Journal of Germanic Linguistics **26.4** (2014):361–394
doi: 10.1017/S1470542714000166

Modeling Intelligibility of Written Germanic Languages: Do We Need to Distinguish Between Orthographic Stem and Affix Variation?

Wilbert Heeringa, Femke Swarte,
Anja Schüppert, Charlotte Gooskens

University of Groningen

We measured orthographic differences between five Germanic languages. First, we tested the hypothesis that orthographic stem variation among languages does not correlate with orthographic variation in inflectional affixes. We found this hypothesis true when considering the aggregated stem and affix distances between the languages. We also correlated the stem and affix distances within the cognate pairs in each language pair. We found low correlations, the lowest of them being not significant. Second, we tested the hypothesis that orthographic stem variation among languages is larger than orthographic variation in inflectional affixes. This hypothesis was also found to be true. Orthographic distance is likely to be a potential predictor of written intelligibility, but our results suggest that when modeling written intelligibility, a distinction needs to be made between stem and affix distances.*

1. Introduction.

Sometimes readers are presented with texts written in a language unknown to them. When a text is written in a language closely related to the native language of the reader, he or she may be able to understand the text to some extent. When the reader is trying to understand the text, several factors may play a role, including, but not limited to, lexical, orthographic, morphological, and syntactic differences.

On the level of the lexicon, cognate words generally facilitate the reading process. Cognates are words that also exist in the reader's native language, have the same or a similar shape, a similar meaning, and the

* We thank two anonymous reviewers for providing us with a wealth of comments and constructive criticism.

same etymological origin. The fewer cognates found in the text, the larger the lexical distance between the reader's native language and the language this text is written in, and the more difficult it is for the reader to understand it.

The intelligibility of the texts also depends on orthographic differences, that is, the extent to which the written form of the cognate words in the text differs from that in the reader's native language. Orthographic differences are the result of differences in spelling conventions (for example, Dutch *se**c**tor* versus German *Se**k**tor*) and historical developments of the pronunciation (for example, Dutch *hel**p**en* versus German *hel**f**en*). In particular, the latter suggests the question whether readers use phonological cues when reading a text written in a Germanic language they do not understand. If the reader is not familiar with the language, he or she does not know the spelling rules and interprets a pronunciation according to the spelling rules and phonological system of his or her own language.

However, when there is (some) knowledge of the spelling system, phonological cues may help. For example, English *nay* and Dutch *nee* are written differently but pronounced approximately the same. This may help the reader to understand that the words have the same meaning. In contrast, words may be pronounced differently and have the same orthography, as for example, English *school* versus Dutch *school*. In that case, phonological cues would not help the reader to understand that the words have the same meaning. In this paper, we are interested in the intelligibility of languages that the reader is not familiar with. In such a situation, we do not expect phonological cues to play a significant role in helping the reader to understand the text.

In orthography, we distinguish between stems and inflectional affixes. A stem is obtained on the basis of a root. Matthews (1991:64) defines a root as "a form that underlies at least one paradigm or partial paradigm, and is itself morphologically simple." For example, *burn* is a root, which underlies at least one paradigm, *burner* with the derivational affix *-er*. The derivational affix changes the meaning and often also the class of a word. In this case, the verb *burn* becomes a noun. A stem is "a form that underlies at least one paradigm or partial paradigm; but it is itself morphologically complex" (p. 64). For example, *burner* is a stem that underlies at least one paradigm, the plural *burners* with the inflectional affix *-s*. An inflectional affix does not change the class of the stem. We focus on stems (root or root plus derivational affix) and inflectional affixes, but we do not consider derivational affixes individually.

Bauer (2003:202) writes that "stem traditionally refers to that morpho-logical unit to which inflectional affixes are added, so that a stem is a subtype of base" and "we can call anything we attach affixes to, whether it is just a root or something bigger than a root, a base" (p. 13). He defines an inflectional affix as "one which produces a new word-form of a lexeme from a base" (p. 14). The stem *show* has paradigms *shows* (3rd person singular, simple present), *showed* (past tense) and *showing* (participle) with inflectional affixes *-s*, *-ed*, and *-ing*.

The intelligibility of a written text may also be influenced by morphological differences. For example, the word *hand* has the same spelling in English (*hand*), Dutch (*hand*), and German (*Hand*). However, the plural forms are *hands*, *handen*, and *Hände*, respectively. While the stems are (almost) the same and hence easy to understand for speakers of each language who are not familiar with the other two, it may be more difficult for a native speaker of English to understand the Dutch and German plural forms, as the plural affixes in these languages differ from the one in English.

Yet another factor is syntax. A great number of studies report that small syntactic differences or ambiguities can slow down the readers' or listeners' brain responses (Frazier & Rayner 1982, Ferreira & Henderson 1991, Osterhout & Holcomb 1992, Joseph & Liversedge 2013). However, in most of these studies, participants were presented with ambiguous or ungrammatical sentences in their native language. In line with these findings, Hilton et al. (2013) showed that when readers are presented with a text in a closely related language, and the word order in the sentences in the text differs from the order in which they would have been written in the reader's native language, the reader has more difficulty understanding the text.

Our goal in this study was to determine to what extent each of the aforementioned factors—lexical, orthographic, morphological, and syntactic—affect the understanding of a text written in a language closely related to the native language of the reader. As far as we know, to date there has been no quantificational study that examined this issue, in any case, not on the basis of a sufficiently large sample. This paper reports partial results of a larger study that aims to find out the exact "weights" of the linguistic factors that play a role in the mutual intelligibility of closely related languages. Intelligibility scores of written languages are obtained in the course of a large-scale web-based experiment. By means of a multiple regression model we estimate the extent to which the linguistic factors explain the intelligibility scores obtained in this web-based experiment.

In this paper, we focus on orthographic distances as a potential explanatory factor of intelligibility scores. In the simplest case, we would calculate orthographic distances between languages on the basis of whole words. Above we mentioned the distinction between stems and affixes. Affixes largely represent morphological information, for example, telling the reader whether a word is in its plural or diminutive form. The information represented by stems and affixes is not the same. Hawkins & Cutler (1988:306) define the difference as follows:

> [T]he stem favors the most salient initial position of a word, and the affix the less salient end position, because in the compositional process of determining the entire meaning of a word from its parts, the stem has computational priority over the affix.

They motivate this as follows:

> Consider, for example, *sad* + *ness*. We can paraphrase the meaning of *sad* as 'having an unhappy state of mind', and that of *-ness* as 'the abstract quality of X', where X is the thing that *-ness* combines with, much as a function category applies to an argument category within a categorial grammar to make a derived expression. The effect of the suffix cannot be determined without knowing what stem it has combined with. (p. 306–307)

Hawkins & Cutler (1988) give an extensive list of studies providing evidence that "inflected words do not have lexical representation independent of their base form, and that base word and inflection are separated in language processing" (p. 301). We mention a few of them. Stanners et al. (1979) and Fowler et al. (1985) found that regular inflected forms (for example, *pours*) show a repetition priming effect on their base words (for example, *pour*) as strong as that of the base word itself. Fowler et al. (1985) showed that pretraining with an inflectional variant (for example, *sees*) significantly facilitates later learning of a word (for example, *seen*) compared to no pretraining or pretraining with a word having as much visual similarity to the target word as the morphological relative (for example, *seed*). Plural morphemes tend to get detached in memory representations (van der Molen & Morton 1979). Inflectional suffixes of all kinds tend to be overlooked in script scanning tasks (Drewnowski & Healy 1980, Smith & Sterling 1982). Jarvella & Meijers (1983) primed target verbs with differently inflected forms of the same stem and with similarly inflected forms of different stems. Subjects performed same-different stem tasks significantly faster than same-different inflection tasks. Therefore, we would expect that stem differences play a more significant role in intelligibility than affix differences.

