Computational Linguistics in the Netherlands Journal 5 (2015) 81-102 Submitted 07/2015; Published 11/2015

The rise of the verbal weak inflection in Germanic An agent-based model

Dirk Pijpops *† DIRK.PIJPOPS@KULEUVEN.BE
Katrien Beuls § KATRIEN@AI.VUB.AC.BE
Freek Van de Velde * FREEK.VANDEVELDE@KULEUVEN.BE

* *University of Leuven, Blijde Inkomststraat 21 Leuven, Belgium*

† *Research Foundation Flanders, Egmontstraat 5 Brussel, Belgium*

§ *Vrije Universiteit Brussel, Pleinlaan 2 Brussel, Belgium*

Abstract

The verbal weak inflection, one of the defining innovations of Proto-Germanic, currently holds a dominant position in the verbal inventory of most remaining Germanic languages. This has not always been the case, though. This paper investigates how the weak inflection could have grown to overthrow its competitor, the strong inflection, even if (i) the strong inflection was still more regular and (ii) the weak inflection had to start from a position vastly inferior in frequency to any strong ablaut class. As opposed to earlier work, which focused on language acquisition in models of iterated learning, our focus lies on language usage, which is why we have composed an agent-based model. This enabled us to test a number of minimal assumptions needed to explain an ascent of the weak inflection, of which several have been proposed in the literature. It was found that the weak inflection's functional advantage of general applicability is sufficient by itself already. That is, the weak dental suffix is in principle applicable to all verbs, while each separate strong ablaut class is not. This is shown to put the strong inflection at a crucial disadvantage, even if (i) the strong system as a whole is applicable to all verbs, and (ii) each separate ablaut class starts out as dominant in both type and token frequency over the weak dental suffix. There is no need to assume that the strong system has irregularized for the weak inflection to get airborne; this irregularization may rather be the result and subsequent catalyst of the rise of the weak inflection.

1. Introduction

Setting analytic perfect formations aside, two main morphological strategies are pervasive in the Germanic language family to form the past tense. On the one hand, there is the dental preterite or weak inflection, which uses a dental suffix to mark past tense (e.g. Dutch *werk-te*, English *work-ed*). On the other hand, there is the ablauting preterite or strong inflection, which uses root-vowel apophony as a way to mark the preterite (Dutch *zing* ∼ *zong*; English *sing* ∼ *sang*). The Germanic dental preterite is the innovative strategy. It is assumed to derive, most probably, from a fused form of the Indo-European verb *do* (Bailey 1997, p. 8-33; Harbert 2007, p. 276) and is unique among the Indo-European languages. Conversely, the strong inflection goes back to the originally aspectual Indo-European perfect, and is less unique among the Indo-European daughter languages. A similar system existed in Ancient Greek, for instance, where the same ablaut – taking into account the effect of sound laws and ignoring the reduplication affix – can be found in present *leíp-ō* ('I leave') versus its perfect *lé-loip-a*.

Despite its somewhat misleading name,[1] the 'weak' inflection has been overtaking the strong inflection in the history of the Germanic languages. It has managed to become the default strategy for forming past tenses in most Germanic languages (Ball 1968, p. 162; Harbert 2007, p. 277).

[1] The terms strong and weak inflection have been coined by Jacob Grimm, the idea being that the weak inflection needs an additional affix to mark the past, while in the strong inflection the verb accomplishes this feat on its own (Grimm 1819).

The weak inflection is productively applied to new verbs, and many originally strong verbs yielded to the weak inflection over time. Meanwhile, the remaining strongly inflected verbs have become increasingly irregular, to the point where the originally relatively transparent ablaut system has become obscured by the working of several sound laws in the individual Germanic languages. This process has proceeded furthest in English, and many English grammars now distinguish between a ‘regular’ and an ‘irregular’ inflection, with all strong verbs belonging to the latter category, together with a few non-regular weak verbs (e.g. *keep* ∼ *kept*, *buy* ∼ *bought*). In the irregular class, there are islands of subregularity (e.g. *drive* ∼ *drove*, *rise* ∼ *rose*, *ride* ∼ *rode*, w*rite* ∼ *wrote*), but they do no longer transparently map on the historical classes that can be fairly easily discerned in, say, Gothic. In other West-Germanic languages like Dutch and German, the ablaut classes of the strong inflection preserve more of their original morphological motivation (Ball 1968; van Bree 1987; Carroll et al. 2012, p. 163-165).

The present paper is not the first to address Germanic verbal inflection competition in modeling terms. However, earlier computational models either focused exclusively on the acquisition of the past tense system, rather than the changes it has gone through throughout history (Rumelhart and McClelland 1986; Pinker and Prince 1988; Macwhinney and Leinbach 1991; Plunkett and Marchman 1991, 1993; Ling and Marinov 1993; Marcus et al. 1995; Plunkett and Juola 1999; Taatgen and Anderson 2002; van Noord 2015), or explicitly chose to ignore the regularity of the strong inflection (Colaiori et al. 2015; Pijpops and Beuls forthc.), or had the weak inflection start from an already well-established position in the verbal inventory (Hare and Elman 1995; Yang 2002). However, at the outset in Early Germanic language stages, the strong inflection still maintained its morphological transparency, and held both greater type and token frequency than the weak inflection (Bailey 1997, p. 8). This then leaves us with the conundrum of how, under such conditions, a fledgling weak inflection could possibly grow to overthrow a stable strong system as the default strategy of past tense formation. This will be the central concern of this paper. We will apply agent-based modeling to get a grip at the possible causes of this rise of the weak inflection.[2]

In the next section, we will introduce the research questions in detail. Section 3 then presents the design of the model, while in Section 4, the results of some experimental runs of the model are described. Finally, Section 5 summarizes the conclusions.

2. Research Questions

This study is situated on two tracks: one more general and one more specific. These correspond to two research questions, the general one dealing with what has been called the Threshold Problem (Nettle 1999), while the specific one focuses on the explanations for the rise of the weak inflection that have been proposed in the historical literature. We will introduce each of these issues in Subsections 2.1 and 2.2. The research questions can be formulated as follows.

1. Can a specific grammatical innovation overcome the threshold of frequency which is needed to become the dominant variant in a community, through functional selection?

2. Does the general applicability of the weak inflection suffice to explain its rise?

2.1 The Threshold Problem

As for the first question, the rise of the weak inflection constitutes a concrete example of the Threshold Problem, as formulated by Nettle:

> *New mutants can only become fixed in a language if they can pass a threshold of frequency which in the early stages they never have* (Nettle 1999, p. 98).

[2]This means we will deal with the spread of the weak inflection, not its possible origins. For an overview of the available literature on this problem, see Ball (1968), Tops (1974) and Hill (2010).

An important factor in overcoming this threshold is social selection. Under social selection, the new variant is able to gain critical momentum either through the network structure of language users (Milroy and Milroy 1985; Trudgill 2008), the higher social status of the early adopters (Rogers 1995; Nettle 1999; Blythe and Croft 2012, p. 274-275) or a differential social weighting of the variants through overt or covert prestige (Labov 1966, p. 108; Labov 2001, p. 511-518; Baxter et al. 2009, p. 269-270). Conversely, under functional selection, the new variant is able to gain ground because it offers an inherent advantage over the old variant, either in language acquisition, processing or use. The role of functional selection is generally considered much more limited than that of social selection in driving the ascent of new variants. Functional selection is assumed to only affect the direction of language change, not its occurrence (Nettle 1999). To summarize, social selection is the engine driving language change, functional selection the steering wheel.[3]

Still, the modeled competitions in which functional selection is evaluated, are mostly of lexical nature (e.g. Nettle 1999; Croft 2000). For grammatical phenomena, the precise form of a variant's functional advantage may be quite diverse. For instance, in comparing the Old-High-German and New-High-German case systems, van Trijp (2013) was able to show that while the Old-High-German case system holds the functional advantage of cue reliability, the New-High-German system boasts the functional advantages of processing efficiency and ease of articulation. The difference in nature between these advantages may have been decisive in their competition. Changes in the linguistic environment may cause cue reliability to become less instrumental for disambiguation purposes, clearing the way for the case system to evolve to the new variant. Another example can be found in the competition between phrasal names and compounds. Here, Landsbergen (2009, p. 15, 47-74) showed that the phrases' functional advantage of lexicalization sufficed for them to overcome the threshold of frequency.

In this paper, the focus will also lie on a specific functional advantage. Still, we certainly do not mean to downplay the importance of social selection in language change. Social selection may very well be the most common mechanism for overcoming the threshold of frequency. We only mean to show that for a specific grammatical phenomenon such as the strong-weak competition, functional selection may also play a role. An important advantage of computational simulation is precisely that the influence of various factors, such as social and functional selection, can be studied in isolation (Landsbergen et al. 2010, p. 367-368). As such, the model presented here will contain no social structure, or a differential social weighting of the variants. Once the individual influence of these factors is properly understood, more advanced models can be developed which may investigate their interplays.

2.2 Explanations for the rise of the weak inflection

Concerning the second question, we will propose that the crucial functional advantage of the weak inflection lies in its general applicability. This functional advantage is the one of three possible explanations which can be found in the historical literature, summed up below (Ball 1968, p. 164; Bailey 1997, p. 17).

