Probing Paul's principle: paradigm heterogeneity and homogenization

Abstract

This paper analyzes the temporal dynamics of the leveling of vocalic and consonantal patterns of allomorphy in Middle and Early New High German. Bayesian modeling is employed to establish reliable relative chronologies between analogical changes affecting different parts of the verbal system. While clear chronologies cannot be established for all change types, results robustly indicate that changes in past tense form of certain verbs precede changes in present tense forms. This finding is compatible with a view where paradigmatic heterogeneity inhibits simplification processes even when its locus is outside of the area targeted by change.

Keywords: historical linguistics, analogical change, paradigm leveling, morphological change, Germanic, Bayesian statistics

1 Introduction

In the domain of analogical change, few linguists' names adorn as many concepts as Hermann Paul's, such as PAUL'S PROPORTION for four-part analogy (Palmer 1972) and PAUL'S DUALISM between proportional and non-proportional mechanisms in analogical change (Gaeta 2007). A third eponymous observation describes circumstances which lead — or rather fail to lead — to paradigm leveling, a type of morphological change that involves the elimination of irregularity within morphological paradigms: paradigms of high irregularity resist leveling, while less heterogeneous paradigms are prone to undergo even more simplification. This idea, termed PAUL'S PRINCIPLE, is not always central to discussions on analogical change, despite the interesting questions that it raises. For instance, while heterogeneous paradigms may resist leveling for a long duration of time, they often succumb nonetheless to processes of homogenization, and the circumstances under which such change takes place remain unclear.

This paper investigates the dynamics of paradigm leveling in certain Middle and Early New High German strong verbs, a process during which several inflectional classes saw a gradual loss of highly irregular allomorphy involving alternating consonants as well as vowels. I inspect variation among diachronic trajectories displayed by different analogical change types using a hierarchical regression model, with an eye to understanding whether the breakdown of regularity was an incremental process where vowel and consonantal allomorphy was neutralized in sequence, and whether change was constant across different parts of the paradigm. Some clear generalizations emerge from the results: it is not possible to date vowel and consonantal leveling changes in certain past tense

forms with respect to each other, but these developments reliably precede certain leveling changes in the present system. I discuss the implications of this finding for our understanding of inter- and intra-paradigm relationships in language change, and outline directions for research further investigating this and related issues.

2 Background and rationale

2.1 Analogical change, irregularity, and the status of leveling

Analogical change encompasses a wide range of processes which involve the restructuring of morphological systems in languages. A common type (or consequence) of analogical change is intra-paradigm leveling, where speakers undo patterns of irregular (i.e., phonologically unpredictable) lexical allomorphy within paradigms, as shown in Table 1. A commonly held view is that language users possess a drive toward paradigm leveling, as the process of leveling achieves biuniqueness, phasing out many-to-one mappings between form and meaning (Wurzel 1984).

[Table 1 about here.]

The view that a preference for regularity shapes analogical change has been contested. Garrett (2008) argues that cases of analogical change traditionally viewed as intra-paradigm leveling are just as easily explained as proportional instances of inter-paradigm extension (namely of patterns from non-alternating paradigms to paradigms which previously contained alternating patterns), vitiating the notion of leveling as a non-proportional process requiring a special status in analogical change. Many concur with the idea that a drive toward level paradigms does not play a major role in analogical change, but question a consequence of this view, namely the implication that proportional processes are predominant in analogical change. Sims-Williams (2016) adduces examples of leveling in Ancient Greek that lack a proportional basis, and Fertig (2016) emphasizes the role of non-proportional mechanisms in analogical change more generally.

Others argue not only that there is no drive toward level paradigms, but that allomorphy benefits language users, both from the standpoint of production and comprehension. For instance, phonologically and semantically unmotivated morphomic alternations are claimed to aid speakers in deducing inflectional realizations of verbs, account for their stability in Romance languages (Maiden 2004). Nübling (2000) argues that irregular allomorphy is beneficial to listeners, as it aids in the recognition of frequent forms that may be uttered with shorter duration due to frequency effects. This view has support from the literature on comprehension; allomorphy renders forms within a paradigm maximally discriminable from each other (Blevins et al. 2017). Baayen and del Prado Martín (2005) note that irregular verbs in Germanic languages tend to have more synonyms, populating denser semantic space, and that this property of irregular verbs facilitates the production and perception of the patterns of allomorphy they display.

There is thus considerable support for the idea that language users benefit from maintaining complex patterns of allomorphy within paradigms. At the same time, the benefits of irregularity may be limited: De Smet and Van de Velde (2019) find that moderately irregular Dutch verbs are less likely to undergo regularization than highly irregular ones.

2.2 Paul's principle: the allure of irregularity

In his *Prinzipien der Sprachgeschichte*, Hermann Paul makes numerous observations regarding tendencies in analogical change, including the circumstances under which leveling (*Ausgleichung*) tends or tends not to occur (Paul 1880:161ff.). He states that leveling is inhibited in the presence of transparent, productive phonological alternations, but may occur when the original conditioning environments for such alternations are lost. At the same time, leveling is hindered by greater amounts of irregularity.

This idea is illustrated using examples of the operation and non-operation of Verner's Law (VL), whereby in most Germanic languages, the Proto-Germanic fricatives **f*, **þ*, **x(ʷ)*, **s* underwent voicing to *β, *δ, *γ, *z ($>$ *r*) when they directly followed unstressed syllables. This development resulted in considerable allomorphy in Germanic strong verb paradigms (a phenomenon termed *Grammatischer Wechsel* 'grammatical alternation'), which ceased to be phonologically predictable once the original prosodic conditioning environments were no longer apparent to speakers. Accordingly, affected verb paradigms underwent extensive analogical restructuring. This breakdown started at an early date, and its effects are already visible in the oldest Germanic attestations.

