in the entire range (see appendix A for their etymology).

To test (v), additional research is necessary, because my survey did not record pronunciation data, and the collected written data is based on the respondents' own interpretations of the sounds. It is easily imaginable that the written answers may have been influenced by the standard spelling and thus may diverge from the *actual* pronunciation. There is an online database with pronunciation data from different regions called MAND (van den Berg, 2003; accessed online via <http://www.meertens.knaw.nl/projecten/mand>), but these unfortunately do not include ordinals (only cardinals).

As for (vi), Dutch phonologists are in agreement that orthographic -ie-, even if sometimes pronounced as a centralizing diphthong [iə], is always one vowel (Ben Hermans, p.c.). There is no phonologically sound way to put a syllable break between [i] and [ə]. Historically, the diphthong comes from long [e:] of which the second half was reduced to schwa while the first half was lifted to a close vowel [i] (Ben Hermans, p.c.; Van Bree, 1987. See also appendix A on the etymology of the number words). In some areas, the schwa was subsequently dropped – possibly explaining the (compensatory) lengthening of the close vowel – rendering the following sequence of sound changes:

(14) te:n > tiən > ti:n

Thus, having a presuffixal schwa-syllable in the ordinal stem for 10-*ste* in the Flemish region seems impossible. However, perhaps the weight of the *cardinal input's* rightmost syllable rather than that of the ordinal's presuffixal syllable can provide a cleaner analysis.

Of the points mentioned in (i)-(vi) in the previous subsection, (i), (ii), (v) and (vi) seemed the most problematic. As I said there, I will not go into (i) and (ii) at this moment, but I can reevaluate (v) and (vi) by addressing the cardinal forms in lieu of the ordinals.

To evaluate the pronunciation of 10, 11 and 12, I made use of the online database MAND (van den Berg, 2003) which was mentioned in the previous subsection. Unfortunately, it does not

contain information from all of the locations represented in my survey data; therefore, I restricted the endeavor to the locations that were represented in both sources.

In MAND I found that the cardinal forms for 11 and 12 are either pronounced with an epenthetic schwa (/ˈɛl.ləf/ and /ˈtva:l.ləf/ or similar to these), or as /ˈɛl.və/ and /ˈtva:l.və/, or even trisyllabic as /ˈɛ.lə.və/ and /ˈtva:l.lə.və/ or variations thereof, but notably **ending on a schwa-syllable across the board**. This is in favor of the hypothesis.

According to what was stated in (vi), the form for 10 should end in a schwa-syllable in the Flemish region and it should be phonetically monosyllabic in the transition zone. I found data that corroborates this prediction, with forms such as [teːɪnə], [tinə] and [tiːənə] in West and East Flanders, and forms like [tin], [teɪn] and [tiŋ] in North Brabant (SD area) and in Antwerp and Flemish Brabant (transition zone).³⁷ Were this the case for 10 only, then it would look very good indeed for the hypothesis currently under evaluation. However, in the areas of West and East Flanders, it seems to be the case that *all cardinals* or almost all are pronounced with a schwa-ending.³⁸ Which cardinals exactly differs per speaker, but there are many examples in my survey data of *viere* and *vijve* ‘4’ and ‘5’ and in fewer speakers, but nonetheless quite many, *tweje* and *drieje* ‘2’ and ‘3’. If our hypothesis would hold, then this would predict forms like *twee-ste*, *drie-ste*, *vier-ste* and *vijf-ste* but these are strikingly not found for any speaker in the survey data.

It is unfortunate for hypothesis (13) that the schwa occurs in too many cardinal forms. However, it may be the case that the schwa appears only in the citation form of each cardinal. Future work should determine whether attributive use of the cardinals can set 10-11-12 apart from the others.

³⁷ I copied the notations as they were in the online version of the MAND database. As they bear little relevance for the current point, I do not elaborate on them; more information on the notation can be found at <http://www.meertens.knaw.nl/projecten/mand/EGTRPkipatabel.html>.

³⁸ Interestingly, the data that I have from Zeeland also feature schwa endings for all cardinals. Zeeland is not as well-represented in the survey data as are the Flemish and the transitional area. This is unfortunate, because while the Dutch (North) Brabant area shows little variation – almost all cardinals and ordinals from all locations closely resemble the standard Dutch forms – Zeeland is much more interesting. Out of nine Zeeuwse respondents, two showed non-standard patterns. This may mean that there is more variation going on in that area than what has surfaced in my survey. Compared to the other areas in the Netherlands, Zeeland is geographically close to the Flemish area. This may mean that it is part of the transition zone in some respects, giving rise to the occasional Flemish-like *ste*-ordinal, and having Flemish-like schwa-cardinals (even if these Flemish-like Zeeuwse cardinals do not necessarily co-occur with Flemish-like ordinals within the same respondent).

4.2.4 Parsing syllables into feet

In the previous section about hypothesis 3, we found that the cardinals 7, 9, 11, 12, and 20, 100, 1,000 group together: their second syllable is always a schwa-syllable (extra light); the corresponding ordinal stems (minus the suffix) have the same form. The only difference between cardinals and ordinals is in *drie* - *der*- '3' and in the fact that (almost) all cardinals in Flemish speakers end in schwa while this schwa is not present in the ordinal stem.

The grouping of 7, 9, 11, 12, and 20, 100, 1,000 does not correspond to any of the three systems described in chapter 2, but it does closely resemble the Flemish and Transitional patterns. The only numbers missing would be 8 (which is explained with an alternative motivation, see 2.1) and million and milliard, which are the newest additions to the number list and thus may have been formed analogously to the other higher numbers (see appendix A). We are left with the following questions:

- (i) Why is 10 treated differently in the transition zone compared to the Flemish zone?
- (ii) Why does a presuffixal schwa syllable result in *-ste* rather than *-de*?
- (iii) Why are the ordinals in SD insensitive to the operation in (ii)?

As was mentioned before, the Dutch stress system is based on syllable weight, and therefore, it only makes sense to extend 4.2.3 into a section that examines the word stress of the numerals. One could also say that the words highlighted in table 48 above were all (phonetically) disyllabic. Kager (1996: 156) shows that Estonian 'has several morphological categories whose allomorphs depend on the number of syllables of the base'. He then shows that it is not necessarily the sheer **number of syllables** that are to blame, but rather that the suffix selection depends on whether the number of syllables of the input is odd or even: because an even number of syllables in the output can be parsed into disyllabic feet, while an odd number of syllables needs the addition of one extra syllable in order to avoid leaving one syllable out at the end: 'feet, rather than raw syllables, are the stuff that the allomorphy is computed on' (Kager, 1996: 158).

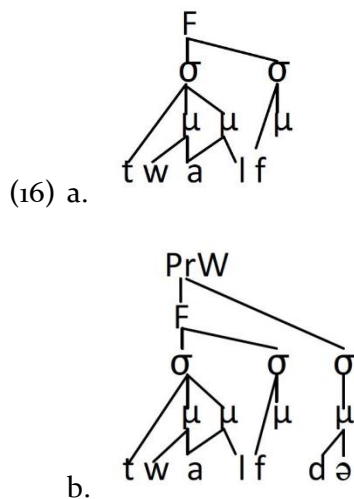
Achtste, *elfde* and *twaaifde* sound phonetically disyllabic, consisting of a suffix attached to a monosyllabic stem form, the corresponding cardinal. What distinguishes the stem forms - or cardinals - 8, 11 and 12 from the other monosyllabic cardinals, is that the former have very full codas. A syllable is normally divided in an onset and a rhyme. Following again the rule that all levels of the structure should be binary, both of these should contain at most two positions, leaving only two options for the rhyme: either there is a long/tense vowel and no coda (VV) or a short/lax vowel with one coda consonant (CV). Now, in the case of *acht* and *elf*, we have one

lax vowel and *two coda consonants*. Oostendorp (2000: 184) analyzes this as follows: the left-over consonant, not able to fit inside the coda of the syllable it phonetically seems to be in, is in fact **in the onset of a new syllable** that lacks a rhyme, making it a headless syllable, or an ‘empty’ syllable.

This can be schematized as follows, whereby $[\]_{\sigma}$ denotes which elements are inside the syllable, and $[\]_{\omega}$ denotes the scope of the prosodic word. Thus, /f/ is included in the prosodic word, but not in the syllable that includes /ɛl/.

- (15) ε l f
 [[V C]_σ C]_ω (attachment to the prosodic word)

The same structure can be drawn for /ɑxt/ and /twa:lf/ mutatis mutandis.³⁹ Given that there are now two syllables present, a strong syllable with a full vowel and then a weak one without a head, the strong plus weak syllable can form a disyllabic trochee. This foot structure is illustrated in (16a), this time for /twaalf/; and adding the ordinal suffix *-de* makes for a structure with an extrametrical suffix syllable in (16b).⁴⁰ The motivation for the foot structure in (16b) is parallel to the motivation for e.g. (ze:..vən)_F(də)_σ.



³⁹ Readers who have been sharply following the argumentation so far will object to this, because the /a:/ in ‘twaalf’ is already taking up two positions. However, as Oostendorp (2000: 33) argues, /a/ is a bit of a special vowel in the sense that it appears to be able to behave just like a lax vowel in situations like this one. The structure in (16) includes moras, and in this kind of configuration we might say that /a/ is special, because it is one of only a few tense vowels that can project into two moras *and* be followed by two more consonants. The other exceptions include a small amount of past tense forms to strong verbs: *hielp*, *zwierf*.

⁴⁰ My gratitude goes out to Marc van Oostendorp and Ben Hermans for helping me by patiently explaining phonological theory. The structures in (x) and (x) are provided by them respectively. Any errors occurring in the structures as presented here are completely mine.

Let us consider the foot structures of the ordinals of Standard Dutch, as proposed by me, in the third column of table 4.9. A subscript σ indicates a syllable that is not parsed into a foot, thus being extrametrical. Whenever this applies to the ordinal suffix, I have indicated this by marking an 'x' in the final column. The ordinals that have an extrametrical suffix are: 7, 8, 9, 11, 12, ≥ 20 .

As is common practice in present-day stress phonology, I assume that feet consist of no more than two syllables. 'Any reference to syllable parity should be reducible to grouping of syllables in binary feet (Kager, 1996; citing Halle & Vergnaud 1987; Hayes 1995).' This is in line with a general preference for binarity in phonological structures. Combining this assumption with the knowledge that Dutch prefers parsing its syllables into **disyllabic trochees** (Booij, 1998: 144) it becomes clear that we can relatively straight-forwardly analyze the disyllabic ordinals 2-6 and 10 as exactly that. The remaining forms, however, consisting of three or more syllables, give us several parsing options to choose from.

Booij (1998) explains how the choice of plural suffix in Dutch (-s or -*ən*) is determined by prosodic constraints on the output form: the strong preference of Dutch for disyllabic trochees collaborates with a number of other constraints to arrive at the most optimal foot structure. In the paper, he distinguishes between several candidate foot structures by drawing up a tableau within the framework of Optimality Theory, in which three constraints are ranked relative to each other: FOOTMAX, FOOTMIN and PARSESYLL. FOOTMAX I have already mentioned above, without calling it a constraint: feet must be no larger than two syllables maximum. FOOTMIN guards the other end of the spectrum: feet should preferably also consist of *minimally* two syllables. In fact, Dutch wants its feet to consist of *exactly* two syllables. However, FOOTMAX is ranked considerably higher than FOOTMIN: while the former is hard or even impossible to violate - Booij does not label it as inviolable, but has no examples of trisyllabic feet to show, either - the latter, FOOTMIN, may be violated in certain cases. For example, if there is a superheavy syllable but no lighter syllable to combine with, the superheavy syllable will be parsed as a monosyllabic foot. Superheavy syllables must receive stress, and a stressed syllable must always be the head of a foot - this is in line with the Strict Layer Hypothesis (Nespor & Vogel 1986: 7, Selkirk 1984).

Table 4.9: Foot structures of the SD ordinals

Number	Form	Foot structure	Suffix extrametrical
2	'tweede	(tʷe:.də) _F	
3	'derde	(dɛr.də) _F	
4	'vierde	(vi:r.də) _F	
5	'vijfde	(vɛif.də) _F	
6	'zesde	(zɛs.də) _F	
7	'zevende	(ze:.vən) _F (də) _σ	x
8	'achtste	(ɑχ.t) _F (stə) _σ	x
9	'negende	(ne:.ɣən) _F (də) _σ	x
10	'tiende	(ti:n.də) _F	
11	'elfde	(ɛl.f) _F (də) _F	x
	'ellefde	<i>variant:</i> (ɛl.ləf) _F (də) _σ ⁴¹	x
12	'twaalfde	(twal.f) _F (də) _σ	x
	'twalefde	<i>variant:</i> (tʷa:.ləf) _F (də) _σ	x
13, 14, 15, 16, 18	'dertiede	(dɛr) _F (ti:n.də) _F	
17, 19	'zeven, tiende	(ze:.vən) _F (ti:n.də) _F	
20, 30, 40, 50, 60, 80	'twintigste	(tʷɪn.təɣ) _F (stə) _σ	x
70, 90	'zeventigste	(ze:.vən) _F (təɣ) _σ (stə) _σ	x
100	'honderdste	(hɔn.dərd) _F (stə) _σ	x
1,000	'duizendste	(dœy.zənd) _F (stə) _σ	x
1,000,000	mil'joenste	(mil) _F (ju:n) _F (stə) _σ	x
1,000,000,000	mil'jardste	(mil) _F (jɑrd) _F (stə) _σ	x

'[O]utputs obey the constraint that a foot consists of two syllables. Dutch feet are preferably disyllabic trochees. However, **monosyllabic feet are also allowed if necessary, provided that the vowel of that monosyllabic foot is a full vowel: a schwa can never function as the head** of a monosyllabic foot.⁴² Since a foot is maximally disyllabic, syllables headed by schwa may be left over, i.e. they cannot be parsed into a foot. In that case, they will be dominated directly by the prosodic word node of the relevant word.'

Booij (1998: 144)

⁴¹ I orthographically indicate ambisyllabic consonants as geminates. See footnote 35.

