

John Benjamins Publishing Company



This is a contribution from *Reorganising Grammatical Variation. Diachronic studies in the retention, redistribution and refunctionalisation of linguistic variants.*

Edited by Antje Dammel, Matthias Eitelmann and Mirjam Schmuck.

© 2018. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

Plural inflection in North Sea Germanic languages

A multivariate analysis of morphological variation

Arjen P. Versloot and Elżbieta Adamczyk

University of Amsterdam & Fryske Akademy / University of Wuppertal & Adam Mickiewicz University, Poznań

The present study explores the variation in the plural inflection of eight varieties of Frisian and English, focusing on irregular plural formations. The aim of the study is to identify and assess the significance of the factors which contributed to the preservation and emergence of irregular plural patterns in the investigated varieties. Multivariate regression analysis, which was employed to weigh the significance of individual factors, reveals that the retention and emergence of irregular plural patterns is primarily determined by a combination of three major factors: relative frequency, absolute frequency and salience. These factors can be reduced to underlying cognitive processes, including the storage and retrieval of linguistic information in the human brain. Another relevant factor, semantics, affects the lemma frequency profile, but can gain momentum as a factor in its own right due to analogical co-activation in language processing.

Keywords: irregular plurals, frequency, salience, English, Scots, Frisian, multivariate analysis

1. Introduction

The focus of the present study is on the diachronic development of irregular plural formations in the two representatives of the North Sea Germanic branch, i.e. English and Frisian.¹ These irregularities are partly vestiges of once regular

1. The concept of '(ir)regularity' can be defined in different ways, often depending on the theoretical framework adopted. It has often been associated and analysed with reference to terms such as 'productivity', 'type frequency', 'rule based', or 'default'. In our study the association with *type frequency* is the most prominent, and accordingly 'irregular plural patterns' can be best defined as plural patterns with a low type frequency. At the same time, we do not take any stance with

inflectional patterns and partly a result of more recent phonological or lexical developments. The Old Germanic languages were characterised by a wide range of plural formation patterns, inherited from the Proto-Indo-European inflectional system and subsequently reshaped in Proto-Germanic. These plural formation patterns have their source in the historical inflectional classes, originally organised on etymological grounds, i.e. according to the stem-type criterion. The diversity of declensional classes, which can be found in the individual Old Germanic languages (e.g. *a*-stems, *s*-stems, *n*-stems, root nouns), is reflected in a varied inventory of plural morphemes whose distribution was highly lexical at the earliest attested stage of Germanic (i.e. the pattern of inflection was not predictable from the nominative sg.), e.g. OE *dæg* – *dagas* ‘day(s)’ (*a*-stem), OE *cynn* – *cynn* ‘kin(s)’ (*ja*-stem), OE *cealf* – *cealfru* (*s*-stem) ‘calf/calves’, OE *fōt* – *fēt* (root noun) ‘foot/feet’. While some of them continue as productive patterns of forming the plural in present-day English and Frisian (*-(e)s*, *-en*), others are only sporadically retained as vestiges of the historical plurality markers. Already at the earliest stage of attestation the declensional systems of the Germanic languages were undergoing a gradual process of reorganisation, induced by an interplay of phonological and morphological developments, including phonological reduction and analogical levelling. These tendencies gradually led to a confusion and merger of class-specific inflectional markers in individual Germanic languages and dialects, and contributed to a reduction in the diversity of the inherited inflectional exponents, including those marking specifically the plural (for Old English see, e.g., Kastovsky 1995, 1997; Hogg & Fulker 2011; Adamczyk 2010, 2013, 2014, 2018). With time, the distribution of plural markers tended to become increasingly determined by phonology and gender (as was the case with the masculine nouns in *-e* in Middle Dutch, which became feminine, in contrast to feminine nouns ending in a consonant, which became masculine (Van Loey 1976: 19, 23)), or guided by semantics (as in the case of the *s*-stems, representing predominantly agrarian vocabulary, e.g. OE *æg* ‘egg’, *cealf* ‘calf’, *hrīðer* ‘horned cattle’, *ēar* ‘ear of grain’, *gycer* ‘acre’, cf. Dammal, Kürschner & Nübling 2010: 591).² Nevertheless, in most present-day Germanic languages the distribution of the plural allomorphs still contains predominantly a lexical component.

The emergence of the system of stem-type inflection in Old Germanic was a relatively recent development, resulting from a fairly uniform inventory of inflectional endings in Proto-Indo-European (Ringe 2006: 269, 272, 279–280). Most of

respect to the theoretical implications that different definitions of ‘(ir)regularity’ may have. For a recent discussion of irregularity and its relation to other concepts, see, e.g., Stolz et al. (2012), Gardani (2013), Strik (2015: 13–24).

2. For a discussion of the interaction between declensional class and gender in Germanic languages, see, e.g., Enger (2004), Schwink (2004), Kürschner (2008).

the phonological processes, such as *i*-mutation, gemination, monophthongisation, apocope and syncope, which had their share in the emergence of the varied set of fused inflectional endings, did not take place earlier than between 300 and 700 AD (Nielsen 2000: 289). The complex interaction between phonology and morphology eventually determined the shape and dynamics of the changes in the nominal system, including plural inflection. More specifically, the changes in the phonological system functioned as the catalyst of morphological restructuring, and the later process of analogical levelling must be seen as a natural response of language users to the changing phonological constellation (Paul 1920: 198; Versloot 2008: 78, 276–294). While the paradigms of most of the nouns affiliated with minor (unproductive) classes were exposed to extensive analogical reshaping which resulted in transfers to other declensional types, some tended to escape this widespread tendency, retaining the historical inflectional pattern until the present day.

As language history shows, new morphological patterns, comprising limited sets of lexical items and lexically idiosyncratic formations, tend to emerge. They co-exist with irregularities whose origins can be sought in historical inflection patterns. Both English and, in particular, the dialects of Frisian abound in such formations, which, from the synchronic perspective, form new inflectional irregularities, entailing, e.g., suppletion or phonological alternations. The focus of the present study is on both – the inherited and novel inflectional irregularities, as attested in the diachronic and dialectal varieties of English and Frisian.

The shape of nominal inflection as it is known now in present-day English and Frisian results from the interaction between several factors of diverse nature (e.g. Kastovsky 1997; Kürschner 2008; Dammel, Kürschner & Nübling 2010). The complex interaction between them was an essential part of the mechanism underlying the restructuring of the early Germanic (plural) inflection. The aim of the present study is to identify and assess the significance of the factors which contributed to the preservation and emergence of irregular plural patterns in the North Sea Germanic languages. The interplay of these conditioning factors could potentially result in three different paths of development, whereby a noun could (1) preserve the historically-developed plural form, (2) become regular, fitting into the prevailing pattern of plural formation, or (3) develop a new irregular form. The present investigation consists of two stages, including (a) the identification of the individual irregular plural formation patterns in several dialects belonging to the North Sea Germanic group, and (b) applying a multivariate analysis to estimate the weight of the individual factors contributing to the preservation and emergence of irregular plurals. As a substantial number of irregular patterns found in the present-day varieties of Germanic can be traced back to the earlier historical stages, an investigation into the historical shape of inflection constitutes the starting point for the quantitative study. The North Sea Germanic languages were selected as the testing

ground for the analysis, since they depart from a relatively recent common ancestor; in other words, Old English and reconstructed Proto-Frisian of ca. 900 were (nearly) identical with respect to inflectional morphology until a relatively late stage of development (compare the juxtaposed Old Frisian and Old English nominal paradigms in Steller 1928: 36–45; also Nielsen 1981).³ The underlying tendencies which caused the (re)grouping of the plural marking patterns and whose origin is to be sought in the Old Germanic period were so persistent that even in the 16th century, English and Frisian, having drifted apart for more than 500 years by then, show a highly similar inventory of nouns with irregular plural forms. In the subsequent 500 years, these North Sea Germanic dialects follow essentially divergent paths of development. In English, the stock of nouns with irregular plurals has been highly resistant to the generalisation of the productive plurality patterns, and new nouns have been added to the group by semantic analogy only occasionally (e.g. *fish*, which follows the endingless neuter plural pattern despite its original affiliation with the masculine *a*-stems). The dialects of modern Frisian testify to a much greater diversity in this respect, and, accordingly, some varieties, such as West Frisian, witness a gradual reduction in the number of historical irregular plurals, occasionally ‘compensated’ by the emergence of new irregular patterns. Others, such as the North Frisian dialect of Amrum and Föhr, testify to a considerable resistance to inflectional innovations. These divergent patterns need to be attributed to language-specific (or dialect-specific) circumstances, which, however, stay beyond the scope of the present investigation. In a broader perspective, the findings from the present study can hopefully afford some more insight into the underlying mechanisms of morphological change.

The present paper is organised in the following way: Section 2 is devoted to an overview of irregular plural formations in the investigated varieties and a discussion of a range of factors potentially determining the shape of the plural forms. Section 3 focuses on the methodological considerations and the characteristics of the corpora used in the study: The selection of the linguistic material used for testing is accounted for, followed by a discussion of the statistical methodology applied in the study; also, the factors which were considered crucial for the preservation and the emergence of irregular plurals are operationalised. Section 4 presents the findings from the examination of the plural formation patterns in minor declensional classes at the earliest attested stages of English and Frisian (4.1), as well as the results of the statistical tests for various data configurations in the selected modern English

| | | | |
|--------------|----------------------------|------------------------------------|--------------|
| 3. OE/Ofris. | <i>u</i> -stem <i>sunu</i> | OE/Ofris. root noun <i>fōt</i> | |
| NOM/ACC | <i>sunu</i> | <i>sunu</i> | <i>fēt</i> |
| GEN | <i>sunu</i> | <i>fōtes</i> | <i>fōta</i> |
| DAT | <i>sunu</i> | OE <i>fēt</i> , OFris. <i>fōte</i> | <i>fōtum</i> |

and Frisian varieties, including both historical and novel irregular plural forms (4.2). The focus of Section 5 is on the theoretical implications of the findings, and finally, the argumentation is completed with a conclusion in Section 6.

2. State of the art

2.1 The origin of present-day irregular plurals

The regular endings of nouns in present-day varieties of West Germanic can be traced back to the Proto-Germanic declensional classes of the *a*-stems and *n*-stems. The *a*-stems are the source of the Modern English and Scots plural ending *-(e)s*, as in *days* < OE *dagas*. Likewise, the North Frisian ending *-er* of masculine nouns is rooted in the same declensional class, e.g. *fasker* ‘fish (PL)’ < OFris. *fiskar*. The dominant plural ending *-en* in West Frisian developed from the earlier *-an*, which was possibly a contamination of the ambiguous ending *-a* from the Old Frisian *n*-stems and the less ambiguous Middle Dutch morpheme *-en* from the same class (cf. Bremmer 2009: 60). Occasionally, some varieties, such as the West Frisian Schiermonnikoog dialect, still retain a distinct feminine inflection, being a continuation of the old feminine *ō*-stem pattern, e.g. *bane* < OFris. *bāna* ‘beans’.

The present-day irregular plurals are predominantly vestiges of the historical minor and unproductive declensional types, including *i*-stems,⁴ *u*-stems, root nouns, *r*-stems (kinship terms), *s*-stems and *nd*-stems (*nomina agentis*). The widespread presence of analogical inflections in the minor paradigm, extended from the productive patterns (*a*-stems, *ō*-stems and *n*-stems), can be observed already in the Old English and Old Frisian period. As the inflections of *i*-stems and *nd*-stems have not left any traces in the modern irregular plural forms, these classes were taken into account only in the examination of the earliest stages of attestation of English and Frisian. The plural forms in these minor declensions were often typified by the presence of allomorphic variation, i.e. the *i*-mutated vowels in the nominative and accusative plural. Additionally, minor declensions are characterised by a specific quantitative profile: Although they are not very large classes, many of the nouns affiliated with them show a high (token) frequency of occurrence/use. Some examples

4. The class of *i*-stems can be viewed as unproductive only from the Anglo-Frisian perspective; the closely related Ingvaemonic Old Saxon testifies to a wide spread of the *i*-stem inflections in the minor stem paradigms, and to a steadily growing productivity of this class. The historical class affiliation of nouns stays in compliance with the classifications found in standard historical handbooks and dictionaries. The complete inventory of lemmas affiliated with the minor declensional classes, as far as they are attested in the modern varieties, is provided in the Appendix.

of such historically-defined irregular plural forms found in Modern English, Scots and Frisian are presented in Table 1, alongside their corresponding forms in Old English and Old Frisian.