Paul states that VL alternations in stem-final consonants were likely to be leveled in Old High German past participles sharing their stem vowel with other principal parts (e.g., **gileran* 'read' $\rightarrow$ *gilesan* on the model of the infinitive *lesan*), but less so when there was a stem vowel alternation, and moreover, that leveling of joint alternations in stem vowels and stem-final consonants was unlikely to occur when it served to enhance semantic distinctions (e.g., in number, tense/aspect/mood [TAM] categories, Bybee cf. 1980). For Paul, this twofold marking imbues paradigms with a psychological salience, and the greater the difference between the original form and a potential analogical creation, the less likely the new form is to catch on.

These observations regarding irregularity have not been fully integrated into later views of analogical change, with some exceptions (Garrett 2007, Weiss 2009, Fertig 2016), although Paul's insights here are prominent enough to have been refashioned as an eponymous principle. This passage goes unmentioned by Kuryłowicz (1945), even though the author's generalization that that bipartite morphological markers (as in German *Buch* 'book', *Bücher* 'books', where plural number is marked by both suffixation and simulfixation via an alternating vowel) replace simple markers can be seen as a corollary to the message expressed by Paul: just as twofold markers are likely to replace simple ones in marking functional contrasts, they are unlikely to be replaced by simple markers via leveling. An exception is Garrett's claim (Garrett *apud* Weiss 2009:121–2, fn. 25) that heterogeneity within paradigms serves as a deterrent to analogical change, even when the locus of heterogeneity is outside of the area targeted by analogical change; it is argued that stems of Latin compound verbs, where regular sound change should result in vowel raising, tend not to restore their vocalism if their perfects are formed via relatively simple, affixal processes, even though on its surface, the formation of the perfect is independent of the vocalism in other TAM categories. The argument that apparently unconnected surface patterns occurring across different TAM categories within verbs have mutual predictive inference is relevant here (Herce 2020) — heterogeneity in one part of a paradigm may be predictive of ostensibly unrelated patterns of heterogeneity in the same paradigm.

In what follows, I seek to shed further light on the diachronic dynamics of leveling. Of the

major questions that emerge from Paul's observations, I provide a preliminary attempt to understand why, if bipartite marking is so salient, it ultimately succumbs to analogical change. To accomplish this, I employ quantitative modeling in a quasi-exploratory fashion in order to establish a relative chronology between analogical change types affecting highly irregular verbal paradigms and uncover consistent pathways of change in the breakdown of irregular allomorphy.

3 Middle and Early New High German verbal morphology

The High German verbal system was in considerable flux during its history. As mentioned by Paul, analogical changes started to operate by the time of Old High German; however, the developments that resulted in the verbal system observed today were far from complete, taking place gradually over a number of centuries through the Middle (MHG) and Early New High German (ENHG) periods (approximately 1050–1700 CE), and additional analogical changes are currently underway (Krause-Lerche 2019). These changes involved verb stems as well as verbal endings, and are described in detail in Fertig 1994, 2000; a comprehensive description of Middle High German verbal morphology can be found in (Jones and Jones 2019).

In particular, strong verbs of inflectional classes I–III underwent a great deal of change in their stem vowel allomorphy between MHG and NHG, key points of which can be summarized as follows (these points are adapted from Fertig 1994, chapter 3):

1. In the preterite of Class I verbs, stems of 1/3sg indicative forms originally contained a diphthong while remaining forms contained a monophthong (e.g., MHG *reit* 3sg.pret.ind, *riten* 3pl.pret.ind 'ride'). The diphthongs in the singular stems were eventually leveled in favor of the monophthongal variant (e.g., NHG *ritte* 3sg.pret.ind, *ritten* 3pl.pret.ind 'ride').

2. In the present of most Class II verbs, singular indicative and imperative forms originally contained a diphthong, usually written *iu* in MHG and *eu* in ENHG, while the remainder of present forms contained the monophthong *ie* (e.g., ENHG *zeuhe* 1sg.pres.ind, *ziehen* 1pl.pres.ind 'pull'); the monophthong found in the plural was eventually generalized to singular forms. In the preterite, *ô* was originally found in the 1/3sg indicative and *u* in all other preterite indicative forms (e.g., ENHG *zoch* 1/3sg.pret.ind, *zugen* 1/3pl.pret.ind), but the singular variant eventually came to dominate the preterite in most verbs.

3. Class III verbs had *a* in 1st and 3rd person singular preterite indicative forms and *u* in all other preterite indicative forms. This allomorphy was eventually leveled in favor of *a* in most but not all verbs.

In Classes IIIb–V, there was sporadic variation in the stem vowels of present forms, but the status of these changes as morphologically motivated is unclear, and they may be due to dialect admixture. Additionally, classes IV and V showed alternations in vowel length that are not consistently marked in written sources and cannot be analyzed systematically in the same way that the quality differences mentioned above can.

In addition to the changes in vocalism outlined above, stem-final consonant alternations were in a state of great variability during this period. Most of this variation is attributable to VL and its gradual

undoing. Under the traditional view, the preterite 2nd person singular, preterite indicative plural, preterite subjunctive, and preterite participle originally showed voiced reflexes of Proto-Germanic fricatives in West Germanic, while the remainder of the verbal paradigm was not affected by VL. However, voiced as well as voiceless fricatives are frequently found in environments where they are not expected, and in some verbs, a single reflex has been generalized to the entire verb paradigm (e.g., *g* in NHG *fangen* 'catch' < MHG *vâhen*, *schlagen* 'hit' < MHG *slahen*). Table 2 provides examples of leveling developments that took place in three verbs from classes I–III in preterite indicative forms (as well as present indicative ones in class II, where present forms underwent considerable analogical change).

[Table 2 about here.]