⁴² In the tableaux below, I follow Booij's example in not considering any candidates with schwa-headed feet, because they are prohibited by an inviolable constraint (Booij, 1998: 146).

I take a top-down approach in the sense that I assume the stress exactly the way it surfaces in the output forms, and then use those observations to determine where the foot boundaries must lie. According to the Strict Layer Hypothesis (Nespor & Vogel 1986:7, Selkirk 1984) a stressed syllable must always be inside a foot; thus, any syllable that we find to be carrying stress, must be in a foot. The other factors that will help us determine the foot structure will be provided by Booij (1998) and his constraint ranking of output constraints that determine the choice of suffix in Dutch plurals.⁴³

'dɛr.tiːndə	FootMax	ParseSyll	FootMin
(dɛr.tiːn.də) _F	*!		
(dɛr.tiːn)(də) _σ		*!	
☞ (dɛr) _F (tiːn.də) _F			*
(dɛr) _F (tiːn) _F (də) _σ		*!	**

The first syllable has primary stress, and following the principles of headedness, if a syllable carries the main stress in the word, it has to be the head of the prosodic word and thus it should also be the head of the Foot (it can't skip a layer according to the Strict Layer Hypothesis: Nespor & Vogel 1986: 7, Selkirk 1984). Thus: /dɛr/ cannot be extrametrical; it has to be in a Foot. This does not appear in the tableau because it is the result of inviolable Headedness. *Honderdste* and *duizendste* will be parsed the same way as *twintigste*.

'tʊɪn.təy.stə	FootMax	ParseSyll	FootMin
(tʊɪn.təy.stə) _F	*!		
☞ (tʊɪn.təy) _F (stə) _σ		*	
(tʊɪn) _F (təy.stə) _F			
(tʊɪn) _F (təy) _σ (stə) _σ		**!	*

⁴³ 'The computation of the correct plural suffix presupposes that the stress pattern of the stem is already given, and cannot shift rightward in the plural form' (Booij, 1998: 146). Van Oostendorp, although never mentioning the ordinal suffixes /-də/ and /-stə/, groups all Dutch suffixes consisting of only coronal sounds and schwa together as 'type A suffixes' (2002: 12). This kind of suffix is always stress-neutral. Booij (1998: 146) says this stress-neutrality can be achieved either by allowing different levels in the derivation, or by using output-output correspondence constraints; but see Oostendorp (2002: 6-7) for discussion on this matter. He argues that in order to capture the phonological behavior of all the different types of derivational suffixes in Dutch, we must assume that the derivational suffixes are adjoined to the stem, which is a Prosodic Word on its own. If we were to follow this line of analysis, it would mean that none of the ordinals can form a foot with their suffix.

¹ ze:.vən.təy.stə	FootMax	ParseSyll	FootMin
(ze:.vən.təy.stə) _F	*!		
(ze:.vən.təy) _F (stə) _σ	*!	*	
☞ (ze:.vən) _F (təy) _σ (stə) _σ		**	
(ze:) _F (vən) _σ (təy) _σ (stə) _σ		***!	*

/mɪl.ju:n/ and /mɪl.jɑrd/ also cannot parse the suffix in a foot. They are of the same type as (αχ.t)_F(stə)_σ, (ɛl.f)_F(də)_F and (twal.f)_F(də)_σ: /ju:n/ and /jɑrd/ both have one consonant, the final one, that cannot be parsed tautosyllabically and therefore they are parsed into the onset of a new, otherwise empty syllable; thus forming their own disyllabic trochee; /mɪl/ is most likely parsed as a monosyllabic foot if we consider the ranking that not parsing syllables is penalized sooner than having monosyllabic feet (PARSESYLL is ranked higher than FootMin. The suffix, however, must be left unparsed: it cannot join (ju:n)_F because the phonology does not allow trisyllabic feet, and it cannot be parsed as a monosyllabic foot, because its head is a schwa, and schwa-headed syllables are banned by an inviolable constraint. Thus, the structure must be (mɪl)_F(ju:n)_F(stə)_σ.

In table 4.9, then, we see that the ordinals with an extrametrical suffix are: 7, 8, 9, 11, 12, ≥20. Recall, however, the abstract representation of *vier* ‘4’ in section 4.2.3 and the syllable weights attributed to the number words in table 4.8. According to the analysis given there for superheavy syllables, their underlying structure should be the same as that of a phonetically disyllabic word with schwa as the second syllable: in *vier* the coda consonant /r/ no longer fits in the first syllable and consequently heads a second syllable, with an empty nucleus. If we follow this analysis, it means that 4 and 5 and also 10 should behave the same way as the phonetically disyllabic ordinal stems in terms of suffix choice, which we know not to be true for 4 and 5, and for 10 should be some way to distinguish between the Flemish and the Transitional pattern. In other words, we must conclude at the end of 4.2.4 that for the Dutch, having a phonologically disyllabic stem does not in all cases mean that the presuffixal syllable is a schwa syllable.

4.3 Conclusion

I repeat once again the research questions of this thesis:

- (i) What are the suffix choices in the ordinal number lines of the different varieties of Dutch (including the standard variety)?
- (ii) How can these patterns be motivated?
- (iii) Do the suffixes share the same underlying representation?

In chapter 4, I answered question (iii) and I gave several answers to question (ii). The Transitional pattern (-*ste* for all ordinals above and including 7, except 10) sits intermediate between the Flemish pattern (-*ste* for all ordinals above and including 7) and the SD pattern (-*ste* on ordinals above and including 20), not only with respect to the system for suffix choice, but also geographically. Thus, one possible motivation for the way that the patterns relate to each other is that the SD pattern ‘moves further away’ from the Flemish pattern grammatically as the distance becomes larger.

Before looking into language-internal motivations, I examined the allomorphic status of the two suffixes. Hinted by their different historical origins and based on the lack of evidence of [d]/[st] alternations elsewhere in the Dutch language, I concluded that the suffixes are most likely suppletive allomorphs and do not share a common UR. Neither is the suffix choice phonologically conditioned by the final consonant of the ordinal stem (or of the cardinal input form).

I found one language-internal motivation: an extra-light (schwa) syllable directly preceding the suffix predicts -*ste*. This goes only for the intermediate pattern. I leave it to future work to uncover the role of 10 in this. It is possible that the Flemish pattern can be explained in a similar way, but with some (diachronic or synchronic) attribute or feature in 10 that is behind the exceptional status of 10 in the Transitional pattern.⁴⁴

Extrapolating the syllable weight to stress patterns by way of parsing the syllables into trochaic feet does not yield a desired grouping of the ordinals. Because superheavy syllables (*vier*, *vijf*) underlyingly represent two phonological syllables, they behave the same way as a phonetically disyllabic syllable (*zeven*, *negen*).

This rouses the suspicion that some number forms may be underlyingly more complex than they appear to be. For this reason, I have compiled information about the diachronic processes

⁴⁴ A suggestion from Johan Rooryck is that 10 has some kind of special feature because it is the counting base in Dutch. Researching this lies outside the scope of this thesis, but might provide a solution for the problem.

that shaped the cardinals and ordinals. Von Mengden (2010) stresses the importance of taking a diachronic perspective when attempting to give a synchronic analysis. He gives the example of Old English *endleofan* and *twelf* '11' and '12' which do not follow the 'generally systematic character' of the teens 13-19 formed through addition; in the synchronic system, they are idiosyncrasies, but they are complex from a diachronic point of view.

'Whenever idiosyncratic simple forms interfere with a regularly systematic sequence of numerals, it is possible, if not likely that compound forms are historically underlying. (...) [T]he existence of idiosyncratic simple forms does not contradict the generally systematic character of a numeral system. In spite of the explanatory force of a diachronic analysis of idiosyncratic forms, an isolated, i.e. noncomparative, analysis of one particular numeral system requires a synchronic analysis of the system and its idiosyncrasies.'

(Von Mengden, 2010: 82)

My recommendation for future research, therefore, is to look into the complexity of the cardinals and ordinals in order to find the motivations for the dialectal variation in ordinal suffix choice.

5. Summary, conclusions and recommendations

In this thesis I presented novel empirical data describing the microvariation in Flemish dialects concerning the Dutch ordinal numerals. I have also given a synchronic analysis of these microvariational data. Based on the literature - especially on Barbiers (2007) - I proposed the following analysis of the Dutch ordinal system in (1), repeated here as (17):

- (17) (a) The default morpheme for ordinal formation in Standard Dutch is *-ste*. This morpheme is applied to **all cardinals** except the lower cardinals 2-12. Some varieties in Flanders (East Flemish and West Flemish) extend the use of the *ste*-morpheme all the way down to 7.
- (b) In Standard Dutch, the lower cardinals 2-19 get a different morpheme: *-de*. In the other dialects, it is applied only to the cardinals 2-6. *Achtste* '8th' is formed with *-de* but the surface form has an inserted *-s-*; we know this on the basis of knowledge of Middle Dutch sound laws and historical forms of 8.
- (c) The ordinal for 1 is unique compared to all other ordinals: its *-ste* is a superlative suffix, which is different from the larger ordinals.

I conducted an online survey to collect novel data about the synchronic status of the ordinal suffixes. The survey shows that the distribution on the ordinal number line of the two competing ordinal suffixes in certain dialects of Flanders differs from the distribution in Standard Dutch (chapter 3). I have analyzed these distribution patterns, both in Standard Dutch as well as in the Flemish of East and West Flanders, and finally in a group of dialects in Antwerp and Flemish Brabant displaying characteristics of being a transition zone between the two previously mentioned patterns (chapter 4).

The ordinal number lines of the Flemish pattern, of the Transitional pattern and of Standard Dutch (SD) were first presented in table 3.1, repeated here for the reader's convenience as table 5.1. Recall that the dialectal forms have been abstracted to look like the SD forms; the suffix remains recognizable throughout my survey data.

Table 5.1: Ordinals of the three main patterns

	Flemish pattern (West and East Flanders)	Transitional pattern (Antwerp, Flemish Brabant)	Standard Dutch pattern (rest of Flanders and Netherlands)
1	eer-ste	eer-ste	eer-ste
2	twee-de	twee-de	twee-de
3	der-de	der-de	der-de
4	vier-de	vier-de	vier-de
5	vijf-de	vijf-de	vijf-de
6	zes-de	zes-de	zes-de
7	zeven-ste	zeven-ste	zeven-de
8	acht-ste	acht-ste	acht-ste
9	negen-ste	negen-ste	negen-de
10, 13-19	tien-ste	tien-de	tien-de
11	elf-ste	elf-ste	elf-de
12	twalf-ste	twalf-ste	twalf-de
20-99	twintig-ste	twintig-ste	twintig-ste
100	honderd-ste	honderd-ste	honderd-ste
1,000	duizend-ste	duizend-ste	duizend-ste
1,000,000	miljoen-ste	miljoen-ste	miljoen-ste
1,000,000,000	miljard-ste	miljard-ste	miljard-ste

The following conclusions were drawn in chapter 4:

The variational data can be explained as the result of geographical factors: the further away from the core region (East/West Flanders) where the use of *-ste* is least restricted, the more restrictions are added. First 10 becomes an exception (in Belgian Brabant and the Antwerp region): the Transitional pattern strongly resembles the Flemish pattern but has an exception for 10 which is formed like *tiende* in the SD pattern. Further research is needed to arrive at a better understanding of this. Possibilities include something to do with 10's status as the numeral base (Johan Rooryck, p.c., 2017) or finding a key distinguishing feature in the 10 of the

Flemish region versus the 10 of the transition zone. The SD pattern is most restricted, as *-ste* is used only for 20 and up.

The two ordinal suffixes of Dutch *-ste* and *-de* are most likely suppletive allomorphs. A phonologically conditioned suffix selection on the basis of final stem consonant fails to capture all the facts. I find a positive effect of syllable weight on suffix choice in one of the three patterns: an extra-light (schwa) syllable preceding the suffix predicts *-ste* in the Transitional pattern. Why it selects *-ste* in those cases is not clear. Extrapolating syllable weight to foot structure shows that 4 and 5 are suddenly falsely predicted to take *-ste*, thus explaining the *-ste* preference by means of syllabification cannot be done satisfactory. Perhaps it can be done with a different perspective on the underlying complexity of superheavy syllables.

Furthermore, considering the etymologies of some number words (*elf* '11' and *twaalf* '12' are untransparently but quite uncontroversially considered bimorphemic, meaning something like 'one left' and 'two left'; after subtracting the numeral base, 10, that is), it may be interesting to compare the complexity of the number words and test whether the suffix selection may be somehow connected to it.

References

- Barbiers, S. (2007). Indefinite numerals one and many and the cause of ordinal suppletion. *Lingua*, 117(5), 859-880.
- Barbiers, S. (2014). Syntactic doubling and deletion as a source of variation. *Linguistic variation in the Minimalist framework*, 197-223.
- Barbiers, S., Bennis, H., De Vogelaer, G., Devos, M., and van der Ham, M. (2005). *Syntactic Atlas of Dutch Dialects. Volume 1: Pronouns, Agreement and Dependencies*. Amsterdam: Amsterdam University Press.
- Barbiers, S., van Koppen, M., Bennis, H., & Corver, N. (2016). Microcomparative MORphosyntactic REsearch (MIMORE): Mapping partial grammars of Flemish, Brabantish and Dutch. *Lingua*, 178, 5-31. <http://doi.org/10.1016/j.lingua.2015.10.018>.
- Backstrom, L., Boldi, P., Rosa, M., Ugander, J., & Vigna, S. (2012). Four degrees of separation. *Proceedings of the 4th Annual ACM Web Science Conference*, 33-42.
- Beekes, R. S. P., & Vaan, M. A. C. de. (2011). *Comparative Indo-European linguistics: an introduction* (2nd ed). Amsterdam / Philadelphia: John Benjamins Pub. Co.
- Bennett, R., Harizanov, B., & Henderson, R. (2015). *Prosodic smothering in Macedonian and Kaqchikel*. Ms., Yale University, New Haven, CT, Stanford University, Stanford, CA, and University of Arizona, Tucson.
- Berg, B.L. van den. (2003). *Phonology & Morphology of Dutch & Frisian Dialects in 1.1 million transcriptions. Goeman-Taeldeman-Van Reenen project 1980-1995*, Meertens Instituut Electronic Publications in Linguistics 3, Amsterdam: Meertens Instituut.
- Bloemhoff, H. & Streekstra, N. *Basisboek historische taalkunde van het Nederlands: met afzonderlijke aandacht voor het Brabants, Fries, Limburgs, Nedersaksisch, Vlaams en Zeeuws*. Groningen: Uitgeverij kleine Uil.
- Booij, G. (1998). Phonological output constraints in morphology. *Phonology and morphology of the Germanic languages*, 143, 63.
- Booij, G. (2010). Constructions and lexical units: An analysis of Dutch numerals. *New impulses in word formation*, 81-100.
- Booij, G. (2012). Allomorphy and the architecture of grammar. *Phonological Explorations. Empirical, Theoretical and Diachronic Issues*. Berlin: De Gruyter, 9-24.
- Booij, G. (2016, January 26). Allomorphy. *Taalportaal*. Retrieved from <http://taalportaal.org/taalportaal/topic/pid/topic-14020545829511269>. (accessed 11 June 2016).