1. General applicability of the weak dental suffix
2. Restrictions on the strong system
3. Irregularization of the strong system

The first explanation states that the applicability of the strong system is divided amongst the different ablaut classes. For example, hearing an instance of the first Dutch ablaut class *grijp* $\sim$ *greep* does not help in the acquisition or entrenchment of the second class *lieg* $\sim$ *loog*. Each of the classes

[3] In Cultural Evolution, a related distinction is made between 'content biases' and 'model-based biases', which respectively compare to functional selection and social selection, see Mesoudi (2011, p. 57).

has to be acquired and reinforced individually. Conversely, once the weak dental suffix is acquired, it can be indiscriminately applied to all verbs.[4]

The second explanation goes further to posit that even the strong system as a whole, i.e. the combination of all ablaut classes, was not applicable to some particular verbs, such as the preterite-presents (Bailey 1997, p. 578). These verbs would then form a safe haven for the weak inflection, allowing it to mature out of reach of competing strong forms.

Lastly, the third explanation seeks the cause of the weak inflection's upsurge in the disintegration or irregularization of the strong system due to sound changes (Ball 1968, p. 164; Bailey 1997, p. 17). These irreguralities would render the strong system vulnerable to competition from seemingly more regular patterns, such as the weak inflection. These three explanations are not seen as in conflict with each other, but rather as all contributing to the rise of the weak inflection. In this article, however, we will attempt to show that explanation (1) already suffices to account for the ascent of the weak system.

If explanation (1) suffices, the irregularization of the strong system might not be the original cause, but rather the result and subsequent catalyst of the rise of the weak inflection. By this, we mean the following. Regularity is most strongly needed by the low frequency verbs; the higher frequency verbs are quite capable of sustaining reasonable levels of irregularity (Bybee 2006, p. 715; Lieberman et al. 2007; Carroll et al. 2012; Cuskley et al. 2014). If the weak inflection was already taking over the low frequency verbs, there would be less need to uphold the regularity of the strong system. As a result, the injurious effects of sound changes would no longer need to be halted or rectified by maintenance and repair mechanisms such as paradigmatic leveling.[5]

A first important argument against the claim that explanation (1) suffices, is that the addition of the weak inflection only further complicates past tense formation. It implies that we leave a local maximum of transparency based on an ablaut system with several classes, to reach an improved maximum based on a single, generally applicable dental suffix. However, after thousands of years, most Germanic languages have yet to reach this improved maximum. And those that did manage to get rid of the strong system, like Afrikaans, only succeeded in doing so by eradicating synthetic past tense formation in general. Moreover, Cuskley et al. (2014) showed that in recent years, the number of strong verbs in English has hardly decreased. During this protracted transition, learners not only need to learn the different ablaut classes, but also for which verbs to switch to the weak inflection. The speakers and learners of Proto-Germanic thus seem to have had nothing to gain from this language change.

A second argument against explanation (1) is that general applicability is only useful if the speaker often needs to express the past tense of a verb for which he has yet to acquire an applicable strong class. For example, a speaker needs to express the past tense of the Dutch verb *vragen* ('to ask'), but he has yet to acquire a vowel alternation for the /a/. Luckily, he can still do so by applying the dental suffix *-de*, which would create the weak form *vraagde* ('asked'). Such a form is easily recognizable as a past tense of *vragen* by any native speaker. However, once the speaker has acquired the applicable ablaut pattern <aa> → <oe>, and learns that the conventional form is *vroeg* ('asked'), not *vraagde*, the weak inflection no longer offers an advantage to him and he will stop using the weak form *vraagde*. Moreover, in a situation where each ablaut pattern is individually still more frequent than a nascent weak inflection, a language user will probably have acquired an applicable ablaut pattern such as <aa> → <oe> before acquiring the weak dental suffix. We will come back to these arguments in Section 4, where the results of the model are presented.

[4]This first explanation is actually not mentioned explicitly in Ball (1968, p. 164) or Bailey (1997, p. 17), but implicitly assumed in the second.

[5]Still, paradigmatic leveling did improve the regularity of the disintegrating strong system at several occasions. An example is the *Ausgleich* in the singular-plural vowel quality distinction, as in singular *band* – plural *bundum* in Gothic to singular *bond* – plural *bonden* in Dutch.

3. Model Design

The goal of agent-based modeling or computational simulation in general is not to design a model to be as realistic as possible. It is rather to test the results of a number of minimal assumptions (Nettle 1999a: 103). Because of this, it is advised to always keep the model as simple as possible, even if this simplicity is unrealistic (Landsbergen 2009, p. 18-19). Even then, agent-based models may be criticized because the results are built-in or a clear evaluation is lacking (de Boer 2012). Therefore, Subsection 3.1 will explicitly define (i) the assumptions that our model is allowed to make, which should be minimal, and (ii) the criteria by which the model will be evaluated. In Subsection 3.2, these assumptions are molded into a concrete model.

3.1 Conceptual level of model design

Our model is composed of the following elementary building blocks, which represent both the constraints on the model and the assumptions it is allowed to make. Only if it is shown that the required results cannot emerge under those conditions, may these constraints be relaxed or more assumptions be added.

1. There is a competition between a single weak, generally applicable construction versus multiple strong constructions.[6]

2. The strong system as a whole is generally applicable and still clearly regular. This means that (i) while an individual strong construction may not always be applicable, at least one of the strong constructions should be applicable at all times; and (ii) there are no irregular verbs in the model, or ways for verbs or constructions to irregularize.

3. The strong system starts absolutely dominant. At the start, even the least frequent strong construction should start more numerous in both type and token frequency than the weak construction.

4. The occurrences of the verbs show a realistic (Zipfian) frequency distribution (Zipf 1932).

5. Agents age and are gradually replaced.

The first building block implements explanation (1) into the model. It states that the weak system is represented by a single construction that can be indiscriminately applied to all verbs. Still, for Proto-Germanic, 3-4 weak classes are assumed (Bailey 1997).[7] However, all of them share the same dental suffix: they only differ in stem vowel (Ball 1968, p. 165-167; Hare and Elman 1995, p. 65-66). As such, we hold it defensible to implement the dental suffix as a single construction. This implementation as a single construction entails that each time an agent hears a weak form, the weak construction will become more entrenched in its memory. As for the strong inflection, each vowel alternation is represented by a separate construction. For instance, this means that each time an agent hears the <ij> → <ee> construction being used, this construction will become more strongly entrenched in its memory, while the <ie> → <oo> construction will not become more entrenched (cf. Subsection 2.2). Together, these ablaut constructions make up the strong system.

The second building block then prescribes that the strong system as a whole should also be generally applicable, and that there can be no irregular verbs. Put more concretely, the agents should

[6] By *construction*, we mean an operation to form the past tense which is amenable to the influence of usage, just like lexical elements (Goldberg 2003; Bybee 2010). That is, language users will more often use constructions which they hear more often. In the remainder of this paper, both the strong classes and the weak dental suffix are named constructions.

[7] Also, the weak dental suffix has both a *-te* and *-de* realization in Present-day Dutch. These realizations present a blind phonetic alternation however, which is based solely on the voicedness of the stem's ending vowel (Haeseryn et al. 1997, p. 72).

be able to conjugate each verb in the model according to one of the strong constructions, without ever needing to use the weak construction. In this way, the second and third explanations for the rise of the weak inflection are excluded from the model. This exclusion is of course absolutely necessary, because we want to safely attribute a possible rise of the dental suffix to the first explanation. For the same reason, any other factors which may cause such a rise have to be excluded from the model. For one, this requires that the model almost abstracts completely away from language acquisition. If a realistic learning algorithm were to be implemented, it would be hard, if not impossible to attribute a possible rise of the weak inflection to the properties of the learning algorithm or to the functional advantage of general applicability. For another, we will refrain from implementing any form of social selection. No specific network structure will be assumed for the agents' population. All agents will have an equal chance of interacting with one another. There will be no agents with more social prestige than others, and the agents will not show any (socially attributed) preference for one of the variants, nor will they ever strive toward 'extravagant' language use (Haspelmath 1999). Lastly, the agents will never forget a strong verb form, nor fail to retrieve it from memory (Taatgen and Anderson 2002, p. 124).

In all other design choices, we will adhere to the maxim of simplicity in agent-based modeling (Landsbergen 2009, p. 18-19), which entails three more important simplifications with regard to reality. The first two are that our model only looks at finite past tenses, not participles, and disregards any influence of phonetic resemblance between verbs. While both are certainly important in reality, we are currently not directly interested in their effects, and thus need to leave them out of the model. Lastly, a hearer will never fail to correctly interpret an utterance. For instance, an 'incorrect' strong past tense form such as *vrieg*, formed by applying the <aa> → <ie> pattern (e.g. *slaap* ∼ *sliep*) to the verb *vragen* ('to ask'), will be recognized just as easily as an 'incorrect' weak past tense form like *vraagde*. Communication will only fail if the speaker does not know any way of forming a past tense of the verb in question.

We now turn to the evaluation criteria. To begin with, these criteria define the results that should emerge from the model in order to positively answer the research questions. This of course includes a rise of the weak inflection. However, not just any rise will be acceptable. For one, a next generation of agents shouldn't suddenly switch to an entirely weak system. For another, the weak inflection should not take over the high frequency verbs before taking over the low frequency verbs. In addition, the evaluation criteria should also delimit what the model is trying to explain. For instance, the criteria for the present model contain no specifics about the stages learners should go through in the acquisition of the past tense. This is simply because we do not claim that general applicability is capable of causing the emergence of such stages.[8] The same goes for other intricacies of the competition that do not fall under these criteria, such as the precise course of development for a specific verb. Lastly, the emergence of these criteria should be quite robust, and not be dependent on specific parameter settings (Blythe and Croft 2012, p. 291-293).

