

# John Benjamins Publishing Company



This is a contribution from Linguistic Approaches to Bilingualism 10:3  
© 2020. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible only to members (students and faculty) of the author's/s' institute. It is not permitted to post this PDF on the internet, or to share it on sites such as Mendeley, ResearchGate, Academia.edu.

Please see our rights policy on <https://benjamins.com/content/customers/rights>

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: [www.copyright.com](http://www.copyright.com)).

Please contact [rights@benjamins.nl](mailto:rights@benjamins.nl) or consult our website: [www.benjamins.com](http://www.benjamins.com)

# How well can intelligibility of closely related languages in Europe be predicted by linguistic and non-linguistic variables?

Charlotte Gooskens<sup>1</sup> and Vincent J. van Heuven<sup>1,2</sup>

<sup>1</sup>University of Groningen / <sup>2</sup>Pannon Egyetem

We measured mutual intelligibility of 16 closely related spoken languages in Europe. Intelligibility was determined for all 70 language combinations using the same uniform methodology (a cloze test). We analysed the results of 1833 listeners representing the mutual intelligibility between young, educated Europeans from the same 16 countries.

Lexical, phonological, orthographic, morphological and syntactic distances were computed as linguistic variables. We also quantified non-linguistic variables (e.g. exposure, attitudes towards the test languages). Using stepwise regression analysis the importance of linguistic and non-linguistic predictors for the mutual intelligibility in the 70 language pairs was assessed.

Exposure to the test language was the most important variable, overriding all other variables. Then, limiting the analysis to the prediction of inherent intelligibility, we analysed the results for a subset of listeners with no or little previous exposure to the test language. Linguistic distances, especially lexical distance, now explain a substantial part of the variance.

**Keywords:** linguistic distances, intelligibility of closely related languages, non-linguistic factors

## 1. Introduction

In this chapter we report on the results of a large internet-based investigation on the mutual intelligibility of 70 closely related language pairs in Europe. The results may be of interest to language policy makers and language teachers. In Europe, a large number of languages are spoken and there are often communication problems when speakers from different native language backgrounds meet. The default strategy is to use English as a *lingua franca* or to depend on one of the

speakers to have learned the language of the other speaker. The alternative type of communication is that which we study in the present chapter, and which is often referred to as *receptive multilingualism* or RM (Zeevaert, 2004; Ten Thije & Zeevaert, 2007; Braunmüller, 2007).<sup>1</sup> RM is based on the fact that some language pairs are so closely related that interlocutors are able to communicate with one another when each interactant continues to use his (or her) own language, without prior (formal or informal) instruction in the interlocutor's language. The speakers only need to discover that they can profit from their own language when trying to crack the L2 code and that it is not necessary for them to actively acquire new grammatical constructions, words and pronunciation habits. The fact that both participants in a conversation can speak the language they master best, their native language, results in an inherent fairness and equality between the speakers since they both have to make an effort to understand the other language. This makes it an attractive way of communicating and the next step towards an active command of a language will often be small.

RM is widely used by speakers of the three mainland Scandinavian languages, Danish, Swedish and Norwegian (Maurud, 1976; Delsing & Lundin Åkesson, 2005), and may also be a useful strategy for other European language pairs. However, to be able to give advice on whether RM is feasible, we need to know more about the variables that determine the degree of mutual intelligibility. The results would also be interesting from a theoretical perspective. They may allow us to determine under what conditions RM works and what its preconditions and its limits are. This will give us an estimate of how deviant a language can be, and on which linguistic levels, before it ceases to be intelligible to listeners from a related language background.

We distinguish inherent from acquired cross-language intelligibility. The former relies on language features that are available to interlocutors *a priori* because of the close genealogical relationship between L1 and L2, whereas the latter presupposes learning through exposure and instruction. For example, a Dane and a Swede can understand each other even if they have never heard the other language before (inherent intelligibility) while speakers of Dutch and Spanish can only do so if they learned each other's languages because these belong to different language families (acquired intelligibility). We are interested in predicting both kinds of intelligibility. Therefore, we included both linguistic and non-linguistic variables in our analysis. In addition, we look at two sets of data. The first set involves the

---

1. RM is an ambiguous term, which could refer (i) to situations where two interlocutors speak each their language and are still able to communicate because of receptive understanding of the language of their interlocutor or (ii) to attrited grammars in communities and individuals where only comprehension (to some degree) is possible (see e.g. Sherkina-Lieber, 2015).

intelligibility results of a selection of 1833 listeners representing the mutual intelligibility between young, educated Europeans from the same 16 countries where the test languages are spoken. Some of the listeners had learned the test language at school. However, also exposure to a language in daily life can improve intelligibility considerably. We refer to this data set as *acquired intelligibility* (even though it may include inherent intelligibility). Next, we present the data from a sub-group of listeners who had not learned the test language and had received minimal exposure to it. This allows us to investigate how well the listeners understand the test language only on the basis of structural similarities between their own language and the test languages (*inherent intelligibility*).

According to Tang and Van Heuven (2015, p. 285) ‘An adequate theory of language should be able to predict the approximate degree of intelligibility of a language A for a native listener of a (related) language B by means of a systematic comparison of the similarities and differences between the languages concerned in terms of their vocabulary, syntax, morphology, phonology and phonetics.’ Gooskens and Van Heuven (2018) present the spoken and written intelligibility data from a large project set up to investigate degree of mutual intelligibility of 16 closely related languages<sup>2</sup> within the Germanic, Slavic and Romance language groups in Europe. In the present paper we investigate the extent to which the intelligibility of spoken language can be predicted from linguistic and non-linguistic variables.

In previous research lexical and phonetic distances have been found to correlate substantially with experimentally-determined intelligibility (e.g. Gooskens, 2007; Tang & Van Heuven, 2009). Linguistic differences at other linguistic levels have been shown to affect intelligibility as well (Gooskens & Van Bezooijen, 2006; Doetjes & Gooskens, 2009; Hilton Gooskens & Schüppert, 2013). The present project is the first to exploit dialectometric distance measurements at various linguistic levels (lexicon, phonetics, orthography, morphology, syntax) as predictors of intelligibility. It may seem odd to include orthographic distances when predicting spoken language intelligibility. However, since research has shown that orthography-related knowledge enhances spoken word recognition (e.g. Perre & Ziegler, 2008; Schüppert, 2011), we decided to include orthographic distances in our statistical model.

However, in the case of the 16 languages in our investigation, many listeners are familiar with the test language, so that non-linguistic variables may influence the scores on the intelligibility test. The non-linguistic variables quantified in our investigation are the amount of exposure to the test language, number of years

---

2. We define ‘closely related’ as belonging to the same subgroup of a language family, in our case the Germanic, Romance or Slavic subgroups of the Indo-European language family.

that the listeners learned the test language and attitudes towards the test language. We expect exposure and learning of the test language to be important predictors of intelligibility. The more exposure listeners have had to a language, the more likely they are to understand it. For example, Golubović (2016) showed that a short teaching intervention of four and a half hours of Croatian to Czech listeners improved their understanding considerably. Positive attitudes may motivate listeners to try to understand a non-native language. However, experimental support for this claim has been rather weak (e.g. Delsing & Lundin Åkesson, 2005; Gooskens, 2006; Gooskens & Van Bezooijen, 2006; Impe, 2010; Schüppert, Hilton & Gooskens, 2015).

Summarizing, we will address the following research questions:

1. How well can acquired intelligibility of closely related languages in Europe be predicted by means of a combination of linguistic and non-linguistic variables?
2. How well can inherent intelligibility of closely related languages in Europe be predicted by means of exclusively linguistic variables?

Previous research, for example on Scandinavian languages, has shown significant correlations between intelligibility scores and various linguistic and non-linguistic measures (e.g. Gooskens, 2006, 2007). We expect to find similar relations when expanding our research to a larger language area. However, different language combinations and language families are likely to show different relations to the various measures. Especially exposure is expected to be an important predictor of acquired intelligibility. In the literature, attitude is often mentioned as an important predictor of intelligibility, but experimental evidence is scarce. We expect to find higher correlations with linguistic measures in the case of inherent intelligibility than in the case of acquired intelligibility because non-linguistic factors may overrule linguistic factors in the case of acquired intelligibility.