- Bowern, C. (2008). *Linguistic Fieldwork. A Practical Guide*. Hampshire / New York: Palgrave Macmillan.
- Braune, W. (1891). *Althochdeutsche Grammatik* (2nd ed.). Halle: Max Niemeyer Verlag.
- Braune, W. & Ebbinghaus, E. A. (1989). *Abriss der Althochdeutschen Grammatik*. Tübingen: Max Niemeyer Verlag.
- Bye, P. (2008). Allomorphy: Selection, not optimization. In: Sylvia Blaho, Patrik Bye, and Martin Krämer (eds.). *Freedom of Analysis*. Mouton de Gruyter, pp. 64–92.
- Carstairs, A. (1988). Some implications of phonologically conditioned suppletion. In: Geert Booij and Jaap van Marle (eds.). *Yearbook of Morphology*. Dordrecht: Foris. pp. 67–94.
- Chambers, J. K., & Trudgill, P. (1998). *Dialectology*. Cambridge University Press.
- Daan, J. & Blok, D.P. (1969). *Van randstad tot landrand: toelichting bij de kaart: Dialecten en naamkunde [in de Atlas van Nederland, blad X-2]*. Amsterdam: Noord-Hollandsche Uitgevers Maatschappij.
- De Vogelaer, G. (2008). (De)grammaticalisation as a source for new constructions: the case of subject doubling in Dutch.
- Dollinger, S. (2015). *The Written Questionnaire in Social Dialectology. History, theory, practice*. Amsterdam / Philadelphia: John Benjamins Publishing Company.
- Halle, M., & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In: K. Hale and S.J. Keyser, (eds.). *The View From Building 20*. Cambridge, Mass.: MIT Press, 111–176.
- Hurford, J., (1987). *Language and Number. The Emergence of a Cognitive System*. Oxford: Blackwell.
- Kager, R. (1996). On affix allomorphy and syllable counting. In: Ursula Kleinhenz (ed.), *Interfaces in Phonology*, 155–171. [Studia Grammatica 41.] Berlin: Akademie Verlag.
- Kortmann, B. (2010). 46. Areal variation in syntax. In: Auer, P. & Schmidt, J.E. (eds.). *Language and Space. An international handbook of linguistic variation. vol. I: Theories and Methods*. pp. 837–865.
- Kroonen, G. (2013). *Etymological dictionary of Proto-Germanic*. Leiden: Brill.
- McColl Millar, R. (2007). *Trask's Historical Linguistics*. London: Hodder Education.
- Nespor, M., & Vogel, I. (1986). *Prosodic phonology* (Studies in generative grammar ; 28. 790318873). Dordrecht [etc.]: Foris.
- Payne. T. (1997). *Describing morphosyntax*. Cambridge: Cambridge University Press.
- Ross, A. & Berns, J. (1992). Germanic. In: Gvozdanović, J. (Ed.). *Indo-European numerals*. Berlin / New York: Mouton de Gruyter, pp. 555–715.

- Render, W. (1805). *A complete analysis or grammar of the German language: being a philological view, both theoretical and practical, of its construction, analogies, and various properties*. London: Cox, Son, and Baylis. Accessed via <http://babel.hathitrust.org/cgi/pt?id=umn.31951t000885790;view=1up;seq=6> (14-10-2015).
- Røyneland, U. (2010). 15. Vertical convergence of linguistic varieties in a language space. In: Auer, P. & Schmidt, J.E. (eds.). *Language and Space. An international handbook of linguistic variation. vol. I: Theories and Methods*. pp. 259-275.
- Seiler, G. (2004). On three types of dialect variation and their implications for linguistic theory. Evidence from verb clusters in Swiss German dialects. In: Kortmann, B. *Dialectology Meets Typology. Dialect Grammar from a Cross-Linguistic Perspective*, 367–399. Berlin / New York: Mouton de Gruyter.
- Selkirk, E. (1984). *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, Mass.: MIT Press.
- Sleeman, P. (2010). Superlative adjectives and the licensing of non-modal infinitival subject relatives. *Adjectives. Formal analyses in syntax and semantics*, 233-264.
- Sleeman, R. (2016). Ordinals in Standard Dutch and Dialects of Dutch. In: D. Batenburg, M. Droogsmā, G. Minnema, V. Tomson & D. van der Weit (Eds.), *Proceedings of the TWIST Student Conference in Linguistics*. Leiden: Lulu.com, pp. 57-68.
- Sleeman, R. [Ruby]. (2016, March 24). Hoi! Ik zoek sprekers van dialecten uit Zuid-Holland, Noord-Brabant, en overall in Vlaanderen! Mijn vraag is of je dit vragenlijstje zou willen invullen. Het duurt 5-10 minuten. Deel dit bericht als je denkt dat je mensen in je vriendenkring hebt [Facebook status update]. Retrieved from <https://www.facebook.com/ruby.sleeman/posts/1256779251004600>.
- Sleeman, R. [Ruby]. (2016, March 30). Hoi! Ik zoek sprekers van varianten van het Fries, en sprekers van dialecten uit gebieden van Zeeland die grenzen aan Brabant of Vlaanderen! Ook dialectsprekers uit Zuid-Holland, Noord-Brabant, en overall in Vlaanderen zijn van harte uitgenodigd om te reageren! [Facebook status update]. Retrieved from <https://www.facebook.com/ruby.sleeman/posts/1263115337037658>.
- Smith, B. W. (2015). *Phonologically Conditioned Allomorphy and UR Constraints* (Doctoral dissertation).
- Stichting Nederlandse Dialecten. (2016, March 25). [Shared Sleeman, R. [Ruby]'s March 24 Facebook Status update]. [Facebook status update]. Retrieved from <https://www.facebook.com/nederlandsedialecten/posts/825292407575349>.

- Taeldeman, J. (2010). 20. Linguistic stability in a language space. In: Auer, P. & Schmidt, J.E. (eds.). *Language and Space. An international handbook of linguistic variation. vol. I: Theories and Methods*. pp. 355-375.
- Van Beers, J. (1854). *Nederduitsche Spraekleer*. Antwerpen: Drukkery der gebroeders Peeters. p. 83.
- Van Bree, C. (1987). *Historische Grammatica van het Nederlands*. Dordrecht: Floris Publications. Retrieved from http://www.dbnl.org/tekst/bree001histo2_01/.
- Van den Abeele, D. [Dirk]. (2016, March 26). [Shared the link to the questionnaire on Google Forms]. [Post in Facebook group]. Retrieved from <https://www.facebook.com/groups/belgeninhongarije/permalink/1004882769585787/>.
- Van der Hulst, H. (1984). *Syllable structure and stress in Dutch* (No. 8). Foris Pubns USA.
- Van der Sijs, N. (Ed.) (2010). *Etymologiebank*. Retrieved from <http://etymologiebank.nl/>.
- Van Loey, A. (1970). *Schönfelds Historische grammatica van het Nederlands. Klankleer. Vormleer. Woordvorming*. 8e dr. Zutphen.
- Von Mengden, F. (2010). *Cardinal numerals: Old English from a cross-linguistic perspective*. Berlin / New York: De Gruyter Mouton.
- Van Oostendorp, M. (2000). *Phonological projection: A theory of feature content and prosodic structure* (Vol. 47). Walter de Gruyter.
- Van Oostendorp, M. (2002). The Phonological and Morphological Status of the Prosodic Word Adjunct. In: Fanselow, G. & Féry, C. (Eds.). *Resolving Conflicts in Grammars: Optimality Theory in Syntax, Morphology, and Phonology*. Hamburg: Buske, 209-235. Retrieved from: <http://researchgate.net> (5-11-2016).
- Van Oostendorp, M. (2006). Topics in the phonology of Dutch. *Unpublished manuscript*, Meertens Instituut, The Netherlands.
- Van Oostendorp, M. (2012). Quantity and the Three-Syllable Window in Dutch Word Stress. In: *Language and Linguistics Compass* 6/6. 343-358, doi: 10.1002. retrieved from: <http://researchgate.net> (18-9-2016).
- Veselinova, L. (1998). Suppletion in the derivation of ordinal numerals. A case study. MITWPL 31.
- Williams, E. (1981). On the notions 'Lexically Related' and 'Head of a Word', *Linguistic Inquiry* 12, pp. 245-274.
- Winter, W. (1992). Some thoughts about Indo-European numerals. In: Gvozdanović, J. (Ed.). *Indo-European Numerals*. pp. 11-28. Berlin / New York: Mouton de Gruyter.

Zonneveld, W. (2007). Issues in Dutch devoicing: Positional faithfulness, positional markedness, and local conjunction. *AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE SERIES* 4, 286, 1.

Appendix A: Diachronic overview of Dutch cardinals and ordinals

This appendix may be used in future research into the possible morphological complexity of the Dutch cardinals. Some parts of it have been used in the main text, but are repeated here nonetheless for the sake of completeness.

The following section gives a concise etymology of the cardinals and ordinals of Dutch.⁴⁵ There are several different subsections, one for each cardinal (or a group of cardinals that are formed the same way, e.g. 13-19). Each section includes a short etymology for the Dutch cardinal, going back as far as PGm and PIE; the etymology for the corresponding Dutch ordinal generally does not go further back than the Germanic phase, because the ordinal formation is a Germanic formation, not something shared by other branches of PIE.

Because we are working with historical data, research is limited to whatever data is available. This is why one has to step outside of the language under investigation (Dutch), and also take into account data from well documented and researched closely related languages. In the case of Dutch, Gothic, Old Saxon and Old High German are examples of such languages. Gothic (Got) is the only East Germanic language available to researchers today, and although it is from a different Germanic branch, the similarities to West Germanic languages (for example in word forms) is striking. A lot of Gothic material has been found and it has been widely studied. Old Dutch (ODu), Old Saxon (OS) and Old High German (OHG) are all members of the group of continental Western Germanic languages, and they too have been studied extensively.⁴⁶

⁴⁵ '1' and '1st' are not included, because the ordinal is not based on the cardinal.

⁴⁶ This section is based on a closed set of works and to improve readability, I refer to the relevant works by abbreviations of the authors. All references in the paragraph below to 'Br' refer to Braune (1891); 'BV' refers to Beekes & de Vaan (2011: 237-240); 'Mn' refers to Von Mengden (2010); references to 'Ph' refer to Philippa et al. (2003-2009, as cited in Sijs, 2010). 'L' refers to van Loey (1970: 153-5). 'RB' is Ross & Berns (1992); 'Vr' refers to de Vries (1971 as cited in Sijs, 2010); 'Wn' refers to Winter (1992: 12-17). Ph and Vr are both frequently cited in the online accessible *Etymologiebank*, edited by Nicoline van der Sijs (2010).

The following abbreviations are used:

Table A

ordinal	Dutch (AN)	Gothic ⁴⁷	OHG
1 st	eerste	<i>fruma</i>	êristo
2 nd	tweede	<i>anþar</i>	<i>ander</i>
3 rd	derde	þridja	dritto
4 th	vierde		feordo, fiordo
5 th	vijfde	fimfta	fimfto, finfto
6 th	zesde	saihsta	sêhsto
7 th	zevende		sibunto
8 th	achtste	ahtuda	ahtodo
9 th	negende	niunda	niunto
10 th	tiende	taihunda	zêhanto
11 th	elfde		einlifto
12 th	twaaifde		zwelifto
15 th	vijftiende	fimftataihunda	finftazêhanto > funfzêndo
20 th	twintigste		zweinzugôsto
100 th	honderdste		zêhanzugôsto

***Twee* ‘2’**

ODu has different forms for different genders and cases; in MDu the neuter form *twee* is mostly used (Ph). *Tweede* ‘2nd’ is a relatively new form - the older form was the suppletive *ander* ‘other’ as can be seen from the forms in Got. and OHG in table 3.2. *Tweede* was formed analogously to *derde*, *vierde* ‘3rd’, ‘4th’ according to Vr; *derde* and *vierde* with *-de* must have been older than *tweede*. The form *tweeste* was made after *eerste* 1st (Vr). According to Ph, the PIE word **duoh₂*, **duo-ih₂* was inflected as a dual, which can be an indication of a complex structure, the specifics of which cannot always be retraced, as is the case here.

***Drie* ‘3’**

Got: Gothic; HG: High German; MDu: Middle Dutch; ModDu: Modern Dutch; ModEn: Modern English; ModGe: Modern German; ModFr: Modern Frisian; ModHG: Modern High German; ModSw: Modern Swedish; ODu: Old Dutch; OE: Old English; OFr: Old Frisian; OHG: Old High German; ON: Old Norse; OS: Old Saxon; OSw: Old Swedish; OWGm: Old West Germanic; PGM: Proto Germanic; PIE: Proto Indo-European.

⁴⁷ The forms that are missing from this table have not been attested according to BV and Braune (1891).

