Readability for foreign language learning: The importance of cognates

Lisa Beinborn, Torsten Zesch and Iryna Gurevych

Ubiquitous Knowledge Processing Lab (UKP-TUDA)
Department of Computer Science, Technische Universität Darmstadt

Ubiquitous Knowledge Processing Lab (UKP-DIPF)
German Institute for International Educational Research

http://www.ukp.tu-darmstadt.de/

Abstract

In this work, we analyze the differences between L1 acquisition and L2 learning and identify four main aspects: input quality and quantity, mapping processes, cross-lingual influence, and reading experience. As a consequence of these differences, we conclude that L1 readability measures cannot be directly mapped to L2 readability. We propose to calculate L2 readability for various dimensions and for smaller units. It is particularly important to account for the cross-lingual influence from the learner's L1 and other previously acquired languages and for the learner's higher experience in reading.

In our analysis, we focus on lexical readability as it has been found to be the most influential dimension for L2 reading comprehension. We discuss the features frequency, lexical variation, concreteness, polysemy, and context specificity and analyze their impact on L2 readability. As a new feature specific to L2 readability, we propose the cognateness of words with words in languages the learner already knows. A pilot study confirms our assumption that learners can deduce the meaning of new words by their cognateness to other languages.

Keywords: cognates, lexical readability, readability, second language learning, L2

1. Introduction

> *"At its most basic, [readability] is the ability to make readers continue from the top to the bottom of the page and then turn that page; and then make them do that 200 times in the course of any [...] book."* Curran (2010)

The description given by Curran is probably the most ambitious form of readability, referring to the ability of bestsellers to captivate their readers. For language learners, the readability aims are usually more modest, focusing on the learners' ability to simply comprehend the text. However, if the learner actually can turn the page 200 times without any problem, the text is probably too easy and the learning effect remains very small. In language learning, reading activities are a major means to facilitate the acquisition of new vocabulary (Ellis, 1994) because the learner is necessarily confronted with unknown words. The vocabulary in texts is usually more diverse and more specific than the words occurring in common speech. Therefore, it does not come as a surprise, that limited vocabulary knowledge is the major obstacle for reading comprehension of language learners (Laufer & Ravenhorst-Kalovski, 2010). If the ratio of unknown words is too high, the learner fails to build a semantic representation of the text and very likely gives up to attempt comprehension. In order to support language learning, it is thus very important to select texts with adequate level of readability for the learner: difficult enough to trigger the necessary learning, but easy enough not to frustrate the learner.

Previous work on readability has focused mostly on predicting the readability of texts for native speakers (L1 readability). As a naïve approach, the measures developed for L1 readability could simply be applied on text data for language learners (L2 readability). However, this approach neglects the fact that L1 acquisition usually precedes L2 learning. This sequential aspect has the consequence that L2 learners can already build on previously acquired knowledge and skills. Hence, L1 acquisition and L2 learning are two different processes which results in large differences in the challenges readers face during text comprehension. These differences should be reflected in measures for L2 readability.

If reading activities were coupled with instructive exercises, the knowledge gain for the learner could even be bigger. In order to generate suitable supporting exercises for a text, it is important to not only be aware of the readability of the text but also of the actual elements in the text that cause the difficulties.

> *"In conclusion, although reading for meaning appear to produce significant results in vocabulary acquisition, such reading supplemented with specific vocabulary exercises produces greater gains for the targeted words. This suggests that although instruction makes a difference, more focused instruction is desirable when the learning period is limited and specific vocabulary outcomes are sought."* Paribakht & Wesche (1997)

In order to identify the difficult elements, readability needs to be assessed on different levels. A text with a low readability on the lexical level can then trigger vocabulary exercises and if syntactic challenges are identified, they can be supported with grammar training. In addition, conceptual elements that are unknown to the learner (e.g. cultural traditions) should also receive special attention.

In this article, we summarize the differences between L1 acquisition and L2 learning and analyze the resulting differences between L1 readability and L2 readability. When discussing individual features that contribute to readability, we focus on lexical readability as it has been shown to have a strong influence on L2 reading comprehension (Lervåg & Aukrust, 2010).

2 L1 acquisition and L2 learning

The acquisition of the native language (L1) and the process of learning a second language (L2) progress in different ways. The process of language learning is a strongly debated topic in psycholinguistics. Many terminological distinctions have been attempted and discussed (see DeKeyser (2009) for an overview). In this article, we focus on L2 learning that starts after the L1 has already been acquired or as Jiang (2000) defines it "second language learning with insufficient natural exposure and/or with established L1 linguistic and semantic systems". We do not account for L2 acquisition by bilingual children but focus on more conscious processes as in standard L2 learning settings. In order to emphasize the aspect of consciousness in second language learning (Ellis, 1994), we will distinguish between *L1 acquisition* and *L2 learning*. In this section, we elaborate on several differences with a focus on the lexical comprehension of texts.

2.1 L2 input

The differences between L1 acquisition and L2 learning already emerge from the general setting; the basic L1 knowledge is learned from the unstructured input children receive from their environment. Therefore, their language knowledge is based on experience with direct day-to-day communication and is centered around very concrete situations. The L2, on the other hand, is usually learned gradually by instruction (Cook, Long, & McDonough, 1979) following a more conscious process that also requires more structured input (Schmidt, 1995). The "poverty of the stimulus" is a term coined by Chomsky (1965) initially pointing to the quality of language input children receive from their environment during L1 acquisition. It describes the fact that in spoken language humans often use incomplete or even ungrammatical sentences. In language learning settings, it is generally assured that learners are presented with a more controlled use of language and also receive direct error feedback through corrections. However, Jiang (2000) argues that L2 learners also suffer from the "poverty of the input". In his argument, Jiang does not refer to the quality of the input but to the frequency. Native speakers are confronted with the language they are about to learn all the time and cannot verbally communicate using another language. L2 learners usually only receive input in the foreign language during a few hours per week (Nation, 2003) and are less dependent on it.

It can be concluded that L2 learners have to learn from less frequent, but more controlled language input than native speakers. Thus, the learning process needs to follow more conscious patterns in order to extract more information from less input. For this aim, L2 learners have access to a wider range of learning strategies than children that acquire their L1 (Meara, 1988) and can thus modulate and adapt their learning process. L2 readability should consider these more conscious strategies of dealing with language input and account for linguistic inferences.

