Şăineanu, L., 1900. *Influenţa orientală asupra limbii şi culturii române.* Bucharest.
Vasiliu, E. 1968. *Fonologia istorică a dialectelor dacromâne.* Bucharest: Editura Academiei Republicii socialiste România.
Vennemann, T. 1978. "Phonetic and Conceptual Analogy". In: P. Baldi and R. Werth (eds), *Readings in Historical Phonology.* University Park Philadelphia: Pennsylvania State University Press, 258–74.
Vincent, N. 1978. "Is Sound Change Teleological?". In: J. Fisiak (ed.), *Recent Developments in Historical Phonology*, 409–29.

University of Oxford
Trinity College Oxford
Oxford OX1 3BH
UK

www.trinity.ox.ac.uk/staff

The development of "junk". Irregularization strategies of HAVE and SAY in the Germanic languages

DAMARIS NÜBLING

1. INTRODUCTION[1]

Although it is a wellknown fact that the most frequent verbs are the most irregular ones (if not suppletive), it is rarely asked how they became irregular. This article deals with the irregularization process of two originally regular (weak) verbs, HAVE and SAY in the Germanic languages, e.g. *have*, but *has/'s* and *had/'d* (instead of regular **haves/*haved*) or *say* [sei], but *says* [sez] and *said* [sed] in English. Other verbs, such as DO, GO, STAND, BE, COME, and so on, also tend to irregularizations again and again without any apparent reason. In contrast to HAVE and SAY these verbs have always been rather irregular, at least dating from their first written records.

Because little attention has been paid to this "regularity of irregularity", this fact has hardly been integrated in morphological theories (for such an integration see, however, Maiden 1991, Janda 1996, Lass 1990, Werner 1987a and b). This sort of irregularity cannot be subsumed under so-called "exaptation" (Lass 1990), that is, the functionalization of morphs which became nonfunctional, but rather under what Lass somewhat derogatively calls "junk" or "marginal garbage":

> "Say a language has a grammatical distinction of some sort, coded by means of morphology. Then say this distinction is jettisoned, prior to the loss of the morphological material that codes it. This morphology is now, functionally speaking, junk; and there are three things that can in principle be done with it:
> (i) it can be dumped entirely;
> (ii) it can be kept as marginal garbage or nonfunctional/nonexpressive residue (suppletion, 'irregularity');
> (iii) it can be kept, but instead of being relegated as in (ii), it can be used for something else, perhaps just as systematic.
> [...] Option (iii) is linguistic exaptation." (Lass 1990: 81/82)

In this paper, we will deal with option (ii), but without sharing the characterization of nonfunctionality: firstly, this type of irregularity correlates too often with highest token frequency; secondly, it is produced too often using different strategies so that it cannot be considered undesired, nonfunctional, accidentally developed "waste from the past".[2] Many morphological theories, such as naturalness theory (see Mayerthaler 1981, Wurzel 1984), consider irregularities to be the result of a "naturalness conflict", a (morphological) price for optimizations at the phonological level. Usually such phonological waste is regularized by analogical processes, but, in the meantime, it has been recognized that high token frequency can have preserving effects in particular if the phonological rule is no longer productive (cf., for example,

Geert Booij and Jaap van Marle (eds), Yearbook of Morphology 1999, 53–74.

Verner's law, which was originally regular and whose last remainders can only be observed in some few verbs: New High German (NHG) *ziehen – zog – gezogen* 'draw – drew – drawn', English (Engl.) *was – were*).

Only the linguistic economy theory,[3] which recognizes the benefit of such irregular relics, considers irregularity including suppletion to be increased formal distinctivity. Distinctivity permits a minimum expression without producing syncretism. Shortness of expression is extremely functional, especially under highest token frequency (see Zipf's law). Here, performance requirements such as short, simple sound chains move to the fore, while competence requirements such as rule-based, additively and transparently structured paradigms fade into the background. These competence requirements develop their effect at intermediate and lower frequencies and therefore concern most verbs, but not the peak level of frequency.

This article aims to demonstrate that irregularity not only develops (passively) by the preservation of nonfunctional relics (in the sense of Lass 1990), but can also be "created" actively by innovative processes. It focuses in particular on this second way of development. Considering that as yet no systematization and, correspondingly, typology of irregularization has been carried out, this article aims to trace the development of the irregularization of two important verbs, namely HAVE and SAY in ten Germanic languages, through the analysis of a relatively small sample. This is also intended to demonstrate that this phenomenon is valid crosslinguistically. Finally, we will consider the costs and benefits of morphological irregularity.

2. IRREGULARIZATION PROCESSES IN THE GERMANIC LANGUAGES: DIFFERENT PATHS TO "JUNK"

Irregularization processes become the most visible in cases where the verb was very regular before its irregularization (that is, it was not yet relatively differentiated, such as the strong verbs or the modal verbs). We will therefore analyze two originally (and, in some languages still) weak verbs, HAVE and SAY. Weak verbs are characterized by stable, uniform, agglutinatively connected morphs, such as NHG {lach}-{en} – {lach}-{t}-{e} – {ge}-{lach}-{t} 'laugh – laughed – laughed.' For this reason, in all Germanic languages, weak verbs represent the largest and most productive inflectional class into which many strong verbs transfer and to which all newly created and borrowed verbs are assigned. In contrast, strong verbs work with fusional (flectional) strategies and have – in particular in the modern Germanic languages – a clear affinity for irregular verbs since their tense ablaut (vowel change) is no longer predictable. The former seven Germanic ablaut classes have more or less split up to different degrees; today, New High German has about 50 different vowel alternations. Most of these alternation classes have only one member (see Augst 1975, Hempen 1988). Here, a high degree of irregularity has already been reached through natural processes. If this degree of irregularity correlates with a high token frequency, it tends to be preserved; otherwise, such "overdifferentiated" verbs transfer to more regular classes, mostly to the weak verbs.

In regard to diachronic depth, we will go back to the first written evidence. Since we must reckon with irregular developments at any time, it would be contradictory to use reconstructed forms which are always based on regular developments. In the case of the irregularizations, we will limit our analysis to only the most important strategies (for more details, see Nübling 2000). For reasons of space, we cannot list every individual paradigm (see, however, Tables 2 and 3, which contain the most important forms of HAVE and SAY).

2.1. HAVE *in the Germanic languages*

In all Germanic languages, a periphrastic perfect formed with the auxiliary HAVE (in some languages also with BE) developed about 1000 years ago. During its process of grammaticalization, HAVE became very frequent; today it shares the top region of the frequency scales with BE. As a consequence of this rapid increase of token frequency, HAVE has been extremely reduced and irregularized in all languages.

In the **New High German** paradigm of *haben*, a longer stem *hab-* [ha:b]- alternates in a unique way with a shorter stem *ha-* [ha]. In the present tense these "allo-stems" have "changing inflection" in the 2. and 3.sg.pres. of many strong verbs (NHG "Wechselflexion", i.e. the raising/leveling of $e > i$ in NHG *gebe* (1.sg.) vs. *gibst/gibt* (2./3.sg.) and the umlaut forms *falle* (1.sg.) vs. *fällst/fällt* (2./3.sg.)); that is, the finite form of NHG *haben* are distributed structurally analogous to the strong verb *fallen* with *falle* vs. *fällst/fällt* in the singular and *fallen, fallt, fallen* in the plural: *ich habe* vs. *du hast/sie hat* (present singular) vs. *wir haben, ihr habt, sie haben* (present plural). The preterite and the conditional II have only short *ha-*: *hatte* 'had,' *hätte* 'would have.' In addition, the conditional II *hätte* is characterized by irregular (analogical) umlaut, since weak verbs were never affected by umlaut in the conditional. Here, analogy with the strong verbs must be assumed.[4] These partially suppletive (since isolated) forms developed due to the Early New High German mixing of two paradigms which were still complete in Middle High German, namely of *haben*, which developed normally, and the reduced, irregularized *hân*.[5] While Middle High German (MHG) *haben* preserved the old lexical meaning 'have, possess,' the formal reduction to *hân* correlates with the semantic reduction to tense forming 'have' as a perfect auxiliary. In Early New High German, this Middle High German paradigmatic splitting was abolished in favor of the creation of only one partially suppletive paradigm which now carries both meanings or functions again (see figure 1). Spoken German has further reductions and irregularizations: On the one hand, long [a:] in the *hab-* forms is shortened to [a] (such as *gehabt* [gə'hapt] 'had'), while, on the other hand, bisyllabic *haben* contracted to monosyllabic *ham* [ham]. Such contractions are completely impossible in comparable verbs such as *graben* (→**gram*). Contracted *ham* has become an unanalyzable portmanteau morph, which in turn increases the degree of allomorphy. At this point already, the duality of shortness and irregularity becomes obvious.