Luckily, there are already several empirical studies available which accurately describe how the weak inflection has been creeping into the past tense formation of the Germanic languages. These studies form the basis of the model's evaluation criteria. In order to positively answer the research questions, the following results are required to emerge from the model.

[8] The emergence of these stages is probably due to the learning algorithm employed by language learners. For readers who are interested in the acquisition of the Germanic past tense, many studies are available, including Rumelhart and McClelland (1986); Pinker and Prince (1988); Macwhinney and Leinbach (1991); Plunkett and Marchman (1991, 1993); Ling and Marinov (1993); Marcus et al. (1995); Plunkett and Juola (1999); Taatgen and Anderson (2002) and van Noord (2015).

1. **Rise of the weak inflection**: we should observe a rise of the weak construction on both token and type level (Carroll et al. 2012; Cuskley et al. 2014).
2. **Gradual rise**: This rise should be gradual; it should typically take more than one generation. It is possible to have prolonged co-existence of both variants, i.e. some verbs may be conjugated according to the strong system, while others are weak (Cuskley et al. 2014).
3. **Conserving Effect**: The weak inflection should first take over the low frequency verbs, and only later the high frequency verbs (Bybee 2006, p. 715; Lieberman et al. 2007).
4. **Class resilience**: Apart from frequency, class membership should be pivotal in whether and when a verb yields to the weak inflection (Mailhammer 2007; Carroll et al. 2012, p. 163-164; Cuskley et al. 2014). Given a number of equally infrequent strong verbs, the verbs that best resist the weak inflection should be those that belong to the more frequent classes.

3.2 Implementational level of model design

The model is an application of the Babel2-framework for agent-based modeling of language (Loetzsch et al. 2008).[9] The first steps in coding the model were to build a fully regular strong system, consisting of several ablaut classes (i.e. building blocks 1 and 2), and to select the verbs that should be conjugated by the agents (building block 4). It is important that all verbs in the model can be conjugated strongly, and that the verbs and ablaut classes exhibit a realistic frequency distribution (cf. Subsection 3.1). As long as these two criteria are met, it is possible to just make up these classes and verbs. Still, we preferred the more straightforward way of directly extracting them from a corpus. As for the classes, we selected all vowel alternations that appeared with more than one verb in the Corpus of Spoken Dutch (CGN, Oostdijk et al. 2002) and did not alter the consonants. This entails that we abstract away from classes sui generis or problematic categories and discern 11 strong vowel-changing patterns, presented in Table 1.[10] They are not just numbered 1 to 11, but carry an identifying number that refers to the original classification in seven classes that is customary in historical grammars (e.g. Prokosch 1939; van Bree 1987), and we distinguish further by using additional letters (II-a, II-b. . .).[11] These identifying numbers are no more than names however; as far as the agents are concerned, the classes III-b and III-c are just as distinct from one another as III-a and V-b, or any other classes.

For the selection of the verbs, all monosyllabic verbs were taken up which solely exhibited these alternations to form their past tense in the Netherlandic component A of the CGN, i.e. the spontaneous conversations.[12] All irregular verbs were discarded (cf. Subsection 3.1). The frequencies

[9] Babel2 is free software downloadable from www.emergent-languages.org/Babel2.

[10] Some clarification is needed here. Class III-c contains Class III verbs that have taken on the preterite ablaut vowel <ie> of Class VII. Class IV-a and V-a look the same in the preterite, but they have a distinct vowel in their participle. Because we disregard participles (cf. Subsection 3.1), they are put together in a single class for present purposes. Concerning the spelling of Dutch, the grapheme for the long vowel is reduced in open syllables: <ee> becomes <e> in *geven*, e.g.

[11] See Knooihuizen and Strik (2014, p. 176) for a similar procedure, and van Bree (1987) and Van Loon (2014) for good overviews of the intricacies of the history of the strong verb classification in Dutch.

[12] The choice for only using monosyllabic verbs was made in accordance with the maxim of simplicity (Landsbergen 2009, p. 18–19). Although there are multisyllabic verbs that can be conjugated strongly, many of them trace back to monosyllabic stems (e.g. *ervaar* ∼ *ervoer* 'to experience' from *vaar* ∼ *voer*) and their status as separate past tense forms can therefore be questioned. Since this discussion is not of direct relevance to our research questions, it should be kept out of the model.

The Netherlandic component A of the CGN already yielded plentiful verbs. Using a larger subset of the corpus would have yielded even more verbs, causing many to be extremely infrequent in the simulations. As a result, a computationally impossible amount of interactions would have to be executed in order for any of these verbs to be used at least once during the lifetime of an agent (see below).

Table 1: Ablaut classes in Dutch

Class	**Phoneme**		**Grapheme**		**Example verb**
	present	preterite	present	preterite	
I	/ɛɪ/	/e/	<ij>	<ee>	*krijgen* → *kreeg*
II-a	/i/	/o/	<ie>	<oo>	*vliegen* → *vloog*
II-b	/oey/	/o/	<ui>	<oo>	*kruipen* → *kroop*
III-a	/ɪ/	/ɔ/	<i>	<o>	*vinden* → *vond*
III-b	/ɛ/	/ɔ/	<e>	<o>	*trekken* → *trok*
III-c	/ɛ/	/i/	<e>	<ie>	*sterven* → *stierf*
IV/V-a	/e/	/ɑ/	<ee>	<a>	*geven* → *gaf*
V-b	/ɪ/	/ɑ/	<i>	<a>	*zitten* → *zat*
VI	/a/	/u/	<aa>	<oe>	*dragen* → *droeg*
VII-a	/a/	/i/	<aa>	<ie>	*laten* → *liet*
VII-b	/ɑ/	/ɪ/	<a>	<i>	*hangen* → *hing*

of these verbs in the simulation is directly correlated with the frequency of their singular past tense in the same subcorpus. This selection procedure yielded the verbs and frequencies of the simulated world in Appendix A.

There are several reasons why we have made use of a corpus of contemporary Dutch to tackle a problem dating back from Proto-Germanic. First, there are no corpora of Proto-Germanic. Using a corpus of Middle Dutch or Gothic would arguably be as 'wrong' as using one from Modern Dutch. Moreover, the CGN has the advantage of being annotated and being more representative of frequency distributions in spoken language than the written material from Gothic and Middle Dutch. Second, in principle, it doesn't matter. Any model that strictly complies with the five elementary building blocks listed above and is capable of generating the four evaluation criteria, will do. Third, the use of actual language from modern day Dutch on the one hand makes the model intuitively interpretable, while it on the other hand makes explicit that the model is not a realistic representation of Proto-Germanic as a living language. This is not the goal of the model and the model is far too simplistic to be used for this end. Still, it should be noted that the choice for Dutch instead of another Germanic language that still retains a fully functioning strong system, is essentially arbitrary. What is crucial here is only to obtain a strong system consisting of several ablaut classes, and a number of verbs that can be conjugated using them, with both classes and verbs exhibiting a realistic frequency distribution.

The flowchart in Figure 1 represents a single interaction between two agents. An event happens in the world, which needs to be expressed by a randomly selected speaker agent. To do so, the speaker agent *ag1* has at his disposal a lexicon L_{ag1} and a grammar G_{ag1}. L_{ag1} directly links the event with verb forms the agent has heard before. G_{ag1} contains all the constructions the agent knows to form past tenses, i.e. the strong classes and the weak suffix. He first looks up the event in L_{ag1}. Minimally two possible verb forms exist for each event, a strong one and a weak one, though they may not always be known to the agent. If the agent knows more than one past tense form, they enter in competition. The probability of each form is given by its relative count, i.e. the number of times the agent has heard that form.[13] If the agent can't find the event in L_{ag1}, because he doesn't know any past tense forms to express it, he will create a new past tense form using G_{ag1}. If the speaker knows more than one construction that can be applied to the verb in question, the applicable constructions enter into competition in the same way as the verb forms. If he does not know any applicable constructions, the speaker remains silent, communication fails and the turn passes without anything being said. If the speaker does utter a verb form, the randomly selected hearer looks it up and adds 1 to the counts of the used form and construction. The hearer is always

[13] This implementation of competition is taken over from Beekhuizen et al. (2014, p. 48), except that no smoothing is employed: if an agent has never heard a form or construction, he will not be able to use it.