Table 1. The irregular plurals continuing the historical inflectional patterns in modern varieties of English and Frisian*

| Historical class | Old English/Old Frisian | Mod. English, Mod. Dutch |
|------------------|--|--|
| root noun | OE <i>brōc</i> : <i>brēc</i> 'trousers' | E. <i>breeches</i> |
| root noun | OFris. <i>gōs</i> : <i>gēs</i> 'goose' | NFris. <i>gus</i> : <i>ges</i> |
| root noun | OE <i>lūs</i> : <i>lȳs</i> 'louse' | E. <i>louse</i> : <i>lice</i> |
| root noun | OE <i>mūs</i> : <i>mȳs</i> 'mouse' | E. <i>mouse</i> : <i>mice</i> |
| <i>u</i> -stem | OFris. <i>lith</i> : <i>litha</i> 'limb, member' | WFris. <i>lid</i> 'limb, member' : <i>lea</i> 'body' |
| <i>s</i> -stem | OE <i>cild</i> : <i>cildru</i> 'child' | Scots <i>child</i> : <i>childer</i> |
| <i>s</i> -stem | OFris. <i>clāth</i> : <i>clāthar</i> 'cloth/clothes' | NFris. [<i>kleet</i>] : <i>kluaser</i> |
| <i>s</i> -stem | OE <i>æg</i> : <i>ægru</i> 'egg' | NFris. <i>ai</i> : <i>aier</i> |
| <i>s</i> -stem | OE <i>cealf</i> : <i>cealfru</i> 'calf' | Scots <i>calf</i> : <i>cair</i> |
| <i>s</i> -stem | OE <i>lomb</i> : <i>lombor</i> 'lamb' | NFris. <i>lum</i> : <i>lumer</i> |

* The following abbreviations were used in Table 1 and 2: E. = English, *EFris.* = East Frisian, *WFris.* = West Frisian, *NFris.* = North Frisian, *infl.* = inflectional.

Being relics of the minor historical patterns, these present-day irregularities are idiosyncratic and entirely lexicalised. In Old English and Old Frisian these plural formations were characteristic of substantially more nouns than the few isolated instances found in the modern varieties (cf. root nouns with 26 lemmas in Old English vs. 6 lemmas in present-day English). As many of the nouns affiliated with the minor classes were highly frequent lemmas, these minor patterns of forming the plural are still relatively easily recognisable in the language.

Apart from the irregularities that must be ascribed to an early historical development, all the investigated modern varieties attest to a range of inflectional innovations whose origin must be traced back to a later stage. Analysing the inventory of historical lemmas and tracing their development to modern times allowed identifying a number of lemmas with irregular plural forms (i.e. irregular from the present-day perspective), which did not represent a continuation of the Old English and Old Frisian plural formation patterns. These novel irregularities often have an indirect linkage to the minor paradigms of Old English and Old Frisian, which means that they are further modifications of the historical patterns, frequently involving mechanisms such as vocalic alternations or suppletion. Table 2 presents selected instances of such irregular plural formations, including their equivalents in the historical classes.

Table 2. Examples of irregular plurals which do not continue the old inflectional patterns in modern varieties of English and Frisian

| Historical class affiliation | Present-day variation | Type of plural marking |
|------------------------------|---|--|
| <i>u</i> -stem | WFris. <i>soan</i> : <i>soannen</i> ‘son’ | vowel alternation + infl. ending |
| <i>r</i> -stem | Scots <i>brother</i> : <i>brether</i> | analogical <i>i</i> -mutation |
| <i>r</i> -stem | E. <i>brother</i> : <i>brethren</i> | analogical <i>i</i> -mutation + new infl. ending |
| <i>s</i> -stem | E. <i>cloth</i> : <i>clothes</i> (vs. <i>cloths</i>) | vowel alternation + infl. ending |
| root noun | WFris. <i>foet</i> : <i>fuotten</i> ‘foot’ | vowel alternation + new infl. ending |
| root noun | WFris. <i>man</i> : <i>manlju</i> | partial suppletion |
| <i>a</i> -stem (n.) | WFris. <i>bern</i> : <i>bern</i> ‘child’ | zero ending |
| root noun | EFris. <i>ku</i> : <i>bäiste</i> ‘cow’ | suppletion |

The new irregularities often involve a combination of the historical inflection (e.g. the *i*-mutated vowel, *r*-formative) and a novel dominant (productive) ending (e.g. *child* : *children*). The emergence of these new irregular patterns was motivated by a number of circumstances and factors, to some extent overlapping with those responsible for the retention of the historical inflectional patterns (cf. Section 4.2.2).

2.2 Controlling factors

The central aim of the present study, as declared in the introduction, is to identify and estimate the significance of the factors which determined the preservation and emergence of irregular plural patterns in North Sea Germanic languages. The factors which will be examined in order to account for the attested plural patterns have been referred to in a number of earlier studies, including, among others, Tiersma (1982), Bybee (1985), Diessel (2007), Dammel, Kürschner & Nübling (2010), Carroll, Svare & Salmons (2012). It is not the intention of our study to refine or extend the list of factors explored in the mentioned publications, but rather to specify which of them have the most significant bearing on the dynamics of the inflectional system in the investigated varieties. The factors are introduced here in a largely random order, and their relative importance will follow from the statistical analyses. The following determinants have been considered relevant for the development of the plural patterns in the investigated languages:

- a. *Frequency of occurrence/use.* Of particular relevance here is token frequency, whose role in morphological change has been defined as a “conserving factor”, i.e. high token frequency is believed to contribute to the conservation of irregular morphology (Greenberg 1966; Bybee 1985; Bybee & Hopper 2001).

- b. *The phonological salience of plural markers.* Phonological salience is defined here in terms of acoustic weight: A zero-ending is less salient than a vowel, which in turn is less salient than a VC-ending (Goldschneider & DeKeyser 2001: 22–23).
- c. *Morpho-phonological complexity of plural marking.* Morpho-phonological complexity is understood here as the complexity of formal marking, i.e. of inflectional exponents. The complexity hierarchy of morphological marking (Corbett et al. 2001: 212–214; Dammel & Kürschner 2008: 248–256) is amply documented in the investigated languages. The types of plural formation attested in the investigated varieties can be ordered on the scale of complexity (from high to low) as follows: lexical suppletion (EFris. *ku* – *bäiste*) > partial lexical suppletion (WFris. *man* – *manlju*) > consonantal root allomorphy (OE *brōc* /bro:k/ – *brēc* /bre:ʃ/) > vocalic root allomorphy (*i*-mutation) > suffixation (cf. Table 1 and 2 and (b) above).
- d. *The interaction of irregular forms with the default (productive) plural markers.* This interaction may be expressed in various ways. In English, the formation of an irregular plural in a noun such as *fish* is most likely triggered by a combination of phonological and semantic factors: a phonological sibilant-sibilant clash (**fish-es*, **goose-s*: push factor), resulting from the attachment of the regular *-es* ending, combined with the semantic content (i.e. game animals) (pull factor).⁵ In various East and North Frisian dialects, the retention of the *-er* marker, pertinent to the historical *s*-stems, was enhanced by the presence of an identical *-er* ending (of a different origin) in the masculine *a*-stems. In both cases, the interaction between the competing markers had some bearing on the introduction or retention of an irregular plural pattern.
- e. *Syllable count.* Most of the nouns showing irregular plural forms are monosyllabic, both in English and Frisian, which implies that the syllable count may be a relevant feature. Accordingly, the number of syllables was included as one of the independent variables in the statistical analysis.
- f. *Semantics.* Various declensional classes show a clustering of semantically related items, such as *r*-stems, covering kinship terms, or *s*-stems, representing predominantly agrarian vocabulary, including animals and other livestock.
- g. *Gender.* Gender is a factor which can potentially play a significant role in the reorganisation of any nominal system, which can be illustrated very well by the developments in the history of German. However, this factor was not considered relevant for the present study for three major reasons: (1) most of the minor classes comprised nouns of different genders, and thus no direct relation existed between a specific inflectional pattern and gender; (2) although gender

5. The terminology (pull vs. push factor) seems justified given that the changes in phonology are considered to have been the primary incentive for the morphological developments. Semantics is accordingly interpreted as a secondary factor, enhancing the activity of the phonological one.

played a significant role in determining the direction of analogical transfers of nouns to regular inflectional classes in the development of Old English and Old Frisian (Adamczyk 2014: 55), its function in the preservation of the archaic forms seems less obvious; (3) gender plays a limited role in the later stages of Frisian and English, where no relation between gender, as far as it is still present, and the choice of a plural ending exists.⁶

- h. *Language contact.* The contact between closely related languages may potentially work as a factor facilitating changes in morphology. All dialects of Frisian have been exposed to extensive language contact for several centuries. West Frisian has been in direct and intense contact with Dutch, which retains no irregular plurals of the type found in Old Frisian, while North and East Frisian have been in close contact with Low German and Danish, both characterised by similar irregular and multiple plural formations. Likewise, English was exposed to intense language contact with Celtic, Norse and French. The aspect of language contact was not included in the model for two reasons; firstly, as both languages experienced fairly intense contact situations over the centuries, it was not expected to be a distinctive feature across the analysed varieties, and secondly, the quantification of a variable as complex as language contact poses many methodological questions.

An interaction of all these factors is expected to have exerted considerable influence on the changes in the inflectional system. Given the potential statistical correlations between them, some of the observed interactions may be instances of confounders (being epiphenomenal in nature), which in a combined statistical model turn out to have no impact. The following two examples illustrate the complexity of potential interactions:

- a. Irregular plurals are the result of historical phonological developments and therefore by definition are relatively old. The older the noun/word, the greater the likelihood that its length was over time reduced by historical reduction processes (apocope). Both irregularity and prosodic structure can be a result of the age of a word. This implies that if we find a correlation between irregularity and prosodic structure, it may be a confounder effect of the age of the noun/word, which is itself not tested as a factor for irregularity. A typical example concerns the strong verbs in Germanic, continuing the Proto-Indo-European verb formation patterns, whose preterite and participle forms are shorter and more irregular than their corresponding forms in the younger and regular weak verbs, e.g. *drink – drank vs. blink – blinked*.

6. One of the anonymous reviewers suggested to evaluate the performance of a statistical model that includes gender as an independent variable. We re-ran the statistical test (see Section 4), including the three genders, but none of them provided a significant contribution to the model.

- b. The effect of the influence of semantics is often difficult to distinguish from frequency effects. For instance, a number of nouns denoting body parts are used predominantly in the plural (e.g. *feet*, *teeth*). Given the fact that absolute and relative frequencies of occurrence of forms have a direct impact on the development and preservation of irregularity, these nouns which are more often used in the plural tend to retain the irregular plural pattern. Accordingly, it is not semantics itself that generates irregularity, but the frequency effects that come with it. In other words, semantics induces high token frequency of these plural forms and high frequency serves as a factor conserving the irregular pattern.

The estimation of the exact role that the individual factors play in the preservation of irregular plural patterns poses some difficulties due to such methodological pitfalls as the existence of confounder factors. This fact certainly needs to be taken into consideration both when designing the statistical procedures and when interpreting the findings.

3. Methodology and database

3.1 Corpus data

The examination of plurality patterns and the evaluation of the conditioning factors was carried out in the present study for English, Scots and a number of varieties of Frisian, at different stages of their development. Accordingly, the investigation covered Old and Modern (standard) English, Early Modern and Modern Scots, and six varieties of Frisian, representing East, West and North Frisian. Scots was included in the study as a variety of English owing to the fact that, in contrast to many existing varieties of English, it shows some inflectional peculiarities, including some inherited archaisms. The basis for the investigation of the Old English material was the data from the *Dictionary of Old English* (electronic corpus) (Healey et al. 2009), covering the period between the 7th and 12th century. The information about present-day English comes from the *Oxford English Dictionary* (1989) and the frequency figures for Modern English were based on the *British National Corpus* (BNC) (2007). The data for Early Modern and Modern Scots come from the *Dictionary of the Scots Language*, comprising two major historical dictionaries of the Scots language (available at <http://www.dsl.ac.uk/>), and were supplemented by the information from a printed edition of the *Concise Scots Dictionary* (Robinson 1987). The analysis of the frequency effects in Scots relied on the frequency figures for English based on the BNC.

As regards the Frisian material, the label “Old Frisian” has been commonly applied to the Frisian attestations from the period between the 13th century and the

early 16th century. The present analysis of Old Frisian was confined to the so-called ‘classical’ period, dated between ca. 1275 and ca. 1410, and attested in 6 manuscripts. This stage of Frisian, despite its relatively young age, is compatible with the other Old Germanic languages in terms of phonological and morphological archaism (Versloot & Adamczyk 2014: 547–548). Present-day Frisian can be essentially viewed as a group of historically related but mutually unintelligible dialects, spoken in various language islands amidst varieties of German and Dutch. These dialects are traditionally grouped into three main branches: West Frisian, used by approximately 450,000 speakers in the Dutch province *Fryslân*; East Frisian, spoken by maximally 1000 inhabitants of the municipality of Saterland, near Oldenburg; and North Frisian, used by ca. 4–5,000 speakers on the western coast and adjacent islands of Schleswig-Holstein. The data from 6 varieties (two from each group) were included in the study. The West Frisian sample included the standard variety which comes close to the idiom of the majority of the West Frisian speakers, and the dialect of Schiermonnikoog, a fairly archaic and dying variety with probably fewer than 100 speakers. East Frisian is represented by the dialect of Saterland, and the dialect of Wangerooge, which died out in the 20th century and was in various respects one of the most archaic varieties of Frisian (albeit not with regard to the formation of the plural). North Frisian, which itself shows a very strong dialectal differentiation, is represented by the dialect of the islands of Föhr and Amrum and by the Mooring-dialect, spoken on the coast.⁷ The frequency figures for modern stages of Frisian were obtainable only for West Frisian and they come from the Frisian Language Database (Versloot & Nijdam 2011). They are based on the language material from the period 1550–1800. As all the varieties are closely related genetically, geographically and culturally, we found it justified to use the robust West Frisian figures as a proxy for the frequencies in the other Frisian varieties (see also Section 3.4).