As mentioned above, orthographic variation among languages represents partly variation in pronunciation. If languages within one particular language group—for example, Germanic—originate from a common root, that is, from a proto-language, the diversification of the pronunciation of the stems does not necessarily run (completely) parallel to the diversification of the pronunciation of the affixes. Venezky (2004:146) writes:

> Since most inflectional endings and many derivational ones do not undergo as extensive phonological change as root morphemes, this principle [of constancy] appears to apply primarily to the latter. For English, the main exceptions are the various <s> inflections (plural, possessive, contraction) and the past tense <d>.

Therefore, we expect that stem variation among languages does not correlate with orthographic variation in inflectional affixes. We also expect affix distances between languages within a language group to be smaller than stem distances. In the process of language change, affix reduction leads to more uniformity; stems are usually not reduced. In addition, a set of sounds that make up inflectional affixes is much smaller than that for stems, so one and the same diachronic or synchronic phonological rule would lead to more diversification in stems than in affixes. An example is final obstruent devoicing or terminal devoicing in German, which affects inflectional affixes only marginally because there is only one inflectional affix with a final voiced obstruent, *-end* for the present participle (compare Brockhaus 1995).

Note also that the principle of morphological constancy is traditionally formulated for stems only, that is, where it is needed most. According to this principle, the spelling of a morpheme remains the same despite pronunciation changes that may occur when the morpheme is combined with others. For example, the English word *health* retains the *ea* spelling of its base form, *heal*, even though the vowel of *health*, /ɛ/, differs from the vowel of *heal*, /i/ (Bourassa & Treiman 2008). The principle of constancy indirectly supports our hypothesis that affix distances between languages within a language group are smaller than stem distances.

In this study, we investigate whether a distinction needs to be made between orthographic stem distances and orthographic affix distances as explanatory factors in the written intelligibility model. We test the following hypotheses:

1. Orthographic stem variation among languages does not correlate with orthographic variation in inflectional affixes.

2. Orthographic stem variation among languages is larger than orthographic variation in inflectional affixes.

For the purposes of this paper, a stem can be a root, a compound, or a derivational complex. For example, the English word *hands* contains the stem *hand*, to which the suffix *-s* is attached to form the plural form. An example of a compound is *motorman*, which contains the nouns *motor* and *man*.[1] A derivational complex is a word formed by derivation, that is, the process by which a new word is created on the basis of an existing one. An example of a derivational complex is *friendship*, which is derived from *friend.* Compounds and derivational complexes can be inflected similarly to stems. In this study, inflectional affixes are usually suffixes (as, for example, in Dutch *regel**s*** versus German *Regel**n***) and in a few cases prefixes (as, for example, in Dutch ***ge**zien*).

We focused on the Germanic language group, and in particular on Danish, Dutch, English, German, and Swedish. We tested the two hypotheses on these languages as a group, and we also broke them up into pairs. When the hypotheses were tested on the languages as a group, the aggregated stem differences among the five languages were compared to the corresponding aggregated affix distances. When the

[1] In many English noun-noun compounds, the nouns are separated by a space, for example, *air force*, *bus driver*, etc.

hypotheses were tested on language pairs, the individual stem distances for the word pairs were compared to the corresponding individual affix distances for each of the language pairs.

This paper focuses exclusively on crosslinguistic variation. We are aware of the fact that intralinguistic variation usually leads to crosslinguistic diversification, that is, enhanced crosslinguistic distance. Studying intralinguistic variation and the relation between intralinguistic variation and crosslinguistic variation is a subject for further research. In section 2, we give a brief overview of related research. Section 3 describes the data source. In section 4, we describe the way in which orthographic distances were measured for stems and affixes. The results of the distance measurements are presented in section 5. In section 6, each of the hypotheses is tested. Finally, some general conclusions are drawn in section 7.

2. Previous Research.

We do not know of any studies that measured and compared orthographic stem distances between languages. However, several studies have measured orthographic distances on the basis of whole words. We mention some examples in section 2.1. Likewise, we do not know of any studies that measured and compared orthographic affix distances between languages. However, as mentioned in section 1, variation in affixes represents morphological distances. Therefore, in section 2.2 we focus on some studies that measure morphological distances between language varieties.

2.1. Measuring Orthographic Stem Distances.

Baroni et al. (2002) presented an algorithm that, by taking a raw corpus as its input, produces a ranked list of morphologically related pairs as its output. The algorithm finds morphologically related pairs by looking at the degree of orthographic and semantic similarity between words from the input corpus. Experiments with German and English inputs gave encouraging results. In this study, orthographic similarity is calculated using EDIT DISTANCE, which is also known as LEVENSHTEIN DISTANCE (Levenshtein 1966). The Levenshtein distance between two strings is calculated as the “cost” of the total set of insertions, deletions, and substitutions needed to transform one string into another (Kruskal 1999). The algorithm calculates orthographic distances by finding the minimum number of letters that need to be inserted, deleted, or replaced when changing the spelling of a word in one language into the spelling of the corresponding word in another language (see section 4).

Van Bezooijen & Gooskens (2005) considered orthography as an explanatory factor of intelligibility between Afrikaans and Dutch, and between Frisian and Dutch. They calculated orthographic distances by means of Levenshtein distance. The authors found that the orthographic distances of cognates that are related—directly or via a synonym—are much smaller between Afrikaans and Dutch than between Frisian and Dutch.

Zulu et al. (2008) measured orthographic distances between 11 South African languages. Levenshtein distances were calculated using existing parallel orthographic word spellings in sets of 50 and 144 words from each of the 11 official languages of South Africa. These data were manually collected from various multilingual dictionaries and online resources. The authors concluded that statistical methods based solely on orthographic transcriptions are able to provide useful objective measures of language similarities.

Doetjes & Gooskens (2009) studied the role of orthography in the mutual intelligibility of Danish and Swedish spoken languages. They measured phonetic distances between the languages and took into consideration the help that the listeners can receive from the orthography when listening to the neighboring language. Both phonetic and orthographic distances were measured by means of Levenshtein distance. The authors concluded that Danish listeners indeed seemed to make use of the additional information that the orthography can provide.

2.2. Affix Distances.

A simple way of measuring morphological distances between language varieties is counting the number of differences. Let us compare English morphology with Dutch morphology on the basis of a small sample of seven words in table 1 (see below). The affixes in the words appear in bold. We find that five out of seven words have different affixes in English and Dutch. The distance can be calculated as 5/7×100=71.4%. Reversely, the similarity can be calculated as 2/7×100=28.6%.

English	Dutch	Difference	Similarity
house**s**	huiz**en**	1	0
sheep	schap**en**	1	0
clock**s**	klokk**en**	1	0
apple**s**	appel**s**	0	1
seen	**ge**zien	1	0
calv**es**	kalv**eren**	1	0
ox**en**	oss**en**	0	1
		5	2

Table 1. Comparing English and Dutch morphology.