4 Data collection, coding, and visualization of change patterns

To investigate these patterns on a large scale, I compiled a data set of attested verb forms from existing dialectally diverse diachronic German corpus resources, all freely available online. I augmented data from the Bonn corpus of Early New High German (Schmitz et al. 2013), a corpus of 40 dialectally diverse texts spanning from the 14th to 17th centuries, with data from the Anselm corpus (Dipper and Schultz-Balluff 2021), comprising 70 versions of the *Interrogatio Sancti Anselmi de Passione Domini* written between the 14th and 16th centuries; the Fürstinnenkorrespondenz corpus (Lühr et al. 2013), containing 600 letters written between members of the nobility between 1546 and 1756; and the Referenzkorpus Mittelhochdeutsch (Klein et al. 2016), containing 398 Middle High German records from the years 1050 to 1350. Document dialect affiliations were reconciled with the dialect schema found in the Bonn corpus and dates of transcription were extracted or approximated on the basis of document-level metadata. This yielded a data set of 331650 verb tokens, coded for DOCUMENT ID, YEAR, DIALECT, LEMMA ID, INFLECTIONAL CATEGORY, and VERB CLASS.

In order to assemble a dossier of verbs affected by VL, I searched Kroonen 2013 for Proto-Germanic strong verbs with stem-final fricatives (and other relevant items, e.g., *hafjan- ∼ *habjan- 'raise, lift') and reflexes in Middle and/or New High German, inspecting these entries with the aid of grammatical descriptions to establish evidence for *Grammatischer Wechsel*. I subdivided these verbs to the inflectional classes of their Middle and New High German descendant forms and reconciled etymologically related MHG and (E)NHG lemma IDs (e.g., MHG *verlièsen*, NHG *verlieren* < PGmc *fra-leusan-). Of the original data set, a subset of 12270 tokens consisted of inflected forms of reflexes of Proto-Germanic verbs affected by VL.

[Table 3 about here.]

I coded forms according to whether their stem-final consonants and stem vowels were EXPECTED under regular sound change or analogically INNOVATED. This process was automated via a scripting procedure designed to glean these properties from orthographic patterns. The codings produced by the scripting procedure were inspected several times to assess overall accuracy. Examples from the resulting dataset can be seen in Table 3, and suffice to showcase the widely varying orthography encountered.

While I tried to capture as much idiosyncratic orthographic variation as possible, some highly irregular transcriptions were not properly accounted for by the heuristics that I used, leading to instances of erroneous coding. Orthographic variation was not the only potential source of error, however. In many cases, it is difficult to tease apart analogical change (or lack thereof) from dialect-specific phonological developments. For instance, many dialects merge *t* and *d*, making it impossible to determine whether a form reflects VL alternants *ð or *þ (Fertig 1994:95, fn. 24). Additionally, some dialects display a sound change *g* > *ch* /kx/ (Fertig 1994:89; Jones and Jones 2019:59), making it difficult to determine whether orthographic *ch* represents a voiceless VL alternant (possibly archaic) or a secondary development from a voiced one.

In order to mitigate error of this sort introduced into the statistical analyses carried out in this paper, I chose to exclude reflexes of verbs with stem-final PGmc *þ, and additionally excluded all Class III verbs, none of which showed unambiguous evidence of any analogical changes in progress due to older leveling changes (cf. Kienle 1969:66–7); inclusion of Class III forms would thus increase the risk of introducing additional errors in the form of spurious archaisms. I excluded lemmata with fewer than 10 attestations across the corpus.

To better generalize about regions of verb paradigms targeted by specific analogical changes, I aggregated individual inflectional categories into inflectional macro-classes representing groups of cells in the paradigm that underwent the same types of change. These classes include the 1/3sg preterite indicative, the 2sg and 1–3pl preterite indicative, the singular present indicative (which includes the 2sg imperative as well as 1–3sg.pres.ind forms), and the plural present indicative (which contains the 1/2pl imperative, present participle, and infinitive alongside 1–3pl.pres.ind forms). Subjunctive forms were excluded due to low attestation.

Finally, in preparing the data for quantitative analysis, it was necessary to take into consideration only inflectional categories targeted by specific types of analogical change and disregard unaffected categories. Concretely, the stem vowel in the singular present indicative of *ziehen* underwent change from a diphthong (usually written *iu* or *eu*) to a monophthong generally written *ie*, but the stem vowel of the plural present indicative of *ziehen* was always a monophthong and never underwent any type of analogical change. Treating an analogically inert element like the stem vowel of the present plural indicative of *ziehen* as a potential locus of change by including it in certain statistical models (e.g., linear regression, with date as a predictor of change) implicitly carries with it the assumption that the category is expected to undergo change at some time in the however distant future. One means of achieving this goal is to include only change types discussed in the literature; however, this can potentially lead to the omission of interesting data points. As an example, the stem-final consonant of present forms of *ziehen* was never a major target of analogical change (Fertig 1994:90), as the majority of forms in MHG as well as NHG show stem-final *h*. At the same time, the data set shows a number of examples of stem-final *g* in present forms that are unlikely to be errors.[1] To avoid further under-use of the data, I discarded change types only where there was no evidence of any innovation as indicated by the coding scheme I implemented. On the off-chance that this strategy inadvertently introduced data points erroneously coded as innovations, my goal

[1] A 1546 letter in the *Fürstinnenkorrespondenz* (Sib_JF1_1546_09_29) reads *das v g eygenner person zu velde nycht tzeyget* (3sg.pres.ind) ... 'that your grace does not take (?) his own person to the field (i.e., engage in combat)'. *Zu Felde ziehen* 'take to the field' is idiomatic, but does not require a reflexive object as seen here, which may serve an emphatic role; alternatively, the possibility that *zeigen* 'show' was intended here cannot be ruled out.

was to use a variety of statistical controls to ensure that these errors did not bias results. Following this procedure, my data set consisted of 9014 potential loci of analogical change in 6165 word forms. The fact that the size of the data set exceeds the number of unique corpus tokens on which it is based stems from the fact that certain individual forms can be targeted by analogical change affecting stem vowels as well as stem-final consonants.[2]