3.2 Interpretative problems

As discussed in the introductory section, the presence of irregular plurals in modern varieties of English and Frisian can be attributed to both inherited archaisms and newly emerging irregularities. The inherited archaisms are not necessarily irregular from a modern perspective, which can be illustrated by the inflections of the nouns originating in the class of the *u*-stems in the present-day dialect of Saterland. In Old Ems Frisian, from which Saterfrisian is derived, the regular

7. For a thorough introduction to all essential aspects of the Frisian language, including its present-day variation and history, see Munske (2001).

plural ending of masculine nouns in the productive vocalic stems was *-ar*. The ending *-a*, which is the historical plural ending of the *u*-stems, was an exception, and Old Ems Frisian witnessed an incidental analogical extension of *-ar* to the *u*-stems. Saterfrisian generalised the ending *-e* < OFris. *-a* to all nouns ending in a consonant by a synchronic phonological rule that now also applies to the historical *u*-stems, such as *sumer* ‘summer’, *sun* ‘son’ and *winter*. In consequence, the plural form *sune*, being a direct continuation of the archaic plural ending of the *u*-stems (< OFris. *suna*), is entirely regular from a synchronic point of view. Another example of the regularisation of a once archaic inflection comes from the *s*-stems. In the strongly gender-based organisation of declensional classes in Old Frisian, the ending *-ar*, found in the neuter *s*-stems but characteristic of the masculine *a*-stem nouns, was anomalous (Versloot 2014). Likewise, a modern Frisian form such as Saterfrisian *loumere* ‘lambs’, a historical neuter *s*-stem, is an irregularity in the synchronic Saterfrisian system of plural inflections, dominated by *-e* and *-en* as markers of plurality. Accordingly, it is classified in the present study as a ‘preserved archaism’. However, in modern dialects of Frisian, the distribution of nominal inflections is often guided by phonological principles rather than by gender. In the dialect of Föhr and Amrum, the ending *-er*, originally pertinent to masculine nouns, is also commonly found in some feminine and neuter nouns. Despite the gender-based preferences for a specific inflection in this dialect, the distribution is at least partly guided by the synchronic phonological structure of the noun, and consequently the marker *-er* appears in various feminine and neuter nouns in this modern dialect. That is the reason why the ending *-er* in nouns such as *lumer* ‘lambs’ and *kualwer* ‘calves’ is not considered to be irregular from a synchronic point of view in this dialect.

Another interpretative difficulty is posed by nouns which retain traces of the original inflection in their phonological shape, but do not display an irregular pattern synchronically. An example is West Frisian *lea* ‘body’ < OFris. *litha* ‘limbs’, showing the inherited *u*-stem ending in *-a*. The form is petrified as a *plurale tantum* with the meaning ‘body’, and is lexically detached from its original singular *lid*, meaning ‘limb, member’. The form was interpreted in the present study as a reflex of the archaic pattern and counted accordingly in the examination of the preserved archaisms. Another example is the paradigm of *brek*, PL *breken* ‘trousers’, as used in the dialect of Föhr and Amrum. The noun is regular from a synchronic point of view, but the root vowel reflects the earlier presence of *i*-mutation, characteristic of the historical plural forms in root nouns, and its subsequent analogical extension to the singular. The Modern English forms *breech* (rare) – *breeches* reflect exactly the same pattern, with analogical levelling of the vowel from the plural to the singular.

One of the essential methodological questions which arose at the outset of the present investigation was which forms should be interpreted as a 'preserved archaism'. Two approaches were adopted in order to interpret the data. In one of them, the diachronic perspective was taken, and accordingly every form that continues the historical plural pattern (or reflexes of it), in compliance with the phonological rules for each specific dialect, was interpreted as a preserved archaism (e.g. Wangerooge Frisian *kiier* 'cows' with the preserved historical *i*-mutated vowel and an additional ending *-er*, cf. Early Modern English *kine*). In the alternative approach, only those forms that do not match the synchronic patterns of regular plural formation were considered to be preserved archaisms (cf. the examples above). As the focus of the study is on the factors facilitating the survival of historical, non-productive endings in present-day varieties of English and Frisian, examining 'archaisms' that coincide with synchronic, productive plural formations was considered less relevant. Both configurations of input data were tested and it was found that the trend in the obtained results was identical. In both instances, the percentage of plural tokens relative to the total number of attested tokens for a given lemma (henceforth 'plural percentage', %PL) turned out to be the dominant factor.

3.3 Statistical methods

In order to investigate the range of factors which could potentially be involved in the reorganisation of the plural formation patterns in the selected Germanic languages and varieties, a multivariate analysis was applied, whereby multiple factors are quantified as independent variables and the correlation with the dependent variable, i.e. archaism or irregularity of plural inflection, is computed. The dataset comprised all nouns belonging to the minor declensional classes as described in Section 2.1, as long as they are still present in modern varieties of Frisian and English (for a full overview see the Appendix).⁸ The statistical tool particularly suited for the data selected for the present study is a logistic regression analysis, where the dependent variable (the phenomenon that is to be explained) is measured on a binary scale: 1 'yes' or 0 'no'. The logistic regression model includes one or more independent variables that are assumed to influence the choice of either of the two

8. The Appendix contains 33 items for Frisian. *Turf* is probably not irregular in Old Frisian anymore, while *kind* 'child' did not survive into the modern varieties, apart from some instances that are generally interpreted as younger loans from Low German. This leaves for Frisian 31 items in 6 dialects (potentially 186 tokens) in the dataset. Due to other lexical gaps, the total number of tokens found in the corpus was 154.

alternatives in the dependent variable.⁹ These independent variables are compared to the actually observed value of the dependent variable. Given the nature of the plural system, the test is applicable to the investigation of the development of plurality patterns: The concept of plurality in nouns is commonly expressed by one plural form; only in isolated instances do languages show multiple plural forms, which very often have different connotations or semantics. In technical terms, this relation can be expressed on a binary scale, namely, a particular archaic form survives ('yes'/1) or does not ('no'/0), and correspondingly, a particular innovation (here inflectional irregularity) emerges ('yes') or does not ('no'). The independent variables can be measured on a binary scale (e.g. semantic category: 'a word denotes or does not denote an animal'), on an interval scale (e.g. 'a word has one, two or three syllables') or a ratio scale (e.g. 'percentage of plural forms ranges from zero to hundred on a continuous scale'). Logistic regression is therefore suitable for analysing the survival of archaic inflectional forms or the emergence of new ones.¹⁰

In the adopted approach, every single lemma in the data set in each investigated language variety provides one data item. Based on these input data, the model computes whether there is in fact a statistically significant relation between each of the supposed independent variables and the dependent variable; in other words, it tests whether an independent variable contributes to the selection of either of the two alternatives, incorporating also the impact of other variables. It should be borne in mind that not every technically significant correlation expresses a causal relationship between the independent and dependent variable (cf. Section 2.2). In some of the cases identified in the analysed data set, the variable may rather be interpreted as a confounder of other controlling factors. Accordingly, a critical evaluation of the various factors entailed eliminating such potential statistical confounders and identifying causal relations between the independent and dependent variables in the model.

The weight of each independent variable is computed and expressed in the variable coefficient. The combination of the input values and their weight produces a likelihood between 0 and 1 for each item in the data set to display a given property (e.g. to be archaic or not). However, this is a probability rather than a discrete prediction: The expectation is that an item with a high probability to be archaic is factually archaic, but it is not necessarily the case. It is an inherent feature of probabilities that also an item with a low probability to be archaic can in fact be

9. The software used for the regression model was an online calculator available at <<http://statpages.org/logistic.html>> (Pezzullo).

10. For an analysis of the actual diversity of plurality patterns which can involve more than two different exponents (e.g. Modern High German), a range of other statistical methods is available; see, e.g., Daelemans (2002).

archaic. The power of a variable is expressed by the so-called Odds Ratio (O.R.). An Odds Ratio expresses the relative increase of chance for an event in the dependent variable. It means that for every single unit increment in the independent variable, the chance for the presence or absence of a property in the dependent variable (the odds of being ‘yes’) increases by the value of the Odds Ratio. For instance, in the statistical analysis of the level of archaism of plural forms in Frisian (Section 4.2.1.1), the Odds Ratio for the semantic feature *animal* is 17.8. This means that, all other variables being equal, the chance to retain an archaic ending is 17.8 times higher for a noun denoting an animal than for a noun denoting any other concept or object.

The statistical analysis procedure was started with a full inventory of potentially relevant variables (enumerated in Section 3.4). After the evaluation in the logistic regression model, the variables whose contribution was not statistically significant were preselected for elimination, and the variable which contributed least to the explanation was removed from the model, and the new model was evaluated again. This procedure was repeated until it reached the stage when only the variables whose contribution to the variation in the dependent variable was statistically significant were included.

3.4 Operationalising the controlling factors

Operationalising the potential controlling factors (see Section 2.2) involved adding the following potential independent variables to the model (discussed here in detail with reference to the Frisian and English material).

- a. *The absolute number of plural tokens, reflecting the intensity of entrenchment of forms in memory.*¹¹ The frequency figures used are supposed to be a proxy for frequency-driven effects in the transmission of plural forms. As no accurate frequency figures are available for all Frisian dialects, the absolute number was computed as a weighted average of the Old Frisian data (14th century) and the early Modern West Frisian data (1550–1800).¹² This mixture was taken as a good reflex of the frequency figures from the last 500 years of Frisian language history. For English and Scots the figures from the (synchronic) *British National*

11. In other corpus-based studies a *corpus percentage* is occasionally used; it refers to an absolute number of tokens for a specific item (absolute token frequency), divided by the total number of tokens in the corpus. The two methods of quantifying the frequency are interchangeable and render the same results.

12. Human perception of quantities follows a logarithmic, relative scale, rather than an absolute, linear scale (Dehaene 2003). Therefore, the logarithm of absolute token frequencies seemed most appropriate for the type of investigation undertaken in the present study.

- Corpus* were used for this and the following variable. The Old English data, dating back to the 8th–12th century, was found too distant to be representative of the frequency proportions that shaped the present-day language.¹³
- b. *The relative plural proportion, expressing the analogical pressure of either of the two paradigm forms: singular or plural.* The phenomenon of an occasional analogical extension of the plural form to the singular in nouns with high plural proportion has been described in qualitative terms by Tiersma (1982: 835, defined as “local markedness”; cf. Haspelmath 2006). In the present study, the plural proportion for Frisian was computed as the average from the Old Frisian and Early Modern West Frisian data to compensate for biases resulting from the small corpus sizes. However, for lemmas where no plural forms were attested in Old Frisian, the plural percentage figures from the larger Early Modern West Frisian corpus were used. This was considered legitimate given that the absence of these attestations in Old Frisian is most likely to be ascribed to the limited size of the corpus rather than the underlying linguistic reality.¹⁴
 - c. *Salience.* While complexity may be expected to induce simplification, the investigated material indicates that it may also contribute to the retention of morphological irregularity on account of the salience of such irregular plural formations. We therefore combined the phonological salience of the ending and its morpho-phonological complexity into one ordered scale of *perceptual salience* (Goldschneider & DeKeyser 2001, cf. Berman 2003: 275), ranging from the Ø-ending through a vocalic ending, consonantal ending to root-alternation.¹⁵ Suppletion was not found in the analysed Old English and

13. The underlying assumption when selecting corpus data is that the frequency of occurrence of most words will be very similar in languages from a similar region. Some words may vary depending on the landscape and the socio-economic characteristics of a society, and of course over time, with material changes. We computed the correlation between the lemma frequencies in the various datasets, expressed in the correlation coefficient r : Old Frisian – Old English: 0.77; Old English – BNC: 0.75; Old Frisian – early Modern West Frisian: 0.77. The lowest correlation was found between Old English and early Modern West Frisian: 0.54. These strong correlations allow us to conclude that the chosen frequency figures are representative.

14. If only the data from the Early Modern Frisian corpus had been used, the figures for the %PL in Table 4 would change from Coefficient 6.8 and O.R. 871 to 4.8 and O.R.123, which does not affect the observed trends.