2.2 L2 mapping processes

The native language is usually acquired in the first years of childhood, while an L2 is generally learned after the L1. This means that a certain level of proficiency in the L1 already exists. Psycholinguistic researchers are especially interested in the differences of language representation in the L1 and the L2. Many experiments with bilinguals are conducted in order to shed light on the question how the different languages are managed in the brain.

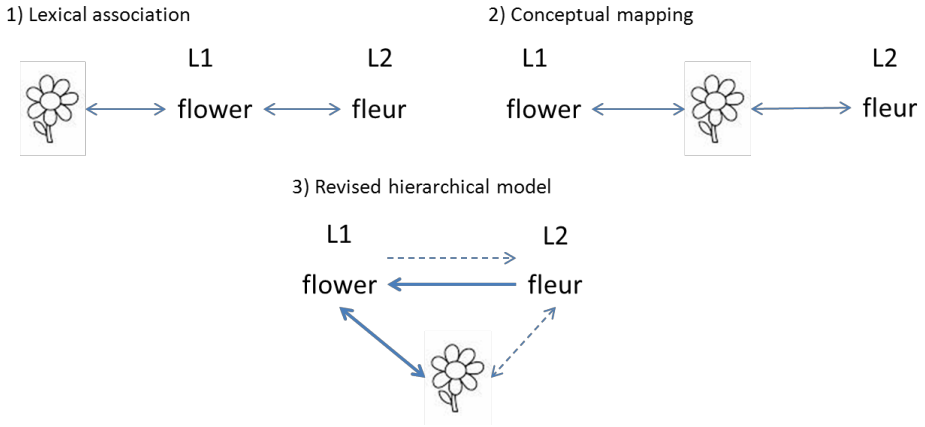

Figure 1: Language representation in language learners. Lexical association (1) and conceptual mapping (2) were both introduced by Potter, So, Eckardt, and Feldman (1984). Kroll and Stewart (1994) combine the two mapping processes into the revised hierarchical model (3) and argue that the different connections have varying strength (the bold arrows signal strong links, dashed arrows stand for weaker links).

Two general mapping models for L2 acquisition can be distinguished: lexical association and conceptual mapping. Lexical association is based on the form of words; learners are assumed to map the L2 form on an existing lexical entry in the L1. The model of conceptual mapping, on the contrary, is based on the idea that the L1 and the L2 word both independently point to the same conceptual representation. Thus, in the first model, learners have to map L2 words on L1 words, while the second model assumes that learners map L2 words on concepts already associated with L1 words. Kroll and Stewart (1994) propose the revised hierarchical model that combines the two ideas. They argue that lexical associations are predominant in early stages of L2 learning and that with increasing L2 competence more direct conceptual links are established (see Figure 1). This revised hierarchical model can explain the finding that translation from L2 to L1 is usually faster than translation from L1 to L2.

> "*Translation from the first language to the second is conceptually mediated, whereas translation from the second language to the first is lexically mediated.*" Kroll & Stewart (1994)

When reading a text in a foreign language, the learner works on L2 words and figures out their meaning in L1. As this is the lexically mediated direction, lexical aspects are particularly important for L2 readability. In a refined version of the revised hierarchical model, Jiang (2000) makes the temporal aspect of lexical development more concrete. He defines three successive stages ranging from pure lexical association at the beginning to conceptual mapping at later stages. Lexical association is thus particularly dominant for less proficient learners. For intermediate to advanced learners both mapping routes – lexical and conceptual – are activated and are used depending on the given cues. Both mapping processes should thus be considered for L2 readability.

2.3 Cross-lingual influence

The mapping models introduced above indicate that L2 learning is influenced by the background knowledge of the learner. As the L1 is already present, basic concepts of languages such as the different behavior of word classes or the syntactic coordination of arguments are already known. In addition, the specific properties of the L1 influence the acquisition of the L2. These phenomena of cross-lingual transfer have been heavily researched on several dimensions. Odlin (1989) discusses cross-lingual effects on grammar, vocabulary and pronunciation acquisition. Jiang (2000) focuses on the negative effects of L1 interference that can be deduced from word choice and usage errors during production.

Ringbom & Jarvis (2009) distinguish between system transfer and item transfer. In system transfer, principles for organizing the elements of the L1 are transferred to the L2. This applies particularly to the morphological behavior of words and to the syntactic combination of words into phrases. Syntactic constructions in the L2 that are parallel to constructions in the mother tongue are more likely to be used by learners (Zobl, 1980). This effect is also reflected in the observation that an L3 (or L4, L5, …) is usually acquired more easily than an L2 (Cenoz, 2003) because multilingual learners already have more linguistic knowledge to rely on. Item transfer refers to the direct mapping of L1 and L2 items. It is mainly a phenomenon in the lexical dimension and particularly important for language comprehension.

> *"The cross-linguistic similarities that underlie item transfer are a concretely perceived similarity of form usually combined with an assumed similarity of function or meaning. […] item learning has a predominantly positive effect on learning, notably on learning for comprehension."* Ringbom & Jarvis (2009)

Cross-lingual item transfer is facilitated for language pairs that contain many cognates (Ellis, 1994)**.** Cognates are words that share a related form with their translation in another language, e.g. *elegance* and *elegancia* in Spanish (see Section "Cognateness of Words" for a more detailed explanation). De Groot and Keijzer (2000) show that cognates are acquired more easily and also persist longer in memory independent of their frequency. They argue that the form overlap of the two words is a strong cue for their successful association.

Cross-lingual item transfer is particularly relevant for recognition tasks where the lexical form is overtly available so that lexical association is strongly facilitated. Reading comprehension is a typical example for a recognition task. Other recognition tasks comprise, for example, the lexical decision whether a word belongs to the language and L2 to L1 translation. However, the cognate facilitation effect persists even across different alphabets, e.g. for Japanese learners of English (Hoshino & Kroll, 2008), and can also be found in the productive picture naming task (Kroll & Stewart, 1994; Lotto & de Groot, 1998). It is assumed that not only the lexical but also the phonetic similarity of cognates contribute to these facilitation effects.

The numerous findings on cross-lingual influence in language learning - and in particular on the cognate facilitation effect - highlight the importance of taking the native language of the learner and also other previously acquired languages into account for the preparation of suitable reading material.