This coding scheme allows us to track analogical changes over the timespan of the corpus and get a feeling as to whether certain change types precede others. As Figure 1 shows, vowel changes in 2sg and 1–3pl preterite indicative *ziehen* appear to precede $s \rightarrow r$ changes in 1/3sg.pret.ind forms, suggesting that in the preterite subparadigm, the erosion of vowel allomorphy may have done away with a barrier against leveling of consonantal patterns. Additionally, in preterite forms of *ziehen* 'pull', the replacement of original *h* by *g* in 1sg and 3sg past indicative forms begins to take place at a much later date that the leveling of the stem vowel from *u* to *o* in the rest of the preterite. However, in preterite forms of *fliehen*, the innovative variant *h* (original *g*) appears to have been leveled to 2sg and 1–3pl indicative forms prior to changes to the 2sg/1–3pl stem vowel *u* (which was eventually leveled to *o*) detracting from the idea that change in vocalism preceded change in consonantism in the preterite of all relevant plural paradigms. Not only is the visual picture unclear, but it is dangerous to rely solely on visual inspection, as many of the patterns we see may be artifacts of dialect-level trends and other biases in the corpus. For this reason, rigorous statistical modeling is needed.

[Figure 1 about here.]

5 Modeling the breakdown of the *Grammatischer Wechsel*

I turn to quantitative methods for a better understanding of the relative chronology of analogical change types in the history of High German. If we can establish that certain changes precede other changes with some degree of credibility, it has the potential to shed important light on the nature of the decay of paradigmatic heterogeneity.

I explore this issue using hierarchical regression modeling in a Bayesian framework, which provides a flexible means of allocating credibility to multiple hypotheses. Here, I am interested in identifying differences among diachronic trajectories of analogical change types. I fit a hierarchical logistic regression model which included DATE of attestation as a linear predictor of the PRESENCE OF AN ANALOGICAL INNOVATION, a binary outcome. Non-linear and non-monotonic temporal trends were captured by including quadratic and cubic terms for date as well.[3] Along with these fixed effects, random effects were included in order to account for the large degree of non-independence in the sample, allowing the intercept and slopes for the linear, quadratic and cubic terms to vary according to the following variables and interactions between variables: DIALECT, DIALECT×LEMMA, DIALECT×INFLECTIONAL CATEGORY, DIALECT×EXPECTED SOUND, and LEMMA×MACRO-INFLECTIONAL CLASS×CHANGE. The inclusion of the first three

[2]Data and code used in this paper are found at `https://github.com/chundrac/verning-man`

[3]Generalized additive models (GAMs) provide an even more flexible means of dealing with non-linearities, but can be computationally costly to fit in a Bayesian framework. On pros and cons of polynomial growth curves versus GAMs, see Winter and Wieling 2016.

random effects ensured that overall dialect-level trends as well as those affecting individual lemmata and inflectional categories were not conflated with the effects of other variables of interest. DIALECT×EXPECTED SOUND served as a safeguard against confusing dialect-specific phonology with analogical change; if a sound undergoes change at unusually high rate in a particular dialect, this behavior should not impact what the model learns about temporal trends in analogical change. LEMMA×MACRO-INFLECTIONAL CLASS×CHANGE represents the scope of analogical changes affecting parts of the paradigms of different lemmata, e.g., *eu* → *ie* in stem vowel of *ziehen* sg.pres.ind. Inference was carried out in *brms* (Bürkner 2018), an R-based front end for the Stan probabilistic programming language (Carpenter et al. 2017), over four chains of 2000 iterations of the No-U-turn Sampler (NUTS). The predictor DATE was centered around zero and divided by two standard deviations (Gelman and Hill 2007). Flat Normal$(0, 5)$ priors were placed over the fixed effects, while Cauchy$(0, 1)$ priors were placed over the random effects, encouraging values to cluster around zero. The inference procedure resulted in a collection of samples from the posterior distributions of the model parameters; for each sample, we can directly compare the parameters of interest while holding all other parameters constant. Model comparison via the approximate leave-one-out expected log pointwise predictive density ($\hat{ELPD}_{\text{LOO}}$; Vehtari et al. 2017) justified the use of this model over alternative models that excluded the quadratic and/or cubic term as well as LEMMA×MACRO-INFLECTIONAL CLASS×CHANGE. Inclusion of a term that aggregated LEMMA×MACRO-INFLECTIONAL CLASS×CHANGE according to verb class led to decreased model fit.

It is worth noting that this model assumes that analogical changes affecting stem vowels and stem-final consonants are independent processes, even though they sometimes affect the same individual forms; however, the rich random effects structure of our models should accommodate any biases that this assumption might introduce. More complex diachronic models, such as continuous-time Markov chains (CTMC), can explicitly model transitions between more than two "states," and could help us assess whether a trajectory of the type *verliuse* → *verliese* → *verliere* is more likely than some alternative. However, the application of CTMC to diachronic corpora as described by Van de Velde and De Smet (2021) involves underusing large portions of the data and does not seem to readily account for different sources of non-independence found in our data (e.g., dialect-level trends).

The standard practice, not employed here, would be to treat LEMMA×MACRO-INFLECTIONAL CLASS×CHANGE, the main variable of interest, as a fixed rather than random effect, which in the Bayesian context would involve use of a prior distribution on coefficients that allocates more density to higher values rather than values close to zero. Gelman et al. (2012) argue that the PARTIAL POOLING strategy employed here allows practitioners to make reliable pairwise comparisons between different groups, since priors which shrink coefficients toward zero reduce the risk of inferring exaggerated between-group differences that lead to false positives. If this view is correct, it allows me to detect whether there are substantial differences in the trajectories of different types of analogical change (e.g., *s* → *r* in *verlieren* sg.pres.ind versus *s* → *r* in *verlieren* pl.pres.ind).

[Figure 2 about here.]