15. Bybee (1995) shows that root alternation is not necessarily more difficult to acquire than affixation, which may be used as an argument against the proposed scale of morpho-phonological salience. However, in the context of Germanic languages, where root vowel or consonant alternation is phonologically distinctive, morphologically unpredictable, and is rather used to distinguish different lemmas, root alternation is highly marked and comes close to suppletion. Cf. *goose ~ geese vs. boost – beast, mouse ~ mice vs. spouse – spice.*

Old Frisian data, and accordingly, it does not appear among the independent variables in the present study. The phonological salience of plural marking and the morpho-phonological complexity of the plural marking in Old English and Old Frisian were combined into one variable, ranked on a scale 0 to 1 from the least to the most salient/complex, where 0 = no explicit plural marker (PL = SG. form); 0.33 = vocalic ending *-a*; 0.67 = vowel-consonant ending (OFris. *-ar*, OE *-or*, *-ur*); 1 = root (vowel) alternation (*i*-mutation). The classification was based on the inflections actually attested in Old Frisian and Old English. The quantification of the salience scale was based on a conversion from a rank order to an interval scale. The application of this procedure is supported by the Old English data in Table 3, where salience and the percentage of innovation in the plural show a correlation of 0.95.

- d. *Syllable count*. The syllable structure was quantified by including the number of syllables of the NOM.SG form in Old Frisian and Old English minor stems.
- e. *Semantics*. Three semantic variables were introduced: (1) persons, such as OFris. *mon* 'man', *mōder* 'mother'; (2) animals, such as OE *lamb* 'lamb' or *gōs* 'goose', and (3) body-parts, such as OFris. *lith* 'limb' and *tōth* 'tooth'. These variables were measured on a binary scale.
- f. *The interaction of irregular plural forms with the default plural inflections*. This aspect was not operationalised as an independent variable. As mentioned in Section 2.2 (d), the interaction between irregular and default plural inflections may be expressed in two ways. One of them involves the lexical distribution of regular and irregular plural endings. This aspect was taken into account in that the historically defined plural forms which overlap with the synchronically productive plural endings were not considered to be 'archaic' in the present study (see earlier discussion in Section 3.2). The second manifestation involves phonotactic complications in the concatenation of the root and inflectional ending. This aspect does not play a role in the Frisian dialects, where the regular plural endings are attached mostly in line with the phonological structure of the noun, causing no phonotactic complications. In English and Scots, the nouns *goose*, *louse* and *mouse* terminate in a sibilant and are the potential targets for a sibilant-sibilant clash: **gooses*, **louses*, **mouses*, which might have potentially contributed to the retention of the irregular plural forms (to avoid such a clash). However, all three plural forms are, on account of their plural percentage and salience of the plural marker, sufficiently identified as potentially archaic (i.e. likely to retain their historical form). Therefore, an independent contribution of phonotactics could not be detected and this factor was eliminated from further statistical testing.

4. Results of the analyses

4.1 Results of the analysis of the historical data

The present section is devoted to a discussion of the findings from the examination of the nominal inflection in the Old English and Old Frisian material. The investigation was confined to minor (unproductive) declensional classes whose relics constitute the source of many irregularities in the plural inflection in present-day varieties of English and Frisian.

Table 3 presents a correlation between the type of inflectional marker and the percentage of innovative inflection in the plural paradigms of minor stems, juxtaposing the data from Old English and Old Frisian. The figures come from a systematic investigation of the distribution of archaic and innovative forms in the minor paradigms in the respective corpora discussed in Section 3.1 (Adamczyk 2018). In order to account for the patterns found in the Old Frisian material, the figures for the relative plural proportion were included as well. They refer to the number of plural tokens as a proportion of all tokens of the analysed lemmas.¹⁶

Table 3. The relation between the type of inflectional marker and the level of innovation in the plural paradigms of Old English and Old Frisian minor stems

| Declensional class | Archaic inflectional marker(s) | Percentage of innovation in the plural | Archaic inflectional marker(s) | Percentage of innovation in the plural | Relative plural proportion |
|--------------------|--------------------------------|--|--------------------------------|--|----------------------------|
| | OLD ENGLISH | | OLD FRISIAN | | |
| <i>r</i> -stems | ∅ | 82% | ∅ | 85% | 18% |
| <i>i</i> -stems | - <i>e</i> | 70% | - <i>e</i> , ∅ | 94% | 34% |
| <i>u</i> -stems | vocalic - <i>a</i> | 71% | vocalic - <i>a</i> | 71% | 23% |
| <i>s</i> -stems | <i>r</i> -formative | 47% | <i>r</i> -formative | 51% | 38% |
| <i>nd</i> -stems* | <i>i</i> -mutation | 9% | ∅ | 24% | 58% |
| root nouns | <i>i</i> -mutation | 8% | <i>i</i> -mutation | 20% | 25% |

* The *nd*-stems include only the few monosyllabic lemmas with a high frequency of occurrence. In Old English, the disyllabic stems, not characterised by *i*-mutation, show a different pattern of restructuring, with the level of innovation in the plural reaching 74% (Adamczyk 2018: 218).

The data from the two investigated languages show largely parallel patterns with respect to analogical restructuring of the plural paradigms of minor stems. The scale and dynamics of the reorganisation of the nominal inflection in the individual declensional classes is comparable in the two languages, being slightly more

16. The grey scale applied to the cells referring to the inflectional markers reflects the quantification of the salience scale as discussed in Sections 2.2 and 3.4. Accordingly, light grey colour corresponds to the markers of low salience, while the dark grey colour refers to the more salient ones.

advanced in Old Frisian (which could be expected given the attestation dates of the two sister languages). Some more significant differences emerge with respect to the relation between the amount of innovation and salience of inflectional markers in the individual classes. This refers in particular to the class of *nd*-stems, where, in contrast to Old English, the plural is not characterised by any inflectional exponent in Old Frisian, and the root nouns, which, although marked by the *i*-mutated vowel in the plural in both languages, show a divergent pattern of restructuring. The Old English data attest to a very strong correlation between the salience of the plural marker and the amount of innovation in the plural paradigm; namely, the less salient exponents, such as vocalic endings or zero markers, are much more prone to analogical levelling than the more salient ones, such as *i*-mutation or the *r*-formative. Accordingly, while the amount of analogical inflection in the *r*-stems, where the plural has no inflectional exponent, reaches 82 percent, the root nouns, marked by the *i*-mutated vowel in the plural, display the analogical pattern only in 8 percent of forms.¹⁷

A more independent role of the plural proportion is found in Old Frisian, where the correlation between salience and innovation percentage is less straightforward. The general pattern found in the analysed data is largely the same as in Old English, with two clear exceptions, namely, the resistance of the plural forms of the *nd*-stems to analogical levelling, and the relatively high percentage of analogical inflections in the root noun paradigm. As regards the latter class, the attested pattern can be explained by the fact that, in contrast to Old English, *i*-mutation was not consistently present in the Old Frisian root noun plural paradigm. In fact, with respect to the salience of inflectional markers, this class is a hybrid one in Old Frisian in that it comprises both nouns which never showed a mutated vowel in the paradigm and others which displayed *i*-mutation more systematically.¹⁸ The absence of a mutated vowel as a salient marker of plurality would explain the

17. Additionally, apart from the direct correlation between the salience of inflectional markers and the innovation level, also the relative plural proportion (%PL) turns out to contribute to the final outcome, as was confirmed in a logistic regression model. In order not to overload the paper with technicalities, we decided not to include the technical description of the results of the statistical tests in this section. The major implications are sufficiently illustrated in Table 3.

18. There is a substantial difference between Old English and Old Frisian in the application of *i*-mutation in the NOM/ACC.PL of root nouns. Old English essentially attests to *i*-mutated forms in all investigated nouns. Old Frisian lacks *i*-mutation in nouns such as *burg*, *bök* and *turf*. Plural forms of ‘mouse’ and ‘louse’ are not attested in Old Frisian; in North Frisian, where a range of the *i*-mutated plural forms is preserved, these nouns show endless plural forms without *i*-mutation. Also the attestation of an endless plural form *mus* in 15th century West Frisian implies that the lack of *i*-mutation may have been an original feature in various root nouns. Likewise, Old Frisian *nd*-stems lack *i*-mutation in the plural paradigm, e.g. *fjand* ‘fiend’ – NOM/ACC.PL *fjand*, cf. OE *fēond* – *fjend*.

relatively high percentage of innovation in the plural paradigm in Old Frisian. In order to account for the lack of a direct correlation between the percentage of innovation and salience in the *nd*-stems, the relative plural proportion can be invoked as a second controlling factor: The low salience of the plural forms in the *nd*-stems is compensated by their relatively high frequency of occurrence in the plural. Consequently, the spread of analogical endings in the plural paradigm is limited by the high token frequency of the plural forms, which works as a factor conserving the historical pattern of inflection, inhibiting the dissemination of analogical endings. The two controlling factors, i.e. salience of the inflectional marker and relative plural proportion, are strongly interwoven in the mechanism of the reorganisation (or formation) of plurality patterns, and this relation will be further elaborated on in Section 4.2.

4.2 Results of the statistical analyses

The present section is devoted to a discussion of the findings from the analyses of the preserved archaisms (4.2.1) and new inflectional irregularities (4.2.2). It must be borne in mind that the investigation of the preserved archaisms was confined only to those forms which do not match synchronic, productive formations.

4.2.1 *Archaisms*

4.2.1.1 *Frisian*

The logistic regression analysis was applied first to the Frisian data, beginning with a model including all the mentioned variables, i.e. plural percentage (%PL), the absolute number of plural tokens (logarithm), salience, three semantic categories (*animals*, *body-parts*, *persons*) and syllable count. The syllable count and the semantic aspect of *person* turned out to be absolute (categorical) in their predictions, and therefore technically not suitable for the logistic regression analysis. These factors will be discussed separately. Accordingly, the second version of the logistic regression model contains only the following variables: plural percentage, the absolute number of plural tokens, salience and semantics, including *animals* and *body-parts*. Subsequently, the model was tested again and variables that did not contribute significantly to the result were stepwise eliminated (beginning with the least significant one). This procedure was repeated until the model contained only those independent variables that make a significant contribution to the observed variation in the dependent variable, i.e. the retention of the archaic plural forms. Table 4 presents the descriptives of the final model.

Table 4. Descriptives of the logistic regression model for the retention of archaic plural forms in six varieties of Frisian

| Descriptives: | | | |
|--|----------|---|---|
| The data set contains 154 relevant instances from 31 lemmas from 6 dialects: 123 instances show $Y = 0$ (innovative ending); 31 cases show $Y = 1$ (archaic ending): | | <i>Avg</i> = average <i>SD</i> = standard deviation <i>p</i> = probability value <i>Y</i> = dependent variable | |
| Variable | Avg | SD | |
| %PL | 0.3309 | 0.2395 | <i>df</i> = degrees of freedom = number of independent variables |
| animals | 0.2273 | 0.4191 | The <i>p-value</i> of the total model is less than 0.05 and therefore the model is considered to be statistically significant. |
| body-parts | 0.1299 | 0.3362 | <i>Coeff.</i> expresses the arithmetic weight factor of the variables in the total model, presented alongside the Standard Error (StdErr) and probability (<i>p</i>). The latter has to be below 0.05 to constitute a significant contribution to the total model. |
| Overall Model Fit: | | | |
| <i>ChiSquare</i> = 58.7274; <i>df</i> = 3; <i>p</i> = 0.0000 | | | |
| Coefficients and Standard Errors: | | | |
| Variable | Coeff. | StdErr | <i>p</i> |
| %PL | 6.7700 | 1.4958 | 0.0000 |
| animals | 2.8814 | 0.7151 | 0.0001 |
| body-parts | 1.6674 | 0.7167 | 0.0200 |
| Intercept | -5.7044 | 1.0411 | 0.0000 |
| Odds Ratios and 95% Confidence Intervals: | | | |
| Variable | O.R. | Low – High | |
| %PL | 871.3082 | 46.4478–16344.753 | For the <i>Odds Ratio</i> (O.R.), see Section 3.3. All variables are on a scale between 0 and 1. Given that the Odds Ratio is defined as the increase of likelihood in the dependent variable per single unit increment, the Odds Ratios express in practice the increase of likelihood from the lowest value of the variable to the highest. This makes the Odds Ratios in our model comparable. |
| animals | 17.8389 | 4.3923–72.4514 | The correlation (<i>r</i>) between each of the three variables is lower than 0.5, which is low enough to exclude collinearity. ^a |
| body-parts | 5.2984 | 1.3004–21.5872 | |

^a '*r*' refers to the *Pearson product-moment correlation coefficient* and is a measure for the correlation between two variables: 0 means no correlation, -1 or +1 means a total correlation. Two independent variables with $r < 0.7$ are generally considered to be acceptable in a regression model, without causing confounder effects.

