



Visible verbal morphology: Morpheme constancy in Germanic and Romance verbal inflection

Nanna Fuhrhop¹

Received: 6 May 2019 / Accepted: 5 November 2020 / Published online: 17 November 2020
© The Author(s) 2020

Abstract In different spelling systems, different grades of morpheme constancy can be found: German has a high degree of morpheme constancy (especially stem constancy, for example *rennen – rennt* both forms with <nn>), while English has comparatively less (*running – run*, only the disyllabic form with <nn>). This paper investigates the interaction between stems and verbal inflectional suffixes in terms of constancy in three Germanic languages (Dutch, English, German) and five Romance languages (French, Italian, Portuguese, Romanian, Spanish). Verbal inflection is always the most widespread inflection, so it is a well-defined area for getting an idea of how spelling systems may function. For the Germanic languages, this analysis will primarily focus on the alternation between monosyllabic and disyllabic forms. For the Romance languages, it will focus on the <c>/ <g>-alternations in interaction with the following vowel. The aim is to describe a scale of morphological spelling: The alternation of <c> and <ç> is not an instance of constancy, but of similarity, something between constancy and non-constancy. Morpheme constancy is no longer a binary feature. Comparing verbal inflection takes us another step towards the development of typological parameters for visible morphology.

Keywords Morpheme constancy · Stem constancy · Affix constancy · Morpheme similarity · Stem alternations

1 Introduction

In different spelling systems, different grades of morpheme constancy can be found: German has a high degree of morpheme constancy (especially stem constancy, for

✉ N. Fuhrhop
nanna.fuhrhop@uni-oldenburg.de

¹ Institut für Germanistik, Universität Oldenburg, Ammerländer Heerstraße 114-118, 26129 Oldenburg, Germany

example *rennen* – *rennt* both forms with <nn>), while English has comparatively less (*running* – *run*, only the disyllabic form with <nn>). This paper investigates the interaction between stems and verbal inflectional suffixes in terms of constancy in three Germanic languages (Dutch, English, German) and five Romance languages (French, Italian, Portuguese, Romanian, Spanish). In all of the languages examined here, there is an interaction between stem and inflectional suffixes in terms of stem constancy or stem variation. This paper focuses first on verbal inflection, and second on special phenomena: for Germanic languages the alternations between monosyllabic and disyllabic verbal forms, and for Romance languages the alternations of <c> and <g>. In a more abstract view, they are more analogous; in Romance languages the phoneme-grapheme-correspondences of <c> and <g> depend on the following vowels. This is true for all Romance languages, and the respective spelling systems have to accommodate this. In the three Germanic languages investigated here, verbal suffixes can be syllabic or non-syllabic. Most of the syllabic suffixes start with a vowel. Thus, in the case of Romance languages, what matters most is the vowel that follows, and in Germanic languages, whether a vowel follows at all.

Stem constancy as a spelling principle means that the stem remains the same in different forms, even though it seems unnecessary from a phonographic point of view. In the German language, for example, forms like *man* – *Mann* (‘one’, pronoun – ‘man’) are pronounced in the same way, with a lax short vowel. Thus, in phonographic terms <nn> in *Mann* is redundant. In *Männer* (‘men’), however, the double consonant letters are necessary, a form like **Mäner* would be pronounced with a tense long vowel. The forms *Männer* and *Mann* are related to each other; in *Mann* the relation is shown by stem constancy.

Verbal inflection was chosen because it is the most differentiated inflection. The verbal suffixes are variable; Germanic languages have verbal suffixes beginning either with a consonant or with a vowel. Romance languages have suffixes beginning with different vowels.

The phenomena chosen here are comparable for the given languages. Germanic languages have ambisyllabic consonants spelled as double consonant letters, and Romance languages have different correspondences of <c> and <g> depending on the following vowel. The phenomena are comparable at least inside the language family, and all the languages can be ordered in a scale. The observations can be transferred to other parts of the spelling system, to nominal inflection as well as derivation morphology.

2 The Germanic languages

In English, German and Dutch, there is less verbal inflection compared to the Romance languages. The main difference is that all these languages have syllabic and non-syllabic verbal suffixes (*running* – *runs*). Dutch and English in particular also have forms without a suffix; sometimes an inflected form has the same form as the stem (*to run* – *we run*). Accordingly, the main difference can be formulated according to whether or not a vowel letter follows, as in *rennen* – *rent* – *ren* in Dutch and *running* – *runs* – *run* in English, or the contrasting *rennen* – *rennt* in German.

The systems can be separated into weak (and regular) inflection and strong (and less regular or irregular); only the weak and mostly regular verbs are examined here (for strong and irregular verbs see Fuhrhop 2017 for English and German).

2.1 English

English weak verbs normally have four inflectional forms.

- a. *want, wants, wanted, wanting*
- b. *call, calls, called, calling*
- c. *beg, begs, begged, begging*
- d. *pack, packs, packed, packing*
- e. *wash, washes, washed, washing*
- f. *echo, echoes, echoed, echoing*

In terms of visible morphology there are a few interesting points:

First the suffix for the third person singular is sometimes <s> and sometimes <es>. If it is pronounced disyllabic (*washes*), it is <es>, though it is also <es> after <o>, unless pronounced monosyllabic.¹ This suffix features two graphemic allomorphs, but no affix constancy. This lack of affix constancy is not based entirely on phonological variation; the variation after <o> also has a graphemic basis.² In summary, the existence of two forms shows less affix constancy, as there are two forms in both spoken and written language.

However, this is not the case for <ed>. For the imperfect and participle suffix *-ed*, the graphemic form is constant, but the phonological form is not; it can be pronounced syllabically or non-syllabically (*repeated – begged*), and be voiced or voiceless (*begged – talked*, C. Chomsky 1970). The affix <ed> is graphemically constant.³ In spoken language, there are monosyllabic and disyllabic forms, whereas in written language there is only one form.