## 2. Material

Section 2.1 describes the experiment we carried out to measure the intelligibility of closely related languages in Europe (the dependent variable). In Section 2.2 we explain how we quantified eight independent variables (linguistic and non-linguistic variables) used to predict the intelligibility results.

## 2.1 Intelligibility of closely related languages

We tested intelligibility between 70 language pairs by administering six functional tests, covering spoken and written intelligibility at the level of (i) single words (word intelligibility test), (ii) detailed sentence intelligibility (cloze test) and (iii) global message understanding at the text level (multiple choice test). In Gooskens and Van Heuven (2017) the results of the six tests are compared with each other and with the intelligibility as perceived by the test persons themselves (judged intelligibility). The spoken cloze test showed the highest correlation with the perceived intelligibility ( $r = .86$ ) and also correlated with word intelligibility ( $r = .73$ ). Since we were interested in predicting intelligibility of whole texts rather than isolated words, we used the results of the spoken cloze test for the present analysis. The cloze test requires the ability to recognise words and to understand context in order to identify the correct words or type of words that belong in the gaps. It is therefore an easy and useful way of testing overall text intelligibility. We will now describe how we established acquired and inherent intelligibility (see Gooskens et al., 2018, for an overview and discussion of the intelligibility results).

### 2.1.1 Test languages

We included the 16 official languages from the three major language families in the EU member states in our investigation, i.e. five Germanic (Swarte, 2016), five Romance (Voigt, in preparation) and six Slavic (Golubović, 2016) languages (see Table 1). If a language is an official language in more than one country we only included the variety from the country with the largest number of speakers. The listeners all came from the same countries as the speakers. Intelligibility was tested only among speakers of languages within the same language family. So, for example, we tested mutual intelligibility between the two Germanic languages Dutch and German and between the two Romance languages, Italian and Spanish but not between Dutch and Italian or between German and Czech. Listeners were

**Table 1.** Germanic, Romance and Slavic languages (with abbreviation and country) included in the investigation

Germanic	Romance	Slavic
Danish (Da, Denmark)	Dutch (Du, Netherlands)	English (En, England)
German (Ge, Germany)	Swedish (Sw, Sweden)	French (Fr, France)
Italian (It, Italy)	Portuguese (Pt, Portugal)	Romanian (Ro, Romania)
Spanish (Sp, Spain)	Bulgarian (Bu, Bulgaria)	Croatian (Cr, Croatia)
Czech (Cz, Czech Republic)	Polish (Po, Poland)	Slovak (Sk, Slovakia)
		Slovene (Sl, Slovenia)

not tested in their own language. Each language was both a listener language and a test language. For example, French listeners were tested in Spanish and Spanish listeners were tested in French. Henceforth we refer to a combination of a listener language and a test language as a ‘language combination’.

### 2.1.1.2 *Texts*

Since we wanted to compare the cross-language intelligibility of 70 different language combinations it was important to use equivalent test materials for all languages. We selected four English texts used to prepare students for the Preliminary English Test (PET) at the University of Cambridge.<sup>3</sup> The texts all have an intermediate level of difficulty (B1 as formulated by the Common European Framework of Reference for Languages, see Council of Europe, 2001) and their contents are culturally neutral. We adapted the texts slightly so that they were uniform in terms of total length (ca. 200 words) and number of sentences (16–17).

The four texts were translated from English into all 16 languages by native speakers with some translation training. All four texts were first translated by one native speaker and then checked by at least two others. The final version was the one everyone agreed upon. Translators and checkers were instructed to stick to the original English texts as much as possible while still producing grammatically correct translations. This yielded texts that were as comparable as possible across languages in terms of content and level of difficulty.

### 2.1.1.3 *Speakers and recordings*

We recorded six female native speakers of each of the 16 test languages. Speakers were between 20 and 40 years old and were standard speakers of their language.

The speakers were instructed to silently read the texts first and then to read them out clearly and at normal speed. We created 16 online surveys, each with sample recordings from one language. Native listeners of the respective languages rated each of the six speakers by answering the question “How suitable is this speaker as a newscaster on national television?” on a five-point scale ranging from “not at all suitable” to “very suitable”. The voices of the four best-rated speakers per language were used in the experiment. From each speaker the recording of a different (randomly chosen) text was used. By using four different speakers we hoped to neutralize the potential influence of voice quality on the results.

### 2.1.1.4 *Listeners*

The listeners were mainly recruited through social media (Facebook), online newspapers and university mailing lists. Since the listeners were tested online, no

---

3. <http://www.cambridgeenglish.org/exams/preliminary>

restrictions concerning their background were set beforehand. We selected listeners for further analysis afterwards by matching the groups according to certain criteria. Since most of the listeners were young adults we focused on this group and excluded listeners younger than 18 and older than 33. The selected listeners all came from the same countries where the speakers hailed from (see Table 1). In total 70 combinations of listener language and test language were tested, 20 Germanic, 20 Romance and 30 Slavic combinations. The selected listeners had all grown up and lived most of their lives in their home country and spoke the national language as their L1. We excluded listeners who spoke another language at home. All listeners followed or had followed a university education. Some of the test languages are also school languages. We excluded listeners who had learned the test language for longer than the maximum period offered during secondary education.<sup>4</sup>

The criteria described above resulted in a selection of 1833 listeners (426 from the Germanic, 581 from the Romance and 826 from the Slavic language area). Sixty-two percent of the Germanic, 51% of the Romance and 43% of the Slavic listeners were male. The mean number of listeners in any language combination was 26.2 (ranging 14–58). The results gained from these listeners represent the intelligibility structure found among younger, educated Europeans and will be referred to as acquired intelligibility.

However, we were also interested in predicting inherent intelligibility, i.e. intelligibility in situations where listeners have had no previous exposure to the test language but are still able to understand it to some extent because it resembles their native language. Therefore we also made a further selection of listeners with little or no prior exposure to the test language. Before the intelligibility test the listeners filled in a questionnaire with questions about their previous exposure to the test language (see Section 2.2.1). We selected a subset of listeners who had indicated that their mean exposure on six five-point scales was below 2.0 (with '1' indicating no exposure) and who had not learned the target language at school. We removed language combinations with fewer than seven listeners, which we regarded as a minimum for a stable analysis. For instance, there are no Dutch listeners between 18 and 33 who have not learned English in school. Nine out of the original 70 language combinations are no longer represented (mostly because the test language was a school language), five of which are in the Germanic group (Danish, Dutch, German and Swedish listeners tested in English, Dutch listeners tested in German), three in the Romance family (Spanish, Portuguese and Romanian listeners tested in French) and one in the Slavic family (Slovak listeners

---

4. No other selection criteria were used than those mentioned here. This means that we did not remove any outliers.



tested in Czech). The total number of listeners selected was 1307. We refer to these results as inherent intelligibility scores.

### 2.1.5 Intelligibility test

We developed a test that could be carried out online and could be scored automatically.<sup>5</sup> We developed a version of the so-called cloze test. The cloze test (Taylor, 1953) has been used extensively for measuring text comprehension in the classroom. In a cloze test selected words are removed from the text and replaced by gaps, i.e. lines or empty spaces of uniform length (in written language) or by beeps (in spoken language). The deleted words are placed above the text; the subjects' task is to reconstruct the original text. The results can be scored automatically, which renders this an efficient and objective way of testing text comprehension.

Each text was divided into twelve sound fragments. In each fragment one word was replaced by a beep of one second (preceded and followed by 30 ms of silence). A schematic representation of this stimulus presentation is shown in Figure 1. For more information on the procedure see Gooskens and Van Heuven (2017, p. 27).