Like ‘2’, there were several ODu forms, due to inflection for gender and case. Analogously to the demonstrative *die* the feminine form *drie* became the most frequent variant (Ph). Ordinal *derde* was formed through metathesis from *drid(d)e*, originally from the PIE cardinal **trei-* (Ph). See Got. and OHG cognate forms *þridja* and *dritto* in table 2.3. Ph also mentions variants *darde*, *dirde*, *dorde*, *dridde*, *drudde*; searches in WNT, MNW and Gysseling have not yielded any results for a *-ste* form (no *drieste*, *derste*, *dridste* or any variants thereof, to my knowledge). None of the works cited here mention any complexity of *drie* or older (PGm/PIE) forms.

Vier ‘4’

ODu *vier*, MDu *vier*. (Ph). *Vierde* ‘4th’ was already present in MDu; MNW also mentions the existence of MDu *vierste*. In table 2.3 we have OHG *fiordo*. Ph mentions Got. *fidwōr*, OS *fiuwar*, *fiwar*, *fior*, OHG *fior*. Pgm **fedwor* comes from an older form **petuōr*, which comes from PIE **k^wétuōr*. L and Ph analyze **p-* < **k^w-* as due to analogy to the form for five (see below). Wn notes that **k^wétuōr* has ‘the **-e.. o-* vowel pattern of a compound’, but that no surviving stems or roots can identify the elements of said potential compound, leaving the formal complexity unmotivated. Kroonen (2013: 133) suggests it was ‘an old collective’ but leaves matters at that.

The ordinal *vierde* was formed by adding the ordinal suffix to the cardinal *vier*; it already occurs in MDu in this form (Vr); alongside a form *vierste* (MNW). There is no older information available, except that the OHG form was *fiordo* with *-d-*; NHG has the form with *-t-*: *vierte* (Vr).

Vijf ‘5’

From Pgm **fimf*, forms like Odu *vinf* and OHG *finf* were formed; then all North Sea Germanic languages (such as Dutch, OHG, OS) lengthened the *i* > *ī* while dropping the *-n-* (Ph). MDu *vijfte* had cognates in Got. *fimfta* and OHG *fimto* (Vr) and changed (at least the spelling) of its suffix to *vijfde* (L), analogously to other ordinals on *-de* like *vierde* and *negende* ‘4th’ and ‘9th’ which have *-d-* suffixes in Got and OHG. Vr mentions *vichte* as the result of MDu sound laws. *Vijfste* and *vichste* are found in the corpus Gysseling (In Brabant, probably Belgium) and West Flanders.

PIE **pénk^we* should have led to **femhwe* if one assumes it followed the common PIE > PGm sound laws; the second *f* was formed most likely after its first *f* (Ph). Wn notes that ‘**pénk^w* might then originally have referred to the hand clenched to form a fist’; this would mean it was a non-complex form.

Zes '6'

ODu *sehs*, MDu *ses*, cognates Got. *saihs*, OHG *sehs* (Ph). Like *vijfte*, MDu form of '6th' was *seste* with voiceless *-te*; cf. the Got. and OHG forms *saihsta* and *sēhsto*. *-de* in the orthography is relatively new. (Pronunciation is still voiceless in speakers of SD today.) It is possible, and hardly, if at all, determinable, that like the surrounding cardinals *vier*, *vijf* and *zeven*, *ses* could combine both with *-te* and with the superlative *-ste*. However, this latter form would be indistinguishable from the former (unless the double *-s-* would return a geminate). I do not have the means to clear this up.

PGm **sehs-* from PIE **(s)uéks*; Wn quotes Szemerényi (1960, as cited in Wn: 14) who posits an underlying **weks* which may be connected to PIE **(H)weks-* 'to grow'; the **s-* in **suéks* may have originated from **séptm* '7'.⁴⁸

Zeven '7'

ODu *sivun*, *sivon*; MDu *seven* (Ph). Got and OHG cognates are *sibun*. Ordinal in OS *sibondo*, *sibotho*, OHG *sibunto* (Ph); Odu *sivondo* (ONW); in Mdu this became *sēvende* (Ph). None of the sources consulted offer remarks on *t > d* in the ordinal form; I assume it to be assimilation to the sonorant *n*.

A variant of the ordinal with *-ste* occurs in one West Flemish author in the corpus Gysseling; formed after *eerste* and *twintigste* according to the WNT. The PIE cardinal was **séptm*; the deletion of the *-t-* in the Germanic languages is unaccounted for. L proposes that it is due to a dissimilation process in the ordinal form with *-to*, but this seems unlikely; one would expect forms like *sibtun*, where the *-to* suffix would not be adjacent to the stem's *-t-*. Wn (p. 13) has difficulties with PIE '7': 'it seems impossible to find a connection of **septm* with any other element of the Proto-Indo-European lexicon; 'seven' just means 'seven' and nothing else. This does not alter the fact that **septm* is a strange form; it remains tempting to identify final **m* with the accusative ending of consonant stems.' He refuses to conclude anything, but adds to his ponderings a theory of borrowing from the literature (citing Möller 1909: 124): 'the possibility of its being a very old loanword from a Semitic language cannot be ruled out entirely' (p. 17).

⁴⁸ BV agrees on the influence from '7' and adds that the **s* is not present in the PIE ordinal **uksó-* '6th'.

Acht ‘8’

ODu *ahto*, MDu *acht* (Ph). According to MNW, *achte* also existed, both as a cardinal and as an ordinal. The ordinal was supposedly formed through reduction from *achtede* (Vr), related to Got. *ahtu-da* and OHG *ahto-do* (see table 3.2). Other ordinal forms in MDu include: *achtende* (with inserted *-n-*, probably analogy from the adjacent forms ‘7th’ (ODu *sivondo*) and ‘9th’ (cf. Got. *niunda*) (Vr); *achtenste* (formed after *zevenste* ‘7th’ and *negenste* ‘9th’, which occurred alongside their *-de* counterparts in MDu (in Flanders) and may have been formed after *eerste* and *twintigste* (Vr)). Thus, in MDu, we find **four** competing ordinal forms so far: *achte*, *achtede*, *achtende* and *achtenste*⁴⁹; of these, if *achtede* comes from *ahtuda*, it is the oldest; *achte* is a reduction thereof; *achtende* was formed after the older forms *zevende* and *negende* and *achtenste* is the newest, formed after *zevenste* and *negenste* which were newer than their *-de* counterparts.

In the (late) MDu era, a fifth form must have arisen: ModDu *achtste*. WNT argues that it occurred in MDu, but was less frequently used than *achte* and *achtende*; according to Vr, it did not occur until the ModDu era.⁵⁰ Although I cannot find a source that argues for its origins, I propose that *achtste* evolved from *achtede* due to a sound change called the Middle Dutch schwa-syncope⁵¹: *achtede* > *achtde*; an *-s-* was inserted to relieve the phonotactically unpleasant combination of two consonants pronounced in the same place of articulation.⁵² The final step is to assimilate *d* > *t* in voice: *achtsde* > *achtste*.⁵³

As for the possible complexity in the cardinal, Ph cites Szemerényi (1960:173, as cited in Sijs, 2010), the PIE form **h₃ekteh₃(u)-* was a dual form. The root **ok-* would mean ‘pointy’ and a word **ok-tom* ‘the four fingers’ was reconstructed, which in its dual form **oktō-* would mean twice four fingers - but Ph explicitly expresses uncertainty considering this analysis. Wn adds that Avestic *ašti-* means ‘width of four fingers’, but notes that there is ‘no connection whatsoever with PIE **k^wetwores* ‘four’. He adds that if PIE ‘nine’ could be shown to be connected to ‘new’, there might be a chance that PIE had a quaternary system - so that ‘nine’ marks the new set of numbers after base 8 is completed. However, he does not consider that to

⁴⁹ In the VMNW we find two additional forms, which are said most likely to have been errors in the texts: *achtechste* and *achstende*.

⁵⁰ WNT says nothing about the origins of *achtste* except that it **differed from the origin of *achte/achtende***. I do not hold this to be true.

⁵¹ This sound change is described in Bloemhoff & Streekstra, 2015: 147-8; it occurred mainly in the domain of verbs, resulting in such changes as *du wonedes* > *du woendes* ‘you dwell’ and *ghi makedet* > *ghi maectet* ‘you make’ in the weak declension paradigms of the preterite; and in participials, for example *ghewonet* > *ghewoent* ‘dwelled’.

⁵² This is also known in the field of phonology as a violation of the Obligatory Contour Principle.

⁵³ A word of gratitude is in order to Peter Alexander Kerkhof, who introduced me to the possibility of this analysis (pers. comm., April 2016).

be adequately proven, and the other attempts at proving the quaternary hypothesis are not strong, in his view.

Negen '9th

ODu *nigun*, but OHD and Got. cognates *niun*. The *-g-* is analysed to come from Pgm **newun* with *w > g* (Ph). No ODu ordinal was present in the ONW or in any of the works in Sijs (2010); OHG *niunto* and Got *niunda* are; and an OFr form *niugunda* which already displays the *-g-* (Ph). I assume voicing assimilation *t > d /_n* as in the case of *sivondo*; see the paragraph on '7' above. MDu ordinal was *neg(h)ende* (Ph); a *-ste* variant *neg(h)en(t)ste* has also been attested.⁵⁴ See also the paragraphs for '7th' and '8th'.

PIE **(h₁)néun-*. Wn is highly critical of a speculative analysis of PIE 9 being related to newness, saying it is not proven (related to the hypothesis for a quaternary system in PIE; see also the paragraph for '7' above), thus a complex formation in this numeral is not supported in the field of Indo-Europeanism. Considering typological evidence, however, it seems that subtraction-based numeral formation occurs frequently just below 10 in a base-ten system. Sjef Barbiers (pers. comm., July 2016) suggests that the forms *ni-un* and *nigun* might be analysed as NEG-1, with insertion of an intervocalic *-g-* as in MDu *ne geen* for 'not one'. This idea is further investigated in the remainder of this chapter.

Tien '10'

Cognates are OS *tehan*, OHG *zehan* and Got. *taihun* (Ph); L (1970: 31) analyses *tien* to have developed from an older Germanic variant *zēhan* by deletion of the *h*: in the case of two different vowels, they would collapse into each other, creating first a diphthong, which then would morph into long *ī*: *ěhă > ea > ia > i ə > ī*, e.g. *zien* from **sēhan*, Got. *sailvan*.⁵⁵ This analysis through a diphthong stage could point in the direction of disyllabism in some dialects; in itself, however, not enough to warrant bimorphemity, of course. As for '9', Sjef Barbiers (pers. comm., July 2016) suggests that the words from OS, OHG, OE and Got period may be interpreted as polymorphemic: the *-an* is reminiscent of the form for '1': OE *an* (Mn p. 75), thus leading to the suggestion that the word may be seen as something like '1x10'. This, too, will be further discussed in the analysis below.

The ordinal is not discussed by Ph; if we look at the forms from OHG and Got., *zēhanto* and *taihunda* respectively, we see a similar *t/d* alternation as we did for 7th and 9th. A variant *tienste*

⁵⁴ In the corpus Gysseling it occurs in a text from Oudenaarde (East Flanders), and in VMNW it is mentioned as exclusively occurring in East Flanders and Eastern Brabant.

⁵⁵ The form *zēhan* is never mentioned by L; I took it from Br.

occurs in the VMNW and in the corpus Gysseling (only in Van Maerlant); L describes it as having been formed after *eerste* and *twintigste* - in the same way as *zevenste*, *achtenste* and *negenste*.

Wn mentions that for PIE **dékmt* 'ten', an interpretation 'two hands' has been suggested; he argues that this sounds persuasive as far as the meaning is concerned, but has some concerns with the form **de-* and with finding a dual form of 'hand' instead of a more likely scenario where 'two' would be expressed by a numeral.

***Elf, twaalf* '11'-'12'**

There is no ODu form for '11' in Sijs (2010), but Ph mentions among others the following cognates: OS *ellevan*, OHG *einlif*, Got. *ainlibim* from PGm **ainlif*. Ph has a reconstructed ODu form **twelif* '12' and cognates OS *twelif*, OHG *zwelif*, Got. *twalif* from PGm **twalif*.

In MDu we find *elf*, the ModDu form, which must be a reduced form of an originally disyllabic word; Ph also mentions a disyllabic MDu form *ellef*.⁵⁶ MDu '12' forms include *twelf*, *twelef*, *tweelf*, *twalef*, *twaelf* among others.

Ph analyses the PGm forms to be compounds of the cardinal for 1 or 2 and PGm. **lib(i)-* < PIE **leik^w* 'to leave', thus: 'one left' or 'two left; or alternatively from PIE **lip*, from *bileiban*, then meaning something like 'one remains' and 'two remains'. These constructions point to a decimal system, because the meaning of 11 and 12 is relative to 10; however, because 13-19 are *not* formed this way, they might also indicate the existence of an old duodecimal system. This potential duodecimal system is also visible in the formation of the decads above 60; see below.

The MDu ordinals are given in VMNW as *elfde* and *elfte*, *twelefte*, *twelefde* and *twaalfde*. Br gives OHG *einlifto*. How the suffix got voiced to *-de* is unclear at this point and is not discussed in any of the works cited. See the section on stem-final consonants in 3.2.1 above. There were also *-ste* ordinals in MDu, reserved to West Flanders and Limburg (VMNW): *elfste* and *twelefste*. VMNW speculates about an analogy to *seste* '6th' and *twintichste* '20th'.