2.4 Reading experience

As L2 learners are older than children acquiring the L1, they also tend to have a more advanced educational background and have already developed higher intellectual abilities (Cook et al., 1979). L2 learners often start reading in the foreign language very early in the learning process while native speakers already have completed a period of several years limited to verbal communication before they face actual texts. Thus, L1 readers still have to learn standard text processing, principles of cohesion and coherence and general linguistic competence, whereas L2 learners can already build on these skills. Meara (1988) argues that most studies on lexical development in children do not manage to disentangle their findings on vocabulary acquisition from cognitive development. For older L2 learners, these processes can be more clearly separated. L2 readability can thus abstract from factors concerned with matching texts to children's intellectual level.

In spite of the intellectual advantages of L2 learners, the vocabulary growth progresses significantly slower than in the L1 (Webb & Chang, 2012). The acquisition of the standard syntactic structures in a foreign language can be successfully completed after a certain period, while the vocabulary acquisition is a continuous process. Lervåg and Aukrust (2010) show that vocabulary size is a critical predictor for reading comprehension by L2 learners. These

findings indicate that texts for L2 learners can be conceptually more difficult than for younger L1 readers but more attention needs to be paid to lexical aspects.

Section summary

As we have seen in this section, L2 learners suffer from a scarcity of language input, which makes it so central to select adequate reading material. As L2 learners have higher cognitive abilities and substantial reading experience, the selected texts might be conceptually more difficult but need to be carefully controlled especially on the lexical level in order to fit the learner's competence level. The importance of the lexical level is reinforced by our analysis of the mapping processes and the cross-lingual influences on L2 learning. Cognates are one lexical element that will play a crucial role in selecting suitable reading material. In the next section, we discuss the consequences of these findings for readability metrics.

3 L1 and L2 readability

We first give an overview of the existing body of work in L1 readability, and then discuss the adaptations that are necessary for L2 readability.

3.1 L1 Readability

Readability measures have a long history, especially in the American education research (DuBay, 2004). The need for these measures is rooted in a very practical task: teachers search for texts that best fit the knowledge level of their students. According to Vygotsky's zone of proximal development (Vygotsky, 1978) the range of suitable texts that a learner can manage without help is very small. Texts, which do not challenge the student, easily lead to boredom, while overly complex language might lead to frustration when no tutoring is available.

Traditional approaches to readability such as the Flesch–Kincaid Grade Level (Kincaid, Fishburne Jr, Rogers, & Chissom, 1975), the Automatic Readability Index (Smith & Senter, 1967) and the Coleman–Liau Index (Coleman & Liau, 1975) date back even until the 1960s and operate only on two main features, word length and sentence length.[i] In addition to the length features, the SMOG grade (McLaughlin, 1969) and the Gunning-Fog Index (Gunning, 1969) also consider the number of complex words defined as words with three or more syllables. DuBay (2004) gives a comprehensive overview of traditional readability measures and Benjamin (2011) evaluates their usability for teachers.

These measures have been incorporated into guidelines for writing standards and are also implemented in standard word processors. However, they have also been subject to criticism as they only capture structural characteristics of the text and can easily be misleading. Consider the two example texts.

1) *There were ten green bottles hanging on the wall,*
   ten green bottles hanging on the wall.
   And if one green bottle should accidently fall,
   there'd be nine green bottles a-hanging on the wall.

2) *This above all: to thine own self be true,*
   And it must follow, as the night the day,
   Thou canst not then be false to any man.
   Farewell, my blessing season this in thee!

Text 1 is a popular children's song and text 2 is a famous quote from Shakespeare's Hamlet. Both texts consist of two sentences and contain 34 words. Most readers would probably agree that the first text is significantly easier to read. However, the traditional readability measures

(see Table 1) actually indicate the opposite. They return lower readability for text 1 (expressed as a higher grade level) because it contains more long words.

Readability measure	Text 1	Text 2
Flesh-Kincaid grade level	6.3	5.6
Automatic readability index	8	4.8
Coleman-Liau index	10.3	6.4
SMOG index	4.4	4.4
Gunning-Fog score	8	8

Table 1: Readability values for text 1 and 2

Recently, progress in the field of text classification has led to a new perspective on readability measures. Supervised learning algorithms automatically combine several text properties extracted from training data and learn to associate them with the corresponding readability class. Naïve Bayes (Collins-Thompson & Callan, 2005), k-nearest neighbors (Heilman, Collins-Thompson, Callan, & Eskenazi, 2007), and support vector machines (Schwarm & Ostendorf, 2005) have all been tested in previous approaches. This automation has led to an explosion of the text features that are taken into account. The features can be grouped in various dimensions focusing on different aspects of the text. In Beinborn, Zesch, and Gurevych (2012), we discuss the lexical, morphological, syntactic, semantic, discourse, conceptual, and pragmatic aspects of readability and give an overview of the features that operationalize them. The features range from simple POS-tag counts over syntactic structures to complex language models and discourse relations. Using these features, it is possible to capture the perceived readability disparity in the two example texts above. More detailed descriptions and results of the individual features can be found in Vajjala & Meurers (2012), Feng, Elhadad, & Huenerfauth (2009), and Pitler & Nenkova (2008). Several researchers also implemented readability calculations for other languages, e.g. French (François & Fairon, 2012), German (Vor der Brück, Hartrumpf, & Helbig, 2008), Portuguese (Aluisio, Specia, Gasperin, & Scarton, 2010), Italian (Dell'Orletta, Montemagni, & Venturi, 2011; Tonelli, Manh, & Pianta, 2012), Swedish (Larsson, 2006), Japanese (Sato, Matsuyoshi, & Kondoh, 2008), and Arabic (Al-Khalifa & Al-Ajlan, 2010). In order to do that, it is not enough to simply adapt the features established for English. The particular characteristics of a given language also need to be considered in the feature selection. Morphological aspects, for example, are less important for English, but crucial for agglutinative languages.