The third verbal suffix *-ing* is always constant, both in spoken and written language. Thus the three verbal suffixes in English behave differently in terms of variation. Only the imperfect suffix shows more morphology in written language than in spoken languages, so this is a part of visible morphology.

The stems themselves may also be of interest. Between the stem and the suffix (especially *-ed* and *-ing*), the ambisyllabic consonant is spelled with a double consonantal letter, for example <gg> and <ll>. The analogue monosyllabic form is sometimes also spelled with the double consonant letter (<ll> in *call* and in *calls*), and sometimes with a singular consonant letter (<g> in *beg* and in *begs*). The rule depends on the letter.

- a. Double: *sniff, call, kiss, buzz*
- b. Single: *rob, wed, beg, dim, ban, step, bet*

¹This is also true for the highly irregular verbs *go – goes, do – does*.

²This is also true for nominal plural forms in *potatoes, echoes*.

³And as Berg et al. (2014) show, <ed> is also highly unique. There are relatively few other words ending in <ed> (like <hundred>), although phonologically, we would expect more words to be spelled with a final <ed> (e.g. *stupid, instead, salad* – the pronunciation of the ending is the same according to CELEX).

c. Single and double:

r: *stir* – *whir(r)* – *err*

z: *quiz* – *whiz(z)* – *buzz*

d. Complex graphemes ‘stay’: *pack*, *wash*, *ring*, *catch*e. Not double: *fix* (*fixing* – **fixxing*)

In this case, English exhibits less stem constancy than is possible because not all double consonants remain constant. Most of them are single in the monosyllabic form. As previously demonstrated, <s>/<es> varies depending on phonological syllabification; if the phonological form is disyllabic, so is the graphemic form, as in the word *kisses*, and vice versa. However, this is not the case with *-ed*, which is always <ed>, regardless of how it is pronounced. The double consonant is interesting in that it behaves much like disyllabic forms such as *begged*, *banned*, *dimmed*.

The verbal suffix *-s* is not used uniformly (there are two variants, *-s* and *-es*, in both written and spoken language), while the *ed*-suffix is used uniformly in written language (<ed>) but not in spoken language; this is a case of affix constancy.

2.2 German

German infinitives are disyllabic with two exceptions (*sein*, *tun*), though this holds only for their graphemic form. This means they have a structure with at least two vowel letters, normally separated by at least one consonant letter, e.g. *knurren* and *drehen*. However, monosyllabic pronunciations are possible for both of these. The ambisyllabic consonant is spelled as a double consonant letter (a); in some cases other combinations (complex graphemes) are used like in English (b):

- a. *schrubben* (‘to scrub’), *hoffen* (‘to hope’), *flaggen* (‘to fly flags’), *bellen* (‘to bark’), *trennen* (‘to separate’), *kippen* (‘to overturn’), *knurren* (‘to growl’), *küssen* (‘to kiss’), *retten* (‘to rescue’)
- b. *lachen* (‘to laugh’), *blicken* (‘to look’), *mischen* (‘to mix’)

As previously mentioned, there are monosyllabic and disyllabic forms in verbal inflection:

- a. *ich schrubbe* – *wir schrubben* (1. ps., ind., pres., act. / sg. – pl.)
- b. *du schrubbst* – *ihr schrubbt* (2. ps., ind., pres., act. / sg. – pl.)

The double consonant letter occurs in every form, in disyllabic as well as in monosyllabic forms. This brings us to another point concerning stem constancy. German phonology has final obstruent devoicing, but the letters remain in this case as well (<bb> and not <pp>). This is also true if there is no double consonant letter:

loben – *lobt* (‘to praise – praises’), *sagen* – *sagt* (‘to say – says’)

For <d> there are only monosyllabic forms with strong verbs:

finden – *fand* (‘to find – found’)

German shows a high degree of stem constancy; this is evident here in the retention of double consonant letters and the tendency to ‘ignore’ final obstruent devoicing consequently in spelling.

2.3 Dutch

Dutch has at least four interesting features that concern the comparison of monosyllabic and disyllabic forms in weak verbal inflection.

1. The behaviour of double consonant letters
2. The behaviour of double vowel letters
3. The behaviour of voiced consonant letters at the stem's end
4. The behaviour of voiced and devoiced letters in present participle.

The double consonant letter is used for a phonological ambisyllabic consonant like in English and German.

tobben ('to worry'), redden ('to save'), blaffen ('to bark'), leggen ('to lay'), hollen ('to run'), trimmen ('to trim'), rennen ('to run'), gissen ('to guess'), pletten ('to crush')

If the consonant letter does not represent the ambisyllabic position, only one letter is used.

tobt ('worries'), redt ('saves'), blaft ('barks') etc.

This is also true for words borrowed from English like *cross* – *hij crost*. To summarize, the double consonant letters are only used in an ambisyllabic position. For pairs like *tobben* – *tobt*, this means less stem constancy. 'Less stem constancy' means that a form like **tobbt* could also refer to the same pronunciation. However, **tobbt* is an ungrammatical form for the spelling of Dutch. There is generally less stem constancy in the Dutch spelling system than in German.

The double vowel letters are only used in closed syllables. They represent a tense vowel which is also long.

faalt (fails), kleedt (dresses), doodt (kills), huurt (rents)

The infinitive forms of these verbs are spelled with only one vowel letter. In open stressed syllables the vowels are normally pronounced long:

falen ('to fail'), kleden ('to dress'), doden ('to kill'), huren ('to rent')

That means that pairs *faalt* – *falen* shows less stem constancy than possible.