Séguy (1971, 1973) and Goebl (1982, 1984, 1993) measured distances between local dialects at several linguistic levels in a similar way, namely, by simply counting the number of differences (Séguy) or similarities (Goebl) between two local dialects, considering a set of item pairs. Séguy focused on 154 local dialects, the data for which appeared in the *Atlas linguistique de la Gascogne*. Among other linguistic levels, he considered morphosyntax and verb morphology. Goebl analyzed *l'Atlas Linguistique de l'Italie et de la Suisse Méridionale* (AIS), compiled by Karl Jaberg and Jakob Jud in the first quarter of the 20th century. He measured lexical and morphosyntactic similarities.

Heeringa et al. (2009) used the same methodology in order to measure morphological distances between Dutch Low Saxon dialect varieties. They used data from the *Morphological Atlas of Dutch Dialects* (De Schutter et al. 2005, Goeman et al. 2009). Morphological dialect variation in the following domains was considered: plural substantives, diminutives, possessive pronouns, verbs, participle pre-fixes, and verb stem alternations. The authors report that their data could be divided into four groups. This classification differs from the one on Daan's map (Daan & Blok 1969), the most recent traditional Dutch dialect map, which is based on the little-arrow method. On the map, places in which, according to the speakers, (nearly) the same dialects are spoken are connected by arrows. In that way, white strips arise where there are no arrows; these are the dialect borders. Daan's map shows nine groups.

In all of the examples mentioned above, categorical data were used. In order to find the right categories, historical knowledge may be required. In the dialect of Dordrecht, for instance, one can find the plural form *huiz**e*** 'houses'. The plural suffix -***e*** may be a realization of either the plural suffix *-en* or *-er*. By using historical knowledge we are able to decide whether the plural suffix *-e* is a reduction of *-en* or *-er* and thus whether it belongs to the category of words pluralized with the suffix *-en* or *-er.* In this case, the suffix is most likely a reduced form of *-en,* since the plural suffix *-er* is more commonly found close to the Dutch/German border.

Another example is the prefix in the past participle form of the verb *werken* 'to work'. In dialects related to Standard Dutch it is ***ge**werkt*, whereas in other Dutch dialects it is ***e**werkt* or *werkt*. In order to answer the question whether the prefixes *ge*- and *e*- are separate categories or not, one needs historical knowledge once again. The two examples show that determining morphological categories is not always easy.

It is our aim in this paper to model intelligibility. Since most language speakers are nonlinguists and, therefore, lack theoretical knowledge of morphology, we assume that intelligibility of affixes is not necessarily determined by historical classes motivated morphologically. Besides, by confining affix variation to a restricted set of classes we lose information that may play a role in intelligibility. For example, the double plural suffix in the English word *child**ren*** and the double plural suffix in the Dutch word *kind**eren*** historically belong to the same class. English *-re* is *-er* that underwent metathesis in English but not in Dutch, as seen in *kind**er**en*. English *-n* and Dutch *-en* have been added after the original plural suffixes (that is, *-re* and *-er*, respectively) had become unproductive. A Dutch nonlinguist, however, would not be aware of the fact that the English and Dutch suffix belong to the same class; rather she or he would perceive the English suffix *-ren* simply as a reduction of the Dutch suffix *-eren*.

A better way of modeling perceptual distances is using the Levenshtein distance. This algorithm offers the most economical way of changing one string of sounds (for example, a word in a target language) into another (for example, the corresponding cognate word in the assumed reader's native language). This algorithm enables one to measure how much effort the reader needs to exert to understand the word in the target language. In order to find the most economical way of changing one string of sounds into another, the two words in question are aligned. When comparing the two suffixes, *-eren* and *-ren*, the algorithm finds one deletion, as shown in table 2.

Dutch	e	r	e	n
English		r	e	n
	deletion	match	match	match

Table 2. Levenshtein distance: *-ren* versus *-eren*.

We assume that this approach better reflects the distance perceived by a reader than the one that relies on a reader's knowledge of historical facts, as only a minority of readers would have such knowledge.

Another example is English *apostroph**es*** versus Dutch *apostrof**s***. The suffixes *-es* and *-s* belong to the same class, but when reading the words, the reader would likely notice that the English suffix has an extra *e* compared to the Dutch suffix. The Levenshtein distance method takes this difference into account. The Levenshtein distance would align the suffixes, as shown in table 3.

English	e	s
Dutch		s
	deletion	match

Table 3. Levenshtein distance: *-es* versus *-s*.

In this paper, we use Levenshtein distance. Compared to earlier studies based on categorical data, our approach does not require historical knowledge in order to find the right categories, and it better reflects the distance perceived by a reader.

3. Data Source.

3.1. Selection and Alignment of Texts.

The basis of our analyses is a set of four English texts at the B1/B2 level as defined by the Common European Framework of Reference for

Languages (CEF).[2] The texts were used as preparation exercises for the Preliminary English Test (PET). The diploma is offered by University of Cambridge ESOL Examinations, England. The texts used in the study were obtained at englishaula.com.[3]

Each text was translated from English into Dutch, Danish, German, and Swedish by native speakers of those languages. The translations were subsequently proofread by two other native speakers. The number of words for each text in each language is given in table 4. There are no significant differences in the number of words across the five languages. In section 5.1, we show that our data set is sufficiently large for an analysis presented in this paper.

Text	Danish	Dutch	English	German	Swedish	Mean
Child Athletes	213	241	223	200	211	217.6
Catching a Cold	219	217	216	207	205	212.8
Driving in Winter	205	217	211	196	189	203.6
Riding a Bike	201	212	223	195	191	204.4
	838	887	873	798	796	838.4

Table 4. Number of words for each text, in each language.

The texts were then arranged in a table with five columns, one per language. Each column contains (the translation of) the four texts, where words are found below each other, each cell containing one word (see table 7). Generally, words are considered separate entities when they are separated by spaces. In case of compounds and verbs, however, groups of words in one language may be aligned with individual words in other languages. Some examples are given in table 5.

[2] See http://www.examenglish.com/CEFR/cefr.php (accessed on September 10, 2014).

[3] Website accessed on October 9, 2011.

English	Danish	Dutch	German	Swedish
school work	lektier	schoolwerk	Hausaufgaben	skolarbete
allow	tillade	toestemming geven	erlauben	tillåtelse ge
elderly	ældre	ouderen	alt[e] Mensch[en]	äldre människor
to stop	at stoppe	stoppen	anzuhalten	att stanna

Table 5. Alignment of compounds and multiword compound verbs.

When a group of words is found in one cell, we changed the order of the words to optimize the matching of a form in one language with the corresponding form in another language by the Levenshtein distance (see section 4.1). This is, of course, a somewhat artificial approach, since in the actual texts these words were presented to the reader in the original order. However, we assume that the Levenshtein distance obtained from the optimized order reflects the readers' effort more realistically. Examples are given in table 6. The second row contains the forms as they appear in the original text. In the third row, the word order for English, Danish, Dutch, and German is changed, so that the forms can be optimally matched with the Swedish form when calculating Levenshtein distance.

Order	English	Danish	Dutch	German	Swedish
original	money for training	peng[e] til træning	geld voor de training	Geld für das Training	träningsbudget
optimized	training money for	træning peng[e] til	training geld voor de	Training Geld für das	träningsbudget

Table 6. Optimalization of word order for cells with multiple words.