Figure 2 shows a number of smooth curves representing the expected trend of each level of LEMMA×MACRO-INFLECTIONAL CLASS×CHANGE, along with 95% credible intervals represent-

ing the uncertainty of these estimates. In some cases, the trend remains close to zero, indicating that the entity in question either does not display any tendencies toward change whatsoever, or this behavior has been absorbed by the other predictors in the model (i.e., the appearance of change may be an artifact of dialect-level phonological idiosyncrasies). For some levels, credible intervals are quite wide, indicating high uncertainty regarding the shape of the curve.

Comparing pairs of change types on a sample-by-sample basis allows for the exploration of differences in their trajectories. If one change type is more advanced than another in 95% of samples or more, this serves as strong evidence that the first change reached completion before the second one. However, assessing differences between smooths of this sort is not entirely straightforward. Certain conventions for interpreting polynomial terms in growth curves (cf. Mirman 2014) are not applicable here. Similarly, while it is possible to visually compare trends for different GAM smooths (Rose et al. 2012), there is no agreement on how to carry out hypothesis testing via visual inspection (Sóskuthy 2017).

Since the major research question here concerns whether certain changes reach completion before others, I chose to compare curves by determining the point at which they reach ceiling level, which I take to be an expected value of .99; this value can be approximated for every sample of every level of LEMMA×MACRO-INFLECTIONAL CLASS×CHANGE using grid approximation. For each posterior sample, one change type is more advanced than another if it reaches .99 before the other.[4] For meaningful inference, I exclude change types where the expected probability of change never exceeds .5, as failing to do so would involve comparison between changes that took place and changes that never took place.

[Figure 3 about here.]

Figure 3 provides a graphical representation of the results of this procedure. An arrow is drawn from one node to another with a solid line if the first change type reaches completion before the second in 95% of samples or more; dashed lines represent the same relationship, but at the 85% cutoff, for which there is only weak evidence. For visual clarity, I include only relationships between change types affecting the same lemmata.

These results show that reliable inferences can be made about the relative chronology of only a fraction of the change types assembled in the corpus, but nonetheless a few interesting points emerge. First, it is not possible to establish a definitive chronology between changes affecting stem-final consonants and stem vowels in class II preterite forms. This is not to say that their trajectories are identical, but simply that there is a great deal of uncertainty as to whether leveling of stem-final consonants precedes leveling of stem vowels, or vice versa (the exception is *fliehen*, where a substantial proportion of posterior samples suggest that the leveling of the stem-final consonant preceded the that of the stem vowel in 2sg and 1–3pl preterite forms, if we allocate any credibility to the less conservative 85% cutoff; *fliehen* itself is exceptional in terms of the direction of leveling). This indicates that the breakdown of irregularity in the preterite was not a series of cascading changes that did away with enough irregularity to bring about wholesale leveling, but rather that analogical pressure was exerting itself on the paradigm even when it was highly irregular.

[4]It is worth noting that while this procedure is in the spirit of Gelman et al. 2012, the quantity being compared is different. Simulation studies are needed to assess the false discovery rate of this procedure.

It is not clear why this is the case, though it is reasonable to suspect the role of changing frequency distributions of lexical items and inflected forms. Secondly, leveling changes in past tense forms precede leveling changes affecting the stem vowel in present singular forms of class II verbs. In the case of *verlieren*, leveling of the stem-final consonant in present forms also appears to precede the leveling change *eu* → *ie*, indicating that a threshold of regularity was achieved throughout the paradigm prior to the leveling of the present singular stem vowel.

This state of affairs is compatible with the view that the conditions needed for the maintenance of paradigmatic heterogeneity in subsets of paradigm cells involve the entire paradigm: homogeneization in the preterite appears to have led the way for loss of heterogeneity in present tense forms. At the same time, some wariness regarding a causal interpretation of this chronology is warranted.[5] Middle High German saw an influx of Latin borrowings adapted to a specific morphological pattern (e.g., MHG *absolvièren*, *jûbilièren*, *confirmièren*, etc.) that would have formed a proportional basis for the *eu* → *ie* change (e.g., *jubiliere* 1sg.pres.ind : *jubilieren* 1pl.pres.ind :: X : *verlieren*, X=*verliere*), though in theory any non-alternating paradigm would serve as a viable proportion. It is therefore entirely possible that Class II singular present forms changed their vocalism solely on the basis of these patterns, independent of changes elsewhere in the verbal paradigm. A closer look at the chronology of Latin loans relative to changes affecting both tense systems of Class II verbs is needed to better understand this issue: if Latin loans were abundant prior to changes affecting preterite forms, it lends support to the notion that changes to the preterite were a prerequisite for changes in the sg.pres.ind vowel.

6 Conclusion

This paper investigated patterns of leveling of stem vowel alternations and stem-final consonant alternations in verbal paradigms during the history of Middle and Early New High German. I used statistical modeling in an attempt to establish a reliable chronology between change types, while controlling for a number of confounding variables and sources of variation across the data set. At the moment, the exact nature of and reasons for the breakdown of bipartite analogical marking remains obscure, though the findings presented here suggest that the trajectory of change did not involve a cascading series of changes that incrementally brought about homogeneity in paradigms, but rather, that changes in type frequencies of paradigms in use by speakers opened the way to analogical change, most likely proportional.

A logical next step will be to turn to the corpora on which this study is based in order to gain a better understanding of the implicational patterns available to speakers during this period (cf. Bonami and Beniamine 2016), a procedure that will be possible when the orthography in these resources is fully normalized. Additionally, more flexible applications of Bayesian modeling than the regression model employed here may shed additional light on this issue. Extending this approach to closely related languages will help us understand why different patterns of change took place in languages like Middle High German, and guard against extrapolating any single-language findings advanced here to analogical change construed more broadly.

[5] On causal analysis in linguistic time-series data see del Prado Martín and Brendel 2016, Kanwal 2018, Rosemeyer and Van de Velde 2021.

References

Baayen, R. H. and del Prado Martín, F. M. (2005), 'Semantic density and past-tense formation in three germanic languages', *Language* **81**(3), 666–698.