Like German, Dutch also has final obstruent devoicing. There is a difference between letters which represent plosive letters (a) and those which represent fricative letters (b):

- a. kleden – kleedt ('to dress – dresses'), redden – redt ('to save – saves'), melden – meldt ('to report – reports')
- b. lezen – leest ('to read – reads'), leven – leeft ('to live – lives')

<d>, <g>, stays, <z> and <v> changes. The examples in (a) show stem constancy, but the examples in (b) do not.

For past tense there are two suffix variants: *-de* and *-te*. The suffix variant is chosen by the preceding consonant: after voiced consonants *-de* is chosen (a), after devoiced *-te* (b).

Table 1 Double consonant letters – depending on the following letter

	One consonant letter	Two consonant letters
German	.	<i>jobb-t, klaff-t, knall-t, schwimm-t, renn-t, schnapp-t, knurr-t, küss-t nimm-t – *nimmen/ nehmen (genommen), tritt</i>
Dutch	kis-t, cros-t (from engl. cross), bel-t, blaf-t, bak-t, ren-t, zwem-t	
English	rob, wed, jog, swim, run, step, split	kill, sniff, buzz, kiss

- hoor-de ('heard'), red-de ('saved'), daag-de ('wear'), meld-de ('reported'), tob-de ('worried'), hol-de ('ran'), trim-de ('trimmed'), ren-de ('ran')
- maak-te ('made'), dank-te ('thanked'), dans-te (danced), damp-te ('fumed'), blaf-te ('barked')
- vrees-de ('feared'), leef-de ('lived')

(c) shows obstruent devoicing with <s> and <f>, but <de> shows the reference to a voiced consonant (/z/ and /v/) in a related form. It is also pronounced like this for the imperfect. The participle forms are as follows:

- gehoor-d ('heard'), gered-d ('saved'), dedaag-d ('wear'), gemeld-d ('reported'), getob-d ('worried'), gehol-d ('ran'), getrim-d ('trimmed'), geren-d ('ran')
- gemaak-t ('made'), gedank-t ('thanked'), gedans-t (danced), gedamp-t ('fumed'), geblaf-t ('barked')
- gevrees-d ('feared'), geleeft-d ('lived')

The <d> in (c) shows morphology – it shows a reference to a related form.

To summarise:

The double consonant letters and also the double vowel letters are not regulated by stem constancy. Obstruent letters at the stem's end show partial morpheme constancy, and the plosives exhibit complete stem constancy. In some cases there is no stem constancy for stems ending with the letters <z> and <v> (*vrezen – vrees*), but a morphological relation is shown in the participle perfect form: in *gevreesd* the combination of -s and -d shows the relation to a form with <z>.

2.4 Summary remarks on the Germanic languages

Each of the three languages examined here uses double consonant letters for phonologically ambisyllabic consonants: *rennen*, *running*, *rennen*. In related monosyllabic forms, the writing systems of the languages differ; German retains the double consonant letters nearly all the time (*rennt*), English retains it depending on the particular consonant (*run – spell*), and Dutch only uses the double consonant letter for the ambisyllabic consonant (see Table 1).

An interesting feature of English is that it has twelve consonants that are doubled when they correspond to ambisyllabic consonants. In morphologically related monosyllabic forms, seven of them omit the second consonant, three of them retain it, for <r> and <z> both are found (*stir; err; quiz, buzz*). That is just an observation about the

types. But English is, regarding this point, not as consistent as the other Germanic languages.

German and Dutch are more consistent in terms of stem constancy; while German shows the stem, Dutch refers very clearly to vowel pronunciation, with the double consonant letter and with the double vowel letter.

3 Visible morphology in Romance languages

Romance verbal inflectional morphology is much more differentiated than that of Germanic languages. Each of the languages investigated here has at least three different infinitival suffixes, a variety of personal endings, and so on. But this holds true for both written and spoken language. So in terms of graphemic affix and stem constancy, two points are of particular interest. The first point concerns all Romance languages examined here, and the second is only relevant for French.

The first point concerns the stem constancy of stems ending in <c> or <g>; the pronunciation depends on the following vowels. The Romance languages have different strategies to deal with this fact, and these strategies will be described here in relation to morpheme constancy. The second point is the visible morphology in French verbal inflection. The other Romance languages have shallow graphemic systems; the verbal suffixes are spelled as they are pronounced. French has a deep graphemic system (Meisenburg 1996: 161), and inflection in written French is much more differentiated than it is in spoken French.

3.1 <c> and <g> in Romance languages

In the Romance languages, the correspondences of <c> and <g> differ depending on the following vowel. Meisenburg (1996) showed that Early Latin had three different spellings for /k/: <C> before <E> and <I>, <K> before <A> and <Q> before <O> and <V>. The letter <K> was later lost in Latin spelling; for Meisenburg, these are too many letters for one sound. The history is complicated, but what is relevant for the investigation here is: 1. Early Latin had the <k>, 2. <q> functioned once without the <u> and 3. the first correspondence of <c> was to /k/. In today's Romance spelling systems, the correspondence of <c> depends on the following vowel; normally there is a difference between <i, e> on the one side, and <a, o, (u)> on the other side. Additionally, <g> and <z> have a curious history. The most important issue here is that there are different phonological correspondences for <g>, depending on the following vowel.

Table 2 shows the correspondences of <c> and <g> depending on the following vowel in different Romance languages; also see Meisenburg 2014, 8f.

More abstractly the correspondence to <c> and <g> are plosives before back vowels and fricatives or affricates before front vowels (see Table 3).

Italian and Romanian have as well a 'plosive indicator' (<h>) as a fricative indicator (<i>) (see Table 4).