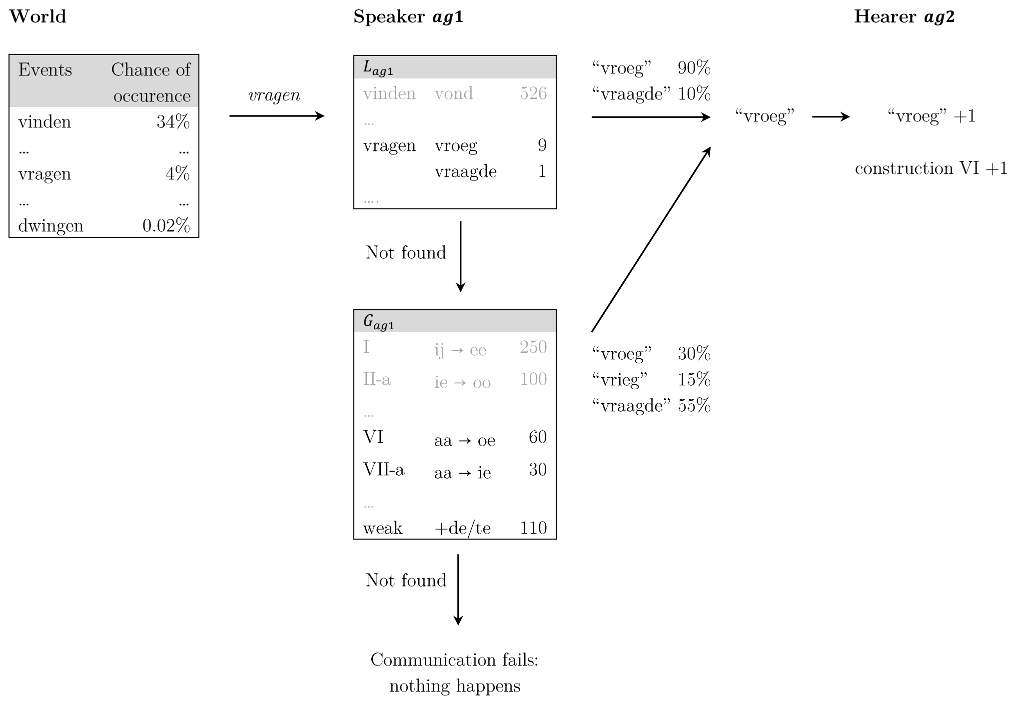


Figure 1: Flowchart of an interaction

able to correctly recognize the used verb form and construction from the utterance (cf. Subsection 3.1). If this is the first time that the hearer hears the verb form or the construction, he adds them to his lexicon respectively grammar with count 1. Once an agent has acquired a construction in this way, he can now apply it to all verbs for which it is applicable. No feedback is given to the speaker. The constructions are implemented using the Fluid Construction Grammar formalism (Steels 2011; van Trijp et al. 2012).[14]

After a certain amount of interactions at the population level, an agent is taken out of the simulation and replaced by an agent who doesn't know any verb forms or constructions, and needs to acquire them as described above. The agent being replaced is always the 'oldest' agent. If several agents are of the same age, one is randomly selected. Once an entire population of agents has been replaced in this way, we will say that one generation has passed.

[14]It may be claimed that it is unrealistic that the constructions only play a role in production if the agent has yet to hear a verb form to express the event. An alternative would be to have the choice between two competing verb forms be decided by a formula which takes into account both the counts of the verb forms and their respective constructions. For instance, this would mean that, if an agents would know the verb forms given in the example in Figure 1, the probability of weak *vraagde* (10%) would be raised because of the high count of the weak construction compared to the competing strong constructions. While we agree that implementing such a formula may be more realistic, it would also make the model more complex. For one, such a design choice would necessitate the addition of two parameters, which would also determine the influence of the counts of the verb forms and constructions in the formula. Moreover, the current implementation of choice-making entails that the agents are highly conservative, only inventing new verb forms if they are left no other option. If anything, this conservatism will impede the rise of the weak inflection, not facilitate it.

4. Results

This section discusses the results of two simulation runs of the model described above, under two different starting conditions, presented respectively in Subsection 4.1 and 4.2. In both runs, all starting agents know perfectly how to conjugate each verb (cf. Appendix A). The starting counts of all strong constructions and verb forms are taken directly from the CGN. The difference between both runs is that in the first, the weak construction is kept out of the model. This is the control condition, in which it can be assessed how the model behaves without the weak inflection. In the second run, the weak construction is then added to the model. Here, the results can be judged as to whether they comply with the evaluation criteria set out in Subsection 3.1.

4.1 Competing strong classes

Figures 2-4 show the development of the type frequencies for each pair of competing strong classes.[15] Only the competing classes, i.e. classes with overlapping vowels in the present, are shown, as only in this case, change is possible. These competing classes compete, in principle, for the same lexemes: a verb that has <aa> in the present, may cause the agent to select either class VI <aa> → <oe> or class VII-a <aa> → <ie>.[16]

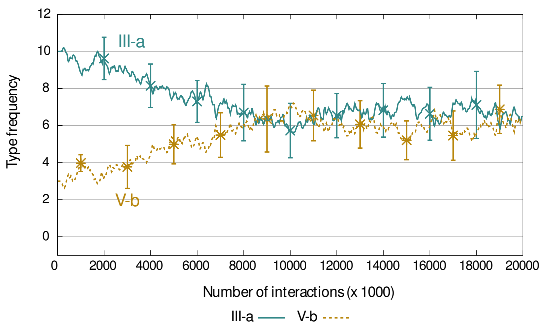


Figure 2: Type frequencies of the competing III-a and V-b constructions

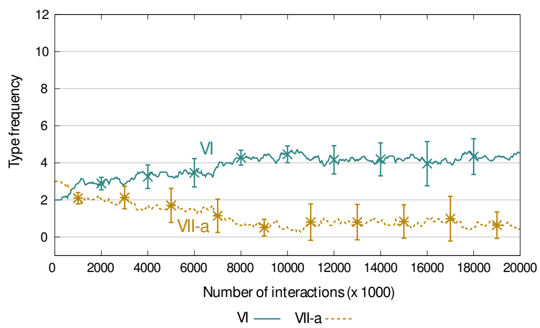


Figure 3: Type frequencies of the competing VI and VII-a constructions

[15]The parameter settings of all graphs shown in this paper are the following: number of agents = 10, replacement rate = 1 agent every 20.000 interactions. Under these settings, the graphs show 8 series of 20.000.000 interactions each, i.e. the passing of 100 generations. The curves depict the running average over all series and the error bars illustrate the standard deviation.

The type frequencies are calculated in the following way. If an agent knows a verb, he will 'vote' as to which construction it belongs to. An agent will always vote for the construction in which he has most often heard it being used, i.e. the construction which he would most probably use himself for expressing the past tense of the verb. The verb is then counted among the construction for which it has received most votes. If the agent has never heard the verb before, he will not cast a vote. If none of the agents knows the verb, the verb is not counted among any of the constructions.

[16]The changes are most clearly visible in type frequencies. For the development of token frequencies, see Appendix B. The token frequencies are calculated by averaging over the counts in the agents' lexica. These token frequencies are set to 0 each time the respective agent is replaced, which causes artificial drops in raw token frequencies. Therefore, the graphs show them in percentages.

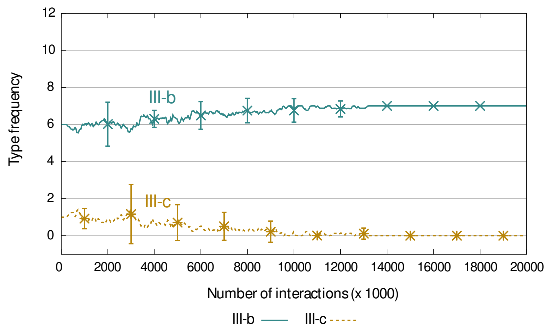


Figure 4: Type frequencies of the competing III-b and V-c constructions

In Figure 2, the high frequency classes III-a and V-b balance each other out. In the first half of the simulation, the V-b class appears to be expanding at the cost of III-a class, but this is simply because variation arises among the verbs of low frequency. This burgeoning variation mostly goes at the expense of the III-a class, since many of the original III-a verbs are of low frequencies, e.g. *dringen, dwingen, springen* etc., while there is only a single original V-b verb of low frequency, *bidden*. Once this difference has been automatically leveled out, the competition stabilizes. In Figure 3, the initially more frequent VI class is clearly taking over from the less frequent VII-a class.[17] Finally, in Figure 4, it can be seen that the III-b class drives the less frequent III-c class to extinction. In a simulation run without the weak inflection, either the competing constructions hold each other in balance or the initially most frequent one prevails.

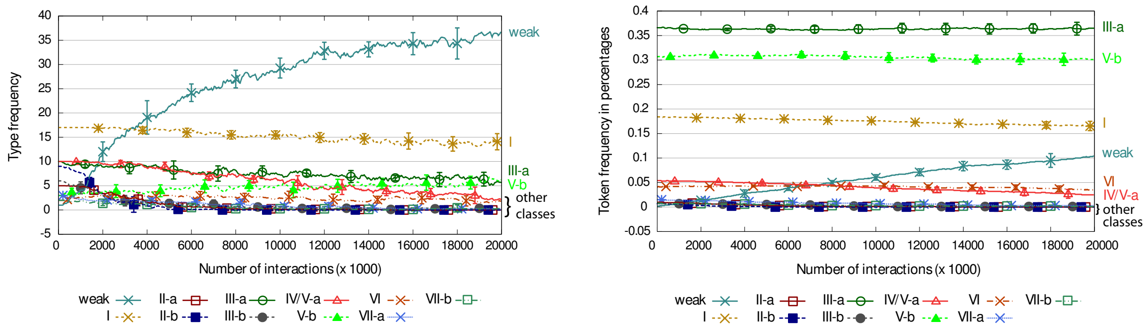
4.2 Weak inflection starting inferior

In the next experimental run, the weak construction is brought into the simulation. As such, the weak construction needs to be assigned a starting position, which should be inferior to the starting position of any individual strong construction in both type and token frequency. In order to secure comparability with the run in Subsection 4.1, we had the weak inflection take the starting position of the feeblest strong construction in the model, i.e. the III-c construction.[18] In the model, III-c starts inferior to the other strong constructions in both token (10) and type frequency (1). Moreover, in the previous simulation, it was driven to extinction by its direct competitor, the III-b construction. This choice of starting position means that at the start of the simulation, only a single low frequency verb, *sterven*, with initial count 10, is conjugated weakly (cf. Appendix A). Note that the only difference between the III-c construction in Subsection 4.1 and the weak construction in the current subsection lies in the weak construction's general applicability. That is, it lies in the weak construction's ability to be, in principle, applicable to any verb.