Table 7 shows the first part of the table and contains the first part of the first sentence of the text "Child Athletes". In the table, affixes appear in square brackets, which mark them as affixes for the algorithm and hence enable us to measure orthographic stem and affix distances separately (see section 4.3).

	Danish	Dutch	English	German	Swedish
1	Forældr[e]	Ouder[s]	Parent[s]	Elter[n]	Föräldr[ar]
2	hvi[s]	wie[ns]	who[se]	der[en]	
3					till
4	børn	kind[eren]	child[ren]	Kind[er]	barn
5					som
6	vis[er]	ton[en]	show		
7				hab[en]	har
8	en		a	ein	ett
9	særlig	special[e]	special	besonder[es]	särskilt
10	interesse	belangstelling	interest	Interesse	intresse
11	inden		in	an	i
12	for	voor			
13	en	een	a	ein[er]	en
14	bestem[t]	bepaal[de]	particular	bestimm[ten]	viss
15	sportgren	sport	sport	Sportart	sport

Table 7. The first part of the table
with the texts in the five languages aligned.

The complete table comprises the four texts, and is a large table that consists of 1259 rows. The table is stored as Microsoft Excel table.

3.2. Number of Word Pairs.

Table 6 contains several empty cells. They are found throughout the table, since not every word appears in every language. For example, the English determiner *a* in the eighth row is not found in the Dutch column, since Dutch does not require a determiner in that position. The English word *show* has equivalents in Danish and Dutch but not in German and

Swedish. The sentences translate less literally into German and Swedish, and, therefore, one finds the words *haben* and *har*, respectively, which generally have a different meaning. Therefore, these words appear in a separate row.

To calculate orthographic distances, our measurements are based on pairwise word comparisons within every pair of languages. Aligned data from five languages yield ten language pairs. However, due to the empty cells, not every word has a match. For example, the English word *show* (sixth row in table 7) does not have a German counterpart, and the German word *haben* does not have an English counterpart. If there is a pair, it is taken into account.

We also restrict our analysis to word pairs, whose members belong to the same word class. For example, in the text "Catching a Cold" the English phrase *the best answer* is translated into Dutch as *het beste*. Being originally an adjective, *best* is inflected as a noun in the Dutch sentence. It is tempting to match English *best* with Dutch *beste*, but since the English word is an adjective and the Dutch word is a noun, we do not match them. In this paper, we focus specifically on affix variation, but it would not be fair to compare affixes of words that belong to different word classes. Therefore, word pairs such as the English/Dutch pair *best*/*beste* were not included in the analysis. Table 8 shows the number of word pairs for each language pair.

	Danish	Dutch	English	German	Swedish
Danish		711	716	609	617
Dutch			742	643	610
English				618	595
German					534
Swedish					

Table 8. The number of word pairs for each language pair.

For each language pair a large number of word pairs is found, varying from 534 (German versus Swedish) to 742 (Dutch versus English).

3.3. Number of Cognate Pairs.

In our orthographic measurements, we distinguish between stems and affixes. When reading a text in a foreign language, the reader would match the words she or he is reading with cognate words in his or her native language on the basis of the stems. For example, when a native speaker of English reads the Dutch word *handen*, she or he might be able to match it with the English word *hands* on the basis of the joint stem *hand*, although the affixes in the two languages differ.

In contrast, when a native speaker of English reads the Dutch word *kevers*, she or he would not be expected to match it with the English word *beetles*, despite the fact that both the Dutch and the English word—having the same meaning—have the same plural affix *-s*. When both the stem and the affix are different—for example, English *ducks* versus Dutch *eenden*—it is less likely that an English reader would relate the Dutch plural suffix *-en* to the English plural suffix *-s*. Considering this, we decided to calculate orthographic distances on the basis of word pairs whose members are cognates.

Focusing on cognate pairs only further reduces the number of analyzed word pairs per language pair. For example, the fourth row in table 7 contains the Danish word *børn* and the Dutch word *kinderen*. Since the Danish and Dutch words are not cognates, this word pair was not included in the analyses. In contrast, Dutch *kinderen* and German *Kinder* are cognates, and the orthographic distances were measured for this word pair.

Table 9 shows the number of cognate pairs per language pair. The numbers are much smaller than the numbers in table 8 and vary from 203 (English versus Swedish) to 435 (Danish versus Swedish).

	Danish	Dutch	English	German	Swedish
Danish		282	273	256	435
Dutch			380	403	225
English				280	203
German					213
Swedish					

Table 9. The number of cognate word pairs for each language pair.

Despite the smaller numbers and apparently large differences, consistent results can still be obtained on the basis of this data set. In section 5.1, we show that the number of cognate pairs for each of the language pairs is sufficiently large to yield consistent results for all of the measurements reported in this paper.

Table 10 shows the percentage of word pairs whose members are cognates. The distance between any two languages obtained on the basis of cognate word pairs is divided by the number of cognate pairs.

	Danish	Dutch	English	German	Swedish
Danish		39.7	38.1	42.0	70.5
Dutch			51.2	62.7	36.9
English				45.3	34.1
German					39.9
Swedish					

Table 10. The percentage of word pairs, whose members are cognates.

The percentages reflect lexical similarity. We conclude that Danish and Swedish are lexically most similar as they share 70.5% of the words in our corpus. However, lexical similarity is not the topic of this paper.

4. Measuring Orthographic Distances.

4.1. Levenshtein Distance.

Orthographic distances between two words are measured with the aid of the Levenshtein distance metric (Levenshtein 1966). Recall from section 2.1 that the Levenshtein distance between two strings is calculated as the "cost" of the total set of insertions, deletions, and substitutions needed to transform one string into another (Kruskal 1999). In our case, the strings to be compared are orthographic transcriptions of words. In table 11 (see below), we illustrate this algorithm by transforming the English word *interest* into the Swedish word *intresse.* This represents a native speaker of Swedish reading English or a native speaker of English reading Swedish, and trying to map the target word to its cognate in his or her native language. The two words are very similar, but in the fourth slot an *e* is deleted, in the eight slot a *t* is replaced by an *s*, and in the ninth slot an *e* is inserted. Therefore, the Levenshtein distance equals three operations. The alignment has nine slots. We calculate the normalized Levenshtein distance as 3/9x100=33%. Many different sequences of operations transform *interest* to *intresse,* but the Levenshtein distance always gives the cost of the cheapest mapping.

	1	2	3	4	5	6	7	8	9
English	i	n	t	e	r	e	s	t	
Swedish	i	n	t		r	e	s	s	e
				1				1	1

Table 11. Levenshtein distance: *interest* versus *intresse*.

For each character, we distinguish between a base and a diacritic. For example, the base of *é* is *e*, and the diacritic is the acute accent. Two characters may differ in the base and/or in their diacritics. We weigh differences in the base as 1. For example, *a* versus *e*, *p* versus *b*. If two characters have the same base but different diacritics, we weigh this as 0.3, for example, *e* versus *é* and *è* versus *é*. We admit that this method is not based on empirical measurements and may appear arbitrary, but our choice is motivated by the idea that diacritical differences should weigh significantly less than base differences, since differences in the base usually confuse the reader to a much greater extent than diacritical differences. When the bases are different, the weight is 1, regardless of whether there are diacritical differences, since differences or similarities between diacritics are meaningless when the corresponding bases are different. Insertions and deletions are weighed as 1.