3.2 L2 Readability

Almost all of the approaches discussed above determine the readability of texts for native speakers. A naïve approach to L2 readability could simply apply the existing L1 models on learner texts to calculate the readability. Greenfield (2004) argues that the traditional formulae already give a reasonable approximation for L2 readability and that their ease of application can make up for other shortcomings. On the contrary, Carrell (1987) and Brown (1998) raise the concern that the traditional work on readability cannot capture the different requirements for L2 readability. They particularly mention the need to consider the reader's background knowledge. François and Fairon (2012), Heilman et al. (2007) and Vajjala and Meurers (2012) added some L2-specific features to their readability frameworks. In the following, we will discuss the differences of L1 and L2 readability in more detail.

As mentioned above, L2 learners process the input they receive using more conscious methods. In language learning scenarios, these strategies should be supported in the complementing exercises. L1 readers already master the basics of the language before they attempt to read texts. L2 readers, on the other hand, often fail to build a representation of a text because they focus on unknown words or constructions. Thus, more detailed information about the difficult elements in the text is required in order to generate suitable support. An appropriate readability measure for L2 learning should give more fine-grained information about the readability on at least three different levels. The lexical and the syntactical level correspond to vocabulary and grammar respectively. In addition, difficult text elements on the conceptual level (i.e. concepts, locations, historical events or traditions that are not part of the learner's culture like, for example, thanksgiving celebrations), should also be detected.

Readability is usually determined for full texts with a minimum length of 100 words. For the readability calculation, many text features are extracted and usually normalized by the length of the text. The normalization conceals the local difficulties and aggregates them into a mean readability measure. For L1 readability, this is a reasonable approach, as readers can consolidate the information in order to facilitate the comprehension of a complex sentence and to dissolve misconceptions and ambiguities. For L2 learners, local difficulties can inhibit the comprehension of the full text. If the mapping processes described above fail to map unknown words on known words or concepts, the learner cannot build a representation of the text content because lexical knowledge is crucial for text comprehension (Laufer & Ravenhorst-Kalovski, 2010). Consider for example the following sentence:

> *"He swam up past the boulders made of spongy gishy-gosh and flew right by the herd of floating feeding fipple-fosh."*

This sentence is taken from the book "The Journey of the Noble Gnarble" [ii] by Daniel Errico which was written for young children. It is thus considered to be adequate for inexperienced L1 readers. However, this sentence could be quite problematic for L2 readers, because it contains two irregular past tense words (*swam* and *flew*). The words *boulder*, *spongy* and *herd* might also be unknown as they are quite domain-specific. In addition, the compounds *gishy-gosh* and *fipple-fos*h are made up by the author. They consist of regular English sound patterns but do not actually exist. L1 readers can easily accept that these new words denote new concepts specific to the world described in the book. L2 learners, on the contrary, will try to map the unknown words on concepts they already know. As this process is bound to fail, they will likely get confused about the sentence and blame their insufficient L2 proficiency instead of simply accepting the new words. In order to avoid that these mapping failures lead to learner frustration, it is important to account for the local difficulties and to provide suitable supporting material. It is thus reasonable to compute the readability on the sentence level (Volodina & Pijetlovic, 2013) or give feedback on even smaller units (e.g. return all words that are likely to be unknown to the reader).

As we explained above, the learner's native language background is very important for determining the difficulty of L2 concepts. Learners with the same mother tongue are more likely to experience the same comprehension obstacles because they follow similar analogies. In order to capture the cross-lingual influence, it is necessary to retrieve L2 texts dedicated to learners with a specific mother tongue. In addition, knowledge from previously acquired languages should also be taken into account.

Consider the following example:

> *"**Definition** of **moral**: **Private** conduct **based** on **strict** adherence to a **sanctioned** or **accepted code** or **dogma** of what is right or wrong, particularly as proclaimed in a sacred **book**, or by a **non-secular group** or **sect**."*

This definition of moral[iii] describes a very abstract concept using mainly academic words and is thus hard (or probably impossible) to understand for beginning L1 readers. For L2 learners, it would also be a challenging sentence but they can rely on a bigger set of comprehension strategies here. First of all, they are probably familiar with the concept of moral and already have an approximate definition in mind. Therefore, the content can be more easily mapped on the known concepts. A German native speaker can easily comprehend the words marked as bold because they are cognates and have a similar or even identical spelling in German. In addition, the underlined words *conduct*, *adherence*, *particularly*, *proclaimed* and *sacred* have Latin roots and are shared among many Western European languages. The learner might have come across them in a previously learned language or in scientific texts. If all these words are successfully resolved due to their lexical similarity to other languages, only very basic vocabulary remains. Together with the conceptual understanding, L2 learners should be able to comprehend the sentence much better than inexperienced L1 readers.

From the above discussion, it should be clear that simply using readability metrics developed for L1 acquisition scenarios will not be possible to reliably select suitable text for L2 learning. Four main differences need to be taken into account:

- L2 readability should be calculated for various dimensions in order to account for the more conscious input processing of L2 learners.
- L2 readability should be analyzed for smaller units, so that local mapping difficulties can be detected.
- L2 readability data should be addressed to learners with a common mother tongue in order to capture effects of cross-lingual influence.
- L2 reading grades for language learners can abstract from intellectual text aspects because L2 learners have a better level of reading experience.

These requirements pose challenging demands on the training data. Datasets commonly used to train L1 readability measures (like the Weekly Reader[iv] Corpus) assign school grades to texts. These school grades correlate with the age classes of the learners, while the target audience for L2 learning is more variable and already builds on reading experience as discussed above. L2 readability should thus be expressed in specific learner levels that take the above defined requirements into account. This implies that training corpora need to be annotated on a very detailed level and for many different languages. This kind of data is not available, and given the large number of possible combinations of L1 and languages already known to the learner, it is hard to obtain. Therefore, it is necessary to gain a deeper understanding of what makes text elements difficult for L2 learners with respect to the different language backgrounds. Hence, we base our analysis in the next section on psychological experiments and other findings in the literature. We focus on the lexical dimension and discuss the role of various features for L2 readability.

4 Lexical readability for language learners

While the process of learning the basic syntactic structures can be considered to be more or less completed at a certain point, vocabulary acquisition is a continuous process that remains important even for advanced learners. The vocabulary growth in L2 learning proceeds significantly slower than in L1 acquisition and might even reach a dead end, called lexical fossilization (Jiang, 2000).