		Swiss Ger.	NHG	→	Spoken Ger.
Infinitive:		**hā**	hāben	→	ham
Present:	sg. 1	**hā**	hābe		hap
	sg. 2	**hesch**	**hast**		**hast**
	sg. 3	**het**	**hat**		**hat**
	pl. 1	**händ**	hāben		ham
	pl. 2	**händ**	hābt		hapt
	pl. 3	**händ**	hāben		ham
Imperative:	sg.	heb/haig!	hāb!		hap!
	pl.	**händ!**	hābt!		hapt!
Preterite:		—	**hatt-**		**hatt-**
Past part.:		**ghā**	gehābt		gehapt
Cond. I:		heig-	hāb-		—
Cond. II:		**hätt(-)**	**hätt-**		**hätt-**

Note: Vowel length is indicated by a cross-bar; the pronunciation of spoken German is only approximately transcribed.

Figure 1: MHG *haben/hân* and its continuations in Swiss German (Basel), New High German, and spoken German

Swiss German (an Alemannic dialect) has reached a relatively high level of irregularization and reduction through completely different routes: Here only the Middle High German short-form paradigm *hân* has been continued. In the present, a sort of "changing inflection" arose, which in general does not hold for other modern Swiss German verbs: *ich hā, du hesch, er het* (present singular) vs. *händ* (uniform present plural). The second and third person (*hesch, het*) contrast by primary umlaut from the rest of the paradigm, while the umlaut in the plural *händ* is based on a later, morphologically conditioned umlauting. As in the case of New High German, the conditional has an irregular umlaut (*hätti*). Finally, the past participle is inflected like a strong verb (*ghā* 'had' < MHG *gehân*) – another adaptation to the strong verbs. In the whole Alemannic dialect area, the preterite has broken down and been eliminated.

Letzeburgish *hunn* continues (like Swiss German *hā*) the Middle High German short-form paradigm, differentiating it later in a similarly strong and also structurally comparable way: with the sg. pres. *hunn, hues, huet* [hʊn, huəs, huət], it has analogously "changing inflection". The present plural is *hunn, huet, hunn*. The

preterite stem *hat-* [ha:t]- diverges from this in that it has ablaut-like vowel change. This tense-specific vowel change emerged through the different but regular development of short *a* in the present versus long *ā* in the preterite. The shortness of the vowel in the present tense, however, is based on irregular reduction. Apart from these unique developments, the presence of a preterite is in itself remarkable since it only exists in the case of 10 to 20 verbs. Only in the imperative *hief!* (sing.) and *hieft* (pl.) was the old stem final consonant preserved.

Dutch has relatively few irregular verbs. As a consequence, the irregularities of *hebben* are all the more important. The Dutch verb is characterized by syncretism in the second and third person of the singular – with exception of *hebben* 'have' (and *zijn* 'be'): *jij heb-t* [hɛpt] 'you have' vs. *hij/zij heef-t* [he:ft] 'he/she has'; here, considerable stem allomorphs exist. The older form *heeft* resisted the usual analogical leveling to the second person singular with the result of a morphological overdifferentiation, which means that *hebben* distinguishes more different forms in the paradigm than usual. The preterite also diverges considerably from the weak pattern by having forms with *a*, that is, vowel-alternating forms (cf. the uniform sg. pret. *had* and the uniform pl. pret. *hadden*). This is no true ablaut, but rather the only preserved case of so-called Rückumlaut (reversed umlaut) in Dutch. The umlaut forms in the present are the regular continuation of Germanic **haƀj-an-*. Since this *j* was missing in the preterite, no umlaut could develop. This vowel change was leveled out on the analogy of all the other verbs with Rückumlaut; Dutch does not even have a residual group of these verbs any longer, as in the case of NHG *kennen – kannte, brennen – brannte*, and so on. Dutch *hebben* thus behaves synchronically like a strong verb. This "strongness" is additionally supported by the (irregular) monosyllabicity of the sg. pret. *had*: if the form had developed according to sound laws, **hadde* should have resulted, as all weak preterites are bisyllabic (cf. *hoorde* 'heard'). Following the pattern of the strong verbs (such as *zong* 'sang'), however, Middle Dutch *hadde*, as the only weak verb, gave up the ending *-e* in the singular. In comparison, the bisyllabic preterite plural *hadden* corresponds to the general bisyllabicity of both the strong and the weak verbs (such as *zongen* 'sang' (strong) or *hoorden* 'heard' (weak)). Finally, the whole preterite – just as in New High German, Swiss German and Letzeburgish – is characterized by the (irregular) deletion of the stem final consonant *f*.

Frisian has more reductions and, at the same time, more irregularizations than Dutch. According to Tiersma 1985 and Sjölin 1969, the following forms are the most important: *ha* [ha:] or *hawwe* ['havə] in the infinitive and in the uniform present plural, *ha(w), ha-st*, and *ha-t* in the present singular, the stem *hie(-)* [hiə] in the preterite, and the past participle *hân* [hɔ:n]. Firstly, the stem final consonant disappeared almost completely – apart from some unique relics – which led to contraction and went in the direction of monosyllabicity. The preterite underwent particularly extraordinary developments: in Old Frisian, the verb abolished its preterite dental suffix and transferred to a small, irregular verb class with extremely high token frequencies: Old Frisian *hade/hede* → Modern Frisian *hie*. The concrete pattern for

this strange process of analogy must have been *dwaan* 'do' or *wêze* 'be'. Table 1 shows this small, interesting group and its preterites, which always contain the stem vowel *ie*:

No.	Infinitive		Pret. sg.	Pret. pl.
(1)	*gean*	'go'	*gie*	*giene(n)*
(2)	*stean*	'stand'	*stie*	*stiene(n)*
(3)	*dwaan*	'do'	*die*	*diene(n)*
(4)	*wêze*	'be'	*wie*	*wiene(n)*
(5)	*ha*	'have'	*hie*	*hiene(n)*

Table 1: Analogous processes in the preterite of Frisian verbs

The formerly weak verb *ha(wwe)* thus transferred entirely to the strong or, more exactly, to the irregular verbs of Table 1. This atypical analogical direction is supported by the strong past participle *hân* 'had' with the nasal suffix *-n*.

English has split up its paradigm into full (independent) and enclitic (dependent) forms: on the one hand, *have, has,* and *had* (all pronounced with short [æ]) and, on the other hand, *'ve, 's,* and *'d*. Even the full forms show irregular reductions (deletion of stem final consonant *-v-* in *has, had*, monophthongization and shortening of the stem vowel to [æ]). The orthography of <have> still implies the pronunciation of *[ei]; that is, *have* has divergent grapheme–phoneme correspondences (cf. the regularly developed *behave*). The bound, enclitic, asyllabic minimal forms *'ve* [v], *'s* [s/z], and *'d* [d] are unique among the Germanic languages. They are used exclusively and accordingly frequently as auxiliaries. In this way, totally suppletive, nonsegmentable minimal portmanteau morphs were created all at once. In English, the other three auxiliaries *be, will,* and *shall* are also affected by cliticization. As Krug (1994) points out, the clitic forms occur much more frequently than their full counterparts.

Starting with **Danish**, we will now deal with the North Germanic languages. Like all continental Scandinavian languages, Danish has very simplified verbal inflection (no person/number morphs). The finite forms of Danish *have* [hæ$^{(ʔ)}$] are the following: pres. *har* [hɑ$^{(ʔ)}$], pret. *havde* [hæ(:)ðə], sup. *haft* [hɑfd̥]. The transcriptions show that the stem-final consonant has disappeared completely, except in the supine which – although being the only regularly developed form – has been isolated from the rest of the paradigm in this way.[6]

In **Swedish**, it is also the preservation of the old [f] in the supine *haft* which uniquely makes the paradigm (as in Danish) more heterogeneous: *ha/har* [ha(:)/ha:r] (inf./pres.) – *hade* [had:e][7] (pret.) – *haft* [haft] (sup.). As in Danish, the loss of [v] is irregular. In addition, a syntactic reduction (ellipsis) is extremely remarkable: In particular in subordinated clauses, the auxiliary *ha* can be totally deleted: *han sa, att han (hade) varit sjuk* (*hade* is usually omitted) (see Holmes 1994: 287).