In French, Portuguese, and Spanish, the <u> may be seen as a plosive indicator, but there is no <cu> but <qu>. For fricativization, Spanish consistently uses different

Table 2 Correspondences of <c> and <g> depending on the following vowel in Romance languages

	French	Spanish	Portuguese	Italian	Romanian
<c> before <i e>	/s/	/θ/ or /s/ (Latin-America)	/s/	/tʃ/ ^a	/tʃ/
<c> before <a o u>	/k/	/k/	/k/	/k/	/k/
<g> before <i e>	/ʒ/	/x/	/ʒ /	/dʒ/	/dʒ/
<g> before <a o u>	/g/	/g/	/g/, inters./ɣ/	/g/	/g/

^a <sc> before <e, i> corresponds to /ʃ/

Table 3 Articulation modus of the correspondences of <c> and <g> in Romance languages

	French	Spanish	Portuguese	Italian	Romanian
<c> and <g> before <i e>	fricative			affricate	
<c> and <g> before <a o u>	plosive				

Table 4 Italian and Romanian ‘allographs’ to <c> and <g>

	Italian/ Romanian
/k/ and /g/ before <i, e>	<ch>, <gh>
/tʃ/ and /dʒ/ before <a, o, u>	<ci>, <gi>

Table 5 West-Romance ‘allographs’ to <c> and <g>

	French/ Portuguese/ Spanish	French/ Portuguese/ Spanish
/k/ and /g/ before <i, e>	<qu>	<gu>
/s/ and /ʒ/ or /x/ before <a, o, u>	<ç>, for Spanish <z>	<j>

letters. In French and Portuguese, there is an allomorph (<ç>), and another letter, (<j>), is used for the other. An interpretation only true for French and Portuguese is that there are two modifications (<gu> for <g> and <ç> for <c>) to ‘retain’ more constancy, so French and Portuguese allow more room for morpheme similarity (see Table 5).

These correspondences are quite interesting in verbal inflection. The Romance languages have a complex verbal inflection system with many different suffixes, often depending on different vowels. The following paragraphs focus on the question: How do stems with final <c> and <g> deal with different suffixes?

3.2 <c> and <g> in Portuguese

Portuguese has the regular infinitive endings *-ar*, *-er*, *-ir*; it has infinitive suffixes with back (*-ar*) and front vowels (*-er*, *-ir*). The classes with *-ar*, *-er* are large, but the *ir*-class is small. For every class, there is at least one personal form with a vowel from the other group. For the *ar*-verbs, there is *-ei* for first person, singular, perfect, indicative (like *ficar – fiquei* ‘to stay – stayed’); for the *er*-verbs *-o* for the first person, singular,

Table 6 Graphemic adaptations for <c> and <g> in Portuguese

Graphemic adaptations	Infinitive, example	Following vowel grapheme	Finite forms, examples
<c> → <qu>	ficar ('to stay'), tocar ('to touch')	<e>	fique, fiquei; ^a toque, toquei
<ç> → <c>	começar ('to start'), dançar ('to dance')	<e>	comece, comecei; dance, dancei
<g> → <gu>	pagar ('to pay'); ligar ('to turn on')	<e>	pague, paguei; ligue, liguei
<c> → <ç>	agradecer ('to thank'); conceder ('to indulge')	<a>/-<o>	agradeço, agradeça; ^b condeço, condeça
<g> → <j>	proteger ('to protect'), tingir ('to dye')	<a>/-<o>	protejo, proteja; tinjo, tinja
<gu> → <g>	erguer ('to lift'); distinguir ('to distinguish')	<a>/-<o>	ergo, erga; distingo, distinga

^a Forms with *-e* are 1. ps., sg., conj., pres.; forms with *-ei* 1. ps., sg., ind., imp.

^b Forms with *-o* are 1. ps., sg., conj., pres.; forms with *-a* 1. ps., sg., ind., imp.

present, indicative *-a* for first person, singular, present, conjunctive (e.g., *conceder* – *condeço* 'to indulge – I indulge'); the same for the *ir*-verbs (*corrigir* – *corrijo* 'to correct – I correct'). So there is always an alternation in the verbal paradigm. The Portuguese spelling system is largely phonographic, so graphemic adaptations are necessary (see Table 6). (This table is modelled after Kessler 2005. The tables for the other Romance languages are modelled after this table construed for Portuguese.)

3.3 <c> and <g> in Spanish

Spanish regular infinitives end in *-ar*, *-er*, *-ir* – very much like Portuguese verbs. The personal endings are also similar, although not the same in all forms. The ending for first person, singular, present, indicative is *-o* for all verb classes, so in terms of back and front vowels there is an alternation for *er*- and *ir*-verbs (*coger* – *cojo* 'to take – I take', *distinguir* – *distingo* 'to distinguish – I distinguish'). An alternation for *ar*-verbs is necessary, for example, for the subjunctive, which has *e* in the first person, singular (*atacar* – *ataquem* 'to attack – I attack'). (See Table 7.)

Compared to Portuguese, there is at least one difference: Portuguese has the alternations <c> → <ç> and also vice versa (<ç> → <c>), where Spanish has <c> ⇔ <z>. Thus Portuguese shows more stem similarity than Spanish at the price of a larger grapheme inventory (+1).

3.4 <c> and <g> in French

The French infinitive suffixes are *-er*, *-ir*, *-re* (*annoncer*, *adoucir/languir*, *vaincre*). Table 8 shows the infinitive and the finite forms. The adaptation strategies shown here are restricted because French infinitive endings do not contain back vowels.