In German, the first letter of every noun is capitalized. If the Levenshtein distance counts lower and upper case letters as different characters, German would become disproportionally distant from the other Germanic languages, since readers do not really distinguish between lower case and upper case letters. Therefore, we do not distinguish between lower case and upper case letters. For example, the distance between the English word *problem* and the German word *Problem* is zero. We also ensure that the minimum cost is based on an alignment in which a vowel matches with a vowel, and a consonant matches with a consonant.

4.2. Aggregated Stem and Affix Distance.

For each language pair, we calculated the aggregated orthographic distance by calculating the average of the normalized Levenshtein distances within the word pairs considered for that language pair. A small sample is given in table 12, which shows the comparison of Dutch and German on the basis of a set of five words. The fourth column shows the Levenshtein distances. When the numbers in the fourth column are divided by the numbers of slots in the alignment (fifth column), the normalized Levenshtein distance is obtained (sixth column).

	Dutch	German	Levenshtein distance	Number of slots in the alignment	Normalized Levenshtein distance
1	helpen	helfen	1	6	0.17
2	monden	Münder	2	6	0.33
3	regels	Regeln	1	6	0.17
4	bakken	backen	1	6	0.17
5	gezegd	gesagt	3	6	0.50
					0.27

Table 12. The aggregated distance
is the average distance (0.27 or 27%).

We specifically focus on stem and affix distances. When calculating the stem distance, we considered the stem of the words only. This is shown in table 13 (see below), where the aggregated stem distance is 28% (the stems appear in bold).

Affix distances are found by considering affixes only. This is illustrated in table 14 (see below). The words in the fifth word pair each contain two affixes: a prefix *ge-* and a suffix *-d* (Dutch) or *-t* (German). The affixes were concatenated to *ged* (Dutch) and *get* (German), and the Levenshtein distance was computed between the two concatenations.

Sometimes words do not have affixes. For example, the plural form of the English word *sheep* is *sheep.* When this word is transformed into, for example, Dutch *schapen*, which has plural suffix *-en*, the Levenshtein distance equals two (two insertions).

	Dutch	German	Levenshtein distance	Number of slots in the alignment	Normalized Levenshtein distance
1	**help**+en	**helf**+en	1	4	0.25
2	**mond**+en	**Münd**+er	1	4	0.25
3	**regel**+s	**Regel**+n	0	5	0
4	**bakk**+en	**back**+en	1	4	0.25
5	ge+**zeg**+d	ge+**sag**+t	2	3	0.67
					0.28

Table 13. The aggregated stem distance is the average stem distance (0.28 or 28%).

	Dutch	German	Levenshtein distance	Number of slots in the alignment	Normalized Levenshtein distance
1	help+**en**	helf+**en**	0	2	0
2	mond+**en**	Münd+**er**	1	2	0.50
3	regel+**s**	Regel+**n**	1	1	1
4	bakk+**en**	back+**en**	0	2	0
5	**ge**+zeg+**d**	**ge**+sag+**t**	1	3	0.33
					0.37

Table 14. The aggregated affix distance is the average affix distance (0.37 or 37%).

These examples show how orthographic language variation in whole words, stems, and affixes can easily be quantified by using Levenshtein distance.

5. Results: Stem and Affix Distances Between Languages.

5.1. Consistency.

In section 4.2, we explained that the aggregated orthographic distance within a language pair is calculated as the average of the normalized Levenshtein distances within the word pairs considered for that language pair. A set of five languages yields 10 language pairs: (5×5–5)/2=10. In table 8, the number of cognate pairs per language pair is shown, which varies from 203 to 435. Tables 15 and 16 present the aggregated orthographic stem and affix distances, respectively.

	Danish	Dutch	English	German	Swedish
Danish		44.5	51.2	47.6	24.1
Dutch			53.3	44.6	45.1
English				59.6	55.0
German					48.1
Swedish					

Table 15. Orthographic stem distances in percentages between Germanic languages measured with Levenshtein distance.

	Danish	Dutch	English	German	Swedish
Danish		32.5	28.5	34.8	22.2
Dutch			20.4	18.1	38.4
English				29.7	26.0
German					40.7
Swedish					

Table 16. Orthographic affix distances in percentages between Germanic languages measured with Levenshtein distance.

For each of the measurements we checked whether the number of words is a sufficient basis for a reliable analysis. We calculated the Cronbach's α value for both the stem distances and the affix distances. Cronbach's α was first introduced in Cronbach 1951 as a coefficient of consistency that can be described as a function of the number of linguistic variables and the average inter-correlation value among the variables. Cronbach's α values range between zero and one. Higher values indicate more reliability. As a rule of thumb, values higher than 0.7 are considered sufficient for consistent results in social sciences (Nunnally 1978). We found Cronbach's $\alpha=0.85$ for the stem distances, and Cronbach's $\alpha=0.95$ for the affix distances, showing that our data are sufficiently consistent.[4]

5.2. Beam Maps and Cluster Analysis.

The distances are visualized in figure 1 by means of so-called beam maps (Inoue 1996). On the maps, the countries are represented by their geographic centers. The geographical centers of the countries are taken from the NGA GEOnet Names Server (GNS).[5] The centers are connected by lines, or "beams", with darker beams connecting orthographically close languages and lighter beams more remote ones. Beam maps were introduced by Goebl (1993). On his maps, only neighboring locations were connected. We use the Groningen-style network maps, where every location can, in principle, link to any other location in the network.[6] In each of the beam maps in this section, the smallest distance is represented by an almost black line, and the largest distance is represented by a white line. On a white background, however, white lines are not visible.

Figure 1 shows a relatively small orthographic stem distance of 24.1% between Danish and Swedish. Orthographic stem distances between other languages are larger. English is found to be especially

[4] Since in our data set, the number of words varies from language pair to language pair, we did not use the commonly used implementation of Cronbach's α, but instead we used Robust Cronbach's α in the coefficient α-package in *R*, which is implemented by Zhang & Yuan (2013).

[5] See: http://earth-info.nga.mil/gns/html.

[6] We are grateful to Peter Kleiweg, whose RuG/L04 package was used to create the beam maps shown in this paper. Examples of the maps developed by Peter Kleiweg can be found in Heeringa 2004.

distant from all of the other varieties. The beam map obtained on the basis of the orthographic affix distances shows a different picture. Relatively small distances are found between Dutch and German (18.1%), Dutch and English (20.4%), and Danish and Swedish (22.2%). English is relatively closer to the continental Germanic varieties than on the beam map obtained on the basis of the orthographic stem distances.

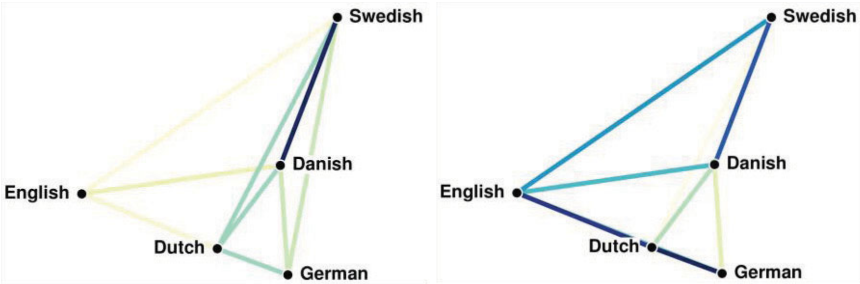

Figure 1. Orthographic stem distances (left)
and affix distances (right) between Germanic languages.