In **Norwegian** (Bokmål and Nynorsk), too, HAVE has been extremely reduced to

ha/har – hadde[8] *– hatt*. Here, however, fewer irregularities can be found than in Swedish. The supine *hatt* also lacks its stem-final consonant, which means that the former *-v-* or *-f-* was consistently abolished, but without any regular basis. Only the devoiced final consonant cluster *-tt* [t(:)] in *hatt* diverges from the small, weak inflectional class into which *ha* has transferred (**hadd* would be regular – cf., for example, *bodd* 'lived').

Faroese *hava* has developed a highly differentiated paradigm. Firstly, it shows a stable instability of the inflectional class, which means that, since Old Norse, this verb stands between two weak classes, the *ē* class and the *ja* class. The infinitive and the uniform present plural, *hava* ['hɛava], the first person singular present *havi* ['hɛavɪ], the conditional *havi*, the imperatives *hav/havið*, and the supine *havt* belong to the *ē* class, while the second and third singular present follow the *ja* class: *hevur* ['he:vʊr] (**havir* would be regular). The preterite has undergone extreme differentiation by splitting up the uniform singular and plural (which, however, is hidden by the orthography): *hevði* ['hɛijɪ] (sing.) and *høvdu* ['hœd:ʊ] (plural). Very complicated reductive and, at the same time, irregularizing phonological and analogical processes (which will not be explained in detail here) led to only the initial [h] remaining as the smallest common denominator. The preterite of *siga* 'say' and *leggja* 'lay' has experienced a similar development. As a result, a new small group has been established (see 2.2.).

As in the case of Faroese, **Icelandic** *hafa* shows "stable instability" by also standing between the *ja* class and the *ē* class, but with different breaks. In Icelandic, the whole present tense (*hef, hefur, hefur*) belongs to the weak *ja* class. In addition, the conditional II contains – similar to German *hätte* – irregular *i*-umlaut: *hefði* instead of regular **hafði*. Finally the supine *haft* lacks the thematic vowel. As a result, Icelandic *hafa* has been isolated; apart from *haft* there are no special reductions (see Table 2).

2.2. SAY *in the Germanic languages*

In contrast to HAVE, SAY distinguishes itself by high token frequency **without** grammaticalization.[9] Its increase in frequency can mainly be attributed to the decline of the strong Germanic verb **kweþ-* (OHG *quedan*, 5th ablaut class) and its replacement by the weak verb **sag-*; interestingly, all Germanic languages underwent this change. Today, SAY belongs to the most frequently used verbs: NHG *sagen* is at Position 5 in the frequency of all verbs, Fris. *sizze* at Position 6, Norwegian (Norw.) *si* at Position 11, and Icel. *segja* at Position 5.

In **New High German**, *sagen – sagte – gesagt* is a regular weak verb. German is thus the only Germanic language in which SAY has not undergone extraordinary developments. This was different in Middle High German and still is different in today's Swiss German and other German dialects: Although OHG *sagēn* (like *habēn*) was actually a verb of the weak *ēn* class, it borrowed features from the weak

		Infinitive	3rd sg./3rd pl.pres.	3rd sg./3rd pl.pret.	PP/sup.
(1a)	Alem.	***hā*** [a:]	***het!/händ!***	—	***ghā***
(1b)	NHG	*haben, (**ham**)*	***hat!/***haben *(**ham**)*	***hatte!/hatten!***	*gehabt*
(2)	Letz.	***hun*** [ʊ]	***huet*** [ʊə]***/hun*** [ʊ]	***hat/haten***	***gehat***
(3)	Dutch	*hebben*	*heeft!* [e:]*/hebben* [ɛ]	***had/hadden***	***gehad***
(4)	Frisian	*hawwe,* ***ha***	***hat****/hawwe,* ***ha***	***hie****/* [i(ə)]	***hân*** [ɔ:]
				hiene(n) [i(ə)]	
(5)	English	*have!, 've!!*	***'s!!*** [s/z], ***has!/*** *'ve!!* [v], *have!*	***'d!!, had***	***had!***
(6)	Danish	***have*** [hæ$^{(ʔ)}$]	***har***	***havde*** ['hæðə]	*haft*
(7)	Swedish	***ha***	***har***	***hade!***	*haft*
(8a)	Bokmål	***ha***	***har***	***hadde***	***hatt***
(8b)	Nynorsk	***ha****/have*	***har***	***hadde***	***hatt***
(9)	Faroese	*hava*	*hevur/hava*	***hevði*** ['hɛijɪ]/ ***høvdu*** ['hœd:ʊ]	*havt*
(10)	Icel.	*hafa*	*hefur (hefir)/hafa*	*hafði/höfðu*	*haft!*

Symbols: ***Boldface***: Short forms, i.e. without stem-final consonant (and thus usually irregular)

"!(!)": Shortness which does not consist of the reduction of the stem-final consonant

Underlining: Irregular form/segment (in regard to phonetics, orthography, and/or morphology)

Boldface + underlining: Short forms + irregular, whereby the irregularity does not result from the reduction of the stem-final consonant.

Table 2: The most important forms of HAVE in the Germanic languages

jan verbs. This can be demonstrated by the umlauted forms OHG *segist/segit* '(you) say/(s/he) says,' *segita* 'said' (pret.), and *gisegit* 'said' (past part.). In Middle High German, *-egi-* contracted to *-ei-*, which explains the monosyllabic forms *seist/seit* (also: *seite* (pret.) and *geseit* (past part.)). These contracted forms were continued in **Swiss German** (Basel German) *du saisch/er sait* 'you say/he says' and *gsait* 'said' (past part.). These *sai*-stems contrast intraparadigmatically with regular *sāg*-forms. In Zurich and Bernese German, analogies to the strong verbs of the 7th ablaut class were made by reforming the conditional with *sieg* 'would say' (cf. also *frāge - frieg* 'ask – would ask', *mache – miech* 'make – would make', and *chouffe – chieff* 'buy – would buy'; see Marti 1985, Weber ^{2}1987, and Nübling 1997).

Letzeburgish has also differentiated *soen*, which, in this case, has been contracted according to sound laws, following the pattern of changing inflection: *ech soen* 'I say' vs. *du sees/hie seet* 'you say/he says'. As in the case of Swiss German, the 2nd and 3rd singular forms in Letzeburgish go back to the contracted umlaut forms of MHG *segist/segit*, which follow the *jan* inflection. Besides the hiatus form *soen* [zo:ən], *son* [zo:n], which is also contracted, exists; this does not hold for less frequent comparative verbs, such as *froen* 'ask' and *kloen* 'moan, complain'. Finally, the pure presence of a preterite (*sot* 'said') must be valued as an exception.

Dutch *zeggen* [zɛχə(n)] continues SAY – as in the case of the remaining languages – as a pure *jan* verb (therefore the gemination of consonants and umlaut in the present tense). Here, the preterite underwent unique, extremely reductive developments, resulting in *zei* [zɛi] (sing.) and *zeiden* [zɛi(d)ə(n)] (pl.). The preterite *zegde(n)* probably developed regularly, which, indeed, is preserved in some derivations, such as *opzeggen* 'hand in one's notice.' In regard to the (irregular) development of the diphthong *ei*, Donaldson (1983: 146/148) writes:

> "In a few words the diphthong *ei* is the result of a contraction of $\check{e}+g+i$ (<*ăgi*) where the *g* has been palatalized after Umlaut and syncopated. [...] *zei* < *zegde* < **zegide*."

The loss of the dental suffix *-de* → *-Ø* (*zeide* → *zei*), however, is not explained; it lacks any regularity and only can be explained by analogy to the strong verbs which always are monosyllabic in the preterite singular. Because it also obviously contains a vowel change, *zeggen* belongs to the mixed verbs with strong preterite and weak past participle (*gezegt*) in Hempen (1988: 27), whereby *zeggen* constitutes a separate subclass, which means that it is isolated from all other verbs. In the preterite plural *zeiden*, the dental suffix is still written, but it is already lost in spoken Dutch ['zɛiə]) (see Haeseryn et al. 1997: 86).