Table 7 Graphemic adaptations for <c> and <g> in Spanish

Graphemic adaptations	Infinitive, example	Following vowel grapheme	Finite forms, examples
<c> → <qu>	atacar ('to attack'), buscar ('to look for')	<e>	ataque, ^a busque
<z> → <c>	almorzar (to have lunch'), curzar (to curl')	<e>	almuerce, cruce
<g> → <gu>	pagar ('to pay'), jugar ('to play')	<e>	pague, ^b juegue
<c> → <z>	vencer (to overcome'), zurcir ('to darn')	<a>	venza, venza, ^c zuro, zurza
<g> → <j>	coger ('to take'), dirigir ('to lead')	<a>	cojo, coja, dirigo, dirija
<gu> → <g> ^d	seguir ('to follow'), distinguir ('to distinguish')	<a>	sigo, siga, distingo, distingia
<qu> → <c>	delinquir ('to offend')	<a>/<o>	delinco, delinca

^a Forms with -e are 1. ps., sg., pretérito perfecto simple

^b 1. ps., sg., pres., subjunctive

^c Forms with -o are 1. ps., sg., pres. ind.; forms with -a 1. ps., sg., pres., subjunctive

^d Also <gu> → <gü> is possible in *averiguar* 'to find out' to *averigüe* – 'I found out'

Table 8 Graphemic adaptations for <c> and <g> in French

Graphemic adaptations	Infinitive, example	Following vowel grapheme	Finite forms, examples
<c> → <ç>	annoncer 'to announce'	<a>/<o>	annonçait – annonçons
<g> followed by <e>	protéger 'to protect'		protège – protégeas (tu)
<gu> stays	conjuguer 'to conjugate'	<a>/<o>	conjugue – conjuguons
<c> → <qu>	vaincre 'overcome'	<a>/<o>	vaincre – vainc – vainquons
<c> in front of consonants (<r>) or nothing			

Verbs with infinitive ending in -ir do not show the alternations we are looking for; the singular forms keep the -i, the plural forms normally insert -iss: *j'alanguis – nous alanguissons* 'I/we devitalize'. There are some verbs ending in -cer, -ger (and -guer), and few verbs ending in -cre (*vaincre – il vainc – nous vainquons* ('vanquish', 'he vanquishes – we vanquish') and none (in Lexique, www.lexique.org) in -gre.

There are fewer adaptations in French: the cedilla, and in one case <c> → <qu>. The price is paid in another way; the primary correspondences become more difficult. A 'silent' <u> is introduced which does not have a phonographic function. The <u> between <g> and <o> does not carry any information for pronunciation; this is different from the <u> between <g> and <e>. The <u> before <o> retains morpheme constancy, as a spelling like **conjugons* would have the same correspondence. In the case of *protéger*, the <e> after <g> retains the fricative of <g> even before <a> and <o>; before <i> in *protégiez* (second person., plural), it is not used. The primary correspondence for /ʒ/ before <a> and <o> is <j> in French (as in *jouer, jaune, jour, jamais*), but a spelling like **protéjons* would not satisfy morpheme constancy. This

Table 9 Graphemic adaptations for <c> and <g> in Italian

Graphemic adaptations	Infinitive, example	Following vowel grapheme	Finite forms, examples
<c> → <ch>, in <i>are</i> -parad.	<i>cercare</i> ‘to search’	<e>/<i>	<i>cerchi</i> ^a
<g> → <gh>, in <i>are</i> -parad.	<i>pagare</i> ‘to pay’	<e>/<i>	<i>paghi</i>
<ci> → <c>, in <i>are</i> -parad.	<i>lasciare</i> ‘to leave’	<e>/<i>	<i>lasci, lascerò</i> ^b
<gi> → <g>, in <i>are</i> -parad.	<i>mangiare</i> ‘to eat’	<e>/<i>	<i>mangerò</i>
<c> → <cqu>, in <i>ere</i> -parad.		<e>/<i>	<i>tacqui, tacque</i> (/k/) ^c
<c> → <cci>, in <i>ere</i> -parad.	<i>tacere</i> ‘to be silent’	<o>	<i>taccio, tacciamo</i> ^d
<c> → <ci>, in <i>ere</i> -parad.		<u>	<i>taciuto</i> ^e

^a Forms with *-i* 1. ps., sg., conj., pres.

^b Forms with *-erò* 1. ps., Fut.

^c Forms of ‘passive remoto’

^d 1. and 3. ps., sg., ind., pres.

^e Pass prossimo

shows again that French has quite a deep writing system, and it retains morphological spelling at different levels (cf. also Meisenburg 1996:200).

3.5 <c> and <g> in Italian

Italian infinitive suffixes are *-are*, *-ere*, *-ire*. The Italian verbal suffixes also contain both back and front vowels. In *are*-paradigms, an <h> is added (<ch>, <gh> to save the plosive before front vowels) like in *cercare* – *cerchi* (‘to search – (you) search’). In *ere*- and *ire*-paradigms an <i> is inserted after <c> (*cuocere* – *cuocio* ‘to bake/ to cook – I bake/cook’, *cucira* – *cucio* ‘to sew, ‘I sew’). Verbs with *-gere* seem to be irregular (*finire* – *fini* ‘similate’). For some endings in *-cere*, an alternation to <s> is possible (*vincere* – *vinsi* ‘win’); these cases are listed as irregular verbs (for example Diaco et al. 2016). (See Table 9.)

The most frequent alternation seems to be the alternation of <c, g> and <ch, gh> or <ci, gi> – the addition or deletion of a letter in a very special environment. And sometimes there is less phonological constancy, as can be seen in *tacere* – *tacque*, with a fricative or plosive before <e>. In summary, the addition or subtraction of a ‘second’ letter shows more stem similarity than is seen in Spanish and Portuguese.

3.6 <c> and <g> in Romanian

Romanian infinitive endings are *-a*, *-ea*, *-e*, *-i*, *-î*. For person forms, there are also some vowel alternations, but we will concentrate on the consonantal endings of the stems (as already done for the other Romance languages). Romanian has the same possibility as Italian with <h> and <i> after <c> and <g>. But in verbal inflection, in many cases the spelling is constant but the pronunciation alternates (*a juca* /k/ – *juca* /ts/ ‘to play – you play’, analogue for *a băga* ‘to plug’). (See Table 10.)