Blevins, J. P., Milin, P. and Ramscar, M. (2017), The Zipfian paradigm cell filling problem, *in* 'Perspectives on morphological organization', Brill, pp. 139–158.

Bonami, O. and Beniamine, S. (2016), 'Joint predictiveness in inflectional paradigms', *Word Structure* **9**(2), 156–182.

Bürkner, P.-C. (2018), 'Advanced Bayesian Multilevel Modeling with the R Package brms', *The R Journal* **10**(1), 395–411.
URL: *https://doi.org/10.32614/RJ-2018-017*

Bybee, J. (1980), 'Morphophonemic change from inside and outside the paradigm', *Lingua* **50**(1-2), 45–59.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P. and Riddell, A. (2017), 'Stan: A probabilistic programming language', *Journal of statistical software* **76(1)**(1), 1–32.

De Smet, I. and Van de Velde, F. (2019), 'Reassessing the evolution of West Germanic preterite inflection', *Diachronica* **36**(2), 139–180.

del Prado Martín, F. M. and Brendel, C. (2016), Case and cause in icelandic: Reconstructing causal networks of cascaded language changes, *in* 'Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)', pp. 2421–2430.

Dipper, S. and Schultz-Balluff, S. (2021), 'Anselm Corpus (Version 1.0)'. Ruhr-Universität Bochum.
URL: *https://doi.org/10.34644/laudatio-dev-AiSSCXMB7CArCQ9CsyMe*

Fertig, D. (1994), Variation and change in verbal inflection in late medieval and early modern Nuremberg, PhD thesis, University of Michigan.

Fertig, D. (2000), *Morphological change up close*, Niemeyer, Tübingen.

Fertig, D. (2016), 'Mechanisms of paradigm leveling and the role of universal preferences in morphophonological change', *Diachronica* **33**(4), 423–460.

Gaeta, L. (2007), Is analogy economic?, *in* F. Montermini, G. Boyé and N. Hathout, eds, 'Selected Proceedings of the 5th Décembrettes: Morphology in Toulouse', Cascadilla Press.

Garrett, A. (2007), 'Paradigmatic heterogeneity', *manuscript, University of California, Berkeley* .

Garrett, A. (2008), Paradigmatic uniformity and markedness, *in* J. Good, ed., 'Explaining linguistic universals: Historical convergence and universal grammar', Oxford University Press, Oxford, pp. 124–143.

Gelman, A. and Hill, J. (2007), *Data analysis using regression and multilevel/hierarchical models*, Cambridge university press, Cambridge.

Gelman, A., Hill, J. and Yajima, M. (2012), 'Why we (usually) don't have to worry about multiple comparisons', *Journal of research on educational effectiveness* **5**(2), 189–211.

Herce, B. (2020), 'Alignment of forms in spanish verbal inflection: the gang *poner*, *tener*, *venir*, *salir*, *valer* as a window into the nature of paradigmatic analogy and predictability', *Morphology* **30**(2), 91–115.

Jones, H. and Jones, M. H. (2019), *The Oxford Guide to Middle High German*, Oxford University Press, Oxford.

Kanwal, J. K. (2018), 'Word length and the principle of least effort: language as an evolving, efficient code for information transfer'.

Kienle, R. (1969), *Historische Laut- und Formenlehre des Deutschen*, Niemeyer, Tübingen.

Klein, T., Wegera, K.-P., Dipper, S. and Wich-Reif, C. (2016), 'Reference Corpus of Middle High German (1050–1350) (Version 1.0)'. Rheinische Friedrich-Wilhelms-UniversitÃd't Bonn, Ruhr-UniversitÃd't Bochum.
URL: *https://doi.org/10.34644/laudatio-dev-xCS3CnMB7CArCQ9C3LRB*

Krause-Lerche, A. (2019), 'Conservation in ongoing analogical change: The measurement and effect (s) of tokenfrequency', *Corpus Linguistics and Linguistic Theory* **1**, 1–38.

Kroonen, G. (2013), *Etymological dictionary of proto-Germanic*, Brill, Leiden.

Kuryłowicz, J. (1945), 'La nature des procès dits «analogiques»', *Acta linguistica* **5**(1), 15–37.

Lühr, R., Faßhauer, V., Prutscher, D. and Henry, S. (2013), 'Fuerstinnenkorrespondenz 1.1 (Version 1.1)'. Universität Jena, DFG.
URL: *https://doi.org/10.34644/laudatio-dev-ZCSVC3MB7CArCQ9CVedt*

Maiden, M. (2004), 'When lexemes become allomorphs — on the genesis of suppletion', *Folia Linguistica* **38**(3-4), 227–256.

Mirman, D. (2014), *Growth curve analysis and visualization using R*, Chapman and Hall/CRC, Boca Raton, FL.

Nübling, D. (2000), *Prinzipien der Irregularisierung. Eine kontrastive Analyse von zehn Verben in zehn germanischen Sprachen*, Niemeyer, Tübingen.

Palmer, L. R. (1972), *Descriptive and comparative linguistics: a critical introduction*, Faber & Faber, London.

Paul, H. (1880), *Principien der Sprachgeschichte*, Niemeyer, Halle.

Rose, N. L., Yang, H., Turner, S. D. and Simpson, G. L. (2012), 'An assessment of the mechanisms for the transfer of lead and mercury from atmospherically contaminated organic soils to lake sediments with particular reference to Scotland, UK', *Geochimica et Cosmochimica Acta* **82**, 113–135.

Rosemeyer, M. and Van de Velde, F. (2021), 'On cause and correlation in language change: Word order and clefting in Brazilian Portuguese', *Language Dynamics and Change* **11**(1), 130 – 166.

Schmitz, H.-C., Schröder, B. and Wegera, K.-P. (2013), 'Das Bonner Frühneuhochdeutsch-Korpus und das Referenzkorpus ‚Frühneuhochdeutsch''.
URL: *https://edoc.bbaw.de/frontdoor/index/index/docId/2163*

Sims-Williams, H. (2016), Analogy in morphological change: The diachronic treatment of irregularity in Greek, PhD thesis, University of Oxford.