Frisian *sizze* ['sɪzə] certainly is phonologically deviant, which, however, is not often treated in Frisian language histories. The regular Old Frisian palatalization of *g-* and *k-* before palatals must have arisen unusually in the middle of the word. The Old Frisian orthography <sed(s)za/sidza> tends to suggest the intermediate stage of affrication. The fronting of *e* > *i*, that is, the development from *seggjan* to *sizze*, has not yet been explained (the same happened with *leggjan* → *lizze* 'lay'). No other examples of *e* > *i* fronting has been documented for Frisian. Here, a so-called "over-palatalization" must have occurred, as in the case of other languages and verbs, such as GIVE. Originating from the extremely palatal stem-final consonant *g*, an exceptionally regressive palatalization must have occurred from *e* > *i*. The present forms in the singular are *sis* [sɪs] '(I) say', *seist* [saist] '(you) say', and *seit* [sait] '(s/he) says', of which the 2nd and 3rd person are based on (irregular) contractions. The uniform plural is homophonous to the infinitive *sizze*. As in the case of Dutch, the weak Frisian preterite has become strong through the irregular loss of its dental suffix and the monosyllabic singular form resulting from this: Old Frisian *seide* → Modern Frisian *sei* 'said (pret.)'. In opposition to Dutch, the Frisian past participle *sein* is also formed in accordance with the strong inflection (c.f. Old Frisian, where the past participle is still *(e)seid* 'said').

English grammars consider *say* to be an irregular verb and, more precisely, a verb with vowel alternation and a dental suffix in the preterite and past participle.

The reason for this is the monophthongization and shortening from [ei] to [e] also in the 3rd person singular present: *say* [sei], but *s/he says* [sez] and *said* [sed]. Engl. *say* shares especially the monophthongization in the 3rd person singular present only with very few, but characteristic verbs:

> "Apart from the three primary verbs *be, have*, and *do*, the only verbs which have an irregular *-s* form are *say* /seɪ/ ~ *says* /sez/, and derivatives of *do* [...]. In the *-s* form, *say* is irregular in pronunciation, but not in spelling. *Gainsay*, historically a derivative of *say*, may have a regular or an irregular pronunciation of the *-s* form: *gainsays* /-seɪz/ or /sez/." (Quirk et al. 1985: 99).

In this respect, *says* [sez] is more marked than *said* [sed].

Danish, like all Scandinavian languages, continues the Old Norse reverse umlaut verb *segja – sagði – sagt* which, already at that time (as still in Modern Icelandic) contained "stable inflectional instability," which means that it united forms of the weak *ja* and *ē* classes (see below). On its way to Danish, *sige* [si:(ə)] has undergone three irregular developments: Firstly, it has become a short verb by consistently eliminating (apart from the supine *sagt* [saɡ̊d̥] 'said') the stem-final consonant and by contracting to monosyllabic forms: *siger* [si:ɐ]/[siɐ̯] (present) and *sagde* [sæ:] 'preterite' (the same holds for the two *s*-passive forms *siges* ['siəs] 'is said' and *sagdes* [sæ:s] 'was said'. Even the preterite suffix disappeared phonologically. Secondly, because the orthography continues the old extended form, there is a discrepancy between the written form and the pronunciation, which is extraordinary even for Danish (for example, see above). Thirdly, Danish (as in Norwegian) must have undergone a so-called "overpalatalization," as is also assumed for Fris. *sizze*: Under highest token frequency use, it can very often be observed that the interactivity between the sounds of a word, that is, contact phenomena, may appear which no longer follow sound laws and which do not occur in phonologically comparable words of lower token frequency. In this case, the palatal *-g-* regressively caused *e* > *i* fronting.[10] Synchronically, Dan. *sige* should be regarded as a strong verb with a slightly suppletive supine.

Swedish has the following stem forms: *säga* ['sɛj:a] – *sa* [sɑ:] – *sagt* [sakt]: Here we immediately notice the short, monosyllabic preterite *sa*, which is in a constructionally counter-iconic relationship with the semantically unmarked, but formally marked present form *säger* ['sɛj:er]. The situation for *lägga* 'lay' with preterite *la* 'laid' is similar. The development was as follows: *sagde* > *sad(h)e* > *sae* > *sa* (see Östman 1992). None of these processes of loss follow sound laws. With the loss of the dental suffix, this verb has become strong as well, as vowel change has also taken place between the present and the preterite. Finally, <säga> ['sej:a] contains an orthographic irregularity, as an application of the grapheme-phoneme correspondence rule, which is very clear and consistent in Swedish, would result in a pronunciation of *['se:ga], that is, the (irregular) [g] > [j] palatalization is not represented graphically.

In **Norwegian** (Bokmål), *si – sa – sagt* is considered to be a strong verb with an irregular supine (Faarlund et al 1997: 186). Except for the bisyllabic present form *sier* ['si:ər], there are no more hints as to its past as a weak verb, not even in the orthography. Both the drastic reductions and the overpalatalization of *e* → *i* before palatals [(g)j] are irregular.

In **Faroese**, *siga* ['si:ja] has also undergone overpalatalization. In the present tense, some forms follow the *ja* class, while other follow the *ē* class. Of particular notice is the number split in the preterite, which has led to similar forms as in the case of *hava*: sg. pret. *segði* ['sɛijɪ] vs. pl. pret. *søgdu* [sœd:ʊ]. Complicated regular and irregular processes of loss and partial analogies to various inflectional classes have led to this split, which is similar to the characteristic number ablaut split of today's strong verbs. If the supine *sagt* and the passive form *sigst* are integrated, one achieves a considerable inventory of differentiated stem allomorphs [si:]-, [sɛi]-, [sœ]-, [sak]-, [sɪk]- for a weak verb.

Since Old Norse, **Icelandic** *segja* has preserved the class instability between the *ja* verb and the *ē* verb, but at a different mixture ratio than Faroese. Except for the missing theme vowel in *sagt* 'said', no special reductions have developed.

Table 3 provides an overview of the most important forms of SAY.

		Infinitive	3rd sg./3rd pl.pres.	3rd sg./3rd pl.pret.	PP/sup.
(1a)	Alem.	*sāge*	***sait/sāge***	—	***gsait*** [ksait]
(1b)	NHG	*sagen*	*sagt/sagen*	*sagte/sagten*	*gesagt*
(2)	Letz.	***son***, *soen*	***seet/soen***	***sot/soten***	***gesot***
(3)	Dutch	*zeggen*	*zegt/zeggen*	***zei/zei(d)en***	*gezegd*
(4)	Frisian	*sizze*	***seit/sizze***	***sei/seine(n)!***	***sein*** [sain]
(5)	English	*say* [sei]	***says*** [sez]/ *say* [sei]	***said*** [sed]	***said*** [sed]
(6)	Danish	***sige*** [si:]	***siger*** [si:ɐ]	***sagde*** [sæ:]	*sagt* [saɡ̊d̥]
(7)	Swedish	*säga* [sɛj:a]	*säger* [sɛj:er]	***sa***	*sagt*
(8a)	Bokmål	***si***	***sier***	***sa***	*sagt*
(8b)	Nynorsk	*seie*	*seier*	***sa***	*sagt*
(9)	Faroese	*siga*	*sigur/siga*	*segði* [sɛijɪ] *søgdu* [sœd:ʊ]	*sagt*
(10)	Icel.	*segja*	*segir/segja*	*sagði/sögðu*	*sagt*

Symbols: ***Boldface***: Short form (and thus usually irregular)
"!": Shortness not due to reduction of stem-final consonant
Underlining: Irregular form/segment (or, rather, phonetics, orthography and/or morphology)
Boldface + underlining: Short form + irregular, whereby the irregularity does not result from the reduction of the stem-final consonant.

Table 3: The most important forms of SAY in the Germanic languages

3. IRREGULARIZATION STRATEGIES AND OTHER PROCESSES AFFECTING HIGH-FREQUENCY VERBS

The described irregularization processes are limited not only to HAVE and SAY, but also affect other high-frequency verbs such as DO/MAKE, BE, GIVE, TAKE, BECOME, and GO. Grammaticalization seems to further these irregularization and reduction processes, but only insofar as they result in increased token frequency, which forms the actual drive behind these processes. Evidence can be found in SAY, which, strictly speaking, is not grammaticalized.