In the Romanian verbal system the phonographic tolerance is higher – or less phonological stem constancy. Therefore there is more graphemic stem constancy at

Table 10 Graphemic adaptations for <c> and <g> in Romanian

Graphemic adaptations	Infinitive, example	Following vowel grapheme	Finite forms, examples
<c>	a pleca 'to leave'	<i>	pleci ^a
<g>	a ruga 'to pray'	<i>	rugi
<c>	a conduce 'to lead'	<a>/<o>	să conducă (conj, 3. ps., sg., pres.) (/k/)
	a juca 'to play'		joci
	a placea 'to like'		place (affricate), plac (plosive)
<c> → <s>	a aduce 'to bring'		aduci, <i>but</i> am adus (perfect)
<g>	a fugi 'to run'	<a>/<o>	să fugă (conj) (/g/)
<gh>	a veghera 'to ward'	<a>	vegھاți (imparative plur.)

^a Forms with -i 2. ps., sg., ind., pres.

Table 11 Graphemic alternations in Romanian

Graphemic adaptations	Infinitive, example	Following vowel grapheme	Finite forms, examples
<t> → <ț> (/ts/)	a cânta 'to sing'	<i>	cântți
<d> → <z>	a crede 'to believe'	<i>	crezi
<s> → <ș>	a lăsa 'to leave'	<i>	lași
<șc> → <șț>	a mișca 'to move'	<i>	miști

some points. In Romanian, there are two interesting additional consonant letters: <ș> and <ț>; their correspondences are <ș> → /ʃ/ and <ț> → /ts/. They often occur in morphologically related forms with <s> and <t>, like shown in Table 11.

The correspondence for the affricate spelling /ts/ is solved graphemically elegantly with <ț>, a minimal modification of the letter form. For the voiced variant, this is not the case (<d> → <z>). The spelling system of Romanian is quite young; Romanian has only been spelled with Roman letters since the 19th century. The first texts in Cyrillic letters are also only from the 16th century (Meisenburg 1996:365). To summarize, the Romanian system has less phonological stem constancy but more graphemic stem constancy; it is phonologically much more tolerant than the other Romance languages. Additionally, it has special letters/graphemes to save stem constancy. Nevertheless, it has a shallow spelling system.

3.7 Summary: <c> and <g> in Romance verbal systems

All of the languages we have examined here have the <c>/<g> alternation. It is evident that the primary correspondence of <c> and <g> differs depending on the following vowel. One aim is to judge the graphemic depth of the systems – if the spelling system is highly phonographic, it is a shallow system, and if there are other principles, such as morphological ones, for example, it is a deeper system. This alternation is even present in French, which has a deep spelling system. Thus, the c-/g-correspondences have to be dealt with in every single language. Logically, there are

Table 12 The spectrum between phonological and graphemic constancy in Romance languages in verbal inflection

	Phonological stem constancy	Graphemic stem alternations	Graphemic stem similarity		Graphemic stem constancy
			a letter is inserted/deleted	diacritic form	
Spanish	yes	<c> ↔ <q> <c> ↔ <z> <g> ↔ <j>	<g> ↔ <gu>		
Portuguese	yes	<g> ↔ <j>	<g> ↔ <gu>	<c> ↔ <ç>	
French	yes ^a	(<c> → <qu>, <i>vaincre</i>)	<g> ↔ <ge>	<c> ↔ <ç>	<gu> stays
Italian	mostly yes	<c> → <s> in irregular forms	<g> → <gh>/ <gi> <c> → <ch>/ <ci>		
Romanian	no	<c> → <s> in irregular forms	not in (regular) conjugation	t ↔ ț s ↔ ș	<c> stays <g> stays

^a Phonological ‘stem constancy’ is only classified for the endings <c> and <g>; the ‘loi of position’ is not taken into account here, thanks to an anonymous referee

three different options: alternation in spelling, alternation in phonology, or a mixed system. Therefore, two new features must be introduced: ‘phonological stem constancy’ and ‘stem similarity’. Stem similarity is less than stem constancy – for example a diacritic is added or just one letter is added or deleted.

In Romanian, for example, the pronunciation alternates within paradigms, so there is – in this case – no phonological stem constancy (*a juca – joci* in Romanian), but in the West-Romance languages, there is (Portuguese *ficar – fique*).

Stem similarity can be construed by increasing the inventory (French and Portuguese, in other cases also Romanian) or by using letters independently of their first correspondence (<i> in Italian and Romanian, <h> in Romanian) or <h> in Italian without any correspondence.

First we want to show the continuum between ‘phonological stem constancy’ and ‘graphemic stem constancy’, as illustrated by the phenomena (see Table 12).

For the variation of <c> and <g>, graphemic stem constancy can be classified into three degrees; there is graphemic similarity between graphemic alternations and graphemic constancy. This is correlated to the phonological behaviour of verbal inflection – it can also alternate like in Romanian. Romanian verbal inflection has the highest graphemic stem constancy, and the smallest phonological stem constancy. The other languages can be classified based upon the possibilities for showing graphemic similarity. Spanish shows less graphemic stem similarity. Portuguese shows more (because of the cedilla). Italian is in-between, with graphemic similarity on par with Portuguese but less strict with regards to phonological constancy.