For language learners, lexical knowledge is considered to be the major factor for text comprehension (Laufer & Ravenhorst-Kalovski, 2010). Lervåg & Aukrust (2010) show that the vocabulary knowledge can predict the growth of reading comprehension skills for learners and argue that additional vocabulary instruction should be given a high priority in L2

learning. In the following, we discuss the five main lexical features influencing readability and how they capture L2 readability aspects.

4.1 Frequency

The relative frequency of a word approximates the probability that a reader knows the word. More frequent words are considered to be more likely known. In traditional readability measures, the word frequency was approximated by the word length. For English and many other languages, it is a useful assumption, as more frequent words tend to be shorter (Sigurd, Eeg-Olofsson, & Van Weijer, 2004). Unfortunately, this is not always accurate, consider for example the words together (length 8, frequency in the American National Corpus (ANC)[v]: 4004) and sag (length: 3, ANC frequency: 27). Thus, it is better to directly determine the relative frequency of a word, if the necessary means, e.g. big corpora, are available. In addition, the frequency calculation can be adapted to different domains and can thus also predict expert vocabulary. However, the frequency effect is more important for L1 acquisition than for L2 learning. As elaborated in Section 2, the L1 is acquired by direct communication and imitation. More frequent words are heard and used more often and are thus more familiar to the learner. For L2 learners it also holds, that the probability of previously having encountered a word is increased, if the word is more frequent, but it is not the only factor. L2 learning is a more conscious process that is usually moderated by the learning material. De Groot and Keijzer (2000) show that the success of learning completely new words is independent of the frequency of the concept. They presented stimuli to the participants consisting of a word from the native language paired with an artificially created word. The participant was supposed to memorize the new word. As the target words did not actually exist, they were all equally infrequent to the participant. The results show that the frequency of the source word did not have an effect on the learning success nor on the retention of the word. Thus, for understanding completely new words, frequency is less important, other factors such as cognateness of the word, domain familiarity and context specificity (which will be discussed in the following sections) also play a role. As a consequence, relying only on the relative frequency of words for measuring L2 readability will be misleading.

4.2 Lexical Diversity

For L1 readability, the difficulty of a text is also determined by the lexical diversity. If the same concept is expressed with different words, the reader has to recognize the similarity relation of the words in order to understand the shared reference. Lexical diversity is usually measured by the type-token ratio, where type is a word and token refers to the different usages of the word in the text. Graesser and McNamara (2004) have discussed the type-token ratio for their readability framework Coh-Metrix and Vajjala and Meurers (2012) experiment with different variants of it. A low ratio indicates that words are frequently repeated in the text. This characteristic might decrease the stylistic elegance of the text but it facilitates text comprehension. Remember the two example texts in the previous section. The type-token ratio of the rather repetitive children's song is 0.5, while the more complex Shakespeare excerpt has a type-token ratio of 0.9. L2 learners are already advanced readers and can thus handle more complexity on the conceptual level. However, high lexical diversity increases the probability of unknown words and it reduces the possibility to deduce a word meaning from the context it appears in. A text containing many difficult words should thus exhibit less lexical diversity in order to be comprehensible for L2 learners.

4.3 Polysemy

Polysemous words are words that have more than one meaning. For L2 learners, a polysemous word can be particularly confusing as they might know one sense of the word but not be aware of the fact that it can also denote something else. This complicates the interpretation of a sentence because the polysemous words have to be identified and

disambiguated. In addition, it has been found that L2 learners cannot successfully apply guessing strategies for the meaning of an unknown word, if the word is polysemous (Verspoor & Lowie, 2003).
Polysemy of a word can also override frequency information as the different senses are not equally frequent. For example, *well* is a highly frequent word, but the occurrence as a noun in the sense of *fountain* is relatively rare. High polysemy of words is an indicator of low readability for both, L1 and L2 readers, but the effect on L2 readers is even stronger. In Graesser and McNamara (2004), polysemy is determined by the number of synsets of a word in WordNet (Fellbaum, 1998).

4.4 Concreteness

In L1 acquisition, the conceptual mapping of abstract words is more difficult as they do not denote a specific object. Concrete words activate a richer network of semantic information (e.g. mental imagery or wider contextual information) than abstract words (Kaushanskaya & Rechtzigel, 2012). Accordingly, a text is more readable for L1 readers, if it contains many concrete words. If the word mapping models described above apply, word concreteness should be less important for L2 learners because they already know the source concept (remember the example with the definition of moral). According to the revised hierarchical model, the L2 word then only needs to be associated with the L1 word. However, experiments by De Groot & Keijzer (2000) do not support this assumption. They show that concrete words are learned faster and retained longer than abstract words. However, if the abstract words are cognates, the learning success is as big as for concrete non-cognates. A possible explanation could be that the two mapping routes influence each other. If the lexical association is facilitated, the influence of the word concreteness is lower for L2 readability than for L1 readability.

4.5 Context Specificity

When reading a foreign text, learners often guess the meaning of words. These informed guesses are called lexical inferencing (Haastrup, 1991). A learner's ability to perform lexical inferencing strongly contributes to her reading comprehension (Zhang & Koda, 2011). According to the distributional hypothesis, the inferred meaning of a word largely depends on the context in which the word has been presented (Firth, 1957). This hypothesis has received a lot of attention in the field of meaning acquisition in the L1 and has caused a theoretical shift towards more statistically motivated methods. It also initiated many experimental studies like McDonald & Ramscar (2001) who show that the similarity of two words is judged differently depending on the context in which they are presented. Distributional properties can provide an explanation how children overcome the poverty of the stimulus and manage the fast acquisition of new words. For L2 learners, their scarce vocabulary knowledge is a hard obstacle for reading comprehension (Lervåg & Aukrust, 2010) and they have to rely even more on contextual information for lexical inferencing. Very specific contexts can narrow the possible conceptual meaning of a word down so that it can be easily guessed by language learners. Consider the following example:

> *The doctor diagnosed him with Ewing's sarcoma and proposed chemotherapy.*

The phrase *diagnosed him with* is almost always followed by the name of a disease. In combination with the clue *chemotherapy*, the reader can guess that *Ewing's sarcoma* is probably a cancer type. This phenomenon can also explain why language learners feel more comfortable with texts from their expert domain. Even though the text might be more complex and the vocabulary very infrequent, their high domain knowledge facilitates the conceptual understanding and thus enables an easier mapping. Higher order language models as applied in (Schwarm & Ostendorf, 2005) can determine the probability of a word in a given

context. Additionally, the context specificity can be operationalized by using a distributional thesaurus (Lin, 1998), collocation frequencies (Anagnostou & Weir, 2006) and semantic similarity methods (Zesch & Gurevych, 2010). The thesaurus provides information about the number of words that are likely to occur in a given context. If all the candidate words are semantically very close, the context is very specific and the meaning of an unknown word can be derived more easily.