Basically, irregularization takes place on various linguistic levels:

- On the prosodic level (cliticization of English auxiliaries)
- On the phonological level (the many reductions, especially those affecting the stem-final consonants)
- On the orthographic level (deviations from the usual grapheme–phoneme correspondence rules)
- On the morphological level (see the various "atypical" analogies to (more) irregular verbs)
- On the syntactic level (see the ellipsis of the Swedish auxiliary *ha* in dependent clauses)

In the process, it is decisive that such irregularities arise not only "passively" through the accumulation of "junk" (as in Lass 1990), which was "overlooked" by analogy, but rather that they are also "actively" created: the latter is impressively substantiated by HAVE and SAY. Irregularities can thus be based on both conservative and innovative processes.

What are the morphological principles, then, which can be realized under very high token frequency? In summary, the following most important processes are let loose; in the process, we will examine the irregularization strategies of other high-frequency verbs (such as BE, GO, GIVE, and COME).[11]

Shortness of expression: Almost all changes in HAVE and SAY are partially connected to extreme reductions; that is, irregularity and shortness of expression usually appear together. Particularly subject to reduction is the stem-final consonant, such as NHG *ha-st* instead of regular **hab-st*, Engl. *ha-s/ha-d* instead of regular **hav-es/*hav-ed*. Entire syllables often disappear, as in the case of the contraction of NHG *haben* → *ham*, Du. *hadde* → *had* 'had,' Fris. *hede* → *hie* 'had,' Swed. *hava* → *ha*, Engl. *has* → *'s*, Old Norse (ON) *segja* → Norw. *si*, and so on.

Reductive phenomena can be ascertained on other levels as well: in the case of Swedish on the orthographic level (Swed. <hade> instead of regular *<hadde>) and even on the syntactic level (the omission of finite *ha* in Swedish dependent clauses). It is very conspicuous that there are hardly any reductions which do not also cause irregularity: in most cases, the reductions spread only partially through the paradigms (and then in the most frequent categories). One of the rare examples

of a systematic reduction (that is, **without** simultaneous irregularization) is the [ei] to [æ] reduction of Engl. *have*.

Irregularization: Almost every Germanic language[12] has irregularized HAVE and SAY. In the process, principally two irregularization strategies have been used: (a) reductive irregularization and (b) nonreductive irregularization.

(a) **Reductive irregularization**: Through their only partial distribution, most reductions cause intraparadigmatic irregularization (whereby even the reductions themselves are also of irregular origin). In the process, stem-final consonant loss dominates. Sonorous consonants are more apt to fall victim to this process than the more consonantal consonants, whereby this again is subject to token frequency: under extremely high token frequency, plosives can also disappear (such as Norw. *la* ← *late* 'let'). Within a paradigm, the stem-final consonant is more likely to disappear (especially in present singular) and/or then when complicated clusters would otherwise arise (see NHG *ha.be*, but *hast/hat* instead of **habst/*habt*). Irregularization often also takes place due to the quantitative reduction of the stem vowel (cf. NHG *hāb-* vs. *ha-*, Letz. *hue-* vs. *hu-*). Only seldom and under high token frequency is the initial sound reduced, which alienates and irregularizes the forms particularly strongly (see Engl. *'s, 've, 'd*). Psycholinguistic experiments confirm the particular prominence of the word-initial sound by means of word recognition tests (see Cutler et al. 1985, Fenk-Oczlon 1989). Other reductions, such as those in orthography, also increase the irregularization balance (such as Swed. <hade> ['had:e] with only one instead of the regular two <dd>). Of particular relevance is thus the observation that almost all morphological (analogical) irregularizations lead to increased shortness in expression; according to traditional thought, this is not the essence or the function of analogy. In the case of the "normal" so-called compensation or regularization analogy, longer expressions arise as usually more transparent structures are created; this is exemplified by the changeovers in class from the strong to the weak verbs (see NHG *buk* → *back-t-e*). On the other hand, the irregularizing analogies ("differentiating analogies") documented in this article are always reductive (Old Frisian *hade* → Modern Frisian *hie* (after *die* 'did'), MHG *haben* → *hân* (after *gân* 'go'). In such cases, it is difficult to decide whether the analogy was possibly motivated by the shortness of the word.

(b) **Nonreductive irregularization**: This includes partial class changeovers, as has been practiced constantly by HAVE and SAY, especially in Faeroese and Icelandic. In the case of another verb, Swed. *ge* 'give,' former versions of the verb containing *e* or *i* combined to a new and thus more strongly differentiated paradigm: *ge/ger – gav – givit* 'give/gives – gave – given'. In Frisian, a unique analogy was carried out with the same verb, *jaan* 'give' to

a suppletive verb *slaan* 'hit' in the preterite: Old Fris. **jef* 'gave' developed to *joech* [ju:χ] (in the entire preterite) under the influence of the suppletive preterite form *sloech* [slu:χ]. This has extremely differentiated the paradigm of *jaan*. In the case of strong suppletion through lexical blending, completely different paradigms are finally combined to a new one. This applies to BE in all Germanic languages and to GO in English: *go – went.*[13] Even in the case of these nonreductive irregularizations, longer expressions are only seldom created. In most cases, the material expense remains the same; that is, irregularization never leads to a longer expression.

This means that, in the case of both irregularization types, there is a clear correlation between irregularity and shortness.

Overdifferentiation: In regard to the formal differentiation of grammatical categories, strong syncretism (homophones) in the high-frequency verbs should thus be expected due to the large number of reductions. Astonishingly, this is not the case at all – on the contrary: more and stronger differentiation is often made, particularly in the high-frequency area than anywhere else. Engl. *be*, therefore, has a triple differentiation of the sg. pres. *am/are/is* and of *was/were* in the usually uniform preterite. In Dutch, *hebben* (and *zijn* 'be') has an exceptional triple differentiation in present singular instead of the usual double differentiation: *ik heb, jij hebt, zij heeft.*[14] Many New High German strong verbs and all weak verbs have syncretism between the 3rd person singular and the 2nd person plural present (such as *macht* '3rd sg./2nd pl.pres.'), but not for *haben* with *hat* [hat] vs. *habt* [ha:pt].

In addition, formal differences can be more strongly contoured and thus more likely to be relocated or transported to the center of the verb than usual; see Engl. *says* [sez], where, except for the usual *-s* flexive, even the root is affected (through monophthongization and shortening).

Intraparadigmatic positions of reduction and (over)differentiation: Clear evidence for the token frequency argument (and against the mere grammaticalization argument) is provided by the position of the paradigm in which the reduction and differentiation takes place: in the process, these are usually the most frequent category combinations, such as the 3rd person singular present indicative active. In keeping with the trend, the following are more likely affected:

- The 3rd person (in comparison with the 1st and 2nd person)
- The singular (in comparison with the plural)
- The present (in comparison with the non-present tenses)
- The indicative (in comparison with the conditional) and so on.

Even chronologically, it can be ascertained that changes appear first in the more frequent categories and later – if at all – in the less frequent categories. This categorial frequency is to be combined with the lexical frequency of the verb; that is, the

preterite of HAVE, for example, is much more frequent than the present of LAUGH. The more frequent a certain verb is, the more possible it is for less frequent categories (such as conjunctive, preterite, 2nd person plural) to develop a deviant expression (see the following preterites: Engl. *was* vs. *were*, Far. *hevði* and *segði* ['hɛijɪ/'sɛijɪ] (sg. pret.) vs. *høvdu* and *søgdu* ['hœd:ʊ/'sœd:ʊ] (pl. pret.), the New High German suppletive forms of *ist* (indicative) vs. *sei* (conjunctive)).