Twelve response alternatives were continually shown at the top of the screen. When moving the mouse over a word a translation of the word into the native language of the participant was revealed. This was done because we wanted to test the intelligibility of whole texts. If some of the response alternatives were unknown to the participants they would not be able to place them in the right gaps, even if they understood the fragments *per se*. The respondents' task was to click on the word they thought had been removed from the place in the fragment where they heard the beep. Listeners heard each fragment twice. This reflects a real-life situation where the listener would be able to ask the speaker to repeat what he said. Inserted words were greyed out in the selection area, in order to help the participants keep track of their choices. In case they wanted to change an answer, they could simply drag and drop a different word into the same gap. Their original word of choice would then re-appear in black in the selection area above the text. The entire task had to be completed within ten minutes.

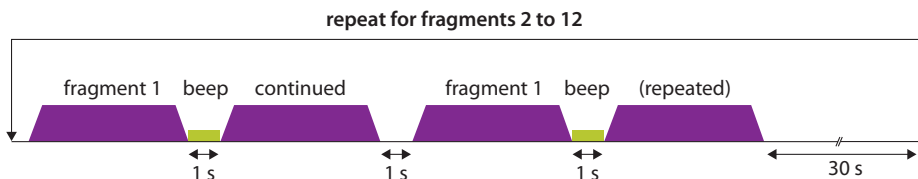


Figure 1. Schematic representation of stimulus presentation of spoken cloze test

5. See <http://www.micrela.nl/app>.

### 2.1.6 Procedure

Listeners first completed a questionnaire about their language background and their attitude towards and exposure to the test language. The responses were used to select listeners with similar profiles when comparing results across listeners groups (see above). Responses were also used for analyses of the effect of various non-linguistic variables on intelligibility (see Section 2.2).

Following the questionnaire the intelligibility test started. The test language was one of the related languages from the language family of the listener (Germanic, Romance or Slavic). In total, there were 64 different tests (4 texts  $\times$  16 languages). The tests were assigned to each listener such that a listener was tested in a randomly selected text and language within his or her own language family (but never the listener's L1). The results were scored automatically and shown on screen on completion of the test. Listeners received no remuneration but had the chance of winning one of a set of prizes. The entire on-line session lasted approximately 15 minutes.

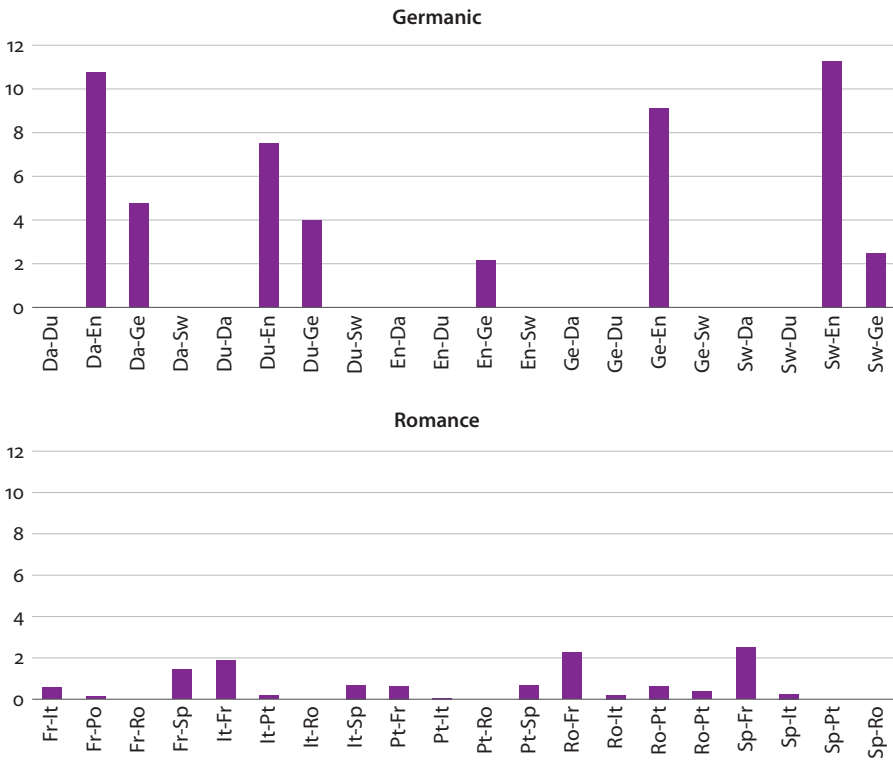
## 2.2 Linguistic and non-linguistic variables

### 2.2.1 Non-linguistic variables

We analysed three non-linguistic variables: (1) number of years the listener had learned the test language, (2) amount of exposure to the test language, and (3) attitude to the test language. The variables were computed from the listeners' responses to the questionnaire. Since these results have not been integrally published before, we will present them here.

**2.2.1.1 Years of learning.** We asked listeners how many years they had learned the test language. The mean results per language combination are presented in Figure 2 and the means per language family in Table 2.

There are large differences among the three language areas. English is a school language for all children in the Germanic area and many children also learn German. All languages except Romanian are learned by at least some children in the Romance area. In contrast, in the Slavic area none of the six Slavic test languages are learned at school as an L2, which explains the absence of a Slavic panel in Figure 2.



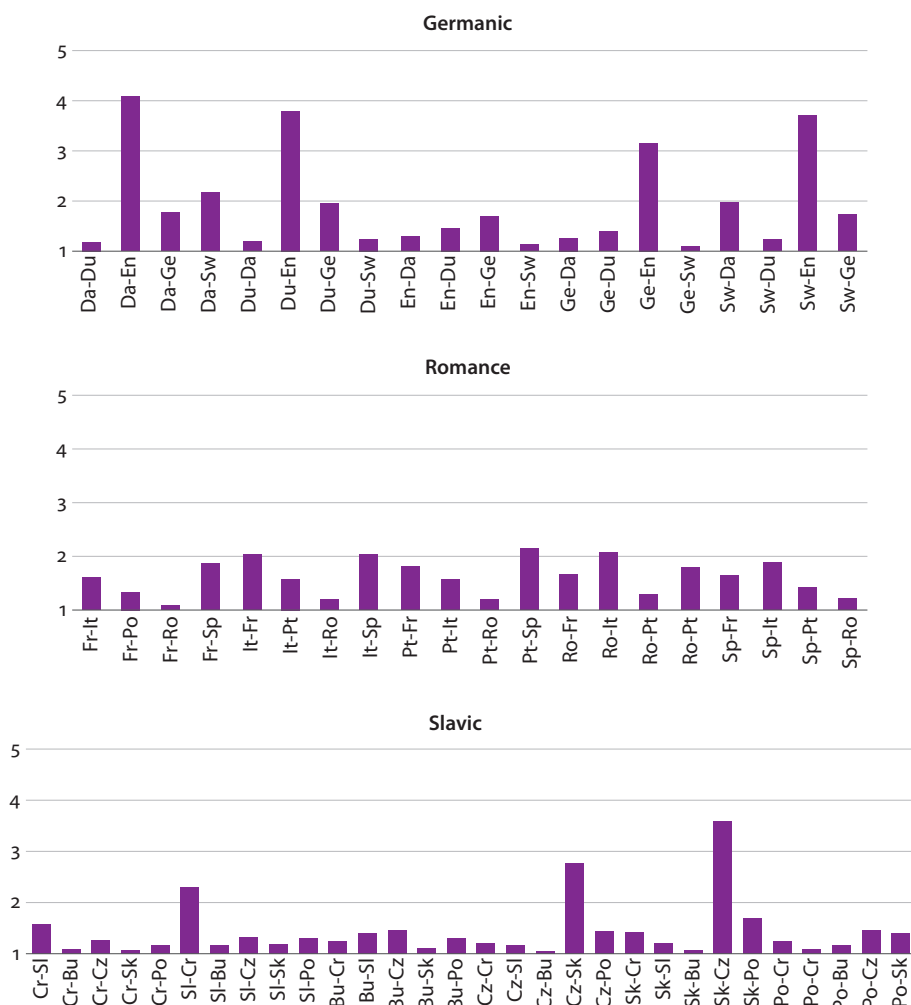
**Figure 2.** Number of years of learning the test language for the Germanic and Romance test areas. In the Slavic language area no listeners learned any of the test languages at school. For each language combination, listener language is presented first, test language second (e.g. Da-Du = Danish listeners tested in Dutch). For abbreviations see Table 1

**2.2.1.2 Amount of exposure.** Listeners indicated how often they were exposed to the test language during the past five years on six five-point scales, ranging from 1 (never) to 5 (every day). They were asked how often they

1. listened to people speaking the test language in their presence (e.g. on vacation, at work, doing shopping, etc.),
2. watched television, DVDs or movies in the test language,
3. played computer games in the test language,
4. chatted or surfed on the internet in the test language,
5. talked to speakers of the test language in person, on the telephone or via Skype,
6. read books, newspapers, magazines and/or text on a computer screen in the test language.