The logistic regression model for the archaic plural forms in six Frisian varieties contains technically three significant, independent variables: plural percentage, *animals* and *body-parts*. The Odds Ratios show that the plural percentage is by far the most relevant one, with a value of 871.3, against 17.8 and 5.3 for the two semantic variables of *animals* and *body-parts*, respectively. The impact of the %PL turns out to be almost 50 times bigger than that of the most influential semantic variable. The evaluation of the relation between semantics and %PL shows that they are strongly interdependent. The average %PL for nouns belonging to the categories of *animals* and *body-parts* is 45%, while for the other nouns in the test set the average

%PL is only 23%. This contrast is significant in a T-test ($p = 0.012$), which means that nouns designating body-parts (e.g. 'eyes', 'legs', 'hands') or animals (e.g. 'geese', 'calves') are relatively more often used in the plural than the other nouns in the data set. This means that the effect of semantics is closely intertwined with that of the plural percentage, which is by far the strongest independent variable in its own right. A conceptual problem in this interpretation is that a causal relation between semantics and *any* irregularity is difficult to account for. More logical from the point of view of language processing is the relation one meaning ~ one form, as in modern English *deer* ~ \emptyset -ending in the plural. Both the group of animals as well as body parts involve various plurality patterns, which undermines the rationale of the relation between semantics and irregularity. Therefore, we conclude that the semantic effect is, in fact, to a large extent a confounder of the plural percentage effect (cf. the discussion in Section 5).

The syllable count can potentially be considered a factor contributing to the retention of archaic plural forms. All nouns retaining synchronically irregular, archaic plural forms are monosyllabic, which makes the effect of the syllable count categorical and, therefore, this factor could not, for technical reasons, be included into the model as an independent variable. It turns out that syllable count and plural percentage correlate strongly ($r = -0.51$): The average %PL in disyllabic nouns in the data set equals 0.09 and corresponds to 0.44 in monosyllabic nouns, which implies that the seeming importance of the syllable count may be a confounder effect of the relative plural proportion. Syllable count also interacts with salience: the disyllabic nouns of relationship (*father*, *mother* etc.; 5 out of the 8 disyllabic instances) lost their Proto-Germanic endings due to their metrical structure, and became largely endingless in Old English and Old Frisian. This made them more susceptible to innovation (see Table 3). Accordingly, it can be concluded that syllable count is a relevant factor that interacts with both frequency and salience, but the extent of this interaction cannot be quantified in this study.

The third semantic category considered in the model was *person*. Its effect was likewise categorical, namely, no items denoting a person retained their historical irregular plural form (although quite a few developed new irregular plurals). Note that five out of the seven nouns which denote the category of person are historical *r*-stems, which were endingless in the plural and therefore susceptible to innovation. Accordingly, as a factor positively contributing to the preservation of archaic forms, this variable was not relevant.

A factor that was expected to make a significant contribution to the statistical model, given its strong correlations with innovation in Old English and Old Frisian, is the salience of the inflectional marker. Its correlation with %PL equals 0.38 ($r = 0.38$). A T-test comparing the plural percentages for nouns with a low

salience to nouns with a high salience shows that the two groups differ significantly ($p = 0.01$): Items with a low salience have on average a lower plural percentage (*average %PL* = 0.22) than items with a high salience (*average %PL* = 0.47). In the versions of the model which included the semantic variables as well, salience was excluded at an early stage of the elimination of the non-significant variables, as its contribution was not statistically significant. Given that the contribution of semantics may conceptually be considered a confounder effect of the plural percentage, we tested an alternative version of the model, which excluded semantics and included only two variables, namely, plural percentage and salience. In this model, which is statistically significant, plural percentage is still by far the most influential variable (*O.R.* = 117.3), while salience shows an Odds Ratio of 5.4 and is significant as well ($p = 0.013$). Table 5 shows the interaction between the salience of inflectional markers, the %PL and the retention of archaic plural forms in the Frisian dialects.

Table 5. The interaction of %PL and salience in the preservation of irregular, archaic plural forms in six Frisian dialects*

| %archaic tokens | %PL | | |
|-----------------|--------|--------|-----|
| salience | < 0.31 | > 0.31 | |
| < 0.4 | 1% | 32% | 11% |
| > 0.4 | 13% | 43% | 36% |
| | 4% | 39% | 20% |

* The thresholds of 0.4 for salience and 0.31 for %PL are the averages in the dataset.

The relatively low contribution of salience to the logistic regression model can be ascribed to the fact that salience itself correlates with %PL ($r = 0.38$): High salience goes with higher %PL. For the variable *plural percentage* the increase of the predictive power is sharper: The amount of archaic inflection increases by ca. 35%-points in the lemmas with %PL below and over the average (39% vs. 4%), but with ca. 25%-points in the lemmas with a more or less salient ending (36% vs. 11%). To sum up, salience turns out to have some intrinsic overlap with plural percentage, and it seems to have had only a limited contribution to the preservation of irregular, archaic plural forms in Frisian. It is, however, conceptually more clearly separable from %PL than semantics: The relation between the meaning of a word and how often it is used in the plural (e.g. *lice* or *teeth*) is much more straightforward than between the relative plural and the phonological shape of the plural form, even when such a relation in the latter case is not entirely absent.

Altogether, it can be concluded that *plural percentage* contributed most substantially to the preservation of archaic, irregular plural forms in the Frisian varieties: When the plural is frequent, the form is simply better remembered. Lemmas with

a low relative plural frequency tend to adhere to the productive patterns of plural formation, because their plural form has too little ‘critical mass’ to be remembered separately from the singular. This plurality effect is most tangible in the two lemmas with the highest plural percentage: ‘clothes’ (0.92) and ‘limb’ (0.84). In both cases, the plural form detached from the singular and became a separate lemma with a *plurale tantum* meaning: ‘clothing’ and ‘body’, respectively. In the dialect of Föhr-Amrum, the noun *brek* shows a similar plural-to-singular transition as the English *breeches*, with further semantic parallels in Modern English *trousers* and Icelandic *buxur* (PL) ‘trousers’.

4.2.1.2 *English and Scots*

The data for Standard English and Scots were examined in a similar way, using the same investigative procedures. The following variables were included in the first version of the model: absolute plural frequency (logarithm), plural percentage, salience and semantics. As both semantic categories of *animals* and *body-parts* had a positive impact on the preservation of archaic forms in the Frisian data set, they were technically combined here into one binary variable. The syllable count was excluded for the same reason as in the examination of the Frisian data, namely, all the preserved archaic irregular forms have a monosyllabic structure (in the singular form), and the effect of this variable was hence categorical, and thus technically not applicable in a logistic regression model (which does not render it meaningless). Just as in Frisian, there is a strong tendency for both the nouns belonging to the mentioned semantic categories and monosyllabic nouns to show relatively high values for plural percentage. For both variables, i.e. semantics and syllable count, this bias towards high values is statistically significant in a t-test ($p < 0.001$). Table 6 presents the descriptives of the model.

The final version of the regression model for English and Scots, just as for Frisian, is much simpler in that it contains fewer variables than the initial version of the model, which included all potential independent variables. In the case of English and Scots, there are two significant, independent variables, namely, the relative plural percentage and salience, both with very high Odds Ratios (%PL: *O.R.* = 85640, salience: *O.R.* = 299). The explanation for the high impact of salience in English and Scots is that, apart from *child(e)r(en)* and Scots *caur* ‘calves’, only root nouns with the *i*-mutated vowel as a marker of plurality (the most salient plural exponent) have been preserved as irregular plurals (at least in this data set containing lemmas originally affiliated with the historical minor classes). Nouns with less salient plural markers in Old English (-∅ and -*a*) have consistently developed an innovative and mostly regular plural ending in Modern English and Scots. Such a strong contrast in the development of nouns with low and high salient markers of plurality produces a high Odds Ratio in the logistic regression model.

Table 6. Descriptives of the logistic regression model for the retention of archaic plural forms in Standard English and Scots**Descriptives:**

The data set contains 56 relevant instances from 28 lemmas from 2 varieties: 38 instances show $Y = 0$ (innovative ending); 18 cases show $Y = 1$ (archaic ending):

| Variable | Avg | SD |
|----------|--------|--------|
| %PL | 0.3550 | 0.2465 |
| Saliency | 0.5711 | 0.4170 |

Overall Model Fit:

Chi Square = 48.9470; *df* = 2; *p* = 0.0000

Coefficients and Standard Errors:

| Variable | Coeff. | StdErr | <i>p</i> |
|-----------|----------|--------|----------|
| %PL | 15.5357 | 6.2965 | 0.0136 |
| Saliency | 5.6988 | 2.4828 | 0.0217 |
| Intercept | -11.6537 | | |

Odds Ratios and 95% Confidence Intervals:

| Variable | O.R. | Low – High |
|----------|------------|--------------------|
| %PL | 85640.4412 | 24.403–377023.0732 |
| Saliency | 298.5154 | 2.2994–38754.3302 |

Avg = average

SD = standard deviation

p = probability value

Y = dependent variable

df = degrees of freedom = number of independent variables

The *p-value* of the total model is less than 0.05 and therefore the model is considered to be statistically significant.

Coeff. expresses the arithmetic weight factor of the variables in the total model, presented alongside the Standard Error (StdErr) and probability (*p*). The latter has to be below 0.05 to constitute a significant contribution to the total model.

For the *Odds Ratio*, see Section 3.3. All variables are on a scale between 0 and 1. Given that the Odds Ratio is defined as the increase of likelihood in the dependent variable per single unit increment, the Odds Ratios express in practice the increase of likelihood from the lowest value of the variable to the highest. This makes the Odds Ratios in our model comparable.

The correlation (*r*) between the two variables is 0.49, which is low enough to exclude collinearity.

The likelihood of preserving an archaic plural form as computed by the logistic regression model, with plural percentage and saliency as independent variables, is presented in Figure 1.

As can be observed, the points in the graph referring to individual nouns follow the so-called S-curve. The steep slope of the curve is an expression of the strong predictive power of the model. The model predicts that nouns with a salient plural marker and a high plural percentage are more likely to retain their historically-defined plural marking than the nouns not characterised by these properties. The predictive power of the model may be strong, but it is not absolute, as can be seen in the case of the two nouns, 'cow' and 'furrow', which, according to the prediction of the model, have a high likelihood to retain their archaic plural forms, yet they show regular plural inflection in modern English. The plural form of 'cow', however, shows archaic traces in Scots and some modern English dialects

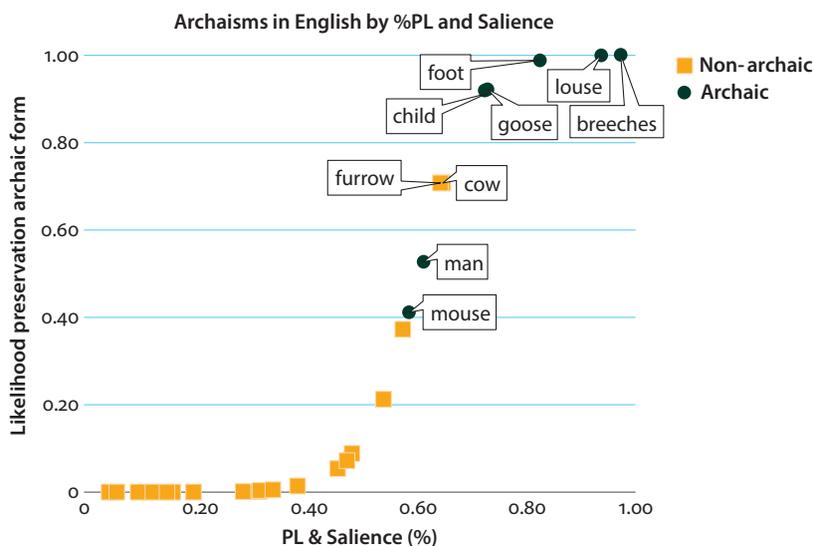


Figure 1. The likelihood of preserving an archaic plural form in English as computed by the logistic regression model, with %PL and saliency as independent variables

(e.g. Cumbria, North Yorkshire): *kine*. The word *furrow*, in turn, violates the monosyllabicity constraint in its modern form (*furrow* < OE *furh*), which is a factor that seems to have had a categorical effect.

To conclude, the archaic plurals in English and Scots are the result of the conserving effect of high plural proportions and the high saliency of the plural marker, i.e. *i*-mutation of the root vowel. Just as it was the case in Frisian, English and Scots testify to the lexicalisation of a plural form, where the plural percentage reaches 85%, i.e. in the ‘plurale tantum’ lemma *breeches*. The saliency of the inflectional exponent turns out to be a much more prominent factor in the regression model for English and Scots than for Frisian, which can be ascribed to the fact that the phonological reduction of endings (apocope) is more advanced in English and Scots than in Frisian. The root alternation induced by historical *i*-mutation was not affected by reduction processes and was therefore the most resistant irregular plurality marker. Not without significance is also the fact that *i*-mutation was more widespread and consistent in Old English than in Old Frisian (cf. footnote 18).