***Dertien, veertien, vijftien - negentien* '13, 14, 15 ... 19'**

13-19 are compound numerals; they are formed through addition: the cardinal '10' is added to a cardinal between 3 and 9 to its left, e.g. *vijf-tien* 'five-ten'.⁵⁷ Neither Ph nor ONW/WNT give

⁵⁶ This form is still heard today and is reminiscent of such pronunciations as *marrek* < *markt* 'market' and *zellef* < *zelf* 'self'. From a phonetic point of view, inserting a schwa into complex coda creates a full new (light) syllable, with a simple coda; and furthermore, a monosyllabic word becomes a disyllabic trochee, which is the 'preferred pattern of organization of syllables into higher prosodic units', as Booij (1998: 146) puts it. See also subsection 3.2.3 and 3.2.4 on syllable weight and foot structure above.

any forms from ODu. We find the respective Got. and OHG forms *fimfta-taihun* and *finftazëhan* ‘15’ in Br. He shows that in OHG the ordinals were originally formed as compounds of two ordinals: 5th+10th: *finftazëhanto* ‘15th’.⁵⁸ However, in the late 10th century, the lower number is no longer an ordinal, but is added in its cardinal form to the ordinal 10th (which incidentally has also lost its *-h-* by this time): *funfzêndo* ‘15th’. In VMNL we find MDu forms *dertiende* as well as *dertienste*; the latter only in West Flanders.

Twintig, dertig, veertig - negentig ‘20, 30, 40 - 90’

The present-day forms are formed by multiplication: a suffix ‘-tig’ representing 10 is added to the cardinal form of a number between 2 and 9.⁵⁹ While for 13-19, there was no mention of ODu forms in the *Etymologiebank* nor in the *GTB*, for 20-90 the former explicitly states that formations with the suffix ‘-tig’ in ODu era have only been found in translations from OHG or in other contexts where they may have been influenced by OHG (Ph). The following forms from ODu are listed: *seszogh* ‘60’, *aghtzhogh* ‘80’ (1100 CE) and *nigenzich*, *nigonzog* (1151-1200). While these *-zhogh* and *-zich* suffixes contain an affricat [tʃ], the Dutch *-tig* suffix doesn’t arise until 1236 MDu *tuintech*, *dertech*, *viftech* ‘20, 30, 50’ (VMNW as cited in Ph).

The suffix, says Ph, comes from a ProtoGermanic noun **tigu* meaning ‘decad’ or ‘an amount of ten’; this PGm noun is the origin of the 10-multiplication suffix in many Germanic daughter languages: OS. *-tig* (mnd. *-tich*); OHG. *-zug*, *-zog*, *-zeg*, *-zig* (NHG. *-zig*); OE. *-tig* (NE. *-ty*), OFr. *-tich* (NFr. *-tich*); ON. *-tigr* (NSw. *-tio*). PGm. **tigu-* < older **tegu-* comes from PIE **dekú-* and the cardinal **dékmt* ‘10’. BV: PIE ‘twenty’ was **dui-dkmt-*, ‘two ten(s)’; all other decads after 20 are formed with the cardinal plus **-dkomt-h₂*.

Recall the observation that 11 and 12 were formed through a different principle than 13-19, pointing to the possibility of a duodecimal system in an older stage of Germanic. In the formation of the decads in several Germanic languages we find a similar observation, namely that the regular formation of decads stops at 60; the decads higher than 60 (70, 80, 90 and 100, still formed as 10-SUFF at that time) are formed after a different rule. The forms in Gothic,

⁵⁷ In the case of 13 and 14, there are small changes in the cardinal root: *der-tien* and *veer-tien*, not **drie-tien* and **vier-tien*. The former has undergone metathesis; the latter a vowel change originally from Eastern MDu, only found in a few other Dutch words. These changes are not of relevance to the current topic, which has to do with the complexity of the cardinal root; this is not affected by these small phonological changes.

⁵⁸ When the compound ordinal was used adjectivally, only the rightmost part, ‘10th’, agreed with the noun.

⁵⁹ Similarly to the cardinals 13-14 (see footnote 55 above), the cardinal roots of 20, 30 and 40 have undergone some slight changes: *twin-tig*, *der-tig*, *veer-tig*, not **twee-tig*, **drie-tig*, **vier-tig*. These three changes do not share the same cause; but they are not relevant to the topic at hand. In the case of 80 there is also a change; this will be revisited below.

OWGm, OS and OE are given in table 3.3.⁶⁰ The suffixes *-zug*, *-zo*, and *-tig*, as well as the Got. syntagm *tigjus*, are descendent from the aforementioned PGm noun **tigu*, originally from PIE **dékmt* ‘10’; the suffix *-tehund* and the prefixes *-ant* and *-hund* all derive from a different PGm word **hunda*, descended from PIE **kmtóm* which is a short form of **(d)kmtóm dékmt* ‘10th 10’ (Ph) - thus, this latter group (*-tehund*, *hund-*, *ant-*) comes, through a different route, from the same PIE word ‘10’.

Table 3.3

	Got.	OHG ⁶¹	OS ⁶²	OE ⁶³
20	<i>twai tigjus</i>	<i>zwein-zug</i>	<i>twēn-tig</i>	<i>twen-tig</i>
30	<i>þreis tigjus</i>	<i>drî-zug</i>	<i>thrî-tig</i>	<i>þri-tig</i>
40	<i>fidwor tigjus</i>	<i>fior-zug</i>	<i>fiwar-tig</i>	<i>feower-tig</i>
50	<i>fimf tigjus</i>	<i>finf-zug</i>	<i>fiþ-tig</i>	<i>fif-tig</i>
60	<i>saihs tigjus</i>	<i>sēhs-zug</i>		<i>syx-tig</i>
70	<i>sibun-tehund</i>	<i>sibun-zo</i>	<i>ant-sibun-da</i>	<i>hund-seofon-tig</i>
80	<i>ahtau-tehund</i>	<i>ahto-zo</i>	<i>ant-ahto-da</i>	<i>hund-eahta-tig</i>
90	<i>niun-tehund</i>	<i>*niun-zo</i>	<i>(*ant-)nigon-da</i>	<i>hund-nigon-tig</i>
100	<i>taihun-tehund</i>	<i>zēhan-zo</i>		<i>hund-teon-tig</i>

The OS prefixed forms, according to L, are visible in a reduced form *t-* in MDu forms *tseventich*, *tachtich*, *tnegentich* ‘70, 80, 90’ and has spread to *tsestich* ‘60’. In ModDu, the form *tachtig*⁶⁴ is the only surviving form (and *tseventig*, *tnegentig* in some dialects), but a phonetic relic can still be found in some dialects, where *zestig*, *zeventig* are pronounced with voiceless [s], assimilated to the [t] from original [tsɛstəχ], as opposed to [zɛstin] ‘16’. A similar distinction is found for [fe:rtəχ, fɛiftəχ] and [ve:rtin, vɛiftin] ‘40, 50’ and ‘14, 15’ (Schröder, 1980 as cited in

⁶⁰ Got. forms are from BV; OHG and OS forms from Braune & Ebenfalls (1989); OE forms from Mn.

⁶¹ For OHG, according to Br (p. 199), the divergent forms 70-100 with *-zo* started following the *-zug* formation of 20-60 in the 9th century.

⁶² Br&Eb report that forms like *siðuntig* ‘etcetera’ existed alongside the divergent forms for 70-90. 60 and 100 are not mentioned.

⁶³ Mn (p. 88-90) describes the formation of 70-100 (and also 110, 120 as will be mentioned below) as formation using a circumfix *hund-__-tig*; he argues for the status of circumfix because if seen as a prefix and a suffix, they cannot be analysed to both add any lexical (numerical) meaning or grammatical function: if *-tig* is analysed to mean ‘x 10’, which it is if we look at 20-60, then there is no meaning left for the prefix; whether Mn is right to assume this, however, is outside the scope of the present thesis.

⁶⁴ The standardized form of the suffix is now spelled *-tig*. However, In the MAND/FAND/GTRP-database [reference], pronunciation of *zeventig* ‘70’ in nearly all 617 locations is transcribed as voiceless [x, χ, or ç]; undoubtedly to do with Dutch phonology, including such sound changes as final devoicing and devoicing of fricatives in general [reference]. Booij (2010: 12) gives the following transcription of *zestig* ‘60’: [sɛs.təχ], while elsewhere in the article (p. 7) he gives the phonological form of the suffix as /-təχ/.

Sijs, 2010); but, although both Schröder and L claim that the voiceless [f] onset in 40 and 50 have spread from 60 and 70, looking at the voiceless forms for 4 and 5 in many Germanic languages (see table 3.2) I would deem it more plausible that *veertig* and *vijftig* have been voiceless from the start but received a voiced *v-* in orthography for reasons unknown (such as standardisation reasons, following analogies from some other element in the language).

Apart from the fact that compound numerals formed by addition start at 13 (thus after 1-12), and that decades have a different formation pattern from 70 onward (thus after 60) there are other indications of an old duodecimal system in Germanic: OE formed not only 70-100 as *hund-CARDINAL-tig*, but formed 110 and 120 in the same way: *hund-endlef-tig* and *hund-twelf-tig* with *endlef* and *twelf* ‘11’ and ‘12’.⁶⁵

Additionally, forms for *hundred* in older Germanic languages would often mean ‘a lot’ before it was standardized to the current amount of 10 times 10, and would in some cases refer to the amount of ‘120’ (like the relics in NHG *Grosshundert* and English *long hundred*) (L). NE *hundredweight* refers to ‘112 pounds’ (McColl Millar, 2007: 284); and there is a ModDu word *schok* ‘60’.⁶⁶

As for the ordinal formations of the decades in Dutch and older Germanic language stages, I find no other mention in the literature than a formation from the originally superlative ending *-ôsto* (RB: 622). ModDu ordinals of decades are formed by adding the suffix *-ste* on the righthand side of the suffix *-tig* (Booij, 2010: 13); for example *dertig-ste* ‘30th’. No results were returned in searches for a formation with the other suffix: *twintig-de*; neither in the corpus Gysseling, nor the GTB, nor in the survey results novel to this study.

***Honderd* ‘100’**

The origins of *honderd* ‘100’ have been mentioned above: before this form came in use, OE, OS, OHG and Got all had compounds of the form ‘ten-ty’, formed in the same way as the decades. These were gradually replaced by this synchronically non-complex form; however, its origin is a complex cardinal formed by the ordinal 10 plus the cardinal 10: **dékmt-óm dékmt* ‘10-th 10’ (BV & Ph).

Thus, from this PIE root for 10 **dékmt*, many different forms have formed, even in the Germanic branch alone:

⁶⁵ One could bring the words for ‘dozen’, such as Dutch *dozijn* and also *gros* and German *Gros*, which also mean ‘dozen’, into this list of indicators of an old duodecimal system; however, these words were all borrowed from French (*grosse douzain* meaning ‘a large 12’ (Sijs 2006 as cited in Sijs 2010)); needless to say, native Germanic words and numeral formation are more indicative of a native Germanic duodecimal system than French borrowings.

⁶⁶ All of these are turning archaic (Van der Sijs 2006 as cited in Van der Sijs 2010).

- forms like *tehan* / *zehan*, developing into Ger *zehn*, Eng *ten*, Dut *tien* etc;
- the PGm noun **tigu*, leading to numerous kinds of multiplication suffixes and syntagms like Ger *-zig*, Eng *-ty*, and Dut *-tig*;
- via PIE **kmtóm*, literally only the ordinal ‘10th’ but analysed as a reduced form representing the entire cardinal **dkmt-óm dékmt* ‘10-th 10’, the PGm word **hunda* was born, in its turn yielding multiplication affixes and syntagms like Got *-tehund*, OS *ant-* and OE *hund-*; and words like Ger *Hundert*, Eng *hundred*, Dut *honderd* ‘100’ (< PGm. **hunda-rada-*, from **hunda-* ‘100, 120’ and **rada-* ‘amount/number’ (Ph)); additionally, **hunda* played a role in the formation of *duizend* ‘1,000’; see the next subsection for this.

As for the ordinal, RB have this to say (p. 642): ‘the ordinals for ‘hundredth’ and higher numeral units are recent, and, therefore, in large part standard. Where the ordinal does not exist, the corresponding cardinal is often used for it (e. g. in Middle English). In East Norse the cardinals are in part inflected adjectivally to form the new ordinals.’

They give the forms *hondertichste* (MDu) and a cardinal variant *hondertich* - most likely formed after the regular decad formation on *-tich*: *tseventich*, *tachtich*, *tnegentich* (see above). VMNW also gives *hunderste*, just the *hundert* formation from **hundarada*, followed by deletion of the *-t-*; this form is reflected today in ModDu *honderdste*; and among the dialectal survey data, we also find that *hunderste* form, often as *honderste*. No forms with *-de* have been found in the GTB or Gysseling, or in the contemporary dialectal forms.

***Duizend* ‘1,000’**

According to Ph and L, *duizend* comes from PGm **būs-hund-i* ‘strong/great hundred’, with the same form *hund* as in the section above, originally from PIE **kmtóm*. Thus, this form was originally complex. It does not originate in the PIE root for ‘1,000’, given by BV to be **g^hes-l-*; the Germanic forms are said to be ‘related to the Balto-Slavic ones’, but this is not elaborated upon.

RB (p. 642) give MDu *dûsentich* and *dûsentichste* with *-tich* formed after the decads, as for ‘100’ above; in this case, too, VMNW reports *dusenste*, again without the *-t-*. No forms with *-de* have been found in the GTB or Gysseling; some instances of *duizen-de* have been found among the data of the survey, but there are too little cases in the current study to be able to rule out errors or effects of idiosyncrasy.

Miljoen, miljard ‘1,000,000; 1,000,000,000’

The words *miljoen*, *miljard* are very young and were borrowed from French. *Million* was formed with an augmenting suffix *-on* from *mille* ‘1,000’, thus meaning ‘a large thousand’; the word was not attested in Dutch until 1510 (Ph), which is normally regarded as the late MDu era [reference].⁶⁷

Miljard is attested in the Early ModDu era, 1578,⁶⁸ as *milliart* (Ph); from French *miliart*; this likely explains the dialectal variation found synchronically between [ˌmɪlˈjɑrt] and [ˌmɪlˈjɑ:r], the latter of which mirrors the French contemporary pronunciation of *milliard*.

Similarly to the data for *duizend*, *miljoen* and *miljard* sporadically show speakers who form a *-de* ordinal from it; however, again, this cannot be ruled out as erratic data. As for the historical development of the ordinals on these three forms, I follow RB in their stance - which was quoted above in the subsection for ‘100’ - in saying that these higher cardinals are younger than the lower ones, and that therefore their ordinal suffix patterns are likely to have been even more susceptible to influence between the forms and analogous formations, than is already normally the case on the numberline.