Sóskuthy, M. (2017), 'Generalised additive mixed models for dynamic analysis in linguistics: A practical introduction', *arXiv preprint arXiv:1703.05339* .

Trask, R. L. (1996), *Historical linguistics*, Arnold, London.

Van de Velde, F. and De Smet, I. (2021), 'Markov models for multi-state language change', *Journal of Quantitative Linguistics* pp. 1–25.

Vehtari, A., Gelman, A. and Gabry, J. (2017), 'Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC', *Statistics and computing* **27**(5), 1413–1432.

Weiss, M. (2009), *Outline of the Historical and Comparative Grammar of Latin*, UMI, Ann Arbor.

Winter, B. and Wieling, M. (2016), 'How to analyze linguistic change using mixed models, growth curve analysis and generalized additive modeling', *Journal of Language Evolution* **1**(1), 7–18.

Wurzel, W. U. (1984), *Flexionsmorphologie und Natürlichkeit: ein Beitrag zur morphologischen Theoriebildung*, Vol. 21, Walter de Gruyter GmbH & Co KG.

List of Figures

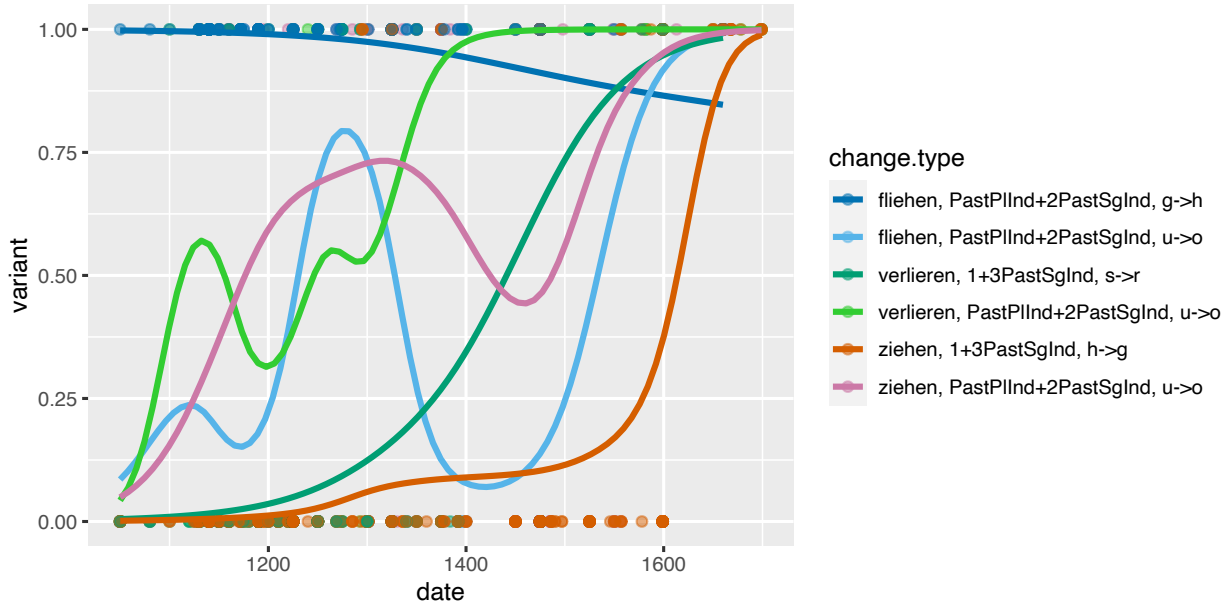

Figure 1: Chronology of selected analogical changes represented by generalized additive model (GAM) smooth curves. Values of 1 correspond to the sound on the right hand side of the arrow.