Lexical inferencing on the basis of context specificity is performed by L1 and L2 learners and is conceptually mediated. In addition, L2 learners can also rely on the lexical similarity of the unknown word to a word in their mother tongue in order to infer the meaning. The cross-lingual lexical similarity of two words is called cognateness and we will elaborate on this feature in the next section. Both guessing strategies – lexical and conceptual – might of course also lead to wrong mappings and thus to severe misunderstanding. In order to provide useful supporting material, it is important to be aware of these processes and be able to predict comprehension facilitation by cognates or context specificity as well as probable misconceptions caused by false friends.

5 Cognateness of Words

In the previous section, we discussed how frequency, lexical variation, polysemy, concreteness, and context specificity contribute to L1 and L2 readability. L2 learners have the advantage that they can also rely on their L1 knowledge in order to facilitate lexical comprehension. The cognateness of words is thus a readability feature restricted to L2 (L3, L4, …) readability.

5.1 Cognateness

Cognates are words that have a quite similar lexical form in the L1 and the L2. A strict definition only considers two words as cognates, if they have the same etymological origin, i.e. they are genetic cognates (Crystal, 2011). Language learners usually lack the linguistic background to make this distinction and will use all similar words to facilitate comprehension regardless of the linguistic derivation. For example, the English word *strange* has the Italian correspondent *strano*. The two words have different roots and are therefore genetically unrelated. However, for language learners the similarity is more evident than, for example, the English-Italian genetic cognate *father-padre*. We thus apply a more operational cognate definition for the language learning setting. We consider all word pairs as cognates that are sufficiently similar to be recognized as related by language learners. This similarity facilitates the word association independent of other features such as frequency or concreteness of the word (De Groot & Keijzer, 2000; Lemhöfer & Dijkstra, 2008). For L1 readability calculations, words containing Latin roots will often be considered more difficult because they are usually longer. In addition, they are mainly used in academic contexts and thus have a low frequency in general domain corpora. In L2 settings, on the contrary, these words can facilitate the comprehension for language learners with a romance language background due to their cognateness. Already, Tharp (1939) criticizes that pure frequency counts "ignore the lack of burden of 'gift' words" such as cognates. Ferreira de Souza (2003) studies how cognates contribute to reading comprehension and finds a positive effect for the tasks of skimming and free written recall. In Uitdenbogerd (2005), the readability of French books was judged by English native speakers based on the first 100 words. The results show a correlation between the number of cognates in the text and its assumed readability. The cognateness cues contribute to the advantage of L3 (L4, L5 …) learners over L2 learners mentioned earlier. Multilingual learners can rely on a wider range of linguistic knowledge and can access a bigger set of potential cognates in order to facilitate lexical association.

The cognateness of words is thus an important feature for L2 readability. However, not all cognates actually have the same meaning. False cognates (or false friends) look lexically similar but mean different things. Famous examples for false friends are the German word *Gift* (meaning *poison* and not *present*) and the Spanish word *embarazada* (meaning *pregnant* and not *embarrassed*). False friends have the opposite effect on readability; they lead the reader towards a wrong interpretation of the sentence. Both, true and false cognates need to be detected and incorporated into readability measures.

Cognateness as readability feature

In order to incorporate the cognateness relation of words into the readability calculation, each content word in the text needs to be checked against a list of known cognates from all the languages that the reader knows. In previous work on cognates, the words are manually selected by human experts. In order to automatize this process, the cognates can either be retrieved from existing cognate resources, be determined by applying a cognateness decision on candidate word pairs extracted from translation lists, or be automatically produced using a generation algorithm.

Many different similarity measures have been applied on candidate pairs to distinguish between cognates and unrelated words. The focus is on string similarity measures that assess the orthographic similarity of two words. The longest common subsequence ratio (LCSR), for example, calculates the ratio of the length of the longest (not necessarily contiguous) common subsequence and the length of the longer word (Melamed, 1999). Kondrak and Dorr (2004) apply generalizations of LCSR to bigrams and trigrams for the detection of confusable drug names. Inkpen, Frunza, and Kondrak (2005) and Montalvo, Pardo, Martinez, and Fresno (2012) show that different variations of the DICE measure (originally introduced by Adamson and Boreham, 1974) return the best results. DICE is defined as "twice the number of shared letter bigrams by the total number of bigrams in both words". XDICE (Brew & McKelvie 1996) uses extended bigrams, those are trigrams without the middle letter. Other measures for cognate identification focus on the prefix similarity (Simard, Foster, & Isabelle, 1992) and the consonant similarity (Danielsson & Mühlenbock, 2000).

The string similarity measures do not capture similarity in the pronunciation of words. Different languages often encode the same phoneme with different spelling (e.g. the English word *rice* is almost equally pronounced as the German cognate *Reis*, but their DICE and XDICE values are 0). The phonetic similarity measures SOUNDEX (Hall & Dowling, 1980), EDITEX (Zobl, 1980) and ALINE (Kondrak, 2000) calculate the similarity of two words based on their phonetic features. The LexStat measure (List, 2012) aligns the phonetic transcription of two words to determine their cognateness. Unfortunately, the phonetic measures do not work for all languages. As each single similarity measure described here can only capture a fraction of cognates, they are usually applied in combination as features for supervised machine learning classifiers (Sepúlveda Torres & Aluisio, 2011).

Most cognateness relations follow regular processes (e.g. *vision-visión*, *tradition-tradición*). Not all of these processes can be captured by similarity measures, for example, the pair *accustomed-acostumbrado* has very low string similarity values but can easily be identified as related by language learners. This phenomenon can be captured by approaches that focus on the production rules that transform a word into its cognate in a foreign language. Mulloni and Pekar (2006) learn the edit operation associations from a list of cognates and generalize them to candidate rules. They then assign statistical scores to each rule to measure the association between the left-hand side and the right-hand side of the rule. Similarly, Gomes & Pereira Lopes (2011) use standard string alignment for the extraction and generalization of substitution patterns and introduce the new measure SpSim.