Summary

I will close off the etymology subsection by giving a summary of the historic development of the cardinals. The ordinals’ history was summarized in section 1.2 above. Cardinals that I believe show some indication of possible morphological complexity are marked in boldface.

The cardinal ***twee*** ‘2’ has a possibly complex PIE ancestor, because it looks like a dual. There are no indications of a complex form in the history of *drie* ‘3’. ***Vier*** ‘4’, like *twee*, looks like a dual (and even more so, it is a bigger form, indicative of a compound), but again, no real motivation can be given. *Vijf* ‘5’ in PIE may refer to a hand, a clenched fist (five fingers); thus there is no complexity here. *Zes* ‘6’ may have been connected to a PIE verb ‘to grow’, making it semantically dependent on ‘5’, but formwise there is only one morpheme. ***Zeven*** ‘7’ seems possibly complex, but of which morphemes is unknown; it seems to be a Semitic loan. ***Acht*** looks like a dual, twice four fingers. ***Negen*** does not seem complex in the PIE era, but its ODu form may possibly be interpreted as complex. The same goes for ***tien***. ***Elf*** and ***twalf*** are quite possibly polymorphemic; they are non-controversially described in the literature as being formed from ‘1 left’ and ‘2 left’ (relative to the base, ten). ***Dertien*** through ***negentien*** are

⁶⁷ Some of the dialectal data give the form *biljoen*, akin to Eng *billion*, for SDu *miljard*; *biljoen*, *billion* and other forms like *triljoen* are backformations from *million* (Payne, 1997: 66).

⁶⁸ Sijs (2001 as cited in Sijs 2010) argues that it was not regularly used until after the French-German 1870-71 war.

compound numerals formed through addition; *twintig* through *negentig* are compound numerals formed through multiplication. *Honderd* and *duizend* were formed as complex words in an older Germanic stage (PGm **hunda-rada* and **þūs-hund-i*). *Miljoen* and *miljard* are loanwords from French; in French, they were formed by suffixation to the stem *mille* ‘1,000’.

There is a possibility of PIE having had a quaternary system, with ‘8’ being ‘2x4’ and ‘9’ meaning ‘new’. What is even more promising, is the indications of an old duodecimal system in Germanic: this can be seen in the divergent forms for 11 and 12, several old daughter words representing amounts of ‘60’, ‘112’ or ‘120’; the fact that in the sister languages of ODu (OE, OHG, OS etc) the decads higher than 60 have a different formation than those below and including 60; and the fact that of these, at least the OE formation carries from 70 all the way to 120.

As for the ordinals: *tweede* is a young formation; there was a suppletive *ander*. *Tweeste* has also been found. *Derde* is a ‘native’ ordinal in the sense that it is relatively old and it is formed with the original Germanic ordinal suffix. There is metathesis in the stem, diminishing the transparency of the formrelation between the ordinal stem and the corresponding cardinal. No forms with the superlative ordinal *-ste* have been attested. *Vierde* is a native ordinal, even more so than *derde* because the cardinal is transparently visible as the stem. There is also a MDu variant on *-ste*. *Vijfde* and *zesde* have older variants with *-t-* and the current form with *-d-* probably has had a change in the spelling to generalize the forms across the paradigm. (*f-* and *s-* in *fimf* and *ses* have also gained voicing.) ModDu *vijfde* and *zesde* are still pronounced with voiceless [t] by many speakers today. *Vijfste* and *vichste* (the latter created by soundlaws) have been found in MDu. It can scarcely be determined whether the MDu form *seste* is the cardinal with a voiceless *-te* ordinal suffix, or a simplification of *ses+ste* > *seste*. *Sevende*, like *seste*, must have had its onset *s-* voiced; and, unlike *zesde*, its *-d-* is a true, voiced, [d] in Standard Dutch today; this is probably due to the adjacent sonorant *-n-*, because cognate forms, for example from OHG, show a *-t-*, but ODu already has *-d-*. *Sevenste* has also been found in MDu. The development of the ordinal ‘9th’ seems to have gone quite similarly to that of ‘7th’; the cardinals look alike and they are close together. This must have benefited mutual influence between both the ordinals.

Acht has as much as five competing ordinals in the MDu era: *achte*, *achtede*, *achtende*, *achtenste* and *achtste*; of these, if *achtede* comes from *ahtuda*, it is the oldest; *achte* is a reduction thereof; *achtende* was formed after the older forms *zevende* and *negende*; then *achtenste* is newer still, formed after *zevenste* and *negenste* which were newer than their *-de*

counterparts; finally, in late MDu the intervocalic schwa in *achtadā* was deleted and replaced with an *-s-* due to OCP.

Tiende follows the same route as 7 and 9 in the respect that OHG had *-t-*, but Got had *-d-*; and so there was already some voicing at play; and in MDu we find *tiende* and *tienste*. *Elfde* and *twaalfde* have *-to* cognate forms, but in MDu we find both the voiceless *-te* suffix as well as the voiced *-de*. It is unclear where the voicing came from (possibly a generalization from the adjacent ordinals?), because *-f-* is not a voiced consonant. The ordinals *dertiende* through *negentiende* were formed with *tiende* as their head, and so they follow *tiende* in their development.

Old cognate forms for ‘20th’, for example from OHG, show use of the superlative suffix *-ôsto*. There is no indication in the historical literature or dialectal data of any other suffix than *-ste* to combine with the *-tig* suffix in the decad formation. This *-tigste* combination has even been found on ‘100th’ and ‘1,000th’ in MDu, with backformation into the cardinal as *hondertich* and *duizentich* instead of *hondert*. The standard ordinal formation for the numbers 100 and up is *-ste*; there is no evidence for a form on *-de* for ‘100th’; but *duizende*, *miljoende* and *miljarde* have sporadically been found among the dialectal data.

Appendix B: Extended survey methods and methodology

Parts of this appendix have appeared in section 3.1 Methods.

The questionnaire was created using Google Forms. This online tool allows you to create a questionnaire and then store it in your Google Drive; upon request it creates a Google Spreadsheet and links the form to it, which is also stored in Drive.⁶⁹ This allows the researcher to watch the responses come in in real-time: as a respondent finishes the survey and hits ‘submit’, the results will be automatically stored in the spreadsheet file: one column for every question, the first column with a timestamp; and one row for each respondent.

The questionnaire was reachable via a link. Participants were not required to log in before they could participate. I did not ask the respondents to write out their answers like they pronounced it, but all participants did this (some to a larger extent than others) even without being instructed to do so.

⁶⁹ This is a file type which presents data in rows and columns; a widely used program of the same type is Microsoft Office’s Excel.

The questionnaire is made up of 4 subparts. The first subpart consists of 5 background questions that ask about the respondent's identity; specifically about their dialect. Subsection 2 asks the participant to give 25 cardinal numeral forms from their dialect; each cardinal is depicted in Arabic numerals, as 1, 2, 3...21, 100, 1,000, 1,000,000 and 1,000,000,000. Section 3 looks very much like section 2, but it asks about the ordinal forms instead. Section 4 is a small closing section with a non-required, open question: does the respondent have anything they wish to say? Each subsection has its own title followed by a short introductory text to explain to the respondent what is asked of them. More details on the design of the questions below.

The questionnaire asks every participant every number only once. What this entails for the data is that whenever something rare occurs or I come across a typo, there is not really any way to check it. However, adding something like example sentences to the questionnaire would have made it longer thus making the threshold for potential respondents higher: the smaller an effort it promises to be, the less of an impact people will assess it to have on their time, the more likely they are to be willing to participate - the higher the response rate. My questionnaire consists of 2x25 number items, plus 5 opening questions and 2 closing questions; that makes 57 questions, which in total take no more than 10 minutes to complete.

The questionnaire specifically and explicitly targets dialect speakers and asks for dialect answers; however, it is written itself in Standard Dutch (SD, or AN for *Algemeen Nederlands*) because SD is the common language in the entire Dutch speaking area, and most if not all dialect speakers also speak SD. This comes with a disadvantage: asking for the forms in SD may prime the respondents into giving the SD forms for the ordinals; this may result for example in an answer that says there is variation between let's say *7-ste* and *7-de*, when in fact the dialect they learned does not have the SD *-de* form, only the *-ste* form, but the speaker also speaks SD and therefore thinks that the dialect allows both options; whereas should I really go do fieldwork among the older speakers of the dialect maybe I would find that there is only one form. However, we cannot ignore the fact that there is language contact between the standard language and the dialects; thus it might not be meaningful to speak of a 'pure' form if it doesn't really exist in any speaker.

Having stated the disadvantages, writing the questionnaire in SD also has some obvious benefits: it saves time not to have to translate the questionnaire in different dialects; some dialects lack a standard spelling, making it difficult to gauge how many versions of the questionnaire should be made; all participants get the same questionnaire and this controls for the factor 'means of data collection' when considering differences in the data, factoring out such complications as translation mistakes or translation gaps between the dialect and SD.

To factor out the disadvantages as well as possible, I stated explicitly what I am looking for in the questionnaire. I asked the respondents to keep the forms of their own dialect in mind and also to report it if they had different forms for the same thing (such as *7-ste* and *7-de*); AND if they did have variation, to also please mention whether or not one of both forms is preferred over the other.

The questionnaire was distributed through Facebook, a social media platform at <http://www.facebook.com>.⁷⁰ It targeted speakers from specific parts of the Dutch language area. For those not familiar with the latter I provide a map in figure (3.1) in the main text with the names of those provinces in the Dutch language area that were the subject of this research. The current study focuses on roughly the same language area as a recent paper on several syntactic phenomena by Sjef Barbiers, Marjo van Koppen, Hans Bennis and Norbert Corver (2016); the dialect descriptions that follow below are based on their descriptions (*ibid.*, 2016: 11). The current study focuses on Flemish⁷¹ (spoken in Belgium, in the provinces of West and East Flanders), Zeeuws (spoken in the Dutch province of Zeeland and the southern-most part of South Holland⁷²), North-Brabantish⁷³ (spoken in the Dutch province of North-Brabant), South-Brabantish (spoken in the Belgian provinces of Antwerp and Brabant), Flemish Limburgish (spoken in the Belgian province of Limburg) and Standard Dutch (the standard language spoken in the Netherlands).

⁷⁰ I chose to use Facebook as the platform for distributing my questionnaire because I believed it would very quickly lead to a large number of responses, an expectation which was indeed confirmed within days. Facebook has many users and a lot of these Facebook users have a large network of contacts. A paper has recently been published about the degrees of separation between Facebook users all over the globe: Backstrom et al. (2012) calculated the average distance between any two users in ‘the entire Facebook network of active users’ to be at 4.74, ‘corresponding to 3.74 intermediaries or “degrees of separation”’. Note that it may very well be the case that some respondents from the same area are related to each other and may even come from the very same background, same community, same street or even same household and therefore if for example in a group of speakers from the same area, 60% shows the same patterns, we should be aware that it is possible that this number is not representative for the entire region because there could have simply coincidentally been a few members of the same family responding with their identical answer sets, even if their pattern is not representative for their entire town or region. This is due to the type of sampling I used to find respondents. Stefan Dollinger (2015:273) calls this snowball sampling: it’s a type of sampling where your contacts lead to other contacts ‘within the same network’. He doesn’t discuss the advantages or disadvantages of this sampling type.

⁷¹ Note that the labels ‘Flanders’ and ‘Flemish’ are often used for the entire Dutch speaking part of Belgium (West/East Flanders, Antwerp, Flemish Brabant and Flemish Limburg).

⁷² Note that the topmost island indicated on the map to be part of Zeeland is officially part of the province of South Holland, but linguistically it is usually taken to be a part of the Zeeuwse dialect area.

⁷³ The paper by Barbiers et al. (2016) starts out with North and South Brabantish as one dialect group, but after analyzing different geographical distributions of four syntactic phenomena, they distinguish between North and South Brabantish; we shall see below that my data also shows reason to distinguish between them.

The original recruitment text (sent out on 24 march) was a Facebook status update and it asked specifically for speakers from South Holland, North Brabant and the five Belgian provinces; it said that filling out the questionnaire should take 5-10 minutes; it asked to share the Facebook post if you know any dialect speakers; and finally, it mentioned that I could be contacted through Facebook's Personal Messaging service if anyone wanted more information; at the bottom it contained a clickable link that lead to the online Google Form. I refer the reader to appendix X for details on the Facebook posts and screenshots.

I originally targeted only the three areas mentioned above because of findings in my previous work on the subject (Sleeman 2016), as mentioned in the introduction above. Zeeuws had not appeared in the previously conducted research, but as it is adjacent to the areas mentioned, I altered the recruitment text to include Zeeland. I did not ask for speakers from Dutch Limburg because it did not appear to show anything other than AN in my previous research, but in a future study I intend to target the entire Dutch speaking area, including Dutch Limburg and the more northern parts of the Netherlands.

It must be noted that the data from South Holland (SH) was more actively sought after, through personal contact via family members and emailing acquaintances; of the responses coming in, very few were from SH. Previous research (Lafeber, 1967 as cited in Sleeman, 2016) had indicated that Gouda would show variation for ordinal suffixes in 7 and 9, and thus I was very interested in these data, but none came in. Speakers of Hollandic dialects are known to deny speaking a dialect and in fact their Dutch often does not sound very different from standard Dutch. This is in sharp contrast with speakers of for example Netherlandic Limburgish or Flemish varieties, who are generally very proud of their language, resulting in eager questionnaire responses.