4.2.2 New irregularities

In Section 2.1, Table 2, new inflectional irregularities present in modern English and Frisian were introduced, discussed and contrasted with the preserved archaisms. They referred only to irregularities within the inventory of lemmas affiliated with the minor classes. These new inflectional irregularities may range from minor phonological allomorphy, such as in English *clothes* ‘clothing’ vs. regular *cloths*, to

far-reaching lexical innovations, such as Saterfrisian suppletive *bäiste* as a plural of *ku* ‘cow’. The second stage of the present investigation involved an examination of these new irregularities, and accordingly the data set included only non-archaic forms, i.e. those which were either regular or irregular, but did not constitute a historical irregularity. The Old Frisian word **gōs*, PL *gēs* ‘goose’ may serve as an example of the divergent development of the plural form in modern dialects, which necessitates a different interpretation of forms in the individual varieties. More specifically, in the evaluation of archaisms, the Föhr-Amrum Frisian plural form of *gus*: *ges* was counted as an archaic feature, while Wangerooge Frisian *goos*, PL *gooz* (with a final voiced consonant from the earlier **gooze*) was considered non-archaic. In the investigation of the new irregularities, the Föhr-Amrum form was discarded from the calculations, being an inherited irregularity, while the Wangerooge form was now counted as a ‘new irregularity’ (**goos* or **goosen* would be the expected regular forms in this dialect). In the dialect of Schiermonnikoog, the noun has a regular plural form of the feminine inflection: *goze*, complying with the inflectional pattern in this specific dialect. This form is counted as ‘regular’, both in the study of archaisms in Section 4.2.1 and in the investigation of new irregularities.

As present-day English and Scots display too few new irregularities in the analysed set of lemmas to make an interesting statistical case, the present investigation was confined only to the data from the varieties of Modern Frisian. This does not imply that such irregularities are entirely non-existent in English or Scots, e.g. the form *clothes* is found alongside the regular form *cloths*, *brethren* alongside the regular *brothers* (Scot. *brether*) (and beyond the current set of lemmas, e.g., Scots *treen* ‘trees’, *breer* ‘eyebrows’). The following variables were included in the first version of the model: the absolute number of plural forms (#PL), plural percentage, three semantic variables (*animals*, *body-parts*, *persons*), salience and syllable count. After repeated elimination of the least significant variable, the total model emerged as significant ($p < 0.001$), with all semantic variables plus syllable count being individually significant ($p \leq 0.02$). Salience did not make a statistically significant contribution this time, which was expected; after all, it is rather unlikely that the salience of historical endings had an impact on the irregularities that developed at a later stage. When tested individually in a Fisher’s Exact Test, only the factor *animals* showed a significant correlation with irregularity ($p < 0.001$). It is the interplay between syllable count and semantics that creates a strong statistical model.¹⁹ The cognitive rationale for such a scheme is hard to imagine. It was concluded in the previous sections that the impact of the semantic factors is potentially a confounder

19. Irregular forms are more likely to be monosyllabic. Of all 12 monosyllabic lemmas with less than 50% of irregular forms, only two of them belong to any of the mentioned semantic categories, and none to the category of *persons*. The correlation (r) between the percentage of irregular forms per lemma and the variable *person* itself is only 0.01.

of plural percentage, while word length (including the number of syllables) was found to correlate with absolute frequency. The relation between semantics and irregularity will be discussed in Section 5.

We tested an alternative version of the model excluding semantics. In this model, syllable count turned out not to be significant. The final version of the model without semantics contained only the absolute number of plural tokens (#PL) as independent variable. The complete results of the statistical model are presented in Table 7.²⁰

Table 7. Descriptives of the logistic regression model for innovative irregular plural forms in six varieties of Frisian

Descriptives:

The data set contains 138 relevant instances from 29 lemmas from 6 dialects: 91 instances show $Y = 0$ (regular ending); 47 instances show $Y = 1$ (innovative irregular ending):

| Variable | Avg | SD |
|----------|--------|--------|
| #PL | 0.5149 | 0.2673 |

(the #PL was normalised to a range 0–1)

Overall Model Fit:

Chi Square = 10.8875; *df* = 1; *p* = 0.0010

Coefficients and Standard Errors:

| Variable | Coeff. | StdErr | p |
|-----------|---------|--------|--------|
| #PL | 2.3306 | 0.7421 | 0.0017 |
| Intercept | -2.5124 | | |

Odds Ratios and 95% Confidence Intervals:

| Variable | O.R. | Low – High |
|----------|---------|----------------|
| #PL | 10.2840 | 2.4016–44.0379 |

The absolute number of plural tokens contributed significantly to the development of new irregular forms. Many of the new irregularities involve phonology-driven innovations, such as the West Frisian and Saterfrisian instances of shortening of the root vowel *luus* [lu:s] ‘louse’, PL *luze* [luzə] ‘lice’ (Saterfrisian) (word-isochrony effect). Such articulatory reduction effects tend to come with high absolute frequency of a form.²¹ High frequency lemmas are more easily and quickly retrieved

20. For the explanation of the abbreviations used in the table, see Table 4 and Table 6.

21. All new irregular plurals in, for example, West Frisian are instances of vowel shortening in the plural, such as *mûs* [mu:s] – *mûzen* [muzən] or *keal* [kr̥əl] – *keallen* [kjelən], a process that is triggered by high frequency of use.

and processed, and hence undergo stronger articulatory reduction, which, precisely due to the high frequency of occurrence, is more acceptable for the listener as these high frequency items are more easily recognised (Ernestus 2000, Diessel 2007: 114).

Figure 2 illustrates the emergence of new irregularities in the investigated varieties of Modern Frisian as determined by frequency, both absolute and relative. The token frequency (#PL) is plotted on the Y-axis and plural percentage (%PL) on the X-axis. The orange circles refer to lemmas which have new irregular plural forms in less than 50% of the relevant forms from the six investigated Frisian dialects. The black dots refer to lemmas which show new irregular plural forms in more than 50% of the relevant forms.

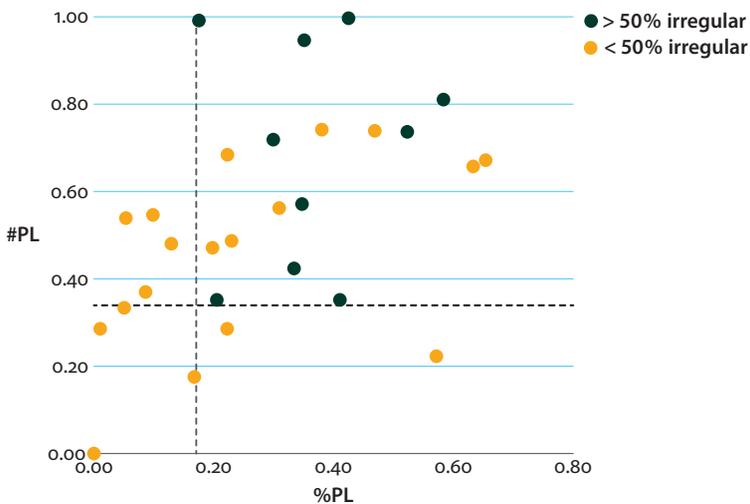


Figure 2. The emergence of new irregular plurals in the Frisian varieties, controlled by token frequency and plural percentage

As the data indicate, there is a clear division in the distribution of lemmas in the data set: 19 lemmas (marked in the graph with an orange circle) developed new irregular plural forms in 40% or less of all the relevant forms in the 6 dialects (disregarding the archaisms). In the remaining 10 lemmas, 66% or more of the relevant forms developed a new irregularity in the plural. The lemmas with many new irregularities are clustered in the upper-right corner of the graph, delimited by a minimum of 0.34 for the #PL and 0.17 for the %PL.²² This implies that the emergence of new irregularities is conditioned by a high plural percentage and in

22. Note that the #PL was rescaled to the range 0–1 to make its impact in the logistic regression model comparable to other binary factors and the %PL-probability, which also ranges between 0 and 1.

particular by a high absolute frequency of plural forms. The lack of a statistically significant contribution of %PL to the model is remarkable. The distribution presented in Figure 2 suggests that a minimum of relative plural frequency is still needed to develop a new irregular plural form. If one compares the impact of high %PL versus low %PL ($low < \%PL < high$) on the emergence of new irregularities, there is a significant impact of %PL in the data ($p = 0.01$, $\phi = 0.22$).²³ A similar procedure for #PL gives $p = 0.28$, rendering it statistically insignificant.²⁴ With some caution, we therefore conclude that both factors make a meaningful contribution to the model. The effect of plural percentage was already noticeable in the analysis of the archaisms: Learning irregular plural forms and storing them in memory alongside the ‘unmarked’ singular forms requires a substantial proportion of plural occurrences of a lemma; otherwise, the analogical pressure will oust any irregular form.²⁵

4.2.3 Synchronic complexity of plural formations

The complexity of the plural formation system results from the combined effect of the historically-defined (inherited) irregularities and the newly emerging ones. All the synchronically irregular plural forms can be considered complex from the perspective of a language learner in that they require more learning effort than the regular plural forms. In Figure 3, the percentage of preserved archaisms is plotted against the overall percentage of forms which are irregular from the synchronic perspective and thus contribute to the complexity of the morphological system, being either inherited or constituting new irregularities.²⁶

A striking feature of the distribution in Figure 3 is an evident lack of correlation between the overall proportion of irregularity (Y-axis) and the proportion of archaic forms (X-axis) found in the investigated material, which is illustrated by the low value of $r^2 (= 0.06)$. One might expect that the preservation of many archaic plural forms will contribute to the overall proportion of irregularities in a language,

23. Fisher’s Exact test. <<http://vassarstats.net/tab2x2.html>>

24. Both variables can be assumed to be normally distributed ($p > 0.05$ in an Anderson-Darling normality test).

25. There are several reasons why %PL does not come out as significant in the regression model, in contrast to #PL. The threshold value for #PL (0.34) is higher and hence more significant than the one for %PL (0.17). Moreover, the lemmas with the highest %PL have few new irregular forms, exactly because quite a few of them have archaic irregular forms. The top 4 in #PL are all strongly innovative, on the other hand. Another technical complication is the high correlation (r) between the two variables: 0.51.

26. The abbreviations used in Figure 3 refer to the following varieties of Frisian: *F-A* = Föhr-Amrum, *Moor* = Mooring Frisian, *Wang* = Wangerooge Frisian, *WF* = West Frisian, *Sat* = Saterland Frisian, *Sch* = Schiermonnikoog Frisian.

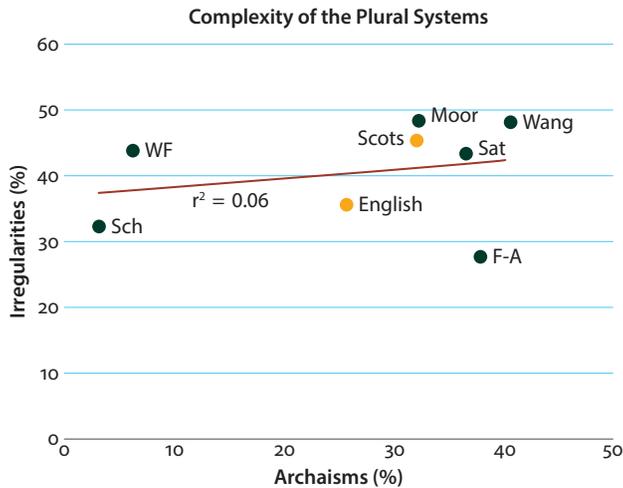


Figure 3. The complexity of the inflectional system reflecting the cumulative effect of preserved and novel irregularities²⁷

reflecting some higher level of complexity of the morphological system. However, the graph illustrates that – at least within the investigated subset of lemmas – there is no correlation between the proportion of preserved archaisms and the total number of irregular forms. In other words, irrespective of the level of preserved archaisms, all investigated varieties, both the Frisian and the English ones, have a similar percentage of synchronically irregular plural forms, ranging roughly between 30 and 50 percent. This level of irregularity resembles a ‘natural’ balance between the regular and irregular forms: When archaic irregularities are not preserved, new irregularities develop to ‘fill the gaps’, i.e. to maintain the equilibrium.²⁸

So far in the study, the new irregular plural forms were not split into any further categories with regard to their complexity or salience. A closer look at the data

27. The X-axis shows the percentage of the preserved archaisms, irregular from the synchronic point of view; the Y-axis shows the percentage of the irregular plural forms, including both archaisms and new irregularities. The relatively low percentage of overall irregularity (Y-axis) for the Föhr-Amrum dialect is a result of the choice to interpret the plurals in *-er* in the historical *s*-stems as regular; otherwise the level of irregularity would reach 34%.