All of the approaches above rely on string alignment and can thus not be applied for language pairs with different alphabets. In Beinborn, Zesch, and Gurevych (2013), we introduce a new algorithm for cognate production (COP) that can be applied to any language pair sharing cognates. In order to account for the regular production patterns of cognates, we apply character-based machine translation. COP is trained on pairs of words and on a language model of the target language. The trained system then applies probable character transformations on the input word in order to produce the target cognate. Combined with a lexicon filter, COP yields very good results for many tested language pairs (e.g. English-Spanish, English-Russian, English-Farsi), even if only limited training data is available.

All the approaches mentioned above can be applied for the identification of cognateness. The choice of an adequate measure for the cognateness decision depends on the available resources and the language pair under study. In order to balance the cognate aspect, it needs to be combined with a feature detecting false friends (Mitkov, Pekar, Blagoev, & Mulloni, 2008; Nakov, 2009). For future research on L2 readability, it would be interesting to see how the cognateness feature interacts with word frequencies and features for context specificity. In order to examine the relevance of cognateness for lexical inferencing, we performed a pilot study with Czech cognates as described in the following section.

5.2 Cognate study

In our pilot study, we examine whether cognates can facilitate the comprehension of unknown words even if no context is available and in the absence of any background knowledge of the target language. We selected 17 Czech-German cognates[vi] and presented them to 15 native German speakers who did not have experience with Eastern-European languages. The participants were asked to name up to 3 guesses for the German translation of the Czech source word. They were not allowed to use a dictionary or any other linguistic material. Table 5 gives an overview of the Czech source words together with the correct German translations and other German associations named by more than one participant.

The results in

Table 2 show that the participants succeed in guessing the right meaning in most of the cases. Some of the Czech words are strongly associated with their correct German translations (e.g. *nudle-Nudel*, *švagr-Schwager*) and do not provoke other associations. For other words, the cognateness relation is not strong enough (*šunka-Schinken*) or the correct cognate is suppressed by stronger false friend associations (e.g. *knoflík –Knoblauch*, *muset-Museum*). The results show that the participants were able to guess the correct meaning of a foreign word without any context information. If they had seen the words within an understandable context, they would probably have an even better intuition for the meaning. Consider, for example, our test word *muset* which is the only verb in the list. Some participants mentioned after the test that the other cognates had primed them for noun meaning, which explains the variety of noun associations named here. In a context that clearly signals the verb property of *muset*, the noun associations would probably have been suppressed.

Another interesting aspect is the influence of languages besides the L1. As we noted before, people rely on all previously learned languages when trying to detect cognates. Our participants all have very good knowledge of English and almost all of them also know French. In addition, each of them (except for one participant) has studied at least one of the languages Spanish, Italian or Latin. This is also reflected in the associations. For example, the German association *Himmel* for the Czech word *cíl* is very likely rooted in the similarity to the French word *ciel* or the Spanish word *cielo*. Both mean *heaven* in English and thus trigger the German translation *Himmel*. The same process might apply when proposing *Salz*

(Spanish: *sal*, French: *sel*, English: *salt*) for *sál*. English translations might have caused the associations for *Aufgabe* (= *task*) for *taška*, *verkauft* (= *sold*) for *žold* and *Bildschirm* (= *screen*) for *skříň*. It should be noted, that all the cross-lingual associations lead to wrong mappings in our examples. These false friends-pairs might cause confusion for the learners and need to be handled separately. Czech is not very close to Western-European languages[vii] which are the background languages of our participants. If they had more experience with Eastern-European languages, the cross-lingual mappings could be more successful.

Czech trigger	**Translation**	**Other Associations**
nudle	Nudel(n) (15)	
švagr	Schwager (13)	
brýle	Brille (12)	brüllen (4)
šlak	Schlag (12)	Schlagsahne (3), schlagen(2)
knedlík	Knödel (10)	gnädig (3), niedlich (2)
žold	Sold (9)	Zoll (5), Gold (2), verkauft (2), Schuld (2)
cíl	Ziel (9)	Himmel (2)
sál	Saal (8)	Salz (13)
taška	Tasche (8)	Aufgabe (4), Tasse (4), Taste (2)
skříň	Schrein/Schrank (5)	Bildschirm/Screen (3), schreien (2)
farář	Pfarrer (4)	Fahrer (7) , fahren (5), Fahrrad (4)
flétna	Flöte (4)	Flotte (4), fliehen (2)
valčík	Walzer (3)	Walze (4), falsch (2)
muset	müssen (3)	Museum (11), Musik (3), Mus (2), Muse (2)
talíř	Teller (2)	Taler (5), zahlen (3), teilen (2)
knoflík	Knopf (1)	Knoblauch (11), knifflig (4)
šunka	Schinken (1)	Sonne (2), schunkeln (2)

Table 2: Results of the pilot experiment - Czech cognates with their correct translations and other associations named by the participants, number of mentions in bracket

The results of the pilot study show that learners are able to guess the meaning of words that are related to their own language (or to other languages they already know) without having received any education in the new language. It confirms the intuition of many language learners and teachers that cognateness of words facilitates their mapping to L1 words and concepts. The ability to guess unknown words also has an impact on the text readability (Zhang & Koda, 2011). A text containing many infrequent words can still be easy to understand for a learner, if the infrequent words exhibit high cognateness.

6 Conclusions

In this work, we analyzed the differences between L1 acquisition and L2 learning and identified four main aspects. First of all, L2 learners receive less language input than children acquiring their L1 but it is more controlled. This allows them to process the input in a more conscious way because they can rely on a wider set of methods. Second, the mapping processes from words to concepts differ for L1 and L2 learners because L2 learners can build on established connections between L1 words and concepts. Third, the existing knowledge in the L1 has a strong influence on L2 learners. Items and structures that are analogous to the L1 are learned more easily and conflicting elements in L1 and L2 can lead to confusion. And finally, L2 learners already have experience with reading activities, so that L2 proficiency is less correlated with intellectual development.