Use of Facebook for distribution

- Full text (24 March): Hoi! Ik zoek sprekers van dialecten uit Zuid-Holland, Noord-Brabant, en overal in Vlaanderen! Mijn vraag is of je dit vragenlijstje zou willen invullen. Het duurt 5-10 minuten. Deel dit bericht als je denkt dat je mensen in je vriendenkring hebt die een relevant dialect spreken. Voor meer informatie: stuur me een persoonlijk bericht ☺
- Full Text (altered, 26 March): Hoi! Ik zoek sprekers van varianten van het Fries, en sprekers van dialecten uit gebieden van Zeeland die grenzen aan Brabant of Vlaanderen! Ook dialectsprekers uit Zuid-Holland, Noord-Brabant, en overal in Vlaanderen zijn van harte uitgenodigd om te reageren! Mijn vraag is of je dit vragenlijstje

zou willen invullen. Het duurt 5-10 minuten. Deel dit bericht als je denkt dat je mensen in je vriendenkring hebt die een relevant dialect spreken. Voor meer informatie: stuur me een persoonlijk bericht ☺

- 30 march: changed the recruitment text to ‘looking for speakers of Frisian, and of dialects spoken in the parts of Zeeland bordering on Brabant or Vlaanderen. Responses from speakers of dialects from Zuid-Holland, Noord-Brabant, and anywhere in Vlaanderen are also very welcome’

- Original post aired on 24 march, 23.35. Put it on ‘public’ in order to enable as much sharing as possible - public posts can be seen by anyone on or off facebook.⁷⁴ Was shared 49 times:

- on the same day, 24/03, shared twice
- I posted it with permission on timeline from someone who lives in Begium (24/03, 23:37, minutes after the original post)
- some notable shares:
 - the post was shared by the facebook page ‘Stichting Nederlandse Dialecten’ (Stichting Nederlandse Dialecten, 2016)
 - it was shared to the group ‘Belgians in Hungary’: *Belgen in Hongarije./Belges en Hongrie*. Public group, 281 members.

⁷⁴ The posts cited in the references below can be viewed by anyone, even without logging on to Facebook.



Ruby Sleeman
24 March at 23:35

Hoi! Ik zoek sprekers van dialecten uit Zuid-Holland, Noord-Brabant, en overall in Vlaanderen! Mijn vraag is of je dit vragenlijstje zou willen invullen. Het duurt 5-10 minuten. Deel dit bericht als je denkt dat je mensen in je vriendenkring hebt die een relevant dialect spreken. Voor meer informatie: stuur me een persoonlijk bericht 😊

- Added a link preview to this post.

 <p>Google Forms</p>	<p>Welkom bij deze enquête over dialect-telwoorden!</p> <p>We beginnen met een paar basisvragen, voordat we op de inhoudsvragen overgaan. Let op: deze enquête is bedoeld voor dialectsprekers. Geef...</p> <p>DOCS.GOOGLE.COM</p>
---	---

Figure 1: Original post, focus on Zuid-Holland, Noord-Brabant and Vlaanderen



Ruby Sleeman
26 March at 21:27

Hoi! Ik zoek sprekers van dialecten uit Zuid-Holland, Noord-Brabant, en overall in Vlaanderen! Ook dialectsprekers uit gebieden van Zeeland die grenzen aan Brabant of Vlaanderen zijn van harte uitgenodigd om te reageren! Mijn vraag is of je dit vragenlijstje zou willen invullen. Het duurt 5-10 minuten. Deel dit bericht als je denkt dat je mensen in je vriendenkring hebt die een relevant dialect spreken. Voor meer informatie: stuur me een persoonlijk bericht 😊

 <p>Google Forms</p>	<p>Welkom bij deze enquête over dia...</p> <p>We beginnen met een paar basisvragen, voordat ...</p> <p>DOCS.GOOGLE.COM</p>
---	---

Figure 2: Post with addition of Zeeland as an area of interest



Ruby Sleeman
30 March at 14:00

Hoi! Ik zoek sprekers van dialecten van het Fries, en sprekers van dialecten uit gebieden van Zeeland die grenzen aan Brabant of Vlaanderen!

Ook dialectsprekers uit Zuid-Holland, Noord-Brabant, en overal in Vlaanderen zijn van harte uitgenodigd om te reageren! Mijn vraag is of je dit vragenlijstje zou willen invullen. Het duurt 5-10 minuten. Deel dit bericht als je denkt dat je mensen in je vriendenkring hebt die een relevant dialect spreken. Voor meer informatie: stuur me een persoonlijk bericht 😊



Welkom bij deze enquête over dia...
We beginnen met een paar basisvragen, voordat ...

DOCS.GOOGLE.COM

Figure 3: Post with addition of Friesland, and focus on Friesland and Zeeland

Ruby Sleeman
24 March · 🌐 ▼

Hoi! Ik zoek sprekers van varianten van het Fries, en sprekers van dialecten uit gebieden van Zeeland die grenzen aan Brabant of Vlaanderen!

Ook dialectsprekers uit Zuid-Holland, Noord-Brabant, en overal in Vlaanderen zijn van harte uitgenodigd om te reageren! Mijn vraag is of je dit vragenlijstje zou willen invullen. Het duurt 5-10 minuten. Deel dit bericht als je denkt dat je mensen in je vriendenkring hebt die een relevant dialect spreken. Voor meer informatie: stuur me een persoonlijk bericht 😊


Google Forms

Welkom bij deze enquête over dialect-telwoorden!

We beginnen met een paar basisvragen, voordat we op de inhoudsvragen overgaan. Let op: deze enquête is bedoeld voor dialectsprekers. Geef...

DOCS.GOOGLE.COM

👍 [redacted] and 3 others · 15 Comments **49 Shares**

👍 Like 💬 Comment ➦ Share

Figure 4: Facebook post with updated contents and total amount of shares since first release



Figure 5: Shared the amended post

Details of questionnaire design

Section 1

In the introduction to the questionnaire, the respondent is reminded that the questionnaire's research focus lies with dialect words and so the respondent is asked explicitly to answer with dialect words rather than SD words.

Section 2 and 3

These two are the subsections essential to the survey. In these sections, the respondent is asked to list their cardinal and ordinal forms. The selection of cardinals and ordinals is asked for is the following, repeated from above:

- 1-21; 100; 1,000; 1,000,000 and 1,000,000,000

Asking for these forms ensured, I believed, the capturing of all relevant data. Many numerals are formed by combining other numeral words together. The variation that I had found in previous work (as described in the introduction above) was limited to the range 7-19, but I wanted to be able to paint a complete picture of the number systems in regional varieties of Dutch. The forms I asked for should be able to form all other numbers on the mathematical numberline.

Section 4

All questions in the survey were marked 'required' and had to be answered before the survey tool would let the participant move on - with exception of the last question, an open question: 'Does anything come to mind that you think could be relevant for the current study?' This question was not marked required and respondents could just leave it unanswered if they so desired.

Screenshots of the questionnaire follow below. Note that some numbers are missing from the screenshots; for both cardinals and ordinals it is the case that 1-21 were asked, 100, 1,000, 1,000,000 and 1,000,000,000 even if only 1-3 are shown.

Welkom bij deze enquête over dialect- telwoorden!

We beginnen met een paar basisvragen, voordat we op de inhoudsvragen overgaan. Let op: deze enquête is bedoeld voor dialectsprekers. Geef alstublieft antwoord in dialect wanneer daarnaar gevraagd wordt.

* Required

1. Wat is uw geslacht? *

- Man
 Vrouw

2. Wat is uw leeftijd? *

- Jonger dan 15 jaar
 15 - 25 jaar
 25-35 jaar
 35-45 jaar
 45-55 jaar
 55-65 jaar
 65-75 jaar
 Ouder dan 75 jaar

3. Welk dialect spreekt u? (Voor Friessprekenden: vul de naam van uw Friese dialect in of gewoon 'Fries'. Er wordt in geen geval gesuggereerd dat Fries een dialect van het Nederlands zou zijn; ook niet in de rest van deze enquête.) *

Your answer _____

4. In welk dorp of welke stad heeft u dit dialect geleerd? *

Your answer _____

5. In welke regio of provincie ligt deze plaats? *

Your answer _____

NEXT

25% complete

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Additional Terms

Google Forms

Figure 6: Subsection 1 of the questionnaire

Welkom bij deze enquête over dialecttelwoorden!

* Required

Hoofdtelwoorden

Geef in dit onderdeel de vormen uit uw dialect voor de hoofdtelwoorden van 1 - 21; verder de telwoorden voor honderd, duizend, miljoen en miljard. Als er geen standaardspelling is voor uw dialect, maakt u zich daar dan vooral niet teveel zorgen over.

1 *

Your answer

2 *

Your answer

3 *

Your answer

4 *

Your answer

Figure 7: Subsection 2 of the questionnaire

The image shows a screenshot of a Google Form questionnaire. It contains five questions, each with a numerical value followed by an asterisk and a text input field labeled "Your answer". The questions are:

- 21 *
- 100 *
- 1000 *
- 1.000.000 *
- 1.000.000.000 *

At the bottom of the form, there are two buttons: "BACK" and "NEXT". To the right of these buttons is a progress bar that is half-filled with blue, labeled "50% complete". Below the progress bar, there is a small text warning: "Never submit passwords through Google Forms."

Figure 8: Subsection 2 of the questionnaire

Welkom bij deze enquête over dialecttelwoorden!

* Required

Rangtelwoorden

Mijn onderzoek houdt zich bezig met de uitgangen -de en -ste in rangtelwoorden van het Nederlands (bijv. eerste, tweede, derde, achtste, twintigste enz). Dialecten verschillen soms in welke uitgang ze gebruiken: -de of -ste. Daarom ben ik benieuwd naar de vormen van uw dialect.

Geef nu uw dialectvormen voor de rangtelwoorden 1e - 21e, honderdste, duizendste, miljoenste en miljardste.

N.B.: Het kan voorkomen dat uw dialect twee of meer verschillende vormen kent voor hetzelfde rangtelwoord (bijvoorbeeld tiende en tienste). Noem dan beide vormen en geef aan of u voorkeur heeft voor één van beide vormen.

Ook nu geldt: let niet teveel op spelling. Probeer wel dezelfde spelling te gebruiken als voor de telwoorden van het vorige onderdeel (dus bijv. 'zwölf > zwölfste' en niet 'twulf > zwölfste').

1e *

Your answer

2e *

Your answer

3e *

Your answer

Figure 9: Subsection 3 of the questionnaire

20e *

Your answer

21e *

Your answer

100e *

Your answer

1000e *

Your answer

1.000.000e *

Your answer

1.000.000.000e *

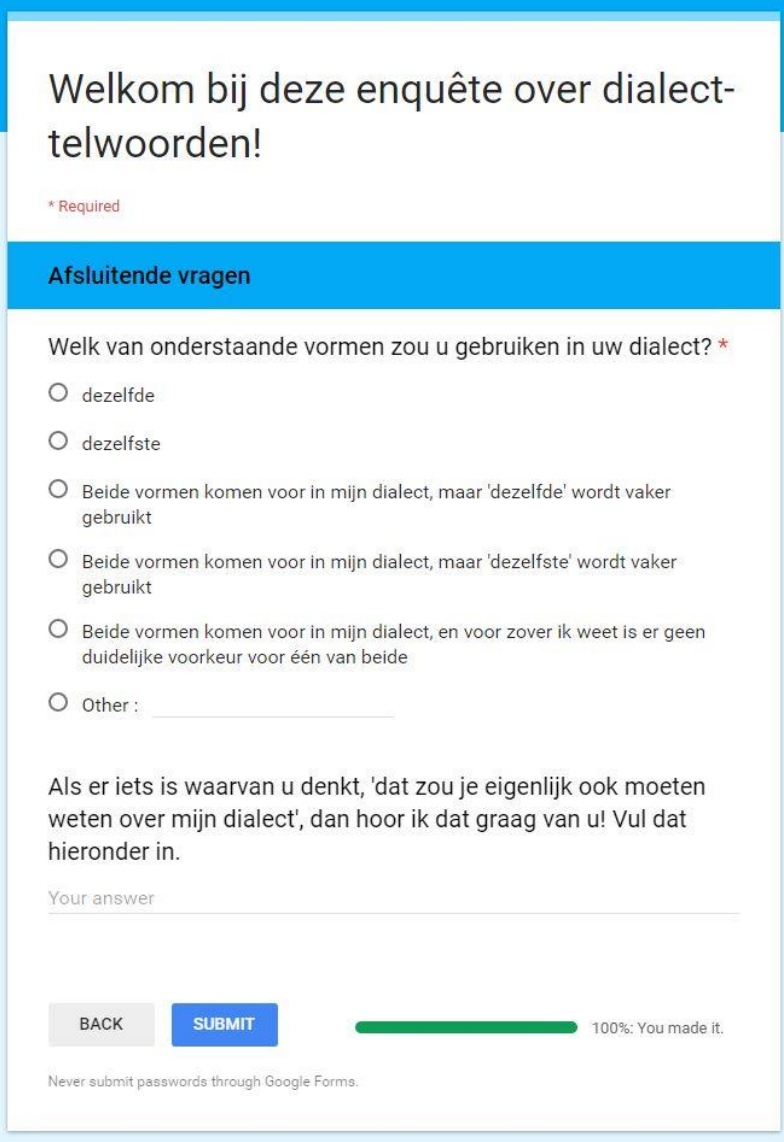
Your answer

BACK NEXT

75% complete

Never submit passwords through Google Forms.

Figure 10: Subsection 3 of the questionnaire



Welkom bij deze enquête over dialecttelwoorden!

* Required

Afsluitende vragen

Welk van onderstaande vormen zou u gebruiken in uw dialect? *

dezelfde

dezelfde

Beide vormen komen voor in mijn dialect, maar 'dezelfde' wordt vaker gebruikt

Beide vormen komen voor in mijn dialect, maar 'dezelfste' wordt vaker gebruikt

Beide vormen komen voor in mijn dialect, en voor zover ik weet is er geen duidelijke voorkeur voor één van beide

Other : _____

Als er iets is waarvan u denkt, 'dat zou je eigenlijk ook moeten weten over mijn dialect', dan hoor ik dat graag van u! Vul dat hieronder in.

Your answer

BACK SUBMIT

100%: You made it.

Never submit passwords through Google Forms.