28. An unpublished parallel study of irregular plural formations in various Scandinavian languages (Eekman 2013) rendered very similar results and testifies to parallel patterns: Of 7 varieties of North Germanic (Icelandic, Faroese, New-Norwegian, Norwegian Bokmål, Swedish, Elfdalian and Danish), Elfdalian exhibited the lowest level of archaisms, being at the same time a language characterised by a very complicated plural morphology, with many inflectional sub-patterns, instigated by synchronic phonological phenomena, such as vowel harmony, vowel balance and level stress (Åkerberg 2012).

suffices to observe that some of the new irregularities are the result of phonological changes, such as shortening of the vowel in the plural (e.g. West Frisian *laam* – *lammen* ‘lamb(s)’); others show rare, new plural inflections (e.g. English *brethren*, being a combination of analogical *i*-mutation and the once productive *-en* marker), or are instances of lexical innovations (e.g. Mooring Frisian *moon* – *kjarls* ‘man – men’). In most cases, the newly emerging irregularities are phonology-driven. This complies with the earlier observation that the absolute number of plural tokens was a statistically significant factor in the model: A high absolute frequency facilitates both the development and lexicalisation of allegro-speech phenomena. In other words, the advancement of phonology-induced changes is enhanced by high absolute frequency.

5. Discussion: Frequency and semantics

The evaluation of the controlling factors evoked in the earlier studies and explored in detail in Section 4 reveals that few of them are actually relevant for the morphological phenomena investigated in the present study, i.e. the preservation of archaic forms as irregular plural formations and the emergence of new irregular plural forms. Some of the presumed factors have no relevance whatsoever for the investigated developments (e.g. the absolute frequency of plural tokens for the preservation of archaisms); others seem to be confounders, especially the semantic variables and the syllable structure.

The findings of the present study have some more theoretical implications, allowing some insight into the way language actually works: how it is learned, stored and retrieved. The factors of prime importance for the mechanism of morphological change are the absolute number of plural tokens (#PL), the plural percentage (%PL) and salience. All three are involved in the strength of the cognitive representation of plural forms. The number of plural tokens is a proxy for the entrenchment of the form in memory, which is shown to have consequences for the accuracy of articulation (e.g. Ernestus 2000, Diessel 2007). The relation between this factor and phonology-based innovative plural forms follows directly from this articulatory routinisation effect. The plural percentage expresses the strength of the form relative to other related morphological forms with the same meaning. The access to a plural form that is not deduced from the (mostly more frequent) singular by default plural endings is easier when the form is relatively frequent and hence more often activated (cf. Haspelmath 2008: 60–63). This stays in compliance with the observation that any knowledge or skill that is regularly activated is less easily forgotten. The case of strong preterite forms, which are retained in the most frequent verbs, can serve as another example (Lieberman et al. 2007). In both manifestations of frequency, the strength of the representation in memory enables easier processing.

Finally, salience of the plural marker is an expression of the qualitative strength of the entrenchment in memory. Just as with every other feature that a human mind remembers, whether a picture or an event, remarkable, outstanding features are more easily remembered, especially when they are relatively rare (which is the case with the more salient plural markers). Another factor which may have some additional bearing on the entrenchment of the forms is related to acoustic salience (which is incorporated in the presented salience-scale), and more specifically, the fact that the acoustically more salient endings are less easily reduced or missed.²⁹ In conclusion, all factors that emerge from our analysis as being statistically relevant can be directly reduced to the underlying cognitive processes, without the necessity to employ any abstract set of linguistic constructs or formalisations.

An intriguing role in the mechanism of the retention and emergence of irregular plurals is played by semantics, which turned out to have a statistically significant contribution to the archaisms and the emergence of new irregularities in Frisian. A major problem for the interpretation of these results is that the contribution of semantics to the formation or retention of the archaisms in terms of cognitive processing effects is far less transparent than for the frequency factors. The correlation between the semantic categories and the plural percentage (%PL) implies that the former can be significant, because the nouns in these categories denote concepts that often appear in the plural. It does not mean, however, that semantics on its own cannot be a factor determining the shape of the plural inflection. An example may be the small group of English endingless plurals denoting game animals, which contains not only historically-determined nouns, such as *deer* and *sheep* (deriving from the neuter *a*-stems), but also a number of newly acquired items, affiliated originally with other declensional classes, e.g. *fish* or *moose* (Ekwall 1975: 92). This brings the concept of analogy into play: The extension of the irregular plurality pattern in the above-mentioned nouns is a result of analogical association. Analogy involves the activation of any type of knowledge in the brain as an effect of any similarity between two entities (Fertig 2013: 12). Given the nature of the human language as an abstract system of signs, composed of the *signifiant* (form) and *signifié* (meaning), as defined by Saussure (1916), this similarity can be found either in the form (ranging from phonetics to syntactic structures) or in the meaning. Activation of one form means automatic co-activation of all items in all categories of which this form is a member (see Skousen (1989) for a model-wise implementation of this concept). In the context of the distribution of archaic or innovative irregular endings, we observed a correlation between high absolute and relative plural percentages on the one hand and the semantic categories of nouns denoting animals,

29. See Goldschneider & DeKeyser (2001: 36) for a discussion of various aspects of salience and their relevance in language acquisition.

persons or body parts on the other, irrespective of the causal relation between them. Originally (in Old English/Frisian) there was no particular association between the irregular plurals and lemmas from specific semantic categories (i.e. body parts and animals). However, the specific semantic profile of some of these lemmas and their high frequency of occurrence (either relative or absolute) created a relative overrepresentation of these semantic categories among the irregular plurals (but not automatically the reverse). The consequence of the strong association of irregular plurals and specific semantic categories is that every time a specific irregular (archaic or new) plural form is produced, these semantic categories (body parts and animals) are co-activated in the brain. In other words, an activation relation is established between the (irregular) plural formation and the semantic category. Even if the original causal direction runs from frequency profile to semantics, analogical association works in a way that this relation, once it emerges, can also be activated in the reverse direction. As the statistical correlations in the regression model are non-directional in terms of causality, the model correctly identifies the existing association between semantics and irregular morphology. For the synchronic (statistical) language learner, the diachronic causal relations are irrelevant; it is rather the model which includes semantics as an independent variable that represents the learner's reality. This implies that semantics is a factor that helps the language learner acquire these irregular forms.

Consequently, non-default (unproductive or irregular) plural formations can obtain a (limited) productivity based on the semantic features of lexical items, as in the case of English endingless plurals denoting animals (*sheep, deer*, etc.). This course of development seems to have had a very limited application in Frisian, because most of the new irregular formations denoting animals in the Frisian dialects are of a phonological nature, involving shortening of long vowels in the plural (word-isochrony). An example representing the mechanism of analogy-driven-by-semantics is found in the Wangerooge dialect, with the extension of the ending *-er* to nouns denoting livestock and human beings belonging to the household. The core is comprised of the former *s*-stems *laumer* 'lambs', *kalver* 'calves' (perhaps also *ooier* 'eggs'), enforced by the original masculine *a*-stem *hingster* 'horses'. The plurals *kiier* 'cows' (OFris. *kī*) and *schaiper* 'sheep', replacing an endingless Old Frisian form (cf. English *sheep*), can be considered the direct result of semantic association with the nouns with a historical plural in *-er*. Other nouns referring to animate members of the household seem to have joined this class: *beener* 'children', *fauner* 'girls', *fenter* 'boys', *knechter* 'servants', *sjeeler* 'men', *wüifer* 'wives'.

Finally, the differences between the patterns of preserved plural irregularities in the varieties of English and Frisian can be attributed to a range of interacting factors, the most significant of them being phonological. They involve the different dynamics of the phonological developments (especially the reduction process),

which have been more advanced in English than in Frisian, and the different scope of implementation of *i*-mutation, which was a more regular and consistent process in early English than early Frisian. The disparate patterns found across the Frisian varieties can be ascribed largely to the contact situation, which differed substantially for individual dialects. Accordingly, while West Frisian dialects tended towards greater regularisation of the inflectional system under the influence of Dutch, the geographical and thus linguistic isolation of the North Frisian varieties allowed them to retain considerable morphological diversity, including the preservation of many inherited plural irregularities. Other factors, such as the size of the linguistic communities or the socio-economic situation (e.g. the isolation of the North Frisian and East Frisian relic areas), must have had further bearing on the attested linguistic patterns. This non-linguistic conditioning of the morphological changes, as much as it can be revealing about the linguistic shape of the individual varieties, stays beyond the scope of the present study.

6. Conclusions

The examination of old and new irregular plural forms in the investigated varieties of English and Frisian allows us to conclude that the developments affecting the plural inflection are guided largely by the frequency profiles of individual lemmas and their paradigm forms (in this study the plural form). The absolute and relative frequencies of occurrence of a specific paradigm form are the best predictors for the survival or the development of irregular forms in the consecutive generations of language learners. The identified frequency effects can be directly linked to cognitive aspects of storage and retrieval of linguistic information in the human brain. The same holds true for the aspect of salience, which emerged as particularly relevant for the explanation of the irregularities continuing historical inflections. The findings of the present study indicate that semantics may have a substantial bearing on the frequency patterns, but it is frequency that determines the chance of a given plural form/pattern to survive or emerge. This does not preclude the fact that analogical levelling based on secondary activation of semantic features can lead to a proliferation of new morphological irregularities.

Acknowledgements

We are very grateful to the anonymous reviewers as well as to the editors of the volume for their insightful remarks and comments, which helped enhance the line of argumentation and thus the quality of the paper.

References

- Adamczyk, Elżbieta. 2010. Morphological reanalysis and the Old English u-declension. *Anglia. Zeitschrift für Englische Philologie* 128(3): 365–390.
- Adamczyk, Elżbieta. 2013. On morphological realignments in Old English root nouns. *Transactions of the Philological Society* 111(3): 274–300. <https://doi.org/10.1111/j.1467-968X.2012.01319.x>
- Adamczyk, Elżbieta. 2014. Patterns of reorganization of the early English nominal inflection. The case of Old English consonantal stems. *NOWELE* 67(3): 51–75. <https://doi.org/10.1075/nowele.67.1.03ada>
- Adamczyk, Elżbieta. 2018. *Reshaping of the Nominal Inflection in Early Northern West Germanic* [Nowele Supplement Series 31]. Amsterdam: John Benjamins. <https://doi.org/10.1075/nss.31>
- Åkerberg, Bengt. 2012. *Älvdalsk grammatik*. Älvdalen: Ulum Dalska.
- Berman, Ruth A. 2003. Children's lexical innovations: Developmental perspectives on Hebrew verb structure. In *Language Processing and Acquisition in Languages of Semitic, Root-Based, Morphology* [Language Acquisition and Language Disorders 28], Joseph Shimron (ed.), 243–292. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.28.13ber>
- Bremmer, Jr., Rolf H. 2009. *An Introduction to Old Frisian: History, Grammar, Reader, Glossary*. Amsterdam: John Benjamins. <https://doi.org/10.1075/z.147>
- BNC: The British National Corpus, version 3 (BNC XML Edition). 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. <<http://www.natcorp.ox.ac.uk>> (1 September 2014).
- Bybee, Joan L. 1985. *Morphology: A Study of the Relation between Meaning and Form* [Typological Studies in Language 9]. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.9>
- Bybee, Joan L. 1995. Are stem changes as natural as affixes? *Linguistics* 33(4): 633. <https://doi.org/10.1515/ling.1995.33.4.633>
- Bybee, Joan L. & Hopper, Paul J. 2001. Introduction to Frequency and the Emergence of Linguistic Structure. In *Frequency and the Emergence of Linguistic Structure* [Typological Studies in Language 45], Joan Bybee & Paul Hopper (eds), 1–26. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.45.01byb>
- Carroll, Ryan, Svare, Ragnar & Salmons, Joseph C. 2012. Quantifying the evolutionary dynamics of German verbs. *Journal of Historical Linguistics* 2(2): 153–172. <https://doi.org/10.1075/jhl.2.2.01car>
- Corbett, Greville, Hippisley, Andrew R., Brown, Dunstan & Marriott, Paul. 2001. Frequency, regularity and the paradigm: A perspective from Russian on a complex relation. In *Frequency and the Emergence of Linguistic Structure* [Typological Studies in Language 45], Joan L. Bybee & Paul J. Hopper (eds), 201–226. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.45.11cor>
- Daelemans, Walter. 2002. A comparison of analogical modeling of language to memory-based language processing. In *Analogical Modelling: An Exemplar-based Approach to Language* [Human Cognitive Processing 10], Royal Skousen, Deryle Lonsdale & Dilworth B. Parkinson (eds), 157–179. Amsterdam: John Benjamins. <https://doi.org/10.1075/hcp.10.13dae>
- Dammel, Antje & Kürschner, Sebastian. 2008. Complexity in nominal plural allomorphy. In *Language Complexity: Typology, Contact, Change* [Studies in Language Companion Series [Studies in Language Companion Series 94], Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds), 244–261. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.94.15dam>