Figure 5 shows the development of the type frequencies of the past tense constructions, as well as the token frequencies, expressed in percentages of the total of occurrences. What we observe in both graphs is a slow rise of the weak inflection. Immediate victims of this rise are the II-b and III-b classes, which have the lowest token frequency and contain exclusively low frequency verbs. The classes with the highest frequency – I, III-b and V-b – offer more stubborn resistance. Even after 20 million interactions and 100 full generation turnovers, these three classes still dominate token frequency.

The weak construction exhibits an initially outspoken and later decelerated rise in type frequency, and a very slow, but steady rise in token frequency. This finding already indicates that the weak inflection is initially expanding among the low frequency verbs, which can be more clearly seen in

[17] Class VI starts less frequent in type frequency, but more frequent in token frequency (cf. Appendix A).
[18] The parameter settings are also kept constant with the previous run in Subsection 4.1.



(a) Token frequency in percentages (b) Type frequency

Figure 5: Development of the constructions

Figure 6. Figure 6 shows the percentage of type frequency of the weak construction among the extreme low frequency verbs, the low frequency verbs, the middle frequency verbs and the high frequency verbs. The (extreme) low frequency verbs are quickly becoming weakly inflected, followed by the middle frequency verbs. The high frequency verbs remain deeply entrenched in their strong forms during almost the entire simulation. The model thus complies with evaluation criteria (1) *Rise of the weak inflection on token and type level*, (2) *Gradual rise* and (3) *Conserving Effect*.

In Figure 7, the development of six equally infrequent verbs can be checked, expressed in percentage of weak token frequency. The verbs *zuigen*, *stuiven* and *vechten*, initially belonging to the infrequent II-b and III-b classes, rapidly yield to the weak inflection, while the verbs *dwingen*, *dringen* and *bidden*, belonging to the III-a and V-b classes, do not. These infrequent verbs enjoy the protection of their highly frequent class mates *vind* ∼ *vond* (III-a) and *zit* ∼ *zat* (V-b). Appendix B contains the development of the weak token frequency of all verbs in the model.[19] Here, it can be checked that the behavior of these six verbs is indicative of the behavior of others, with the most frequent strong classes best retaining even their low frequency verbs. Evaluation criterion (4) *Class resilience*, is thus also accounted for.

The protecting influence of these classes is responsible for two more findings in the results of the model. The first is the flattening of the rise of the weak inflection in Figure 5a and in Figure 6. The easy prey, i.e. the low frequency strong classes, are rapidly taken over by the weak inflection. Once these have fallen however, further expansion is slowed down as the remaining low frequency verbs fall under the protection of the highly frequent strong classes. The second finding is that a lower percentage of the extreme low frequency verbs than of the low frequency verbs seem to be turning over to the weak inflection. The cause of this is that a larger percentage of these extreme low frequency verbs (45,5%) than of the low frequency verbs (33.3%) are protected by the classes I, III-a and V-b (see Appendix A).

The first argument against general applicability being a decisive functional advantage for the weak inflection, made the point that the advent of the weak inflection has only complicated past tense inflection in the Germanic languages. This holds true in our model. At the start of a simulation, it is very easy to predict how to conjugate a verb, looking only at its vowel: all verbs except for *sterven* ('to die') are conjugated strongly. At the end of the simulation, variation is much more rampant. However, it should be noted that the agents – and this generally holds for language users in reality (Keller 1994) – are not actively trying to improve their language system. Their goal is much more local: they simply try to form a past tense of the verb they are confronted with. As the accumulated result of such local usage events, a large, slow, and deeply disruptive language change such as the rise of the weak inflection may take place on the community level. Figure 8 shows the number of

[19] To maintain overview, they are divided among the constructions to which they belong at the start of the simulation.

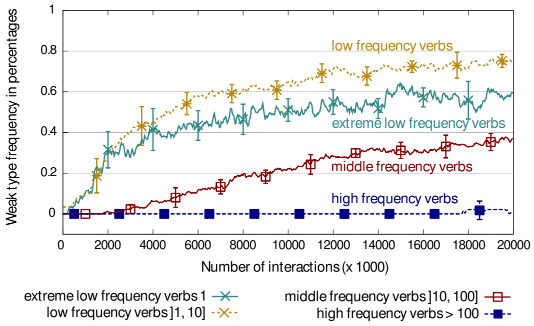


Figure 6: Type frequency in percentages of the weak construction among the extreme low, low, middle, and high frequency verbs

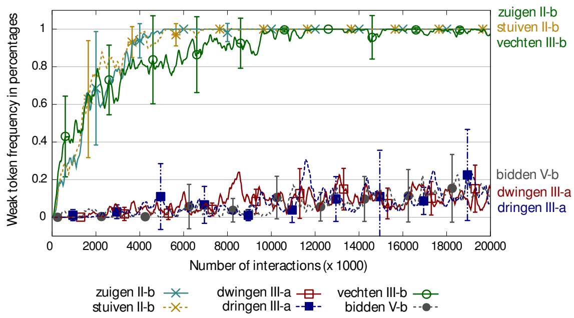


Figure 7: Token frequency in percentages of the weak forms of 6 equally infrequent verbs belonging to different classes

times a speaker fails to communicate with his hearer during 1 million interactions, because he does not know any way to express the past tense of the verb in question, in both experimental runs. In the first experimental run, containing only the strong classes (cf. Subsection 4.1), the number of failures remains more or less constant. In the second experimental run, also containing the weak inflection, the number of failures drops gradually as the weak inflection accumulates more verbs, i.e. rises in type frequency. While it may be argued that in the second experimental run, the system has become more complex by the end of the simulation, the agents also more often succeed in expressing the past tense.

The second argument against general applicability claimed that it is only an advantage if the speaker often needs to express the past tense of a verb for which he has yet to acquire an applicable strong class. Once an agent has properly acquired the system, he will stop using 'faulty' weak forms. This also holds true in our model. Figure 9 shows the number of incorrect weak verb forms on average produced by a single agent and his 'successors' during 1000 interactions in the first million interactions of the second experimental run.[20] Weak forms are counted as incorrect if they are not weakly conjugated in the original starting setup.[21] It shows that once the agent is replaced, the new learner agent does produce a number of such overgeneralized weak forms, but rapidly learns that these verbs should not be conjugated weakly, and acquires the correct past tense inflection. After a peak when the agent has just entered the simulation, the usage of these faulty weak forms quickly diminishes. It is also true in our model that initially, a learner agent confronted with any unknown verb is more likely to have already acquired an applicable strong construction than the weak construction, as each strong construction is initially more frequent than the weak construction. However, once the weak construction is acquired, it is available to form the past tense of any verb, regardless of its vowel. This seemingly tiny advantage already suffices for the weak inflection to start its ascent from a starting position which is massively inferior in frequency.

The model only contains two parameters. One is the number of agents involved. As shown in earlier models (Nettle 1999b; Blythe and Croft 2012: 19), a larger population of agents means change is slower, because it takes longer to spread throughout the entire population.[22] The second

[20] By 'successors', we mean the agents that replace him once he is taken out of the simulation.

[21] In the current simulation, these are the weak forms of all verbs except *sterven* ('to die') (see Appendix A).

[22] The results of these models may thus be seen as in conflict with what can be found in reality, where languages spoken in small communities are often more conservative than those spoken in large communities, as is the case in e.g. the Romance and Germanic language families (also see Lupyan and Dale 2010). This indicates that to explain this tendency, more assumptions should be added. One such possible assumption could be that larger communities have a higher number of potential language innovators, and therefore have a higher potential for generating new language elements. Another assumption could be that larger language communities tend to correlate with a higher number of

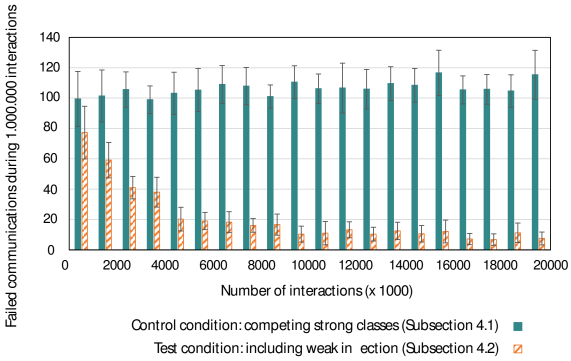


Figure 8: Number of times an agent failed to communicate during 1 million interactions in both experimental runs

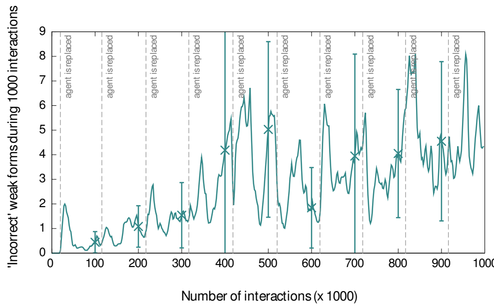


Figure 9: Number of times a single agent used an 'incorrect' weak form during 1000 interactions.

parameter is the replacement rate of the agents. Setting this rate lower, i.e. letting a longer time pass before replacing the next agent, will also slow down the rise of the weak inflection. However, perhaps contrary to intuition, a higher replacement rate is needed for the weak inflection to die out. The reason is that initially, there is only a single low frequency verb that is conjugated weakly. The higher the replacement rate is set, the higher the chance that an entire generation of agents is replaced without ever using this verb, i.e. without passing on the weak inflection to the next generation. Once this happens, the weak inflection has effectively died out, and in the current model, has no chance of reviving. The eventual rise of the weak inflection, as well as the emergence of the other evaluation criteria, are thus not dependent upon specific parameter settings. Only if the replacement rate is set so high that the weak construction is simply never passed on to the following generation, can its rise be prevented.