We computed a mean exposure score per test person across the six scales and the mean exposure score per language combination. The mean results per language combination are presented in Figure 3 and the means are presented in Table 2.

The mean exposure is highest in the Germanic language area (1.9) and lowest in the Slavic area (1.4). The exposure results are likely to correlate with the number of years the listeners had learned the test language. Some listeners may have learned the language at school but not be exposed to the language in daily life. For example many Danes learn German at school but are not exposed to it very often. On the other hand, listeners may not have learned the test language at school but



**Figure 3.** Mean exposure score on a scale from 1 (never) to 5 (every day) in the three language areas (also see Figure 2)

**Table 2.** Mean number of years of learning the language, mean exposure and attitude

Test language	Number of years	Exposure	Attitude
Germanic	2.6	1.9	3.1
Romance	0.8	1.6	3.6
Slavic	0.0	1.4	3.4

may still be exposed to it often like for example in the case of Czech and Slovak. The two scales are correlated at  $r = .92$  ( $p < .01$ ) in the Germanic language family and at  $r = .45$  ( $p < .05$ ) in the Romance language family.



**Figure 4.** Mean attitude score in the three language areas on a scale from 1 ‘ugly’ to 5 ‘beautiful’ (also see Figure 2)

**2.2.1.3 Attitude.** We measured attitudes towards the test language as the rating of how beautiful the listeners found the test language. They first listened to a short sound fragment of the language to make sure that all listeners were familiar with the language before rating it. The fragment was the first article of the Universal Declaration of Human Rights, recorded by the same four speakers we used for recording the testing material. The listeners rated the beauty of the language between 1 ‘very ugly’ and 5 ‘very beautiful’. The mean attitude scores per language family are lowest in the Germanic language family and highest in the Romance family (see Table 2). Figure 4 reveals large differences in attitudes towards the test languages within each of the three language families.

### 2.2.2 Linguistic variables

We computed five kinds of linguistic distances between the native language of the listeners and the test languages in our analysis:

1. Lexical
2. Phonetic
3. Orthographic stem
4. Orthographic affix
5. Syntactic

The lexical, orthographic and syntactic distances are based on the four texts used for the cloze tests (see Section 2.1). A list of the 100 most frequently used nouns in the British National Corpus (BNC Consortium 2007) was used for the phonetic distance measurements.<sup>6</sup> This list was translated into the 16 languages by the same translators who also translated the texts for the cloze tests. Next, broad phonetic transcriptions were made of the 16 word lists by means of pronunciation dictionaries and native speakers with a background in phonetics. Phonetic distances were then computed from these transcriptions. The methods for measuring linguistic distance are discussed in detail and the results of the measurements for the 70 language combinations are presented in Gooskens and Heeringa (in preparation).

**2.2.2.1 Lexical distance.** Following Seguy (1973) we defined lexical distance between the members of a pair of languages as the percentage of non-cognates (historically unrelated words) in the two lexicons.

**2.2.2.2 Phonetic distance.** Phonetic distance was computed for the aligned cognate word pairs in the vocabulary lists of 100 words for each pair of languages. The degree of similarity between cognates was computed by the Levenshtein

---

6. <http://www.natcorp.ox.ac.uk/>

algorithm, which computes the smallest number of string edit operations (i.e. deletions, insertions and substitutions) needed to convert the string of phonetic symbols in language A to the cognate string in B. We illustrate this algorithm by a simplified example ignoring diacritics, comparing English *interest* with its Swedish cognate *intresse* in Figure 5.

Figure 5. Illustration of the Levenshtein algorithm

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

In the fourth slot /ə/ is deleted, in the sixth slot /e/ is replaced by /ɛ/, in the eighth slot /ə/ is inserted and in the ninth slot /t/ is deleted. The total number of penalty points (4) is then divided by the length of the alignment (9, the number of alignment slots) to yield a length-normalised Levenshtein distance, in the example  $(4/9) \times 100 = 44\%$ . To constrain possible alignments, vowels match with vowels and consonants with consonants but [j, w] also with vowels and schwa with sonorants. The overall phonetic distance between language A and B is the mean normalised distance across all cognate word pairs. The simple version of the algorithm uses binary differences between alignments. We, however, used graded weights (between 0 and 1) that express acoustic segment distances so that, for example, the pair [i, o] is seen as more different than the pair [i, ɪ] (for details see Nerbonne & Heeringa, 2010).

2.2.2.3 *Orthographic distance, stem and affix.* We computed orthographic distances on the basis of the cognates in the word lists in a similar way as for the phonetic distances. However, the operation weights are different. For each character we distinguished between a base and a diacritic. For example, the base of é is e, and the diacritic is the acute accent. We weighted differences in the base as 1 and in diacritics as 0.3.

Bulgarian materials, which are written exclusively in Cyrillic, were replaced by Latin transliterations produced by the web application Translit.<sup>7</sup> We computed orthographic distances separately for stems and affixes – which could not be done in the case of phonetic distances since these were based on uninflected dictionary lemmas.

7. The application can be found at: <http://bg.translit.cc/>.

**2.2.2.4 Syntactic distance.** The syntax measures were based on the 66 sentences in the text data using the translations that were described in Section 2.1.2. When calculating the syntactic distances between language A and B we calculated the distance between the original text in language A and its literal translation in language B.

We measured the syntactic distance between two languages by computing the correlation between syntactic trigram frequencies (see Nerbonne & Wiersma, 2006). We defined 14 lexical categories: noun, verb, modal verb, adjective, adverb, pronoun, preposition, conjunction, numeral, determiner, interjection, *to* before infinitive, abbreviations and sentence boundary. All trigrams (different sequences of three lexical category labels) were then inventoried and counted. This yielded different frequencies in languages A and B. Syntactic distance was then defined as 1 minus the Pearson correlation coefficient ( $r$ ) found between the trigram frequencies. For details and examples see Heeringa et al. (2017).

### 3. Results

#### 3.1 Predicting intelligibility from linguistic and non-linguistic variables

In this section we examine how well spoken text comprehension of the 70 language combinations in our investigation can be accounted for. For this purpose we include the eight predictors described in Section 2.2, i.e. three non-linguistic variables (number of years learned, exposure and attitude) and five linguistic variables (lexical, phonetic, orthographic stem, orthographic affix and syntactic distances). Since we examine the mean comprehension scores of all 1833 listeners, the results are based on a compounding of inherent intelligibility and acquired intelligibility gained from exposure and formal learning of the test language.

We first correlated the mean results of each language combination with the eight linguistic and non-linguistic variables. The results are presented in Table 3. We calculated correlations on the basis of all 70 language combinations as well as for each language family separately (20 language combinations for Germanic and Romance and 30 for Slavic).