"Boundary crossings": Under high frequency, the general tendency for various types of structures to become weaker or even disappear altogether has been observed again and again. This is reflected on many levels. On the prosodic level, for example, in the English auxiliary clitic forms of *have, be* and *shall/will* to *'ve, 's, 'm, 'll*, and so on, the word accent was abolished, and then the junctions disappeared; the result was the cliticization of these forms and their subordination under a new prosodic unit: *I have → I've* [aiv], *I am → I'm, I had/would/should → I'd*, and so on. The clitic forms must not necessarily – as in these cases – lose their syllable status; what is important here is only the loss of their status as a word and their word boundaries. On the phonological level, there is often interaction between the individual sounds and their features that goes beyond the normal dimensions, that is, the word-internal assimilation readiness increases. This includes the so-called "overpalatization" of ON *segja* 'say' and *geva* 'give' to Norw. *si, gi*, Dan. *sige, give* (and borrowed from Danish as *give* into English), and Far. *siga*. In all cases, palatal *g* [(g)j] palatalized its vocalic environment beyond the usual dimensions by fronting the vowel *e* to *i* or by palatalizing it more strongly. So-called "overlabiovelarization" can be found in all languages in the case of the verb COME, where the original Germanic **kwem-* was assimilated to *kom-* and the initial consonant cluster was reduced. Bilabial [w] transferred the feature [+labial] to following *e*, which resulted in *o*. The only exception is the Dutch Pret. *kwam* 'came'. Lower-frequency verbs with comparable phonetic conditions, such as OHG *quelan* 'well up' < Gmc **kwel-*, have retained the old vocal quality and initial consonance (NHG *quellen*).

An additional, interesting "overassimilation", which occurs in various Germanic languages and which does not follow any sound laws, has appeared in the verb COME: before a following alveolar, the stem-final consonant *m* underwent an alveolarization to *n* in Alemannic, Letzeburgish, North Frisian, and Icelandic (see Table 4):

In Icelandic, this irregular assimilation is limited to the imperative singular form <komdu> ['kʰɔn̪dʏ] 'come', which is hidden by the orthography.[15]

Not by chance, these irregular assimilations are carried out only in the present, even though the phonetic, but not the frequential prerequisites are the same in the preterite: Letz. **këmms* (Pres.) became *kënns*, but Letz. *koums* (Pret.) did not become **kouns*. The paradigm is differentiated through such irregular and only partially effective processes (increased stem allomorphy). At the same time, through the creation of homorganic or even reduced clusters, articulation is facilitated.

			Alem.	Letz.	North Fris. (Wiedinghard)
Infinitive			*chō*	*komm-en*	*käm-e*
Present	Sg.	1	*chum-e*	*komm-en*	*käm*
		2	***chun-sch***	***kënn-s***	***kän-st***
		3	***chun-t***	***kënn-t***	***kän-t***
	Pl.	1	*chöm-e*	*komm-en*	*käm-e*
		2	*chöm-et*	*komm-t*	*käm-e*
		3	*chöm-e*	*komm-en*	*käm-e*

Table 4: Assimilations in the case of COME (< Gmc **kwem-*) in Alemannic, Letzeburgish, and North Frisian

Under high token frequency, syntagmatic simplifications are carried out paradigmatically; that is, more importance is attached to articulatory economy than to uniform and transparent morphological structures. The respective cost–benefit analysis depends on the token frequency of the affected unit. Additional evidence for the phonological optimization at the cost of morphology and regularity is provided by the irregular assimilation and contraction of NHG *haben* → *ham* (in spoken German). This example clearly shows that such phonological internal compression can lead to the dissolution of morphological structures: {hab}-{en} → {ham}. The highest internal compression at the total dissolution of morphological boundaries is carried out in suppletive forms (NHG *bin* 'am', *ist* 'is', *sind* 'are').

Additional boundary crossings or dissolutions are only hinted at. Even word-externally, increased external sandhi with syntactic neighbouring units may arise especially under high token frequency, without the verb itself cliticizing (such as Alem. *händ#mer* → *hämmer* 'have we'). In regard to interparadigmatic aspects, the formation of strong suppletion due to paradigm mixing (such as *go – went*) represents a boundary crossing. On the semantic level, in the course of grammaticalization, so-called synsemantization takes place; that is, a grammatical category is realized through the combination of a finite auxiliary and a non-finite full verb.

4. THE FUNCTIONALITY OF "JUNK": IRREGULARITY AS FORMAL DIFFERENTIATION WHICH PERMITS A MINIMUM EXPRESSION WITHOUT THE DANGER OF HOMONYMY

The common denominator of all the processes listed in Section 3 is that morphological and intraparadigmatic structures and rules are destroyed and, at the same time, the forms become shorter (or at least never longer). Irregularity and shortness also

correlate to a high degree. Especially the morphologically destructive effects have led a number of morphological theories (such as the naturalness theory) to exclude the core area of these verbs from their general observations or to assign them a special status which isolates them from the rest of the "normal" verbs. This often happens to the verbs BE, HAVE, GO, and so on. In spoken German, only these verbs represent more than half of all verbs appearing in a text. For the further frequency ranking of German verbs, see Table 5:

1	*sein* 'be'	24.11%
2	*haben* 'have'	22.72%
3	*gehen* 'go'	4.77%
4	*kommen* 'come'	3.78%
5	*müssen* 'must'	3.24%
6	*werden* 'become'	2.67%
7	*machen* 'make'	2.58%
8	*sagen* 'say'	2.26%
9	*können* 'can'	2.01%
10	*wissen* 'know'	1.21%
	total:	69.35%

Table 5: The frequency values of the 10 most frequent verbs in New High German (according to Ruoff 1990)

In order to be able to understand the function of irregularity, it is important to replace this negative term, which implies the absence of truly expectable regularity, by the positive term of distinctivity: irregularization always effects a differentiation of the paradigm, that is, the forms drift apart and become more and more dissimilar. This differentiation has the advantage of protecting the forms which become increasingly shorter under the effects of high token frequency from homophony (syncretism). Theoretically, with increasing word shortness, a merging of the forms should be expected, especially since the reductions – as shown – are usually at the end or in the middle of the word, where the most important categories are marked. The more strongly and further forwards the word is differentiated, however, the more strongly it can be reduced without the danger of homonymy. In its pure form, this ideal combination – minimum expression at maximum distinctivity – is realized by strong suppletive forms (cf. the English clitic forms of *have*, which are quantitatively minimal, but qualitatively maximally differentiated: *'d* [d], *'ve* [v], *'s* [s/z]).

An entire scale of various coding techniques thus results – regulated according to token frequency – in accordance with the economy concept of Werner (1987, 1989) (see Figure 2 and the following quotation of Werner 1987b):

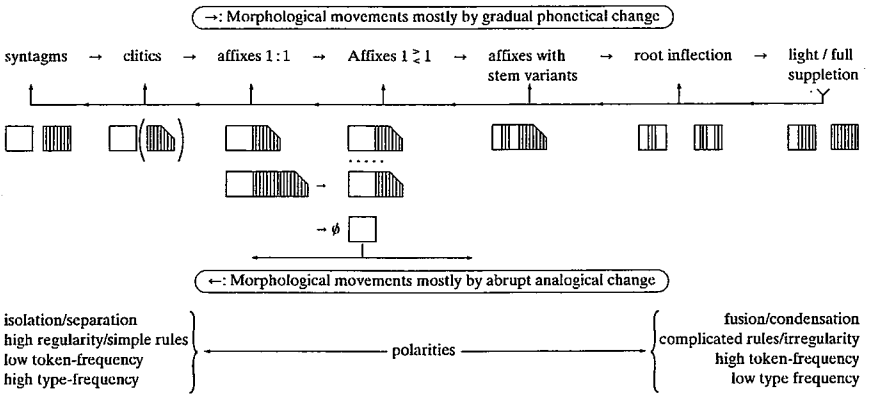

Figure 2: The morphological language change model of Werner (1987a)

> Daraus ergibt sich, daß "Einfachheit der linguistischen Beschreibung", was Gleichmäßigkeit/ Parallelität des Systems voraussetzt, und "Einfachheit/ Ökonomie einer natürlichen Sprache" keineswegs parallel laufen, oder daß der Sprachwandel darin bestünde, die Regeln zu vereinfachen. Im Gegenteil, um eine Sprache möglichst einfach für den Gebrauch zu machen, müssen immer wieder Ungleichmäßigkeiten im System hergestellt [...] werden. (Werner 1987b: 296).

The (syntactic) periphrasis (NHG *hat gelacht* 'did laugh' 'Perfect', *wird lachen* 'will laugh' 'Future') is located on the left pole in Figure 2, while suppletion as the strongest form of information condensing (Engl. *am, is, are, were, been*) is located on the right pole. The scale thus runs from maximum expansion, which correlates with regularity, a low token frequency, and a high type frequency, over additive (agglutinative) morphological interlinkage, to more modulatory, overlapping methods (flection) up to maximum compression (suppletion). This coding scale correlates with increasing usage frequency; that is, every constellation on this scale is economical, as long as the named correlations apply. It would be extremely uneconomical if the present tense of BE were expressed periphrastically and the pluperfect of LAUGH synthetically.