5. Conclusions

In this paper, we set out to answer two research questions. Firstly, we asked ourselves whether a specific grammatical innovation can overcome the threshold of frequency which is needed to become the dominant variant in a community, through functional selection. Secondly, we wondered whether the general applicability of the weak inflection suffices to explain its rise.

Having assessed the behavior of the model, both research questions can be answered positively. As for the first question, a functional advantage of a specific grammatical variant may suffice to overcome the Threshold Problem. This does not contradict, however, that social selection may be an important and widely used mechanism to overcome the threshold of frequency. For one, the functional advantage under scrutiny in this study, general applicability, is specific to the weak inflection. While our findings can be used to show that in principle, a functional advantage such as general applicability may suffice to overcome the threshold of frequency, this does not simply carry over to other grammatical competitions. Other competitions may show very different functional advantages, whose effects deserve to be studied independently, as has been done in e.g. Landsbergen (2009, p. 47-74) and van Trijp (2013).

As for the second question, the general applicability of the weak inflection suffices to explain its rise, even under the conditions that (i) the strong inflection is fully regular and (ii) the weak inflection has to start from a position vastly inferior in frequency. We do not need to assume that

second language learners and more language contact (cf. Lupyan and Dale 2010). This may have sweeping effects on morphology (Dale and Lupyan 2012; Bentz and Winter 2013).

the strong system was already becoming irregular for the weak inflection to step up to the task of taking over past tense formation. The disintegration of the strong system might rather be the result – and perhaps subsequent catalyst – of the weak inflection's advent. That is, with the weak inflection taking over the low frequency verbs, there would be less pressure on the strong system to uphold its regularity. The resulting decline in regularity of the strong system may then in turn fuel the rise of the weak inflection, setting off a vicious cycle of weak ascension and strong decline.

These conclusions of course do not tell the whole story; many questions still remain. For one, we have assumed that the weak dental suffix was already present as a strategy for forming the preterite. We have not attempted to explain how it came into being, which is the topic of a century-old but still ongoing debate (o.a. Collitz 1912; Ball 1968; Meid 1971; Shields 1982; Ringe 2006, p. 179-785; Hill 2010). Still, most accounts of the origin of the dental suffix, e.g. the 'to do' composition proposal (o.a. Loewe 1898, p. 356-357; Tops 1974; Hill 2010), are compatible with the thesis that the suffix was, in principle, generally applicable from its very first appearance in the language.

We hope that computational and historical linguists will continue to work together to shed light on such matters, which are currently still shrouded in history.

Acknowledgments

We would like to thank Remi van Trijp for interesting discussions and excellent advice about the model, as well as the participants of the SLE 48 workshop 'Shifting classes: Germanic strong and weak preterites and participles'. This article has also been improved thanks to the useful comments of three anonymous reviewers, to whom we are most grateful.

References

Bailey, Christopher Gordon (1997), *The Etymology of the Old High German Weak Verb*, PhD thesis, University of Newcastle upon Tyne.

Ball, Christopher (1968), The Germanic dental preterite, *Transactions of the Philological Society* **67**, pp. 162–188.

Baxter, Gareth, Richard Blythe, William Croft, and Alan McKane (2009), Modeling language change: An evaluation of Trudgill's theory of the emergence of New Zealand English, *Language Variation and Change* **21** (2), pp. 257–296.

Beekhuizen, Barend, Rens Bod, and Arie Verhagen (2014), A Usage-Based Model of Early Grammatical Development, *Proceedings of the Fifth Workshop on Cognitive Modeling and Computational Linguistics*, Association for Computational Linguistics, Baltimore, pp. 46–54.

Bentz, Christian and Bodo Winter (2013), Languages with More Second Language Learners Tend to Lose Nominal Case, *Language Dynamics and Change* **3** (1), pp. 1–27.

Blythe, Richard and William Croft (2012), S-curves and the mechanisms of propagation in language change, *Language* **88** (2), pp. 269–304.

Bybee, Joan (2006), From Usage to Grammar: The Mind's Response to Repetition, *Language* **82** (4), pp. 711–733.

Bybee, Joan (2010), *Language, usage and cognition*, Cambridge University Press, Cambridge.

Carroll, Ryan, Ragnar Svare, and Joseph Salmons (2012), Quantifying the evolutionary dynamics of German verbs, *Journal of Historical Linguistics* **2** (2), pp. 153–172.

Colaiori, Francesca, Claudio Castellano, Christine Cuskley, Vittorio Loreto, Martina Pugliese, and Francesca Tria (2015), General three-state model with biased population replacement: Analytical solution and application to language dynamics, *Physical review. E, Statistical, nonlinear, and soft matter physics* **91** (1-1), pp. 12808.

Collitz, Hermann (1912), *Das schwache Praeteritum und seine Vorgeschichte*, Vandenhoeck and Ruprecht, Göttingen.

Croft, William (2000), *Explaining language change: An evolutionary approach*, Pearson Education Limited, Essex.

Cuskley, Christine, Martina Pugliese, Claudio Castellano, Frascesca Colaiori, Vittorio Loreto, and Franscesca Tria (2014), Internal and External Dynamics in Language: Evidence from Verb Regularity in a Historical Corpus of English, *Plos One* **9** (8), pp. e102882.

Dale, Rick and Gary Lupyan (2012), Understanding the origins of morphological diversity: The linguistic niche hypothesis, *Advances in Complex Systems* **15** (3), pp. 1150017–1–1150017–16.

de Boer, Bart (2012), Modelling and Language Evolution: beyond fact-free science, *in* McCrohon, Luke, Tomomi Fujimura, Kazuo Okanoya, Koji Fujita, Reiji Suzuki, Roger Martin, and Noriaki Yusa, editors, *The Evolution of Language: Proceedings of the 9th International Conference*, Evolang-9 Organizing Committee, pp. 83–92.

Goldberg, Adele Eva (2003), Constructions: A new theoretical approach to language, **7** (5), pp. 219–224.

Grimm, Jacob (1819), *Deutsche Grammatik*, Dieterichsche Buchhandlung, Göttingen.

Haeseryn, Walter, Kirsten Romijn, Guido Geerts, Jaap de Rooij, and Maarten van den Toorn (1997), *Algemene Nederlandse Spraakkunst*, Nijhoff, Groningen.

Harbert, Wayne (2007), *The Germanic languages*, Cambridge university press, Cambridge.

Hare, Mary and Jeffrey Elman (1995), Learning and morphological change, *Cognition* **56** (1), pp. 61–98.

Haspelmath, Martin (1999), Why is grammaticalization irreversible?, *Linguistics* **37** (6), pp. 1043–1068.

Hill, Eugen (2010), A case study in grammaticalized inflectional morphology: Origin and development of the Germanic weak preterite, *Diachronica* **27** (3), pp. 411–458.

Keller, Rudi (1994), *On language change: the invisible hand in language*, London Routledge, 1994, London.

Knooihuizen, Remco and Oscar Strik (2014), Relative productivity potential of Dutch verbal inflection patterns, *Folia linguistica historica* **35** (1), pp. 173–200.

Labov, William (1966), Hypercorrection in the lower middle class as a factor in linguistic change, *in* Bright, William, editor, *Sociolinguistics*, Mouton, Den Haag, pp. 84–113.

Labov, William (2001), *Principles of linguistic change, vol. 2: Social factors*, Blackwell, Oxford.

Landsbergen, Frank (2009), *Cultural evolutionary modeling of patterns in language change: exercises in evolutionary linguistics*, PhD thesis, University of Leiden, Utrecht.

Landsbergen, Frank, Robert Lachlan, Carel ten Cate, and Arie Verhagen (2010), A cultural evolutionary model of patterns in semantic change, *Linguistics* **48** (2), pp. 363.

Lieberman, Erez, Jean-Baptste Michel, Joe Jackson, Tina Tang, and Martin Nowak (2007), Quantifying the evolutionary dynamics of language, *Nature* **449** (7163), pp. 713–716.

Ling, Charles and Marin Marinov (1993), Answering the connectionist challenge: a symbolic model of learning the past tenses of English verbs, *Cognition* **49** (3), pp. 235–290.