The correlations of intelligibility (top row) and exposure are significant and high, between  $r = .87$  for Romance and  $.93$  for Germanic and  $.90$  when all language combinations are included. Correlation of intelligibility and number of years participants learned the language is significant only for the Germanic group. Attitude shows rather high significant correlations with intelligibility scores ( $.60 < r < .81$  for the three language families). When looking at the next row, we observe a high correlation between exposure and attitude ( $.68 < r < .78$ ). This suggests that people



**Table 3.** Correlation coefficients *r* between intelligibility scores and the eight predictors across all 70 language combinations and for the three language families separately (20 Germanic, 20 Romance and 30 Slavic language combinations)

	Expos.	Learning	Attitude	Lexical	Phonetic	Orthographic		Syntactic
						stem	affix	
Intelligibility								
All	.90**	.60**	.65**	-.38**	-.14	-.22	-.46**	-.40**
Germ.	.93**	.86**	.60**	-.21	.03	.01	-.44	-.30
Rom.	.87**	.19	.70**	-.36	-.30	-.45*	-.45*	-.40
Slav.	.92**	-	.81**	-.82**	-.83**	-.82**	-.86**	-.62**
Exposure								
All		.76**	.61**	-.18	.07	-.04	-.55**	-.31**
Germ.		.92**	.68**	.07	.30	.21	-.35	-.25
Rom.		.45*	.78**	-.41	-.08	-.35	-.54*	-.40
Slav.		-	.78**	-.70**	-.77**	-.75**	-.81**	-.63**
Learning (years)								
All			.32**	.19	.44**	.38**	-.30*	.04
Germ.			.61**	.30	.28	.44	-.16	.05
Rom.			.34	-.12	.56**	.17	-.22	.06
Slav.			-	-	-	-	-	-
Attitude								
All				-.38**	-.06	-.04	-.33	-.51**
Germ.				-.04	.15	.14	-.16	-.16
Rom.				-.43	.04	-.25	-.42	-.41
Slav.				-.69**	-.63**	-.62**	-.81*	-.69**
Lexical								
All					.34**	.61**	.50**	.52**
Germ.					.54*	.82**	.52*	.47*
Rom.					.08	.75**	.80**	.81**
Slav.					.88**	.78**	.71**	.41*
Phonetic								
All						.64**	-.13	.20
Germ.						.33	.22	.04
Rom.						.23	-.10	.18
Slav.						.83**	.70**	.47**

Table 3. (continued)

	Expos.	Learning	Attitude	Lexical	Phonetic	Orthographic		Syntactic
						stem	affix	
Orthographic stem								
	All						.27*	.22
	Germ.						.15	.48*
	Rom.						.74**	.82**
	Slav.						.84**	.32
Orthographic affix								
	All							.35**
	Germ.							.63**
	Rom.							.63**
	Slav.							.63**

\*  $p \leq .05$ ;\*\*  $p \leq .01$  (two-tailed)

in general are more positive towards languages that they are familiar with than towards languages that they are exposed to less often (and/or *vice versa*).

The other test languages are only school subjects in the Germanic and Romance language areas, and only within the Germanic group does years of learning correlate significantly with intelligibility. Naturally, there is a significant correlation between exposure and years of learning. Correlations are high in the Germanic language family ( $r = .92$ ), but much lower for the Romance languages ( $r = .45$ ), which is due to the fact that many Romance listeners learn French at school but are not exposed to it very often (compare Figures 1 and 2).

The correlations between intelligibility and the linguistic variables are generally low and insignificant. Exceptions are the Romance family showing significant correlations with orthographic stem and affix distances ( $r = -.45$  for both) and the Slavic language family, where all correlations are rather high and significant (between  $-.62$  for syntactic distances and  $-.86$  for orthographic affix distances). As discussed in Section 2, there is generally little exposure to closely related languages in the Slavic language area and the other Slavic test languages are rarely taught at school in the six Slavic countries concerned. So even though the correlation between intelligibility and exposure is high in the Slavic area ( $r = .92$ ) this does not override the linguistic variables, unlike what we find in the other two language families. It is also striking that exposure correlates significantly with linguistic variables in the Slavic language family ( $-.63 < r < -.81$ ). So, in general, Slavic people have more exposure to languages that are similar to their own than

to linguistically distant languages. This may be explained by the fact that linguistically closely related languages are also often geographically close. This relationship is weaker in the other two language families.

Table 3 also shows that there are many significant correlations among linguistic distances. Especially lexical distances correlate significantly with all other linguistic distances (except the phonetic distances in the Romance languages). The other linguistic distances show more complicated relationships per language family. We will return to this issue in Section 3.2.

To investigate how well we can predict our intelligibility scores from the eight linguistic and non-linguistic variables we carried out regression analyses with the mean intelligibility scores per language combination as the criterion and the eight linguistic and non-linguistic variables as predictors. The results are presented in Table 4 for all language combinations and separately for each of the language families.

**Table 4.** Stepwise regression analyses with mean intelligibility score as the criterion and eight linguistic and non-linguistic predictors

Language combination	Predictors	$R^2$	$t$	$p$
All	Exposure	.82	20.086	< .001
	Lexical distance	.86	-3.769	< .001
	Phonetic distance	.88	-2.952	.004
Germanic	Exposure	.86	2.316	.034
	Lexical distance	.93	-9.130	< .001
	Years learned	.98	5.740	< .001
Romance	Exposure	.75	7.814	< .001
	Phonetic distance	.80	-2.170	.044
Slavic	Exposure	.85	5.714	< .001
	Lexical distance	.91	-3.507	.002
	Orthographic affix distance	.92	-2.096	.046

Again, exposure is the most important predictor for all three language families. Years learned is included in the Germanic model, but this non-linguistic variable adds very little to the model, probably because of the high intercorrelation with exposure. Linguistic distances, lexical in the case of Germanic and Slavic, affix in Slavic and phonetic in the case of Romance, are included in the models but add little to the predictive power.

Obviously, if people have learned a language via exposure or formal learning they will understand it better than when they have had little previous exposure to

the language, regardless of the linguistic distances: linguistic variables are overridden here by exposure. However, we are also interested to learn how well we can predict our results from linguistic distances only. We therefore carried out another analysis where we left out the non-linguistic predictors. The result is presented in Table 5.

**Table 5.** Stepwise regression analyses with mean intelligibility scores as criterion variable and five linguistic predictors

Language combinations	Predictors (distances)	$R^2$	$t$	$p$
All	Orthographic affix	.33	-6.088	< .001
	Phonetic	.37	-2.138	.036
Germanic	-	-	-	-
Romance	Orthographic affix	.20	-2.142	.046
Slavic	Orthographic affix	.74	-5.027	< .001
	Phonetic	.84	-4.274	< .001

For the Germanic and the Romance language families, linguistic distances have little predictive power (no predictors are included in the Germanic model; orthographic distances predict 20% in Romance). As for Slavic, we noted high correlation between intelligibility and linguistics distances and high intercorrelations between non-linguistic and linguistic variables in Table 2. This explains why Slavic linguistic (specifically affix and phonetic) distances can predict 84% of the variance.

### 3.2 Predicting inherent intelligibility from linguistic variables only

In this section we investigate how well we can predict inherent intelligibility from linguistic distances only. Ideally, we should correlate linguistic distances with intelligibility scores from listeners who have never been exposed to the test language before. This would tell us how well listeners understand the test language exclusively on the basis of its similarity to their L1. Linguistic distances should predict the intelligibility scores for these listeners better than for the larger group that includes listeners with previous exposure to the test languages. As explained in Section 2.1, we selected a subset of listeners who had indicated that their mean exposure on six five-point scales was below 2.0 (with '1' for no exposure) and who

had not learned the test language at school.<sup>8</sup> This reduced the number of listeners to 1,307 and the number of language combinations to 61.