Largest distances are found between varieties not connected by a line. Orthographic stem distances vary from 24.1% (between Danish and Swedish) to 59.6% (between English and German). Orthographic affix distances vary from 18.1% (between Dutch and German) to 40.7% (between German and Swedish).

We applied hierarchical cluster analysis to both the stem and the affix distances. The result is a binary tree structure known as a dendro-gram (Jain & Dubes 1988), in which one branch is for the stem distances and the other for the affix distances. The varieties are the leaves, and the branches reflect the distances between the leaves. As for the cluster method, there exist several options. We used the Unweighted Pair Group Method using Arithmetic averages (UPGMA), since dendrograms generated by this method reflect distances that correlate most strongly with the original Levenshtein distances (r=0.97 for stem distances and r=0.78 for affix distances; see Sokal & Rohlf 1962).

The dendrograms obtained on the basis of stem distances and affix distances are shown in figure 2. Both dendrograms show a North Germanic group, including Danish and Swedish, and a West Germanic group, including Dutch and German. In the stem dendrogram, the smallest distance is found between Danish and Swedish, in the affix dendrogram, Dutch and German are closest. Note also the position of English: In the stem dendrogram, English is apart from all other varieties, but in the affix dendrogram, English is clustered together with the West Germanic varieties.

Figure 2. Dendrograms obtained on the basis of stem distances (left) and affix distances (right).

The two dendrograms in figure 2 also show that the affix distances are smaller than the stem distances. We come back to this in section 6.2.

The affix dendrogram is consistent with the classification commonly found in Indo-European family trees. The classification shown by this kind of trees is based on the criterion of shared innovation. Assuming that all of the languages in this tree descend from Proto-Indo-European, a shared innovation (or, departure from the proto-language) may take place in a single daughter language. This daughter language in turn has diverse daughters of its own, each of which would inherit, and therefore share the same innovation. In those trees, English is found in the West-Germanic group, together with Frisian, Dutch, Afrikaans, Low German, High German, and Yiddish (see, for example, Campbell 2013:176–177). This is not surprising, since English originates from the fusion of closely related dialects now collectively termed Old English and spoken by Germanic settlers, who came from their ancestral region of Angeln, presently known as Schleswig-Holstein (Baugh & Cable 1978).

6. Stem and Affix Distances in Relation to One Another.

6.1. First Hypothesis.

If the Germanic languages emerged from one common root, that is, Proto-Germanic, it is reasonable to assume that diversification of the pronunciation of the Proto-Germanic stems likely does not run (completely) parallel to diversification of the pronunciation of the Proto-Germanic affixes (see discussion in section 1). We then hypothesized that orthographic stem variation among languages does not correlate with orthographic variation in inflectional affixes. In order to test this hypothesis we correlated the orthographic stem distances with the orthographic affix distances. Figure 3 shows a scatter plot in which orthographic affix distances are drawn against the orthographic stem distances. The correlation between the two measures is r=0.15.

For finding the significance of this correlation coefficient we used the Mantel test (Mantel 1967), a widely used method to account for distance correlations. Classical tests rely on the assumption that the correlated objects are independent. However, values in distance matrices are usually correlated in some way, and are not independent (Bonnet & Van de Peer 2002). By using the Mantel test, we found that p=0.36. Therefore our hypothesis is confirmed: Orthographic stem variation among languages does not correlate with orthographic variation in inflectional affixes.

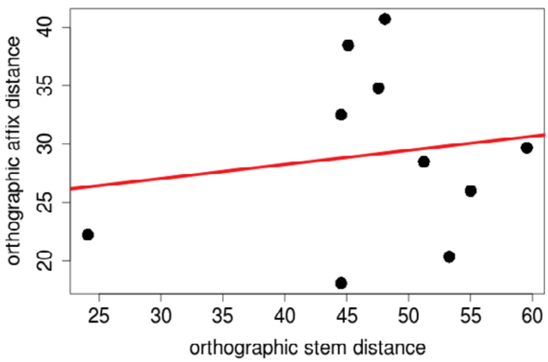

Figure 3. The orthographic affix distances
and the orthographic stem distances.

In figure 3, the two measurements do not correlate. In the scatter plot, the Danish/Swedish pair is found in the lower left corner. It is found distant from the other points. If this pair is excluded, the correlation between stem and affix distances becomes r=-0.24, which is still not significant (p=0.29).

We also calculated correlations per language pair. The number of cognate pairs in each language pair is given in table 8. For each cognate pair, both the orthographic stem distance and the orthographic affix distance were measured. The orthographic stem distances of a language pair are correlated with the corresponding orthographic affix distances. For example, for the Danish/Dutch language pair we have 282 cognate pairs. Therefore, the correlation is calculated between 282 stem distances and 282 affix distances, which is r=0.17. The correlations of each of the language pairs are shown in table 17: ** means p<0.01, *** means p<0.001, **** means p<0.0001. The significant correlations have a small effect size.

	Danish	Dutch	English	German	Swedish
Danish		0.17**	-0.27***	0.03	0.13**
Dutch			-0.18***	-0.16**	0.01
English				-0.07	-0.23**
German					0.03
Swedish					

Table 17. Correlations between orthographic stem distances and orthographic affix distances per language pair.

This table shows either nonsignificant correlations or the significant correlations that have a small effect size according to the guidelines of Cohen 1988. These results confirm our hypothesis that stem and affix distance are not correlated.

6.2. Second Hypothesis.

In section 1, we assumed that the diversification of affixes in Proto-Germanic proceeded slower than the diversification of stems. Assuming that all present-day Germanic languages originate from Proto-Germanic, we expect the affix distances between those languages to be smaller than stem distances. In this section, we test the hypothesis that orthographic stem distances among languages are larger than orthographic affix distances.

The stem and affix distances are shown in figure 4. We compared stem and affix distances using a paired-samples t-test and found that stem distances are significantly larger than affix distances (t=5.10, df=9, p<0.001, r=0.86; large effect size).

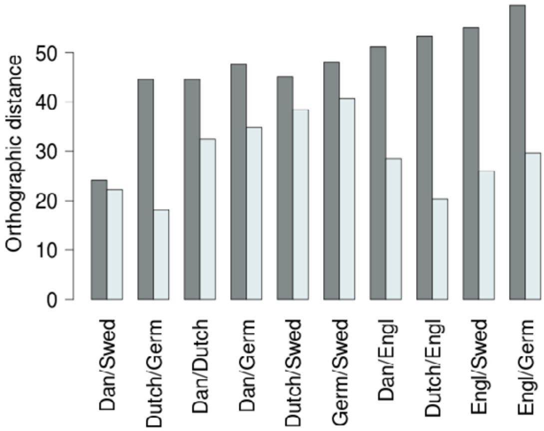

Figure 4. Stem distances (dark)
and affix distances (light) per language pair.

The differences between orthographic stem distances and affix distances are shown in figure 5 and table 18. The smallest distance is found for the Danish/Swedish language pair.