Figure 11: Subsection 4 of the questionnaire

I wanted to reach a broad group, representative of the linguistic community of speakers of Dutch dialects. This means that the language of my questionnaire should not be too formal; that would make it less available to the less highly educated. Secondly, the very nature of the forms I am looking for - non-standard forms - it might help if the language of the questionnaire is also less standard if I want the respondents to think of less standard forms, 'in as much as the social conventions allow' (Dollinger, 2015:261). This means, for instance, that I did stay very polite; this is reflected in the use of the polite you-form *u* rather than *je*; I used the polite form *alstublieft* for 'please';

- 'write the questionnaire with the reading skills of a Grade 9 student in mind.' (same quote as above: (Dollinger, 2015:261)).

- in the same spirit of making potential respondents willing to participate, I made the questions sound like I was very interested:
 - in section 3, where I first explain that it is their ordinal forms that I am after, and also *why* I want to know their ordinal forms, I say literally that I am *curious about your forms*. This way, I express my enthusiasm and also I appeal to them directly by explicitly saying ‘*your forms*’, as if I am triggering them to take on the role of language instructor.
 - The final question of the questionnaire is another example of how I express my eagerness and try to appeal to the respondent directly, again to get them to want to teach me their language.⁷⁵
- The intro text to the questionnaire did not mention the respondent’s anonymity, nor did it provide an estimation of the amount of time the respondent is likely to need. These things were, however, included in the facebook recruitment text. The purpose of the questionnaire was not mentioned until the intro text of subsection 3 about the ordinal numerals. I did not want to overload participants with too much information and instruction at once, and I found it important to make sure that they were properly instructed about the background questions on which dialect they speak and where they learned it. Then for subsection 2 I just wanted to get them started typing out simple counting words, instead of getting a whole lot more information right after the introduction section with questions about their background. This is why I postponed explaining the actual purpose until after section 2 was done.
- There was no consent form in the beginning of the questionnaire. no sensitive information concerning the respondents was stored, and since the questionnaire could be filled out by respondents in their own private homes on their own computers, I didn’t think it necessary to remind them of their right to just click away and not submit their answers. In a setting where the participant comes to a lab for example, and sits down to fill out the task form, they may feel obligated to stay. In such cases, it is much more apparent that they should be made aware of the fact that they can in fact just get up and leave at any time. This was not necessary in the case of an online questionnaire which did not require any logins or passwords.

⁷⁵ Many indeed respond positively to this in the sense that they answer the cardinal and ordinal forms with additional instructions to pronunciation, and at the end of the questionnaire the answers can be quite elaborate on something unrelated to the subject of the questionnaire, but simply something that the respondent finds particularly attractive or interesting about their dialect, for example: ‘in Hasselt the morphological genitive case is still used widely, as in ‘thetens’ = te haren thuis’; some respondents refer me to Youtube for lessons in their dialect or written sources (grammars)...

- Dollinger 2015:269-270 says that background questions at the beginning rather than at the end may scare off potential respondents. I hardly think this posed a problem in this questionnaire. In any case, I got over 250 respondents so it can't have been too bad; but also, I never asked for anything but gender, age and linguistic background; the latter only very superficially, and the age question was a multiple choice question asking for the age cohort, even less of a face threatening question.
- I often ask about the respondent's personal forms. In the second-to-last question, before the open one, I ask about *dezelfde/dezelfste*. This is a different word than all the numbers that we've seen so far. It's possible - I don't know this - that it may be more heavily stigmatized in any or all dialects than the number words are. For this question, then, I specifically ask about what they hear around them from other speakers of their dialect. Dollinger 2015:265 states that in these kind of situations there is a risk of losing face, even if the questionnaire is completely anonymous. People don't want to be 'caught' saying things that are seen as wrong from the perspective of the standard language. Asking about the use or the preference of either of the items in their language (I only asked about existence and preference, not about frequency, neither in a quantitative nor a qualitative way) therefore is safer for them to answer if it doesn't concern them personally.
- A colleague pointed it out to me that if I intended to get Frisian results (which I did intend to), it was important to be clear about my stance on the status of Frisian as a separate language. Frisian speakers are generally very proud and will feel offended when their language is referred to as a dialect - it has obtained the official status of minority language in the Netherlands and can be traced historically to be a different branch in the Germanic language family.⁷⁶ Many Dutch people ignorant of these facts sometimes still refer to Frisian as just a variety of Dutch and Frisian speakers, proud of their heritage and whatnot, get annoyed by this. For this reason I had to choose my words very carefully and in the beginning of the questionnaire I made sure to stipulate that speakers of Frisian should read 'language' anywhere that it said 'dialect', and that in no way did I intend to deny its status of language, and that this questionnaire had no interest in doing anything of the sort.

⁷⁶ Frisian is a member of the Anglo-Frisian subbranch of West Germanic languages, whereas Dutch belongs to the Low Franconian subbranch.

Appendix C - Additional survey findings

Of the original 265 respondents 5 were excluded before starting the data organization.⁷⁷ Each of the remaining 260 datasets then was labeled with the numeric range that the *-ste* suffix took over the ordinal number line. Each new numeric scope was considered a new ‘pattern’, and given its own unique number. When suspected typos occurred (e.g. ‘*achtstw*’, where ‘*tw*’ is not a valid phonotactic combination in Dutch, and incidentally, *w* and *e* are adjacent on the QWERTY keyboard which is commonly used in the Netherlands), or an unexpected pattern emerged (for example, a new palatalized suffix appeared, and only in one speaker) I took one of the following two approaches: either the outlying data could easily be explained, as in the typo-example above, and I labeled it as a new pattern or as an existing pattern; or, if the outlier was less easily resolved - this happened in 14 out of the 260 remaining cases - I marked the pattern simply as ‘o’ in order to prevent falsely extrapolating the data. These 14 speakers’ data thus were not taken into account.⁷⁸ After excluding these 14, a few geographical areas in the Netherlands were so underrepresented in the data that I decided to discard them altogether, deleting another 6 datasets, making the total amount of speakers 240. The reported areas are the Dutch provinces of Drenthe and Gelderland, and Friesland (where Frisian is spoken; this is not a dialect of Dutch but is nonetheless closely related to it and has a history of contact with Dutch). The Dutch province of Limburg also has only two data points, but I decided to keep them in nonetheless, because they are geographically very close to some of the Belgian municipalities represented in the data. However, one should note that not many conclusions, if any at all, can or should be drawn from two data points which both show AN patterns - AN ordinal paradigms occur throughout the entire map and do not seem to predominantly occur in any specific area.

Using statistics and calculating significance of differences between occurrences seemed not to fit the nature of the data; the latter being individual patterns of dialect speakers, sometimes only one speaker’s data ‘representing’ an entire region or subdialect, for lack of other informants from the same area. It is for this reason that I decided to make no claims about significance, but only to calculate the number of occurrences of each pattern throughout the

⁷⁷ These were either submitted doubly (identical answers in addition to near identical time stamps) or unserious replies.

⁷⁸ 9 out of the 14 speakers I discluded from the data organization into pattern clusters were Frisian. This is due to the fact that some phonological things happen in their data which make it difficult for me to gauge which pattern I can file them under, or whether I should create a new pattern for them. I have decided to not include these data at this stage.

data, and the number of occurrences of each stem-suffix combination, and the percentages in comparison to the whole data set, so that these could quantitatively support any claims I want to make about rarity of stem-suffix combinations.

After labeling all different patterns, I concluded that the vast majority of patterns occurred only once in the data; many small differences occurred between two speakers, for example, only one of the ordinals in two sets would differ, but they were initially treated as two different systems, as has been noted above. The next paragraphs are devoted to some general tendencies like these.

Whenever a pattern allows more than one suffix for the same ordinal, there is only one speaker with that exact pattern. Whenever a rare form occurs (e.g. 1-de, 8-de, *duizend-de* (dzd-de) ‘thousandth’, *miljoen-de* (mln-de) ‘millionth’ or *miljard-de* (mld-de) ‘milliardth’, there is only one speaker with that exact pattern. For both these tendencies, there may be more speakers with the same variation in their patterns, but the patterns will never have an exact duplicate in another speaker. For example: patterns 25, 27 and 29 share the property that they have 1-de; but in 25, there is 8-ste; in 27, there are no -ste forms except ≥ 20 ; and in 29, there is no 8-ste but there is 7-ste. Therefore, all three patterns differ in some slight aspect.

In only 2% of all speakers the ordinal for 1 is *not* formed exclusively with the -ste suffix: 4 speakers used -de, 1 speaker had both suffixes. There are only five participants in the entire dataset where the participant’s spelling seems to indicate a different suffix than the AN -ste in *eerste* ‘first’. In these five participants, who are all from Diepenbeek, the form is spelled *josde* or *jozde*, seemingly indicating a -de suffix.⁷⁹ There is only one spelling it with a ‘z’, and AN *zesde* ‘sixth’ is spelled *zizde* by this same participant, leading me to speculate that these forms might have come to share some form properties in this speaker’s grammar. However, the stem of *eer-ste* in AN is uncontroversially taken to be *eer*, making it unlikely for the s in *josde* to have originated as a part of the stem form. A more likely analysis is that the -ste suffix has undergone a form of voicing: *eer-stə* changed to *jɔ-stə*, this was resyllabified into *jɔs.tə* and then it could have analogously to *zes-də* ‘sixth’ been voiced to *jɔs.də*.

To further investigate this, a look into superlatives (-ste is originally a superlative suffix, as was mentioned earlier) might help clear this up. A brief look at some recorded Diepenbeek dialect speech has not revealed any voicing in superlative forms, but the data in question did not contain the speaker’s form for *first*, therefore no conclusions can be made based on this

⁷⁹ It must not be overlooked that these participants could all be related to one another or share some other characteristic; they might for example live in a sub-area of the region with this word in the regiolect that the rest of their municipality does not share. Unfortunately, it is not possible to check these facts in the current study.

attempt.⁸⁰ Searching the MAND and SAND with the MIMORE search tool did not yield any helpful results either. More research with speakers should be done to find answers to this question, which I leave now as outside the scope of this thesis. We will return to the exceptional forms of the data such as *josde*, *achde* and *miljoende* in the analysis in the sections below.

Of the five participants with 1-de, two also have -de on 8, and two others have -de on *miljoen* ‘million’. For 8, these are the only 2 speakers in the entire dataset that have a -de form instead of SD -ste. One of the preliminary conclusions I drew from the previous research done, mentioned in the introduction above, was that although historically some different forms were found for *achtste* (*achtde*, *achtende* etc), **no** synchronical regional variety of Dutch reported anything other than *achtste* or variants thereof (*achste*, *aachste* etc). The results of this questionnaire are in concordance with that conclusion.

11-ste and 12-ste seem to group together. For example, 12-ste (instead of AN 12-de) hardly ever occurs without 11-ste. 11-ste occurs without 12-ste a lot more often: 47,6% of all speakers had 11-ste; 35,4% of all speakers had both 11-ste and 12-ste; 87 out of 117. 74,4% out of all speakers who had 11ste (87/117x100) also had 12ste.

[13-19] behave the same way in 93,5% of all speakers. Whether it be that they all vary, that they all get -de or all get -ste within the same speaker, in 93,5% of all cases a speaker does the same to these numerals all in a row. When 10 gets -ste, or varies in -de/-ste, 13-19 also get -ste or vary. (not vice versa; 10 gets -ste more rarely than 13-19 do).

The ordinals for 20 and beyond are fairly robust in their combining only with -ste.⁸¹ In the set of 240 speakers, there is not a single speaker that forms the ordinal for 20 with -de; there are three that use -de for thousand ‘duizend’, four that use -de for million ‘miljoen’, and two that form the ordinal for milliard ‘miljard’ with -de, one of these two, the only speaker in the entire dataset, forms both *miljoende* and *miljarde* with -de. All these occurrences that diverge from the general tendency to use exclusively -ste for 20 and above, are unique patterns and

⁸⁰ The material in question was found in a videotaped comedy performance, spoken in the Diepenbeek dialect, on YouTube, reachable at <https://www.youtube.com/watch?v=WWv48L1riM> 1:28. The word *veur-ste* ‘front-SUPL’ resembles in morphology the AN word *eerste* but is not historically related (incidentally, it is related to english *first*), and *veur-ste* has not undergone a stem change like *eer-ste* potentially has. *veur-ste* does not in this instance sound voiced. But the man might not voice his *joste* either. This word was not heard in the video.

⁸¹ *Twintig-de* should not pose any problems phonotactically: the preterite singular form of the verb *nodigen* ‘invite’ is *nodigde*. If we look outside of form motivations, perhaps this robust use of -ste is somehow tied to multiplication: all the tens are formed through multiplication, and while 100, 1,000, 1,000,000 and 1,000,000,000 are not formed through multiplication, they can be used as multiplicands: *vijfhonderd* ‘500’, *vijf duizend* ‘5,000’, *vijf miljoen* 5,000,000 and *vijf miljard* ‘5,000,000,000’.

with exception of the one just mentioned, only have one *-de* form among the *-ste* complying forms above (whether that be for ‘duizend’ or for ‘million’).

Let us consider for a brief moment the outliers in the regional distribution of the patterns. As can be seen in tables (2.5) and (2.6) above, the speakers in the database from the Dutch provinces North Brabant (NB), Zeeland (ZL) and South Holland (SH) almost exclusively have AN patterns, save a few exceptions. For NB, this exception is pattern 2c which predominantly occurs in Flemish Limburg and Antwerp, the latter of which is a language area adjacent to NB. The speaker’s municipality, Etten-Leur, is quite close to the border with the Belgian province of Antwerp. SH’s two outliers are from the patterns 2c and 1, which is less easily explainable: the predominant areas for those patterns are Antwerp and Flemish Limburg for 2c, and West and East Flanders for 1, both not adjacent to SH. Because there is so little other data for these regions, I do not at this point see it fit to conclude anything from or about these outliers.

ZL’s two outliers show patterns 2 and 1; the former being found mostly in Antwerp, the latter in West and East Flanders. All are adjacent to ZL, and both the outlying ZL respondents are from the southernmost part of ZL (from the municipalities Sint Jansteen and Axel), and thus I argue must simply fall within the dialect areas of Antwerp and West/East Flanders.