**Table 6.** Correlations between inherent intelligibility and five linguistic distances, for 61 language combinations (upper part), and when Romanian listeners are excluded (lower part)

Romanian	Listeners	Lexical	Phonetic	Orthographic		Syntactic
				stem	affix	
included						
	All	-.63**	-.52**	-.59**	-.43**	-.54**
	Germ.	-.95**	-.28	-.91**	-.49	-.67**
	Rom.	-.39	-.51*	-.53*	-.41	-.49*
	Slav.	-.80**	-.79**	-.77**	-.81**	-.53**
excluded						
	All	-.76**	-.51**	-.68**	-.49**	-.56**
	Germ.	-.95**	-.28	-.91**	-.49	-.67**
	Rom.	-.69**	-.47	-.68*	-.54*	-.77**
	Slav.	-.80**	-.79**	-.77**	-.81**	-.53**

Table 6 (upper half) shows the correlations between inherent intelligibility and linguistic distances. The correlations in Table 6 are typically better than their counterparts obtained for the whole dataset (Table 3, Intelligibility rows). For example, in the Germanic language family the correlations with lexical distances have increased from  $-.21$  to  $-.95$  and with the orthographic stem distances from  $.01$  to  $-.91$ . Correlations with phonetic distances are insignificant. Danish-Swedish and Swedish-Danish intelligibility are outliers. The listeners in these two language combinations understand each other better than would be expected from phonetic distance, probably because there are hardly any lexical differences between the two languages (a distance of 4.6% for Danish-Swedish and 5.8% for Swedish-Danish). This may compensate for the impediment that pronunciation differences may form. Excluding these two language combinations increases the correlation with phonetic distance in the Germanic language family to  $-.87$ .

At first glance the correlations in the Romance family are only slightly higher in the inherent data set than in the full data set. A closer look at the data reveals that the intelligibility scores of the Romanian listeners are outliers for all test languages. The Romanian listeners obtained much higher intelligibility scores than would be expected from the linguistic distances. One possible explanation is that

8. Too few cases remain if we include only listeners with exposure = 1.

most Romanian listeners have learned French at school and apply their knowledge of French to other Romance languages. Of course, other listeners may benefit from knowledge about other languages that they know. However, Romanians also watch a lot of television from Spain and Italy with subtitles. The score on the exposure scale concerning how often they watched television, DVDs or movies in the test language was higher for the Romanians (2.30, Section 2.2.1) than for the other Romance listeners (1.50–1.74). Consequently intelligibility for Romanian listeners cannot be characterized as inherent only. Nevertheless, Romanians were included in the analysis of inherent intelligibility since their mean score exposure score was below 2. Excluding Romanian listeners (Table 6, lower half) yields higher correlations with all distances except for the phonetic distances.

The correlations in the Slavic language family were already high when we included the results from all listeners and they hardly change when filtered for exposure and years of learning the language, presumably because listeners from this language group had little previous exposure to the test languages (see Figures 1 and 2). The Slavic language family shows the highest correlation between intelligibility scores and affix distances. The Slavic languages have a rich inflectional system that probably contains more information that listeners need to understand a text than the languages in the other two families. Lexical, phonetic and orthographic stem distances also correlate highly. Syntactic distances show the lowest correlation, probably because the Slavic languages are characterized by a rather free word order.

Table 7 presents the results of stepwise regression analyses with the five linguistic variables as predictors and inherent intelligibility scores as the criterion. We excluded the Romanian listeners since they must have had a lot of exposure to the other Romance languages (see above). When comparing the results to those of the regression analyses with acquired intelligibility (Tables 4 and 5) we see a much higher predictive power, especially for Germanic, where 93% of the variance is explained by lexical and orthographic stem distances. Since we are dealing with spoken language we would expect phonetic distances to be included rather than orthographic stem distances. As we already discussed above the correlations between intelligibility and phonetic distances are low in the Germanic language family due to the Danish-Swedish outliers. Moreover, orthography is likely to reflect phonological differences to a high degree.

In the Slavic family the percentage of explained variance is also high. Orthographic affix and phonetic distances together explain 80% of the variance.

While the predictive power is high for both the Germanic and the Slavic families, the situation is less clear for the Romance family. Rather unexpectedly syntactic distance is the only variable included in the model. The predictive power is rather low (59%). This can probably be explained by the high correlation

between syntactic and orthographic stem distances ( $r = .82$ , Table 3). Orthographic stem differences also correlate well with lexical and orthographic affix distances ( $r = .75$  and  $.74$ ).

**Table 7.** Stepwise regression analyses with mean inherent intelligibility scores per language combination as the criterion and five linguistic distances as predictors. The four Romanian listener groups are excluded (see text)

Languages	Predictors (distances)	$R^2$	$t$	$p$
All	Lexical	.57	-2.775	.008
	Orthographic stem	.67	-2.610	.012
	Syntactic	.73	-2.668	.010
	Phonetic	.75	-2.929	.005
	Orthographic affix	.77	-2.088	.042
Germanic	Lexical	.89	-4.862	.001
	Orthographic stem	.93	-2.751	.019
Romance	Syntactic	.59	-4.329	.001
Slavic	Orthographic affix	.66	-4.916	< .001
	Phonetic	.80	-4.327	< .001

We have shown that linguistic distances predict inherent intelligibility to a high extent. With the exception of the phonetic distances, which were based on parallel lists of 100 words (Section 2.2.1), our linguistic distance measurements were derived from the four texts that were used in the cloze tests. This means that we have shown that we can predict intelligibility scores by means of linguistic distances that are calculated on the testing material itself. We would also like to know whether we can predict intelligibility scores from data that are independent of the data set underlying our intelligibility measurements. To test the generalizability of our results we therefore measured lexical and orthographic distances in the lists of 100 words that were also used for the phonetic distance measurements (Section 2.2.1).

The lexical and orthographic distances in the text data and the word list data are strongly correlated ( $p < .01$ ). For lexical distances,  $r$  was  $.76$  for Germanic,  $.87$  for Romance and  $.92$  for Slavic. For orthographic distances  $r$  was  $.82$  for Germanic,  $.87$  for Romance and  $.83$  for Slavic. This shows that the two kinds of distance measurements are interchangeable to a high degree.

Table 8 presents the correlations between the inherent intelligibility scores and the three linguistic variables measured on the basis of the list of 100 words. We excluded the language combinations with Romance listeners (see above). We can now compare the correlations with those found in the lower half of Table 6. The correlations with phonetic distances are identical as they are based on the same

word lists. In the Germanic language area, the correlations with lexical distances are lower when correlated with the word list data ( $r = -.75$ ) than with the text data ( $r = -.95$ ), and the same goes for correlations with orthography in the Slavic language area ( $r = -.52$  for word list versus  $-.77$  for text data), but for all other correlations it hardly matters whether we correlate the intelligibility scores with the distances measured from the test materials or from the 100-word lists. Some correlations are even higher, especially with lexical distances in the Romance family ( $r = .69$  when correlated with text data and  $.82$  when correlated with the list data).

**Table 8.** Correlations between inherent intelligibility scores and the three linguistic predictors computed from the word lists. The four Romanian listener groups are excluded (see text)

	Lexical	Phonetic	Orthographic
All	-.52**	-.48**	-.54**
Germ.	-.75**	-.28**	-.89**
Rom.	-.82**	-.47**	-.72**
Slav.	-.86**	-.79**	-.52**

Table 9 presents the results of a regression analysis with the inherent intelligibility (without the Romanian listeners) and distances computed from 100-word lists. This table can be compared with Table 7. The variance explained by the word list data and by the text data is the same for Slavic, 2 points less for Germanic, and 20 points higher for Romance. Orthographic and lexical distances are generally most important and phonetic distances are also included in the Romance model.