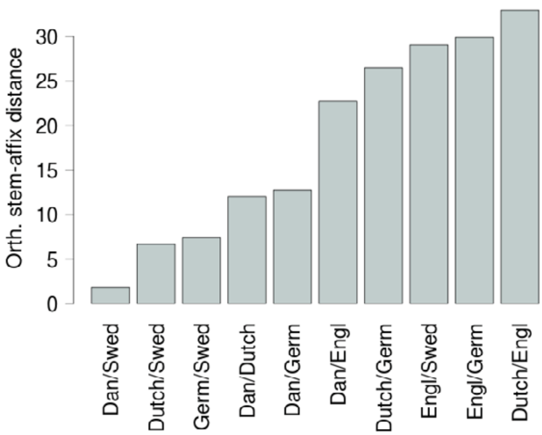

Figure 5. Differences between stem distances and affix distances, ranging from 1.9% (Danish/Swedish) to 32.9% (Dutch/English).

Applying a paired-samples *t*-test, we tested, for each language pair, whether the stem distances are larger than the affix distances. The *p*-

values are given in table 18. For all language pairs, stem distances are almost always significantly larger than the affix distances, which further supports the hypothesis that stem distances are larger than affix distances.

	Danish	Dutch	English	German	Swedish
Danish		12.0****	22.7****	12.8****	1.9
Dutch			32.9****	26.5****	6.7*
English				29.9****	29.0****
German					7.4*
Swedish					

Table 18. Differences between stem distances and affix distances (* means $p<0.05$, **** means $p<0.0001$).

For the Danish/Swedish language pair we do not find stem distances being significantly larger than affix distances. Figure 5 shows that the difference between the aggregated stem distance and the aggregated affix distance is smallest for the Danish/Swedish language pair, and in figure 4 it can be seen that the stem distance within this language pair is smallest. This raises the question whether there is a correlation between the differences between stem and affix distances (as shown in table 18/figure 5) and stem distances (as shown in table 15/figure 4). Do smaller stem distances within a language pair correspond with smaller differences between stem and affix distances? Indeed, we found a significant correlation ($r=0.74$, $p<0.05$).

We also considered the correlation between stem distances and the affix/stem distance ratio. To obtain the affix/stem distance ratio we divide an affix distance by its corresponding stem distance. The ratios are shown in figure 6. We found a significant correlation ($r=0.66$, $p<0.05$). This means that affix distances become proportionally larger when stem distances become smaller.

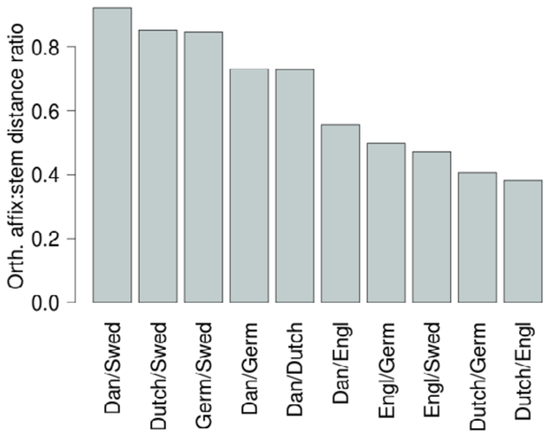

Figure 6. Affix/stem distance ratios,
ranging from 0.92 (Danish/Swedish) to 0.39 (Dutch/English).

To summarize, we found that stem distances are significantly larger than affix distances, and that smaller stem distances within a language pair correspond with smaller differences between stem and affix distances and with proportionally larger affix distances.

7. Conclusion.

We have conducted a large-scale web-based experiment in order to obtain intelligibility scores of written and spoken languages. This experiment enabled us to find the extent to which several linguistic levels play a role in the intelligibility of closely related languages. One of the levels is orthography. In this study, we investigated whether a distinction needs to be made between orthographic stem distances and orthographic affix distances as explanatory factors in the written intelligibility model.

First, we tested the hypothesis that orthographic stem variation among languages does not correlate with orthographic variation in inflectional affixes. We found this hypothesis true when considering the aggregated stem and affix distances between the languages. We also correlated the stem and affix distances within the cognate pairs in each language pair. We found low correlations, the lowest of them being not significant. The results look arbitrary, both positive and negative correlations are found, ranging from -0.27 to 0.17.

Second, we tested the hypothesis that orthographic stem variation among languages is larger than orthographic variation in inflectional affixes. This hypothesis was also found to be true. This is consistent with a study of Heeringa & Hinskens (2014), who studied Dutch dialect change at the lexical level, the level of the sound components, and the morphological level. They found that the morphological level has been affected the least, and therefore is the most stable level. However, our results are based on orthography, and, as mentioned in section 1, ortho-graphic differences are the result of differences in spelling conventions and historical developments of the pronunciation. We have not investigated yet whether spelling differences and pronunciation differences contribute to the same degree to stem distances and affix distances. This may be examined in future research.

When comparing the stem and affix distances within the cognate pairs in each language pair, for nearly each language pair we found the stem distances significantly larger than the affix distances, except for the Danish/Swedish language pair. For this language pair, the aggregated stem distance and affix distance is almost the same (24.1% versus 22.2%). This raises the question whether smaller stem distances correspond with differences between stem and affix distances and with smaller affix/stem distance ratios. In both cases we found a significant correlation. The smaller the stem distance, the smaller the difference between stem and affix distances, and the larger the affix distance relatively to the stem distance.

Having confirmed both of the hypotheses, we conclude that orthographic distances should be split into stem and affix distances, and both, orthographic stem distances and orthographic affix distances should be included in the model aiming to explain mutual written intelligibility of Germanic languages. We wonder whether our results are specific for the Germanic languages. Therefore, in the future we intend to conduct similar analyses for the Romance and Slavic language groups.

REFERENCES

Baroni, Marco, Johannes Matiasek, & Harald Trost. 2002. Unsupervised discovery of morphologically related words based on orthographic and semantic similarity. *Proceedings of the Workshop on Morphological and*

Phonological Learning of ACL/SIGPHON-2002, ed. by Mike Maxwell, 48–57. East Stroudsburg, PA: ACL.

Bauer, Laurie. 2003. *Introducing linguistic morphology*. 2nd edn. Edinburgh: Edinburgh University Press.

Baugh, Albert C., & Thomas Cable. 1978. *History of the English language*. 3rd edn. Englewood Cliffs, N.J.: Prentice-Hall.

Bezooijen, Renée van, & Charlotte Gooskens. 2005. How easy is it for speakers of Dutch to understand spoken and written Frisian and Afrikaans, and why? *Linguistics in the Netherlands 22*, ed. by Jenny Doetjes & Jeroen van de Weijer, 13–24. Amsterdam: John Benjamins.

Bonnet, Eric, & Yves Van de Peer. 2002. zt: a software tool for simple and partial Mantel tests. *Journal of Statistical Software* 7. 1–12. Available at http://www.jstatsoft.org/.

Bourassa, Derrick C., & Rebecca Treiman. 2008. Morphological constancy in spelling: A comparison of children with dyslexia and typically developing children. *Dyslexia* 14. 155–169.

Brockhaus, Wiebke. 1995. *Final devoicing in the phonology of German.* Tübingen: Max Niemeyer Verlag.

Campbell, Lyle. 2013. *Historical linguistics. An introduction*. 3rd edn. Edinburgh: Edinburgh University Press.