Because French is a language with a deep system, a clearer result was expected. There are some points which show more graphemic constancy: 1. The column graphemic alternation in French is empty, which means more graphemic constancy. 2. The column graphemic constancy shows one case, which also shows in the same direction. However, it may be that the result was less clear than expected due to the

Table 13 Correspondences between spoken and written verbal suffixes in French

/ɛ/	<ais>	imperfect 1./2. ps, sg	<i>er-ais</i> , 1./2. ps, sg, conditional
	<ait>	imperfect 3. ps, sg	<i>er-ait</i> 3. ps, sg, conditional
	<aient>	imperfect 3. ps, pl	<i>er-aient</i> 3. ps, pl, conditional
	<ai>		<i>er-ai</i> 1. ps, sg, future
/ɔ̃/	<ons>	present 1. ps, pl	<i>(e)r-ons</i> 1. ps, pl, future
	<ont>		<i>(e)r-ont</i> 3. ps, pl, future
/e/	<er>	infinitive of the largest verb class	
	<é> (<ée>)	participle, sometimes marked for gender	
	<ez>	2. ps, pl, present and future	
∅	<e>	1./ 3. ps, sg, pres, ind/subj	
	<es>	2. ps, sg, pres, ind/subj	
	<s>	1./2. ps, sg, pres, ind/subj – depending on the verbal class	
	<ent>	3. ps, pl, pres, ind/subj	

fact that only verbal inflection was examined. French shows fewer variants within infinitive suffixes, and all infinitive suffixes begin with front vowels.

In terms of depth, all other systems are shallow compared to French; all of the systems have very regular phoneme-grapheme-correspondences. These differences have already been shown here. French is a special case in this respect as well. French is highly tolerant in terms of grapheme-phoneme-correspondences; in French, redundant vowel letters (like <gu> before <o> in *conjuguons*) do not seem to disturb anything. This indicates that French is a deep system. On account of this depth, we should take a closer look at its affixes.

3.8 Verbal inflection in French

The inflectional suffixes in the other Romance languages are written and spoken. Therefore, it does not seem to be an object of interest for the analysis of visible morphology at this moment. We should, however, examine some features of the French system, as it is a prime example of visible morphology.

In French verbal inflection, tense and mood are specified in written as well as in spoken language. However, person and number are normally only a specified in written language.⁴ Thus, many phonological endings are ambiguous as shown in Table 13.

While many Romance languages are pro-drop-languages, French is not. Normally the subject pronoun is overt both in written and in spoken language. For written lan-

⁴“Si les oppositions de temps sont bien assurées à l’oral et à écrit, les distinctions de personne et de nombre sont moins bien effectuées par les désinences verbales à l’oral. Ce sont les formes du sujet, en particulier les pronoms personnels, qui apportent souvent l’indication de la personne et du nombre.” (Riegel et al. 1994:441).

“Since the oppositions of tenses are saved in spoken and written language, the distinctions of person and number are not as easily distinguished by verbal endings in spoken language. The forms of the subjects, especially the personal pronouns, often give the information about the person and the number.” (Translation N.F.)

Table 14 Relative uniqueness of correspondences to /ɛ/ in French

	Total number	Also verbs	Percent of verbs	Exclusively verbs	Percent of exclusively verbs
-er	5687	4738	83.3%	4596	80.82%
-é	7207	5855	81%	2824	39.2%
-ée	5092	4584	90%	1820	35.7%
-ez	4286	4274	99.7%	4248	99.11%

guage, this seems to be redundant because the verbal morphology already indicates person and number.

Let us take a closer look at /ɛ/ and its correspondences to *-er*, *-é*, *-ée*, *-ez*. It shows a large spectrum of uniqueness, which means that a form ending for example in *-ez* will be a verbal form, second person, plural and nothing else. The following numbers are from Lexique, a database of 150,000 words of French. Lexique is a lexical database that allows for counting the types. For example, there are 5,687 types ('total number') ending in *-er*. From these 5,687 words, 4,738 words are classified as verbs in the database and 4,596 are exclusively classified as verbs. So 80.82% of words ending in *-er* can only be verbs and nothing else. That is how Table 14 should be interpreted.

Different counts are possible. Here the classification as a verb is counted. Sometimes a form with *-er* can also be a noun (*écuyer* 'horseman'), a form with *-ée* can also be a noun (*chicorée*), and so on. Very often this is a result of productive conversion (*caramélisée* as participle or adjective). Thus there are different numbers: 'also verbs' means that a lemma is classified as verb and as something else, and 'exclusively verbs' indicates the classification only as a verb. The difference between the two numbers is really high for *ée*, because the participle forms indicated for feminine (the second <e>) are usually also classified as adjectives. This is true in an analogous but less dramatic way for <é>. Any form that ends with *-ez* is very probably a verbal form (second person, plural) – this is a nearly unique representation. For the other word endings, the numbers are high but they are not unique at all. One should keep in mind, however, that the phonological form of all these words ends in /ɛ/. The spelling shows a lot more; particularly for words ending in <ez>, the probability of being a verb form (second person plural) is quite high.

Now the same is done with words phonologically ending in /ɛ/. These can be spelled with *-ai* (*je prendrai* 'I am going to take' – future simple), *-ait* (*il prenait* 'he took'), *-ais* (*je prenais* 'I took'), or *-aient* (*ils prenaient* 'they took') (see Table 15). Verbal forms are over 90% for every form, mostly over 98%. These numbers are quite high.

Table 16 completes the investigation for /ɔ/ and for unpronounced but spelled <ent>. The others for ∅ are not counted here because <s> (and <es>) can also be endings for plural forms of nouns and adjectives and <e> is a quite common ending for French words. The most frequent verbal form (the third person, singular, indicative, present *il prend* 'he takes') is not specifically marked.

To summarize, verbal inflection in French is at least partly nearly unique. It is also quite a lot more differentiated in written language than in spoken language. French (verbal) inflection is a very strong example of visible morphology.