**Table 9.** Stepwise regression analyses with mean inherent intelligibility scores per language combination as the criterion and three linguistic distances calculated from word lists as predictors

Languages	Predictors (distances)	$R^2$	$t$	$p$
All	Orthographic	.29	-5.088	< .001
	Lexical	.63	-8.345	< .001
	Phonetic	.72	-4.175	< .001
Germanic	Orthographic	.79	-6.310	< .001
	Lexical	.91	-3.694	.004
Romance	Lexical	.67	-5.700	< .001
	Phonetic	.79	-2.648	.021
Slavic	Lexical	.74	-8.293	< .001
	Orthographic	.80	-2.823	.009



#### 4. Conclusions and discussion

By means of a series of stepwise regression analyses we have shown that we can predict the intelligibility of closely related languages among young educated Europeans to a high extent. We first analysed the mean intelligibility results for 20 Germanic, 20 Romance and 30 Slavic language combinations with a selection of 1833 listeners and, as expected, found that exposure to the test language is the most important predictor and that it tends to override all other predictors. Trivial though this may seem, these results can be used to raise awareness among policy makers and language teachers of the importance of exposing language learners, and people in general, to languages. Through exposure listeners will get used to the sounds of the non-native language and will learn how these sounds correspond to those in their own language. They are also likely to learn some of the vocabulary. Even when inherent intelligibility is poor, it often takes only a small effort to learn to understand a closely related language well enough to sustain receptive multilingualism (RM). Previous research (e.g. Hedquist, 1985; Golubović, 2016) has shown that in the case of closely related languages, only a short language course that makes speakers conscious of the most important differences and similarities between their native language and the language of the speaker can improve receptive proficiency considerably (see also the chapter about transfer by Rothman in this volume). The predictive power of linguistic distances turned out to be low in the Germanic and Romance language areas because of the high predictive power of exposure (and years of learning in the case of Germanic). In the Slavic area there is so little exposure to the other languages that we are in fact dealing with inherent intelligibility, i.e. intelligibility in situations where listeners received no previous exposure to the test language but are still able to understand it to some extent because it resembles their native language.

Attitude and linguistic distances hardly improve the model. It is possible that a more sophisticated method of eliciting conscious or unconscious attitudes, such as a matched guise experiment, would provide us with attitude measurements that are more precise and show a higher correlation with intelligibility. However, it is also possible that attitudes only play a minor role in an experimental situation because the participants try to perform as well as possible and are therefore not influenced by attitudes towards a language and its speakers in the same way as they would be in a real-life situation.

We wanted to keep our experiments short to attract as many participants as possible. Therefore we limited our choice of non-linguistic variables to the three factors that are most often mentioned in the literature and which can be assumed to play the most important role in predicting intelligibility. It should be noted that even more non-linguistic factors could play a role in predicting intelligibility. The

level of understanding between two interlocutors with different L1s depends on a complicated interaction of speaker and listener competencies and activities. Individual personality traits identified within psychology have been shown to influence language learning and therefore can also be expected to play a role in RM. Examples of such traits are the ability to adapt to new situations, knowledge of the world, sociocultural resources and cognitive resources, age, literacy, plurilingual resources and the mastery of interaction strategies. In Gooskens (submitted) a detailed discussion of linguistic and non-linguistic determinants of RM is provided. In future intelligibility research expanded questionnaires with more detailed questions about non-linguistic factors could be included. This would provide a stronger basis to interpret intelligibility results.

We were also interested in predicting inherent intelligibility. We therefore carried out another regression analysis on a subset of listeners with minimal exposure to the test language. As expected, linguistic distances now explain a substantial part of the variance. Lexical distance is the most important predictor in the Germanic area and also correlate highly in the Slavic area. In Romance, syntactic distances were the only useful predictor. However, lexical and orthographic distances also correlate well with intelligibility in this language family. It therefore seems safe to say that lexical distances are generally the most important predictors of inherent intelligibility – which is also what we would expect from common sense. If a language has too many words that cannot be related to the words in the listeners' L1 they will have no way to understand them unless they have learned them (or know them from some other language). If only the pronunciation of a word in a related language is different than in the listeners' L1, they may still be able to understand the word.

We wanted to know whether our results can be generalized, i.e. whether we would get the same results if we would predict the intelligibility scores by means of linguistic distance measurements based on another dataset. We therefore repeated the regression analysis with distance measures based on lists of 100 nouns. When we include all three language families (Romanian listeners excluded) the explained variance is only slightly less for the list data set (72%) than for the actual text data (77%). Lexical, orthographic and phonetic distances are included as predictors in both data sets and lexical distances are included in the models of all three language families.

We conclude that the percentage of variance explained is virtually the same for text data and word list data. Consequently, we may predict inherent intelligibility from linguistic distances, using a random word list and just three kinds of distance: lexical, orthographic, and phonetic. This conclusion is all the more remarkable considering that random word lists contain only 100 nouns whereas the text data comprise ca. 800 words distributed over multiple grammatical categories.

Moreover, the orthographic distances are less fine-grained in the word lists, since no distinction is made between stems and affixes.

We have shown that inherent intelligibility can be predicted quite well by linguistic distances and that a short word list provides sufficient input for computing the distance measures needed. When objective estimates are desired of how well speakers of closely related languages will be able to understand each other without prior exposure or instruction, it may therefore be an option to rely on distance measurements rather than on costly functional testing or on the subjective opinion of the speakers themselves. In addition to objectivity another advantage of distance measurements is that no intelligibility tests have to be developed and administered. For various reasons such an objective estimate may be important, for instance, to resolve issues that concern language planning and policies, second-language learning, and language contact. Unbiased data about distances and intelligibility can also be crucial for sociolinguistic studies. Varieties that have strong social stigmas attached to them may unrightfully be deemed hard to understand (Giles & Niedzielski, 1998; Wolff, 1959). Advances in the field of linguistic distances and intelligibility measurements provide sociolinguists with objective data to resolve conflicts that arise concerning varieties on a standard-nonstandard continuum. Such knowledge is also needed for standardization and development of new orthographies in communities where no standardized orthography exists. Note, however, that the results that we have presented here are based on means across language combinations. It is much harder to predict the level of intelligibility of single listeners due to individual variation in working memory, general intelligence or language aptitude (see Vanhove, 2014).

Note, finally, that the percentages of explained variance for inherent intelligibility are not optimal. This means that there is room for improvement of our linguistic distance measurements. Such measurements should to a larger extent take into account communicatively relevant distances by weighting linguistic difference that are important for communication more heavily than differences that are less important. Improvements of the algorithm should take into account the human decoding processes. For example, in general consonants are better predictors of intelligibility of than vowels, consonant substitutions are better predictors than insertions or deletions, and word beginnings are more important than later parts of words (Van Heuven, 2008 and references therein). Gooskens et al. (2015) found that minor phonetic details that could hardly be captured by Levenshtein distances, may sometimes have a major impact on the intelligibility of isolated words. In addition to linguistic factors, paralinguistic factors such as pitch, volume, speech rate, fluency, facial expressions, and hand gestures should also be included in a more complete model of intelligibility.

## Funding

This work was supported by a grant from the Netherlands Organisation for Scientific Research (NWO) 360-70-430 awarded to Charlotte Gooskens and Vincent van Heuven.

## Acknowledgements

We thank Jelena Golubović, Femke Swarte and Stefanie Voigt and numerous student assistants for collecting the material for this investigation. We thank Aleksandar Mančić for programming the web application. We thank Wilbert Heeringa for calculating the linguistic distances.