Cohen, Jacob. 1988. *Statistical power analysis for the behavioral sciences.* 2nd edn. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cronbach, Lee J. 1951. Coefficient alpha and the internal structure of tests. *Psychometrika* 16. 297–334.

Daan, Jo, & Dirk Peter Blok. 1969. *Van randstad tot landrand (Bijdragen en Mededelingen der Dialectencommissie van de Koninklijke Nederlandse Akademie van Wetenschappen te Amsterdam 37)*. Amsterdam: N.V. Noord-Hollandsche uitgevers maatschappij.

De Schutter, Georges, Boudewijn van den Berg, Ton Goeman, & Thera de Jong. 2005. *Morphological atlas of the Dutch dialects,* vol. I. Amsterdam: Amsterdam University Press.

Doetjes, Gerke, & Charlotte Gooskens. 2009. Skriftsprogets rolle i den dansk-svenske talesprogsforståelse. *Språk och Stil* 19. 105–123.

Drewnowski, Adam, & Alice F. Healy. 1980. Missing *-ing* in reading: Letter detection errors on word endings. *Journal of Verbal Learning and Verbal Behavior* 19. 247–262.

Ferreira, Fernanda, & John M. Henderson. 1991. Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language* 31. 725–745.

Fowler, Carol A., Sherley E. Napps, & Laurie B. Feldman. 1985. Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory and Cognition* 13. 241–255.

Frazier, Lynn, & Keith Rayner. 1982. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology* 14. 178–210.

Goebl, Hans, with assistance of Wolf-Dieter Rase & Hilmar Pudlatz. 1982. *Dialektometrie: Prinzipien und Methoden des Einsatzes der numerischen Taxonomie im Bereich der Dialektgeographie* (*Philosophisch-Historische Klasse Denkschriften 157*). Vienna: Verlag der Österreichischen Akademie der Wissenschaften.

Goebl, Hans, with assistance of Siegfried Selberherr, Wolf-Dieter Rase & Hilmar Pudlatz. 1984. *Dialektometrische Studien. Anhand italoromanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF* (*Beihefte zur Zeitschrift für Romanische Philologie 191, 192, 193*). Tübingen: Max Niemeyer Verlag.

Goebl, Hans. 1993. Probleme und Methoden der Dialektometrie: Geolinguistik in globaler Perspektive. *Proceedings of the International Congress of Dialectologists* 1, ed. by Wolfgang Viereck, 37–81. Stuttgart: Franz Steiner Verlag.

Goeman, Ton, Marc van Oostendorp, Pieter van Reenen, Oele Koornwinder, Boudewijn van den Berg, & Anke van Reenen. 2009. *Morphological atlas of the Dutch dialects,* vol. II. Amsterdam: Amsterdam University Press.

Hawkins, John A., & Anne Cutler. 1988. Psycholinguistic factors in morphological asymmetry. *Explaining language universals*, ed. by John A. Hawkins, 280–317. Oxford: Blackwell.

Heeringa, Wilbert. 2004. *Measuring dialect pronunciation differences using Levenshtein distance*. Groningen, The Netherlands: University of Groningen dissertation.

Heeringa, Wilbert, Martijn Wieling, Boudewijn van den Berg, & John Nerbonne. 2009. A quantitative examination of variation in Dutch Low Saxon morphology. *Low Saxon dialects across borders/Niedersächsiche dialekte über grenzen hinweg* (*ZDL-Beiheft 138*), ed. by Alexandra N. Lenz, Charlotte Gooskens, & Siemon Reker, 195–216. Stuttgart: Franz Steiner Verlag.

Heeringa, Wilbert, & Frans Hinskens. 2014. Convergence between dialect varieties and dialect groups in the Dutch language area. *Aggregating dialectology, typology, and register analysis: Linguistic variation in text and speech*. (*Linguae et Litterae: Publications of the School of Language and Literature, Freiburg Institute for Advanced Studies 28)*, ed. by Benedikt Szmrecsanyi & Bernhard Wälchli, 26–52 and 452–453. Berlin & Boston: De Gruyter.

Hilton, Nanna Haug, Charlotte Gooskens, & Anja Schüppert. 2013. The influence of non-native morphosyntax on the intelligibility of a closely related language. *Lingua* 137. 1–18.

Inoue, Fumio. 1996. Computational dialectology (2). *Area and Culture Studies* 53. 115–134.

Jain, Anil K., & Richard C. Dubes. 1988. *Algorithms for clustering data*. Englewood Cliffs, NJ: Prentice Hall.

Jarvella, Robert J., & Guust Meijers. 1983. Recognizing morphemes in spoken words: Some evidence for a stem-organized mental lexicon. *The process of language understanding*, ed. by Robert J. Jarvella & Giovanni B. Flores d'Arcais. Chichester: John Wiley.

Joseph, Holly S. S. L., & Simon P. Liversedge. 2013. Children's and adults' on-line processing of syntactically ambiguous sentences during reading. *PLoS ONE* 8. 1–13.

Kruskal, Joseph B. 1999. An overview of sequence comparison. *Time warps, string edits, and macromolecules. The theory and practice of sequence comparison*, ed. by David Sankoff & Joseph B. Kruskal, 1–44. 2nd edn. Stanford, CA: CSLI.

Levenshtein, Vladimir. 1966. Binary codes capable of correcting deletions, insertions, and reversals. *Cybernetics and Control Theory* 10. 707–710.

Mantel, Nathan. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27. 209–220.

Matthews, Peter H. 1991. *Morphology*. 2nd edn. Cambridge: Cambridge University Press.

Molen, Hugo van der, & John Morton. 1979. Remembering plurals: Unit of coding and form of coding during serial recall. *Cognition* 7. 35–47.

Nunnally, Jum C. 1978. *Psychometric theory*. 2nd edn. New York, NY: McGraw-Hill.

Osterhout, Lee, & Phillip J. Holcomb. 1992. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language* 31. 785–780.

Séguy, Jean. 1971. La relation entre la distance spatiale et la distance lexicale. *Revue de Linguistique Romane* 35. 335–357.

Séguy, Jean. 1973. La dialectométrie dans l'Atlas linguistique de la Gascogne. *Revue de Linguistique Romane* 37. 1–24.

Smith, Philip T., & Christopher M. Sterling. 1982. Factors affecting the perceived morphological structure of written words. *Journal of Verbal Learning and Verbal Behavior* 21. 704–721.

Sokal, Robert R., & F. James Rohlf. 1962. The comparison of dendrograms by objective methods. *Taxon* 11. 33–40.

Stanners, Robert F., James J. Neiser, William P. Hernon, & Roger Hall. 1979. Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior* 18. 399–412.

Venezky, Richard L. 2004. In search of the perfect orthography. *Written Language and Literacy* 7. 139–163.

Zhang, Zhiyong, & Ke-Hai Yuan. 2013. *Robust Cronbach's alpha with missing and non-normal data*. Software package description in R.

Zulu, Nicholas, Gerrit Botha, & Etienne Barnard. 2008. Orthographic measures of language distances between the official South African languages. *Literator* 29. 185–204.

University of Groningen
Faculty of Arts, Applied Linguistics
P.O. Box 716, 9700AS Groningen
The Netherlands
[w.j.heeringa@rug.nl]
[f.h.e.swarte@rug.nl]
[a.schueppert@rug.nl]
[c.s.gooskens@rug.nl]