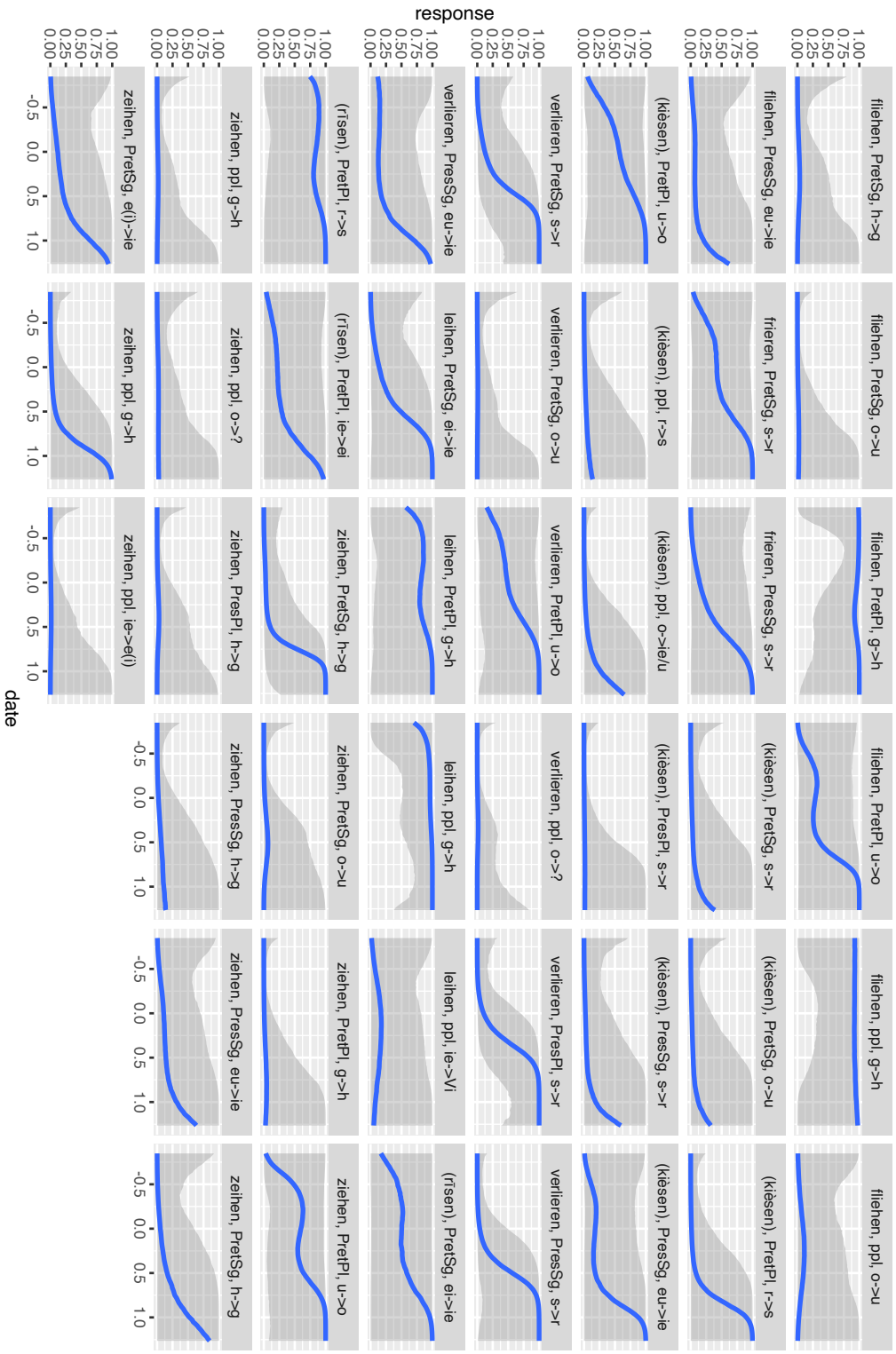

Figure 2: Posterior estimates of trends for levels of LEMMA×MACRO-INFLECTIONAL CLASS×CHANGE. For readability, PretSg represents 1/3sg.pret.ind. and PretPl represents 2sg/1–3pl.pret.ind. Verbs that do not survive paste MHG are in parentheses.

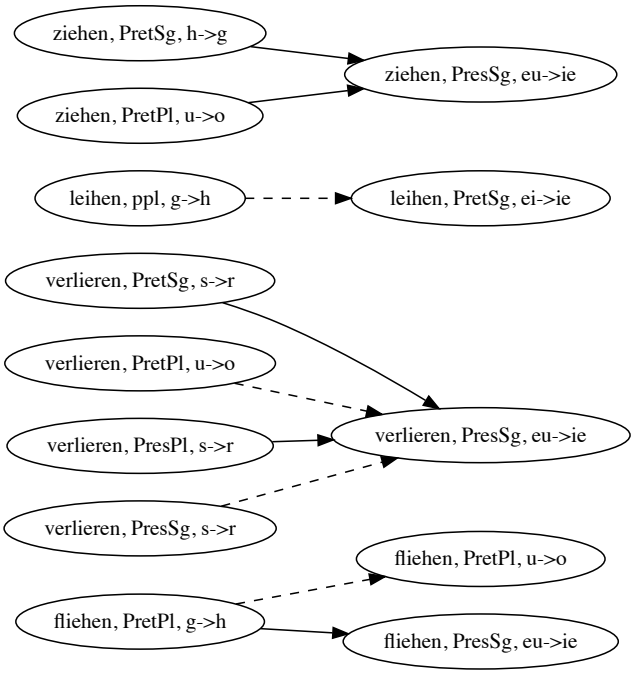

Figure 3: Relative chronologies of change types inferred from model results. For clarity, only intra-lemma chronologies are provided. Solid lines represent precedence relationships supported by 95% or more of samples (85% for dashed lines). For readability, PretSg represents 1/3sg.pret.ind. and PretPl represents 2sg/1–3pl.pret.ind.

List of Tables

1sg *aim*	1pl ***a****mons* → Mod. French *aimons*
2sg *aimes*	2pl ***a****mez* → Mod. French *aimez*
3sg *aimet*	3pl *aiment*

Table 1: Leveling of stem alternation between *aim-* and *am-* in the Old French present paradigm of *amer* 'love', adapted from (Trask 1996, Garrett 2008).

	Class I		Class II		Class III	
Proto-Germanic	*tīhan- ‘accuse’		*fra-leusan- ‘lose’		*finþan- ‘find’	
	MHG	NHG	MHG	NHG	MHG	NHG
inf.	zîhen	zeihen	verliè**s**en	verlieren	vinden	finden
sg.pres.ind			verl**ius**-	verl**ier**-		
pl.pres.ind			verlie**s**-	verlie**r**-		
1/3sg.pret.ind.	z**ê**ch-	zieh-	verlô**s**-	verlor-	vant-	fand-
2sg/1–3pl.pret.ind	zi**g**-	zieh-	verl**u**r-	verlor-	v**u**nd-	fand-

Table 2: Selected Middle and New High German Class I, II and III preterite paradigms, displaying patterns found before and after the operation of leveling of allomorphy in stem vowels and stem-final consonants, with vowels and consonants affected by leveling shown in bold. Note that the $t \sim d$ variation found in MHG *vinden* is unlikely to reflect variation brought about via VL (see Jones and Jones 2019:123ff.).

Year	Dialect	Lemma	Inflection	Form	Coding
1190	C.Bav.	*verlieren*	2sg.pres.ind	verlivses	diphthong, *s*
1050	C.Bav.	*verlieren*	1sg.pres.ind	fliese	**monophthong**, *s*
1300	E.Fr.	*ziehen*	3sg.pret.ind	vberzohe	o, h
1490	E.Swab.	*frieren*	3sg.pres.ind	freürt	diphthong, ***r***

Table 3: Reflexes of *verlieren* 'lose', *ziehen* 'pull', and *frieren* 'freeze' in Central Bavarian, East Franconian, and East Swabian dialects, coded according to expected (unbolded) or innovated (bolded) features that are displayed.