Loetzsch, Martin, Pieter Wellens, Joachim De Beule, Joachim Bleys, and Remi van Trijp (2008), The Babel2 Manual, *AI-Memo 01-08*, AI-Lab VUB, Brussels.

Loewe, Richard (1898), Das schwache Präteritum des Germanischen, *Indogermanische Forschungen* **8**, pp. 254–266.

Lupyan, Gary and Rick Dale (2010), Language structure is partly determined by social structure, *PloS one* **5** (1), pp. e8559.

Macwhinney, Brian and Jared Leinbach (1991), Implementations are not conceptualizations: revising the verb learning model, *Cognition* **40** (1-2), pp. 121.

Mailhammer, Robert (2007), Islands of resilience: The history of the German strong verbs from a systemic point of view, *Morphology* **17**, pp. 77–108.

Marcus, Gary, Ursula Brinkmann, Harald Clahsen, Richard Wiese, and Steven Pinker (1995), German inflection: the exception that proves the rule, *Cognitive Psychology* **29** (3), pp. 189.

Meid, Wolfgang (1971), *Das germanische Praeteritum*, Institut für vergleichende Sprachwissenschaft der Universität Innsbruck, Innsbruck.

Mesoudi, Alex (2011), *Cultural Evolution: How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences*, University Of Chicago Press, Chicago.

Milroy, James and Lesley Milroy (1985), Linguistic change, social network and speaker innovation, *Journal of Linguistics* **21** (2), pp. 339–384.

Nettle, Daniel (1999), Using social impact theory to simulate language change, *Lingua* **108** (2-3), pp. 95–117.

Oostdijk, Nelleke, Wim Goedertier, Frank Van Eynde, Louis Boves, Jean-Pierre Martens, Michael Moortgat, and Harald Baayen (2002), Experiences from the Spoken Dutch corpus project.

Pijpops, Dirk and Katrien Beuls (forthcoming), Agent-gebaseerde modellering in de historische taalkunde. Een model van regularisatiedruk op de Nederlandse werkwoorden, *Handelingen der Koninklijke Zuid-Nederlandse Maatschappij voor Taal- en Letterkunde en Geschiedenis.*

Pinker, Steven and Alan Prince (1988), On language and connectionism: Analysis of a parallel distributed processing model of language acquisition, *Cognition* **28** (1), pp. 73–193.

Plunkett, Kim and Patrick Juola (1999), A Connectionist Model of English Past Tense and Plural Morphology, *Cognitive Science* **23** (4), pp. 463–490.

Plunkett, Kim and Virginia Marchman (1991), U-shaped learning and frequency effects in a multi-layered perception: Implications for child language acquisition, *Cognition* **38** (1), pp. 43–102.

Plunkett, Kim and Virginia Marchman (1993), From rote learning to system building: acquiring verb morphology in children and connectionist nets, *Cognition* **48** (1), pp. 21–69.

Prokosch, Eduard (1939), *A comparative Germanic grammar*, Philadelphia Linguistic society of America, 1939, Philadelphia.

Ringe, Don (2006), A sociolinguistically informed solution to an old historical problem: the Gothic genitive plural, *Transactions of the Philological Society* **104** (2), pp. 167–206, Oxford, UK.

Rogers, Everett (1995), *Diffusion of Innovations*, 4th ed., Free Press, New York.

Rumelhart, David and James McClelland (1986), On learning the past tense of English verbs, *in* Rumelhart, David and James McClelland, editors, *Parallel distributed processing: explorations in the microstructure of cognition*, MIT Press, Cambridge, pp. 216–271.

Shields, Kenneth (1982), The origin of the Germanic dental preterite: A new proposal, *Leuvense Bijdragen* **71**, pp. 427–440.

Steels, Luc (2011), *Design Patterns in Fluid Construction Grammar*, John Benjamins, Amsterdam.

Taatgen, Niels and John Anderson (2002), Why do children learn to say "Broke"? A model of learning the past tense without feedback, *Cognition* **86**, pp. 123–155.

Tops, Guy (1974), *The origin of the Germanic dental preterit*, Brill, Leiden.

Trudgill, Peter (2008), Colonial dialect contact in the history of European languages: on the irrelevance of identity to new-dialect formation, *Language in Society* **37** (2), pp. 241–254.

van Bree, Cornelis (1987), *Historische grammatica van het Nederlands*, Foris, Dordrecht.

Van Loon, Jozef (2014), *Historische fonologie van het Nederlands*, 2nd ed., Deurne Universitas 2014.

van Noord, Rik (2015), Modeling the learning of the English past tense with memory-based learning, *Computational Linguistics in the Netherlands (CLIN)*, Antwerp, 6 February.

van Trijp, Remi (2013), Linguistic Assessment Criteria for Explaining Language Change: A Case Study on Syncretism in German Definite Articles, *Language Dynamics and Change* **3**, pp. 105–132.

van Trijp, Remi, Luc Steels, Katrien Beuls, and Pieter Wellens (2012), Fluid construction grammar: The new kid on the block, *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, ACL, Avignon.

Yang, Charles (2002), *Knowledge and learning in natural language*, Oxford University Press, Oxford.

Zipf, George Kingsley (1932), *Selected Studies of the Principle of Relative Frequency in Language*, Harvard University Press, Harvard.

Appendix A: Model Design

Table 2: World

Rank	Event	Chance of occurrence
1	vinden	34.0435 %
2	zitten	25.9475 %
3	krijgen	8.0511 %
4	liggen	4.6647 %
5	vragen	3.9022 %
6	kijken	2.3324 %
7	lijken	1.9287 %
8	blijken	1.7268 %
9	blijven	1.6596 %
10	geven	1.5026 %
11	nemen	1.4802 %
12	rijden	1.3232 %
13	lezen	1.1213 %
14	laten	1.0316 %
15	schrikken	0.7176 %
16	hangen	0.7176 %
17	spreken	0.7176 %
18	schrijven	0.5607 %
19	trekken	0.5158 %
20	vliegen	0.4485 %
21	zingen	0.4485 %
22	winnen	0.4261 %
23	slapen	0.3813 %
24	klinken	0.314 %
25	schijnen	0.2915 %
26	schieten	0.2467 %
27	stinken	0.2467 %
28	dragen	0.2467 %
29	eten	0.2243 %
30	sterven	0.2243 %
31	drinken	0.2018 %
32	springen	0.1794 %
33	kruipen	0.1346 %
34	steken	0.1346 %
35	rieken	0.1121 %
36	bieden	0.1121 %
37	spuiten	0.1121 %
38	ruiken	0.1121 %
39	sluiten	0.1121 %
40	schuiven	0.0897 %
41	gelden	0.0897 %
42	breken	0.0897 %
43	snijden	0.0897 %
44	grijpen	0.0673 %
45	bijten	0.0673 %
46	knijpen	0.0673 %
47	wijzen	0.0673 %
48	duiken	0.0673 %
49	lijden	0.0449 %
50	zwijgen	0.0449 %
51	kiezen	0.0449 %
52	sluipen	0.0449 %
53	treffen	0.0449 %
54	zwemmen	0.0449 %
55	treden	0.0449 %
56	blazen	0.0449 %
57	vangen	0.0449 %
58	rijgen	0.0224 %
59	glijden	0.0224 %
60	zuigen	0.0224 %
61	stuiven	0.0224 %
62	dwingen	0.0224 %
63	dringen	0.0224 %
64	schelden	0.0224 %
65	vechten	0.0224 %
66	stelen	0.0224 %
67	vreten	0.0224 %
68	bidden	0.0224 %

Table 3: Initial lexicon L_{ag}

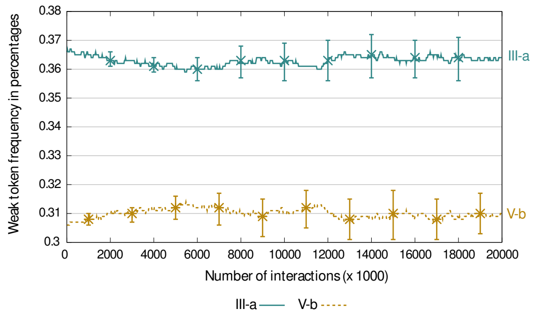
Event	Verb forms	Count
vinden	vond	1518
zitten	zat	1157
krijgen	kreeg	359
liggen	lag	208
vragen	vroeg	174
kijken	keek	104
lijken	leek	86
blijken	bleek	77
blijven	bleef	74
geven	gaf	67
nemen	nam	66
rijden	reed	59
lezen	las	50
laten	liet	46
schrikken	schrok	32
hangen	hing	32
spreken	sprak	32
schrijven	schreef	25
trekken	trok	23
vliegen	vloog	20
zingen	zong	20
winnen	won	19
slapen	sliep	17
klinken	klonk	14
schijnen	scheen	13
schieten	schoot	11
stinken	stonk	11
dragen	droeg	11
eten	at	10
sterven	stierf; sterfde	10
drinken	dronk	9
springen	sprong	8
kruipen	kroop	6
steken	stak	6
rieken	rook	5
bieden	bood	5
spuiten	spoot	5
ruiken	rook	5
sluiten	sloot	5
schuiven	schoof	4
gelden	gold	4
breken	brak	4
snijden	sneed	4
grijpen	greep	3
bijten	beet	3
knijpen	kneep	3
wijzen	wees	3
duiken	dook	3
lijden	leed	2
zwijgen	zweeg	2
kiezen	koos	2
sluipen	sloop	2
treffen	trof	2
zwemmen	zwom	2
treden	trad	2
blazen	blies	2
vangen	ving	2
rijgen	reeg	1
glijden	gleed	1
zuigen	zoog	1
stuiven	stoof	1
dwingen	dwong	1
dringen	drong	1
schelden	schold	1
vechten	vocht	1
stelen	stal	1
vreten	vrat	1
bidden	bad	1