## References

- Braunmüller, K. (2007). Receptive multilingualism in Northern Europe in the Middle Ages: A description of a scenario. In J. D. ten Thije & L. Zeevaert (Eds.), *Receptive multilingualism* (pp. 25–47). Amsterdam: John Benjamins. <https://doi.org/10.1075/hsm.6.04bra>
- Council of Europe (2001). *Common European framework of reference for languages. Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Delsing, L. O., & Lundin Åkesson, K. (2005). *Håller språket ihop Norden? En forskningsrapport om ungdomars förståelse av danska, svenska och norska* [Does the language keep together the Nordic countries? A research report of mutual comprehension between young Danes, Swedes and Norwegians]. Copenhagen: Nordiska ministerrådet. <https://doi.org/10.6027/tn2005-573>
- Doetjes, G., & Gooskens, C. (2009). Skriftsprogets rolle i den dansk-svenske talesprogsforståelse [The role of orthography in the mutual intelligibility of spoken Danish and Swedish]. *Språk och stil*, 19, 105–123.
- Giles, H., & Niedzielski, N. (1998). Italian is beautiful, German is ugly. In L. Bauer & P. Trudgill (Eds.), *Language myths* (pp. 85–93). London: Penguin.
- Golubović, J. (2016). *Mutual intelligibility in the Slavic language area*. Groningen: Center for Language and Cognition.
- Gooskens, C. (accepted). Receptive multilingualism. In S. Montanari & S. Quay (Eds.), *Multidisciplinary perspectives on multilingualism*. Berlin: De Gruyter.
- Gooskens, C. (2006). Linguistic and extra-linguistic predictors of Inter-Scandinavian intelligibility. In J. van de Weijer & B. Los (Eds.), *Linguistics in the Netherlands*, 23, (pp. 101–113). Amsterdam: John Benjamins.
- Gooskens, C. (2007). The contribution of linguistic factors to the intelligibility of closely related languages. *Journal of Multilingual and Multicultural Development*, 28(6), 445–467. <https://doi.org/10.2167/jmmd511.0>
- Gooskens, C., & Bezooijen, R. van (2006). Mutual comprehensibility of written Afrikaans and Dutch: symmetrical or asymmetrical? *Literary and Linguistic Computing*, 23, 543–557. <https://doi.org/10.1093/lc/fql036>

- Gooskens, C., Bezooijen, R. van, & Heuven, V. J. van (2015). Mutual intelligibility of Dutch-German cognates by children: The devil is in the detail. *Linguistics*, 53(2), 255–283. <https://doi.org/10.1515/ling-2015-0002>
- Gooskens, C., & Heeringa, W. (in preparation). Linguistic distances between Germanic, Romance and Slavic languages.
- Gooskens, C., & Heuven, V. J. van (2017). Measuring cross-linguistic intelligibility in the Germanic, Romance and Slavic language groups. *Speech Communication*, 89, 25–36. <https://doi.org/10.1016/j.specom.2017.02.008>
- Gooskens, C., Heuven, V. J. van, Golubović, J., Schüppert, A., Swarte, F., & Voigt, S. (2018). Mutual intelligibility between closely related language in Europe. *International Journal of Multilingualism*, 15(2), 169–193. <https://doi.org/10.1080/14790718.2017.1350185>
- Hedquist, R. (1985). *Nederländares förståelse av danka och svenska. En språkpedagogisk undersökning med utnyttjande av likheterna mellan språken*. Umeå: Institutionerna för fonetik och nordiska språk, Umeå universitet.
- Heeringa, W., Swarte, F., Schüppert, A., & Gooskens, C. (2017). Measuring syntactical variation in Germanic texts. *Digital Scholarship in the Humanities* 33(2), 279–296. <https://doi.org/10.1093/llc/fqx029>
- Heuven, V. J. van. (2008). Making sense of strange sounds: (Mutual) intelligibility of related language varieties. A review. *International Journal of Humanities and Arts Computing*, 2(1–2), 39–62. <https://doi.org/10.3366/E1753854809000305>
- Hilton, N. H., Gooskens, C., & Schüppert, A. (2013). The influence of non-native morphosyntax on the intelligibility of a closely related language. *Lingua*, 137, 1–18. <https://doi.org/10.1016/j.lingua.2013.07.007>
- Impe, L. (2010). *Mutual intelligibility of national and regional varieties of Dutch in the Low Countries*. Leuven: University of Leuven.
- Maurud, Ø. (1976). *Nabospråksförståelse i Skandinavien. En undersøkelse om gjensidig forståelse av tale- og skriftspråk i Danmark, Norge og Sverige*. [Neighbouring language comprehension of spoken and written language in Denmark, Norway and Sweden.] Stockholm: Nordiska rådet.
- Nerbonne, J., & Heeringa, W. (2010). Measuring dialect differences. In J. E. Schmidt & P. Auer (Eds.), *Language and Space: Theories and Methods. Handbooks of Linguistics and Communication Science* (pp. 550–567). Berlin: Mouton De Gruyter.
- Nerbonne, J., & Wiersma, W. (2006). A measure of Aggregate Syntactic Distance. In J. Nerbonne & E. Hinrichs (Eds.), *Linguistic Distances Workshop at the joint conference of International Committee on Computational Linguistics and the Association for Computational Linguistics, Sydney, July, 2006* (pp. 82–90).
- Perre, L., & Ziegler, J. C. (2008). On-line activation of orthography in spoken word recognition. *Brain Research*, 1188, 132–138. <https://doi.org/10.1016/j.brainres.2007.10.084>
- Schüppert, A. (2011). *Origin of asymmetry: Mutual intelligibility of spoken Danish and Swedish*. Groningen: Center for Language and Cognition.
- Schüppert, A., Hilton, N. H., & Gooskens, C. (2015). Swedish is beautiful, Danish is ugly? Investigating the link between language attitudes and intelligibility. *Linguistics*, 53(2), 375–403. <https://doi.org/10.1515/ling-2015-0003>
- Séguy, J. (1973). La dialectométrie dans l'Atlas linguistique de la Gascogne [The dialectometry in the Linguistic Atlas of Gascogne]. *Revue de Linguistique Romane*, 37, 1–24.
- Sherkina-Lieber, M. (2015). Tense, aspect, and agreement in heritage Labrador Inuttitut. *Linguistic Approaches to Bilingualism*, 5(1), 30–61. <https://doi.org/10.1075/lab.5.1.02she>

- Swarte, F. (2016). *Predicting the mutual intelligibility of Germanic languages from linguistic and extralinguistic factors*. Groningen: Center for Language and Cognition.
- Tang, C., & Heuven, V. J. van (2009). Mutual intelligibility of Chinese dialects experimentally tested. *Lingua*, 119(5), 709–732. <https://doi.org/10.1016/j.lingua.2008.10.001>
- Tang, C., & Heuven, V. J. van (2015). Predicting mutual intelligibility of Chinese dialects from multiple objective linguistic distance measures. *Linguistics*, 53(2), 285–312. <https://doi.org/10.1515/ling-2015-0005>
- Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, 30, 415–433.
- Thije, J. D. ten, & Zeevaert, L. (2007). Introduction. In J. D. ten Thije & L. Zeevaert (Eds.). *Receptive Multilingualism* (pp. 1–25). Amsterdam: John Benjamins. <https://doi.org/10.1075/hsm.6>
- Vanhove, J. (2014). *Receptive multilingualism across the lifespan. Cognitive and linguistic factors in cognate guessing*. Fribourg: University of Fribourg.
- Voigt, S. (in preparation). *Predicting the (mutual) intelligibility of Romance languages from linguistic and extra-linguistic factors*. Groningen: Center for Language and Cognition.
- Wolff, H. (1959). Intelligibility and inter-ethnic attitudes. *Anthropological Linguistics*, 1, 34–41.
- Zeevaert, L. (2004). *Interskandinavische Kommunikation. Strategien zur Etablierung von Verständigung zwischen Skandinaviern im Diskurs*. [Interscandinavian communication. Strategies for establishing comprehension between Scandinavians in discourse.] Hamburg: Verlag Dr. Kovač.

### *Authors' addresses*

Charlotte Gooskens  
 Department of European Languages and  
 Cultures  
 University of Groningen  
 Postbus 716  
 9700 AS Groningen  
 The Netherlands  
 email [c.s.gooskens@rug.nl](mailto:c.s.gooskens@rug.nl)

Vincent J. van Heuven  
 Department of Hungarian and Applied  
 Linguistics  
 Pannon Egyetem  
 Egyetem Ut. 10 8200 Veszprém  
 Hungary  
[v.j.j.p.van.heuven@hum.leidenuniv.nl](mailto:v.j.j.p.van.heuven@hum.leidenuniv.nl)

### **Publication history**

Date received: 13 December 2017

Date accepted: 19 November 2018

Published online: 28 January 2019