- Dammel, Antje, Kürschner, Sebastian & Nübling, Damaris. 2010. Pluralallomorphie in zehn germanischen Sprachen; Konvergenzen und Divergenzen in Ausdrucksverfahren und Konditionierung. In *Kontrastive germanistische Linguistik*, Vol. 2 [Germanistische Linguistik 206–209], Damaris Nübling (ed.), 587–642. Hildesheim: Olms.
- Diessel, Holger. 2007. Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology* 25(2): 104–123.
<https://doi.org/10.1016/j.newideapsych.2007.02.002>
- Dehaene, Stanislas. 2003. The neural basis of Weber-Fechner's law: Neuronal recordings reveal a logarithmic scale for number. *Trends in Cognitive Science* 7: 145–147.
[https://doi.org/10.1016/S1364-6613\(03\)00055-X](https://doi.org/10.1016/S1364-6613(03)00055-X)
- Eekman, Stephan. 2013. *The Diachronic Development of Irregular Noun Morphology in North Germanic Languages: A Logistic Regression Analysis* (MA-paper). Universiteit van Amsterdam.
- Ekwall, Eilert. 1975. *A History of Modern English Sounds and Morphology*, Alan Ward (trans.). Oxford: Basil Blackwell.
- Enger, Hans-Olav. 2004. On the relation between gender and declension. A diachronic perspective from Norwegian. *Studies in Language* 28: 51–82. <https://doi.org/10.1075/sl.28.1.03eng>
- Ernestus, Mirjam T. C. 2000. *Voice Assimilation and Segment Reduction in Casual Dutch: A Corpus-Based Study of the Phonology-Phonetics Interface*. Utrecht: LOT.
- Fertig, David. 2013. *Analogy and Morphological Change*. Edinburgh: EUP.
- Gardani, Francesco. 2013. *Dynamics of Morphological Productivity. The Evolution of Noun Classes from Latin to Italian*. Leiden: Brill. <https://doi.org/10.1163/9789004244658>
- Goldschneider, Jennifer & DeKeyser, Robert, 2001. Explaining the 'natural order of L2 morpheme acquisition' in English: A meta-analysis of multiple determinants. *Language Learning* 51: 1–50. <https://doi.org/10.1111/1467-9922.00147>
- Greenberg, Joseph H. 1966. *Language Universals, with Special Reference to Feature Hierarchies*. The Hague: Mouton.
- Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42(1): 25–70. <https://doi.org/10.1017/S0022226705003683>
- Haspelmath, Martin. 2008. Reply to Haiman and Croft. *Cognitive Linguistics* 19(1): 59–66.
<https://doi.org/10.1515/COG.2008.004>
- Healey, Antonette di Paolo, Holland, John, McDougall, David, McDougall, Ian & Xiang, Xin (eds) 2009. *The Dictionary of Old English Corpus in Electronic Form*. Toronto: Toronto University Press. <<http://www.doe.utoronto.ca/pages/pub/web-corpus.html>>
- Hogg, Richard M. & Fulk, Robert D. 2011. *A Grammar of Old English, Vol. 2: Morphology*. Chichester: Wiley-Blackwell.
- Kastovsky, Dieter. 1995. Morphological reanalysis and typology: The case of the German r-plural and why English did not develop it. In *Historical Linguistics: Selected Papers from the 11th International Conference on Historical Linguistics, Los Angeles, 16–20 August 1993* [Current Issues in Linguistic Theory 124], Henning Andersen (ed.), 227–238. Amsterdam: John Benjamins.
- Kastovsky, Dieter. 1997. Morphological classification in English historical linguistics: The interplay of diachrony, synchrony and morphological theory. In *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen* [Mémoires de La Société Néophilologique de Helsinki 52], Terttu Nevalainen & Leena Kahlas-Tarkka (eds), 63–75. Helsinki: Société Néophilologique.

- Kürschner, Sebastian. 2008. *Deklinationenklassen-Wandel: Eine diachron-kontrastive Studie zur Entwicklung der Pluralallomorphie im Deutschen, Niederländischen, Schwedischen und Dänischen*. Berlin: Walter de Gruyter.
- Lieberman, Erez, Michel, Jean-Baptiste, Jackson, Joe, Tang, Tina & Nowak, Martin A. 2007. Quantifying the evolutionary dynamics of language. *Nature* 449 (7163): 713–716. <https://doi.org/10.1038/nature06137>
- van Loey, A. 1976. *Middel nederlandse spraakkunst, Deel I: Vormleer*. Groningen: H.D. Tjeenk Willink.
- Munske, Horst Haider (ed.). 2001. *Handbuch des Friesischen (Handbook of Frisian Studies)*. Tübingen: Max Niemeyer. <https://doi.org/10.1515/9783110946925>
- Nielsen, Hans Frede. 1981. *Old English and the Continental Germanic Languages: A Survey of Morphological and Phonological Interrelations*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Nielsen, Hans Frede. 2000. *The Early Runic Language of Scandinavia: Studies in Germanic Dialect Geography*. Heidelberg: Winter.
- Paul, Hermann. 1920. *Prinzipien der Sprachgeschichte*, 5th edn. Halle (Saale): Max Niemeyer.
- Pezzullo, John C. Logistic Regression. <<http://statpages.org/logistic.html>> (4 January 2015)
- Ringe, Don. 2006. *A Linguistic History of English, Vol. 1: From Proto-Indo-European to Proto-Germanic*. Oxford: OUP. <https://doi.org/10.1093/acprof:oso/9780199284139.001.0001>
- Robinson, Mairi. 1987. *The Concise Scots Dictionary*. Aberdeen: Aberdeen University Press.
- de Saussure, Ferdinand. 1916. *Cours de linguistique générale*, Charles Bally, Albert Sechehaye & Albert Riedlinger (eds). Paris: Payot.
- Schwink, Frederick W. 2004. *The Third Gender. Studies in the Origin and History of Germanic Grammatical Gender*. Heidelberg: Winter.
- Simpson, John A. & Weiner, Edmund (eds). 1989. *The Oxford English Dictionary*. Oxford: OUP. <<http://www.oed.com/>>
- Skousen, Royal. 1989. *Analogical Modeling of Language*. Dordrecht: Kluwer.
- Steller, Walther. 1928. *Abriss der altfriesischen Grammatik, mit Berücksichtigung der westgermanischen Dialecte des altenglischen, altsächsischen und althochdeutschen*. Halle (Saale): Max Niemeyer.
- Stolz, Thomas, Otsuka, Hitomi, Urdze, Aina, van der Auwera, Johan. 2012. Introduction: Irregularity – glimpses of a ubiquitous phenomenon. In *Irregularity in Morphology (and Beyond)*, Thomas Stolz, Otsuka Hitomi, Urdze Aina & Johan van der Auwera (eds), 7–37. Berlin: Akademie Verlag. <https://doi.org/10.1524/9783050059587>
- Strik, Oscar. 2015. *Modelling Analogical Change: A History of Swedish and Frisian Verb Inflection*. Groningen: University of Groningen.
- Tiersma, Peter M. 1982. Local and general markedness. *Language* 58(4): 832–849. <https://doi.org/10.2307/413959>
- Versloot, Arjen P. 2008. *Mechanisms of Language Change: Vowel Reduction in 15th century West Frisian*. Utrecht: LOT.
- Versloot, Arjen P. 2014. Die -ar-Plurale im Altwestfriesischen mit einem Exkurs über die sächlichen Plurale im Westfriesischen. *Us Wurk* 63: 93–114.
- Versloot, Arjen P. & Adamczyk, Elżbieta. 2014. Corpus size and composition: Evidence from the inflectional morphology of nouns in Old English and Old Frisian. In *Directions for Old Frisian Philology*, Rolf H. Bremmer, Jr., Stephen Laker & Oebele Vries (eds). *Amsterdamer Beiträge zur älteren Germanistik* 73: 539–569.
- Versloot, Arjen P. & Nijdam, Han (eds). 2011. *Integrated Frisian Language Database 2.0. Language Database*. <<http://tdb.fryske-akademy.eu/tdb/index-en.html>>

Appendix

The appendix presents a complete inventory of lemmas used in the present study of both historical and new plural irregularities. In addition to the Old Frisian and Old English forms, present-day variants from West Frisian and Scots are included to exemplify the irregular plural patterns and their interpretations. The *italics* in the column WF-sg are used to denote lexical innovations, while in the columns WF-pl and Scots-pl, they refer to innovative irregular plurals. The bold font used in the columns WF-pl and Scots-pl refers to archaic irregular plurals.

Abbreviations

WF = West Frisian ‘stem’ = declensional class ‘root n.’ = root noun

| Stem | Gloss | Ofris-sg | Ofris pl | WF-sg | WF-pl | OE-sg | OE-pl | Scots-sg | Scots-pl |
|---------|------------------|-----------|----------|-----------------|------------------|-----------|----------|----------|---|
| root n. | <i>louse</i> | *lūs | *lūs | lūs [u:] | <i>lūzen</i> [u] | lūs | lȳs | louse | lyse |
| root n. | <i>book</i> | bōk | *bōk | boek | boeken | bōc | bēc | buik | bukis |
| root n. | <i>trousers</i> | brōk | *brēk | broek | broeken | brōc | brēc | breke | brekis |
| root n. | <i>city</i> | burch | *burch | burch | burgen | burg | byrg | burgh | burghis |
| root n. | <i>foot</i> | fōt | fēt | foet | <i>fuotten</i> | fōt | fēt | fute | fete(s) , futis |
| root n. | <i>furrow</i> | furch | *furch | fuorge | fuorgen | furh | fyrh | fur | furris |
| root n. | <i>goose</i> | gōs | *gēs | goes | <i>guozzen</i> | gōs | gēs | guse | geese |
| root n. | <i>cow</i> | kū | *kī | ko | kij | cū | cȳ | kow | kye |
| root n. | <i>man</i> | mon | man/men | man | <i>manlju</i> | man | men | man | men |
| root n. | <i>mouse</i> | mūs | *mūs | mūs | <i>mūzen</i> | mūs | mȳs | mous | mys |
| root n. | <i>night</i> | nacht | nacht | nacht | nachten | neaht | nieht | nicht | nichtis |
| root n. | <i>tooth</i> | tōth | tēth | <i>tosk</i> | <i>tosken</i> | tōþ | tēþ | tuth | teth |
| root n. | <i>turf</i> | (regular) | – | – | – | turf | tyrf | turf | <i>turven</i> |
| r-stem | <i>brother</i> | brōther | brōther | broer | <i>bruorren</i> | brōþer | brōþer | brother | <i>brether</i> |
| r-stem | <i>daughter</i> | dochter | dochter | dochter | dochters | dohter | dohter | dochter | dochters |
| r-stem | <i>father</i> | feder | feder | <i>heit</i> | <i>heiten</i> | fæder | *fæder | fathir | fathiris |
| r-stem | <i>mother</i> | mōder | mōder | <i>mem</i> | <i>memmen</i> | mōder | mōder | mother | *mothers |
| r-stem | <i>sister</i> | suster | suster | suster | susters | sweoster | sweoster | sister | sisteres |
| s-stem | <i>egg</i> | *āi | *āiar | aai | aaien | æg | ægru | eg | eggis |
| s-stem | <i>lamb</i> | *lamb | *lambar | laam | <i>lammen</i> | lamb | lambru | lamb | lambis |
| s-stem | <i>calf</i> | kalf | *kalvar | keal | <i>keallen</i> | cealf | cealfu | calf | cauf , cair , <i>calwis</i> |
| s-stem | <i>child</i> | kind | kindar | <i>bern</i> | <i>bern</i> | cild | cild(ru) | child | childer |
| s-stem | <i>clothes</i> | klāth | klāthar | – | <i>klean</i> | (regular) | – | – | – |
| s-stem | <i>cloth</i> | klāth | klāthar | kleed | kleden | (regular) | – | – | – |
| u-stem | <i>field</i> | feld | felda | fjild | fjilden | feld | felda | feld | feldis |
| u-stem | <i>hand</i> | hond | honda | hân | <i>hannen</i> | hand | handa | hand | handis |
| u-stem | <i>body</i> | lith | litha | – | lea | (regular) | – | – | – |
| u-stem | <i>limb</i> | lith | litha | lid | lidden | (regular) | – | – | – |
| u-stem | <i>hand-mill</i> | quern | *querna | <i>hân mole</i> | <i>hân molen</i> | (cweorn) | – | – | – |
| u-stem | <i>summer</i> | sumer | *sumera | simmer | simmers | sumor | sumora | somer | someris |
| u-stem | <i>son</i> | sune -u | sunā | soan | <i>soannen</i> | sunu | sunā | son | sons |
| u-stem | <i>forest</i> | wald | *walda | wâld | wâlden | weald | wealda | wod | wodis |
| u-stem | <i>winter</i> | winter | *wintera | winter | winters | winter | wintra | winter | winteris |