Table 4: Initial grammar G_{ag}

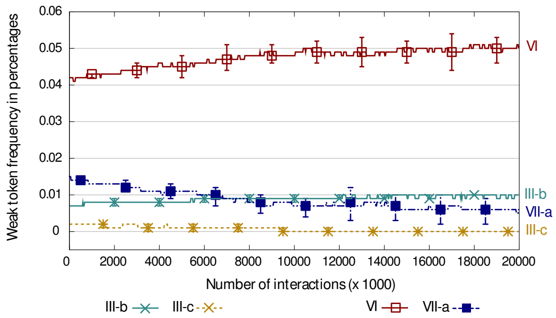
Construction	Effect	Count
I	ij → ee	819
II-a	ie → oo	43
II-b	ui → oo	32
III-a	i → o	1633
III-b	e → o	33
III-c; weak	e → ie; +de/te	10
IV/V-a	ee → a	239
V-b	i → a	1366
VIa	aa → oe	185
VII-a	aa → ie	65
VII-b	a → i	34

Appendix B: Weak token frequencies

Figure 10: Token frequency of the competing strong classes in the control condition, i.e. in the first simulation run described in Subsection 4.1.

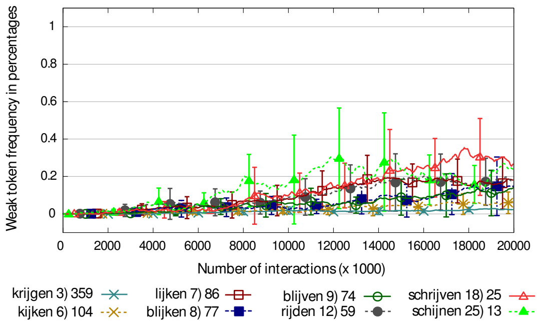


(a) token frequencies of the competing III-a and V-b classes in the first simulation run

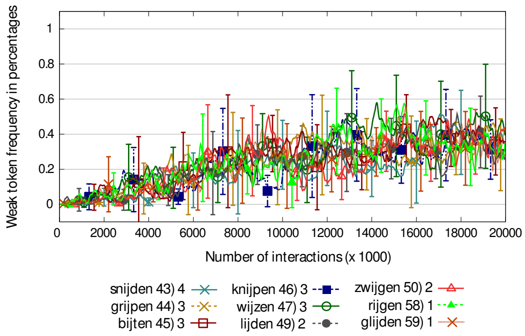


(b) token frequencies of the competing III-b and III-c classes and the competing VI and VII-a classes in the first simulation run

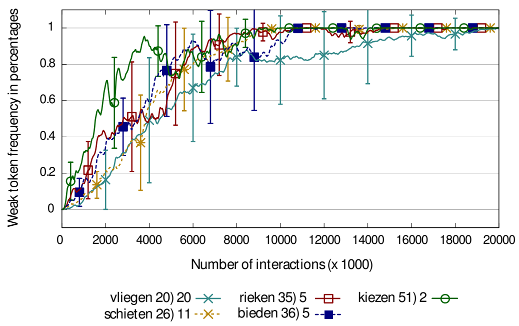
Figure 11: Token frequency of the weak forms of each verb in percentages, for each of the grammatical constructions they start in, in the second simulation run described in Subsection 4.2. Next to each verb, its rank followed by its frequency is shown.



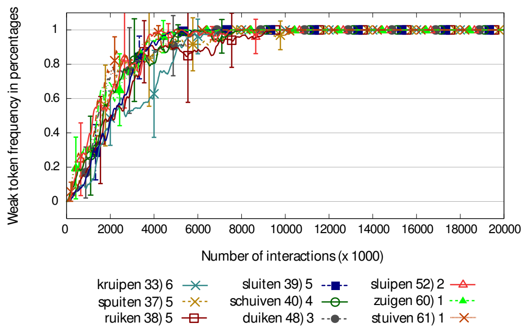
(a) Class I highly frequent verbs



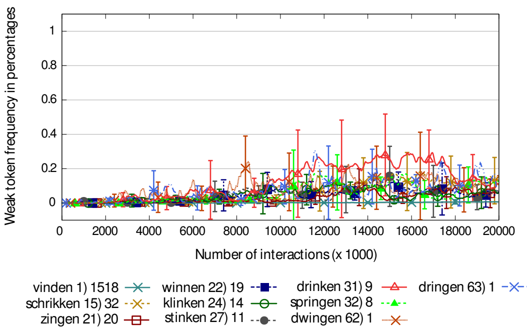
(b) Class I less frequent verbs



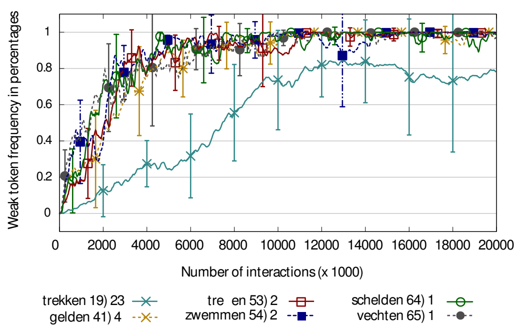
(c) Class II-a



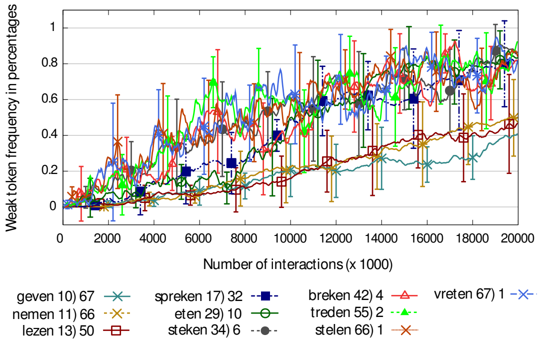
(d) Class II-b



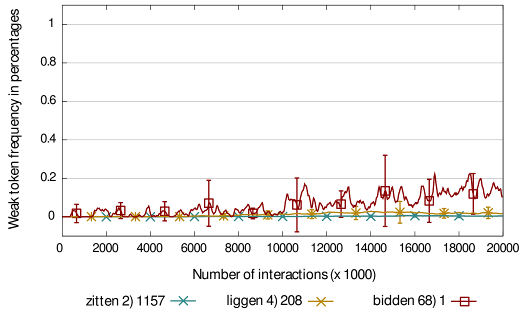
(e) Class III-a



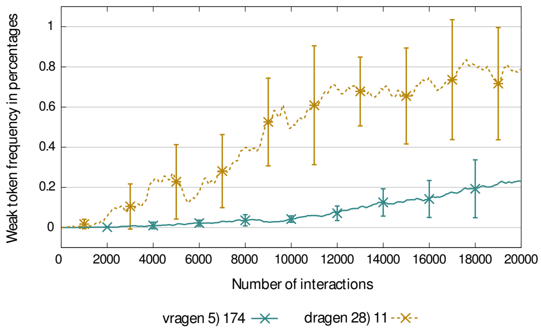
(f) Class III-b



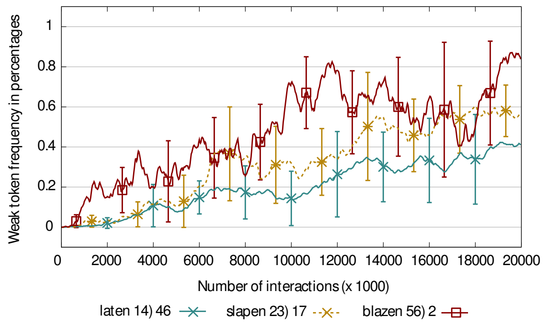
(g) Class IV/V-a



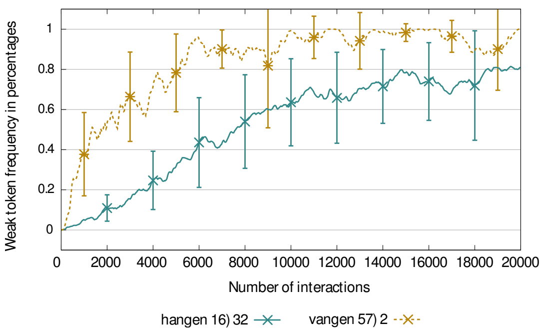
(h) Class V-b



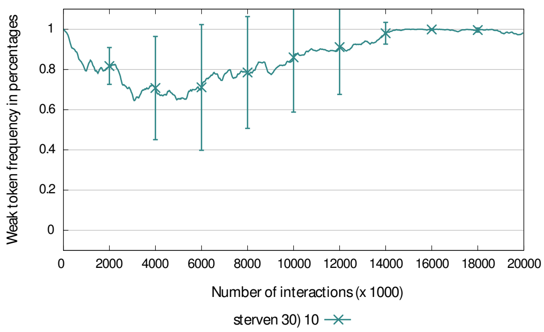
(i) Class VI



(j) Class VII-a



(k) Class VII-b



(l) Weak construction