The expanding methods on the left side of the scale have the competence-related advantage of the applicability of rules and combinatory techniques, while the compression methods on the right side have the performance-related advantage of shortness, which is desired under high frequency of use (see Ronneberger-Sibold 1980). Through constant use, the form is stored as an unanalyzed unit and then recalled, that is, morphological transparency is in this case, from the point of view of the cognitive requirements, not even necessary. Frequency change should also cause a coding change. Indeed, it is the strong verbs decreasing in frequency which transfer

to the expanding class of weak verbs. The opposite case, which is much more seldom, is represented by HAVE and SAY. The observation of Vennemann (1993) explains these correlations as well (1993):

> "The antithetic character of language changes is evident at other levels as well. For example, morphological iconicity is highly valued by speakers of languages; it is evidently related to the principle of uniform linguistic symbolization. But the iconic construction of words with complex meanings creates many long words, and length is a disadvantage on another parameter, especially for frequently used words. Thus we need not be surprised that the most frequent words of many languages are not iconic but suppletive and are, when lost, often not replaced with iconic constructs but with new suppletive forms. Paradigms [...] of words meaning 'to be' and 'to go' provide ample evidence for this in many languages." (323)

The fact that the various coding techniques are conditioned by frequency and that every language has verbs of varying frequency explains why very few languages are typologically uniform. The described mixture principle is the most economical: "The aim of morphological change is a good mixture – not a uniform language type" (Werner 1987a). Only a balanced cost–benefit analysis which places both performance and competence needs in relation to token frequency leads to an adequate evaluation of the morphological relations and change and is able to integrate seemingly chaotic developments such as those presented in this article.

NOTES

[1] I am very grateful to Martin Haspelmath for his useful comments on this paper.

[2] For further criticism of these terms see Vincent (1993).

[3] See Ronneberger-Sibold 1980, Werner 1987a,b and 1989, Fenk-Oczlon 1989 and 1991, Harnisch 1988 and 1990, Nübling (2000).

[4] Already Old High German (OHG) *haben* – although originally a regular verb of the weak *ēn* class – tends to different irregularizations. *habēn* is the OHG verb with the most variants: Apart from various weak forms (OHG still has three weak inflectional classes), it also adopts features of the strong verbs, eventually develops contracted forms, and partially follows the small group of athematic verbs. In view of this excess supply of forms, OHG *habēn* would have had many possibilities to adapt to a uniform inflection pattern, e.g. to the strong verbs as it is expected by Dishington (1980): "One must ask why the result was a mixed paradigm and not simply a strong one" (14). Instead, time and again, it tends to atypical, irregular, mixed paradigms.

[5] The present tense of this short verb has leveled out analogically to the two stem verbs *gân* 'go' and *stân* 'stand', but not the past tense.

[6] As already shown above for English, irregularities can also be found at the orthographic level. Especially Danish is well-known for its strong discrepancy between pronunciation and orthography. This discrepancy particularly affects frequent words, such as *have* [hæ$^{(ɔ)}$]. Even

in Danish, the correlation of <have> → [hæ(ˑ)] is unique and highly unpredictable and must be learned as an exception. The same holds for the preterite *havde* → [hæ(:)ðə]. Only the supine *haft* [hɑfd] has phonologically and graphically preserved the old stem-final consonant. Thus, only in written Danish has morphological uniformity been preserved for the most part.

[7] In Swedish, the grapheme–phoneme correlation is more structured and closer to a one-to-one relationship than in Danish. Here, the only irregularity at the orthographic level concerns the unique writing of the preterite with only one instead of two <dd> (cf. *bodde* 'lived' in regular writing). In Swedish, this orthographic irregularity is at the same time reductive while Danish preserved the writing of the older long forms.

[8] The preterite <hadde> is (contrary to Swedish) correctly written.

[9] Here, grammaticalization in a stricter sense is meant, namely as a qualitative process of the formation of grammatical categories, while Hopper/Traugott 1993, for example, see the pure increase in frequency as a form of grammaticalization: "Frequency demonstrates a kind of generalization in use patterns" (103).

[10] For a case of "overvelarization", see COME < Germanic **kwem-* in all Germanic languages (except the preterite form *kwam* 'came' in Dutch), e.g. OHG *queman* → NHG *kommen*, but OHG *quelan* > NHG *quellen* 'swell' (for further details and examples, e.g. Frisian *jaan* 'give', see Nübling (2000)).

[11] For more details see Nübling (2000), where the reduction and irregularization strategies of ten verbs in ten Germanic languages are examined: HAVE, BECOME, GIVE, TAKE, COME, SAY, BE, DO, GO, STAND.

[12] This finding also applies to the Romance languages. In what way all of these described processes are connected to the individual language or language type would be an interesting topic of research in a contrastive study.

[13] On this point, see the Romance languages, which have mixed three verbs to form the GO paradigm (see Fr. *vais – allons – irai*).

[14] There is usually syncretism between the 2nd and 3rd person singular present.

[15] In many dialects such as Zuger German (*chūsch, chūt*) or the Cologne dialekt (*kütt*), a segmental reduction has even appeared here. In Low German *kümp* (instead of **kümt*), the reverse, a progressive bilabialization of *t* > *p* after *m*, has occurred (Münsterland; see Lindow 1998).

REFERENCES

Augst, Gerhard 1975. "Wie stark sind die starken Verben? Überlegungen zur Subklassifikation der nhd. Verben". In G. Augst (ed.). *Untersuchungen zum Morpheminventar der deutschen Gegenwartssprache*, 231–281. Tübingen: Narr.

Bybee, Joan L. 1985. *Morphology. A Study of the Relation between Meaning and Form.* Amsterdam: John Bejamins.

Cutler, Anne/Hawkins, John/Gilligan, Gary 1985. "The Suffixing Preference: a Processing Explanation". *Linguistics* 23, 4–6, 723–758.

Dishington, James 1980. "The Origin and Synchronic Status of OHG hebist, hebit. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 102, 5–18.

Donaldson, B.C. 1983. *Dutch. A linguistic history of Holland and Belgium.* Leiden: Nijhoff.

Dressler, Wolfgang U. 1985. "Sur le statut de la suppléance dans la morphologie naturelle". In *Langages* 78, 41–56.

DUDEN-*Grammatik der deutschen Gegenwartssprache.* Duden Bd. 4. Mannheim: Dudenverlag [5]1995.

Faarlund, Jan Terje/Lie, Svein/Vannebo/Kjell Ivar 1997. *Norsk referansegrammatikk.* Oslo: Universitetsforlaget.

Fenk-Oczlon, Gertraud 1989. "Geläufigkeit als Determinante von phonologischen Backgrounding-Prozessen". *Papiere zur Linguistik* 40, Heft 1, 91–103.

Fenk-Oczlon, Gertraud 1991. "Frequenz und Kognition – Frequenz und Markiertheit". *Folia Linguistica* XXV/3–4, 361–394.

Haeseryn, W. al. 1997. *Algemene Nederlandse Spraakkunst.* Groningen: Nijhoff.

Harnisch, Rüdiger 1988. "Natürliche Morphologie und morphologische Ökonomie". *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 41, 426–437.

Harnisch, Rüdiger 1990. *Morphologische Irregularität – Gebrauchshäufigkeit – psychische Nähe. Ein Zusammenhang im empirischen Befund und in seiner theoretischen Tragweite.* In J. Méndez Dosuna, C. Pensado (eds). Naturalists at Krems. Papers from the Workshop on Natural Phonology and Natural Morphology, 53–64. Salamanca.

Harnisch, Rüdiger 1994. *Allomorphische Variation und kognitive Kosten. Evidenz aus dem Ortsdialekt von Ludwigsstadt in Oberfranken.* In W. Viereck (ed.). Verhandlungen des Internationalen Dialektologenkongresses Bamberg 1990. ZDL-Beiheft 76, Bd. 3, 283–296.

Hempen, Ute 1988. *Die starken Verben im Deutschen und Niederländischen. Diachrone Morphologie.* Linguistische Arbeiten 214. Tübingen: Niemeyer.