Table 15 Relative uniqueness of correspondences to /ɛ/ in French

For /ɛ/ the same is done	Total number	Also verbs	Per cent of verbs	Exclusively verbs	Per cent of exclusively verbs
-ai	2322	2298	98.97%	2289	98.58%
-ait	5530	5519	99.8%	5473	98.97%
-ais	3240	3057	94.35%	3029	93.49%
-aient	3853	3852 ^a	99.97%	3842	99.71%

^a The only non-verbal example is *oseraient* but also *oser* ('to risk') is a French verb, so this may be a mistake in the database.

Table 16 Relative uniqueness of <ons>, <ont> and <ent> in French

	Total number	Also verbs	Per cent of verbs	Exclusively verbs	Per cent of exclusively verbs
/s/ <i>ons</i>	5073	3638	71.71%	3330	65.64%
<i>ont</i>	1112	1098	98.74%	1090	98.02%
∅ <i>ent</i> (not -aient)	7468	4980	66.68%	4907	65.71%

4 Discussion and conclusion

This analysis examined stem constancy in (regular) verbal inflection. An assumed relation between morpheme constancy and the depth of the system was evident. Deep systems have more graphemic morpheme constancy than shallow systems – it is possible to retain spellings in alternative phonological contexts. Seidenberg (2011) assumes a relation between a rich inflectional system and a shallow spelling system, and he already mentions that French is an exception (see Table 17).

For shallow systems, two points emerge. Firstly, it is necessary to be aware that there is also the possibility of phonological variation, as can be seen in Italian and Romanian. The pronunciation alternates in different forms. Therefore, graphemic stem constancy (in the sense of more than phonological stem constancy) is also possible in shallow systems. Secondly, some languages introduce variants of letters or special combinations of letters. French and Portuguese have <ç>, and Italian has the combination of <c> and <g> with <h> and <i>, letters with a special function in a special environment. More stem constancy is possible, but comes at the price of more complexity.

Does the gradation developed for the Romance languages above also fit for the Germanic languages? One important difference to Romance languages would be the monosyllabic forms in Germanic verbal inflection. The ambisyllabic consonant in disyllabic forms is no longer ambisyllabic in related monosyllabic forms. The final obstruent devoicing only takes place at the end of the syllable (in *lobt* but not in *loben*), and so on. Especially in the last case, there is a highly regular phonological alternation. There is no phonological stem constancy because of final obstruent devoicing. A graphemic alternation is found only in Dutch. An alternation is found for the double consonant in Dutch and English. The forms in Dutch for example differ by

Table 17 The spectrum between phonological and graphemic constancy in Germanic languages, analogue to the Romance languages in Table 12

	Phonological stem constancy	Graphemic stem alternations	Graphemic stem similarity		Graphemic stem constancy
			a letter is inserted/deleted	diacritic form	
German	[-] final obstruent devoicing			(not for regular verbs, otherwise <ä> – <i>lädt</i>)	<i>lobt, legt, reist rennt, klappt</i>
Dutch	[-] final obstruent devoicing	<v> → <f> <z> → <s>	<i>rennen – rent</i>		, <d>, <g> stays
English	yes		<i>running – run</i>		

one consonant letter ('a letter is deleted'). Compared to Romance languages, diacritic forms are few.

This paper concentrates on (regular) verbal inflection. Analogous phenomena are quite useful in comparing language systems, and from an abstract point of view, this was carried out for both language families: In Romance languages, it was important to note which vowel follows the verbal stems, and in Germanic languages, whether a vowel follows or not. In Germanic languages, it was always the ambisyllabic consonant that was of interest, and in German and Dutch, the graphemic consequences of final obstruent devoicing. In Romance languages, it was the behaviour of the letters <c> and <g>. The results are not restricted to verbal inflection. However, verbal inflection is the most interesting part because suffixes of different forms can be found. This is a good starting point for comparing spelling systems. The parameter 'stem constancy' has been widened. Spelling systems can be described more precisely. The next steps would be to investigate these affixes more precisely, to look for 'graphemic patterns', and to extend this study to encompass derivational morphology.

Funding Note Open Access funding enabled and organized by Projekt DEAL.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Berg, K., Buchmann, F., Dybiec, K., & Fuhrhop, N. (2014). The interface between morphology and graphemics. In *Written Language and Literacy* (Vol. 17, pp. 282–307).

- Chomsky, C. (1970). Reading, writing, and phonology. *Harvard Educational Review*, 40(2), 287–309.
- Diacio, M., Kraft, L., & Recchia, G. (2016). *Verbtabelle Plus Italienisch*. Stuttgart: Pons.
- Fuhrhop, N. (2017). Sichtbare Morphologie in der Flexion der starken und unregelmäßigen Verben. In N. Fuhrhop, R. Szczepaniak, & K. Schmidt (Eds.), *Sichtbare und hörbare Morphologie* (pp. 43–76). Berlin: de Gruyter.
- Kessler, I. M. (2005). *PONS Grammatik kurz & bündig. Portugiesisch*. Stuttgart: Ernst Klett.
- Lexique. www.lexique.org/3-4-2019.
- Meisenburg, T. (1996). *Romanische Schriftsysteme im Vergleich. Eine diachrone Studie*. Tübingen: Narr.
- Meisenburg, T. (2014). *Graphematik der romanischen Sprachen*. Ms. Osnabrück.
- Riegel, M., Pellat, J.-C., & Rioul, J. (1994). *Grammaire méthodique du français*. Paris: Presses universitaires de France.
- Seidenberg, M. S. (2011). Reading in different writing systems: One architecture, multiple solutions. In P. McCardle, J. Ren, & O. J. L. Tzeng (Eds.), *Dyslexia across languages: Orthography and the gene-brain-behavior link* (pp. 146–168). Baltimore: Paul Brooke Publishing.