Holmes, Philip/Hinchliffe, Jan 1994. *Swedish. A Comparative Grammar.* London/New York: Routledge.

Hopper, Paul J./Traugott, Elizabeth 1993. *Grammaticalization.* Cambridge: University Press.

Hunnius, Klaus 1989. "Wie schwierig sind die unregelmäßigen Verben? Morphologische Unregelmäßigkeit aus strukturaler und funktionaler Sicht." *Romanistisches Jahrbuch* 40: 44–59.

Janda, Laura 1996. *Back from the Brink: a Study of how Relic Forms Serve as Source Material for Analogical Extension.* LINCOM Studies in Slavic Linguistics 01. München: Lincom Europa.

Krug, Manfred 1994. *Contractions in Spoken and Written English: a Corpus-based Study of Brachychronic Language Change.* Exeter. (Magisterarbeit)

Krug, Manfred 1998. "String Frequency: A Cognitive Motivating Factor in Coalescence, Language Processing and Linguistic Change". *Journal of English Linguistics,* Vol. 26/4, 286–320.

Lass, Roger 1990. "How to Do Things with Junk: Exaptation in Language Evolution". *Linguistics* 26, 79–102.

Lindow, Wolfgang et al. 1998. *Niederdeutsche Grammatik.* Leer: Schuster.

Lockwood, W.B. 1977. *An Introduction to Modern Faroese.* Tórshavn: Dimmalætting.

Maiden, Martin 1991. *Interactive Morphonology. Metaphony in Italy.* London/New York: Routledge.

Marti, Werner 1985. *Berndeutsch-Grammatik.* Bern: Francke.

Mayerthaler, Willi 1981. *Morphologische Natürlichkeit.* Linguistische Forschungen 28. Wiesbaden: Athenaion.

Nübling, Damaris 1995. "Kurzverben in germanischen Sprachen: Unterschiedliche Wege – gleiche Ziele". *Zeitschrift für Dialektologie und Linguistik*, LXII. Jahrgang, Heft 2, 127–154.

Nübling, Damaris 1997. *Der alemannische Konjunktiv II zwischen Morphologie und Syntax. Zur Neuordnung des Konjunktivsystems nach dem Präteritumschwund.* In A. Ruoff,

P. Löffelad (ed.). Syntax und Stilistik. Beiträge der 12. Arbeitstagung zur alemannischen Dialektologie. Idiomatica 18. Tübingen: Niemeyer, 107–121.
Nübling, Damaris 1998. *Zur Funktionalität von Suppletion.* In M. Butt, N. Fuhrhop (eds). Variation und Wortstruktur. Germanistische Linguistik 141–142, 1998, 77–101.
Nübling, Damaris 2000. *Prinzipien der Irregularisierung. Eine kontrastive Untersuchung von zehn Verben in zehn germanischen Sprachen.* Linguistische Arbeiten 415. Tübingen: Niemeyer.
Östman, Carin 1992. *Den korta svenskan. Om reducerade ordformers inbrytning i skriftspråket under nysvensk tid.* Uppsala: Inst. för Nordiska språk vid Uppsala Universitet.
Pind, Jörgen 1991. *Íslensk ordatídnibók.* Reykjavík: Ordabók Háskolans.
Quirk, Randolph et al. [13]1995. *A comprehensive grammar of the English language.* London: Longman.
Ronneberger-Sibold, Elke 1980. *Sprachverwendung – Sprachsystem. Ökonomie und Wandel.* Linguistische Arbeiten 87. Tübingen: Niemeyer.
Ruoff, Arno [2]1990. *Häufigkeitswörterbuch gesprochener Sprache.* Idiomatica Bd. 8. Tübingen: Niemeyer.
Sjölin, Bo 1969. *Einführung in das Friesische.* Stuttgart: Metzler.
Tiersma, Pieter Meijes 1985. *Frisian Reference Grammar.* Dordrecht: Foris Publ.
Vennemann, Theo 1993. *Language change as language improvement.* In Ch. Jones (ed.). Historical Linguistics, 319–344. London: Longman.
Vincent, Nigel 1993. *Exaptation and Grammaticalization.* In H. Andersen (ed.). Historical Linguistics 1993. Selected Papers from the 11th International Conference on Historical Linguistics, Los Angeles, 16–20 August 1993. Amsterdam/Philadelphia: John Benjamins, 433–445.
Weber, Albert [3]1987. *Zürichdeutsche Grammatik. Ein Wegweiser zur guten Mundart.* Zürich: Verlag Hans Rohr.
Werner, Otmar 1987a. *The aim of morphological change is a good mixture – not a uniform language type.* In A. Giacalone Ramat et al. (eds). *Papers from the 7th International Conference on Historical Linguistics*, 591–616. Amsterdam: John Benjamins B.V.
Werner, Otmar 1987b. *Natürlichkeit und Nutzen morphologischer Irregularität.* In N. Boretzky et al. (eds). *Beiträge zum 3. Essener Kolloquium über Sprachwandel und seine bestimmenden Faktoren*, 289–316. Bochum: Brockmeyer.
Werner, Otmar 1989. "Sprachökonomie und Natürlichkeit im Bereich der Morphologie". *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42, 1, 34–47.
Wurzel, Wolfgang Ullrich 1984. *Flexionsmorphologie und Natürlichkeit.* Studia grammatica XXI. Berlin: Akademie-Verlag.
Wurzel, Wolfgang Ullrich 1994. "Skizze der natürlichen Morphologie". *Papiere zur Linguistik* 50, Heft 1, 23–50.
Zipf, George 1929. "Relative frequency as a determinant of phonetic change". *Harvard Studies in Classical Philology* 40, 1–95.
Zipf, George 1936. *The Psycho-Biology of Language. An Introduction to Dynamic Philology.* Massachussets: M.I.T. Press.
Zipf, George 1949. *Human Behaviour and the Principle of Least Effort.* New York: Hafner.

University of Mainz
Germany

nuebling@mail.uni-mainz.de

Paradigm organization and lexical connections in the development of the Italian *passato remoto*

ELISABETTA MAGNI

1. INTRODUCTION

In modern Italian the verb "is a category whose complexity and intricacy is unmatched elsewhere in the grammar" (Maiden 1995: 122). On the whole the Italian verb system displays a historical tension between the tendency for one meaning to be represented by one morphological form (a tendency manifested through the mechanism of analogical leveling of allomorphy), and the amplification and spread of existing alternation patterns.

These phenomena can be observed in particular in the inflection of the *passato remoto* (remote preterite), where two different inflectional patterns, regular and irregular, exist. The second pattern shows stem alternations with highly idiosyncratic vocalic and consonantal allomorphy, and despite its complexity, it is not only preserved but also tends to become productive as time goes by.

The aim of this study is to investigate some unresolved problems connected with diachronic and synchronic aspects of this pattern of inflection. More specifically, on the one hand I examine mechanisms of morphological change of some innovative irregular preterites, and on the other hand, I discuss the strategies of alternating inflection and the possible reasons for its resistance to analogical leveling. I argue that the pathway along which the category of irregular preterites evolves and expands cannot be explained by assuming regular phonological developments but is crucially determined by morphologically and cognitively based generalizations. Following Bybee (1985), I also show that cognitive strategies are relevant in order to understand the way these structures of Italian verb morphology are produced, learned and changed by the speakers.

The proposal advanced here can contribute to the debate "connectionism vs. rules" (Dressler et al. 1997), where the issue under discussion is whether surface distinctions between regular and irregular forms have to be attributed to an underlying distinction in production mechanisms (single-system model vs. dualistic model, see Lazzeroni and Magni, forthcoming). In my opinion, the morphological systems of natural languages cannot be considered only in terms of a sharp dichotomy between absolute regularity and irregularity coinciding with the rule-rote bifurcation in lexical processing, and an alternative model of morphological processing is needed.

In fact, the regularity–irregularity split focuses on the two extremes of a *continuum* that, in synchrony and in diachrony, displays some intermediate areas of subregularity as well. Therefore I assume that, on the one hand, the regular aspects of inflection are covered by affixation that belongs to the system of combinatorial operations, and on the other hand the irregular domain of inflection is encoded in

Geert Booij and Jaap van Marle (eds), Yearbook of Morphology 1999, 75–96.