As a consequence of these differences, we conclude that the approaches to L1 readability cannot be directly mapped to L2 readability. More detailed linguistic feedback on at least three dimensions (lexical, syntactical, conceptual) is required in order to adjust the selection of suitable reading material to the more conscious input processing of L2 learners. For the detection of local difficulties, readability should be analyzed for smaller units. It is particularly important to account for the cross-lingual influence from the learner's L1 and other previously acquired languages. As L2 learners already have solid reading experiences, texts can be conceptually more difficult than for L1 learners.

In our analysis, we focused on lexical readability as it is the most influential dimension for L2 reading comprehension. We discussed the existing features frequency, lexical variation, concreteness, polysemy, and context specificity and analyzed their impact on L2 readability. As a new feature specific to L2 readability, we propose to consider the cognateness of new words with words in languages the learner already knows. True cognates facilitate the lexical mapping and thus enable lexical inferencing. A pilot study confirmed our assumption that learners can deduce the meaning of new words by their cognateness to known words. It is important to note that this cognateness relation holds not only for the mother tongue but also for other previously learned languages. However, the study also showed that false friends can cause wrong mappings and thus reverse the effect for readability.

We have implemented the discussed lexical features and use them for predicting lexical difficulty. In future work, we will analyze their usefulness for appropriate text selection for L2 learners. In addition, we will use the prediction of unknown words to trigger useful vocabulary exercises to complement the text.

Notes

[i] The word length is calculated either by the number of characters or by the number of syllables per word.
[ii] Retrieved from http://www.magickeys.com/books/, 20.6.2013
[iii] Retrieved from http://www.businessdictionary.com/definition/moral.html#ixzz2WqEQybmj, 20.6.2013
[iv] http://www.weeklyreader.com
[v] http://www.americannationalcorpus.org/
[vi] The words were selected from http://de.wikipedia.org/wiki/Tschechische_Sprache#Deutsche_Lehnw.C3.B6rter_im_Tschechischen, 1.6.2013
[vii] A language reform in the 19th century eliminated many Czech words with Austrian or German roots.

Acknowledgements

This work has been supported by the Volkswagen Foundation as part of the Lichtenberg-Professorship Program under grant No. I/82806, and by the Klaus Tschira Foundation under project No. 00.133.2008.

Bibliography

Adamson, G. W., & Boreham, J. (1974). The use of an association measure based on character structure to identify semantically related pairs of words and document titles. *Information Storage and Retrieval*, *10*(7), 253–260.

Al-Khalifa, H. S., & Al-Ajlan, A. (2010). Automatic readability measurements of the arabic text: an exploratory study. *The Arabian Journal for Science and Engineering*, *35*(2C).

Aluisio, S., Specia, L., Gasperin, C., & Scarton, C. (2010). Readability assessment for text simplification. *Proceedings of the NAACL HLT 2010 Fifth Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 1–9). Association for Computational Linguistics.

Anagnostou, N. K., & Weir, G. R. S. (2006). From corpus-based collocation frequencies to readability measure. *ICT in the Analysis, Teaching and Learning of Languages* (pp. 33–46).

Beinborn, L., Zesch, T., & Gurevych, I. (2012). Towards fine-grained readability measures for self-directed language learning. *Proceedings of the 1st Workshop on NLP for computer-assisted language learning* (Vol. 80, pp. 11–19). Linköping University Electronic Press, Linköping Universitet.

Beinborn, L., Zesch, T., & Gurevych, I. (2013). Cognate Production using Character-based Machine Translation. *Proceedings of the Sixth International Joint Conference on Natural Language Processing* (pp. 883–891). Asian Federation of Natural Language Processing.

Benjamin, R. G. (2011). Reconstructing readability: Recent developments and recommendations in the analysis of text difficulty. *Educational Psychology Review*, *24*(1), 63–88.

Brew, C., & McKelvie, D. (1996). Word-pair extraction for lexicography. *Proceedings of the second international conference on new methods in language processing*, 45–55.

Brown, J. D. (1998). An EFL readability index. *JALT Journal*, *20*(2), 7–36.

Carrell, P. L. (1987). Readability in ESL. *Reading in a Foreign Language*, *4*(1), 21–40.

Cenoz, J. (2003). The additive effect of bilingualism on third language acquisition: a review. *International Journal of Bilingualism*, *7*(1), 71–87.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.

Coleman, M., & Liau, T. (1975). A computer readability formula designed for machine scoring. *Journal of Applied Psychology*, *60*(2), 283–284.

Collins-Thompson, K., & Callan, J. (2005). Predicting reading difficulty with statistical language models. *Journal of the American Society for Information Science and Technology*, *56*(13), 1448–1462.

Cook, V. J., Long, J., & McDonough, S. (1979). *First and second language learning*. (G. E. Perren, Ed.)*The Mother Tongue and Other Languages in Education* (pp. 7–22).

Curran, J. (2010). *Agatha Christie's Secret Notebooks: Fifty Years of Mysteries in the Making* (p. 496). HarperCollins.

Danielsson, P., & Mühlenbock, K. (2000). Small but efficient: The misconception of high-frequency words in Scandinavian translation. *Envisioning Machine Translation in the Information Future*, 158–168.

De Groot, A. M. B., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, *50*(1), 1–56.

DeKeyser, R. M. (2009). Cognitive-psychological processes in second language learning. In M. H. Long & C. J. Doughty (Eds.), *The Handbook of Language Teaching* (pp. 119–138). Oxford,UK: Wiley-Blackwell.

Dell'Orletta, F., Montemagni, S., & Venturi, G. (2011). READ-IT: Assessing readability of Italian texts with a view to text simplification. *Proceedings of the Second Workshop on Speech and Language Processing for Assistive Technologies* (pp. 73–83).

DuBay, W. H. (2004). The principles of readability. *Impact Information*, 1–76.

Ellis, N. (1994). Consciousness in second language learning: psychological perspectives on the role of conscious processes in vocabulary acquisition, *11*.

Fellbaum, C. (1998). *WordNet: An electronic database*. MIT Press, Cambridge, MA.

Feng, L., Elhadad, N., & Huenerfauth, M. (2009). Cognitively motivated features for readability assessment. *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 229–237). Morristown, NJ, USA: Association for Computational Linguistics.

Ferreira de Souza, V. (2003). *The role of cognates in reading comprehension*. *repositorio.ufsc.br*. Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Firth, J. (1957). A synopsis of linguistic theory, 1930-1955.

François, T., & Fairon, C. (2012). An "AI readability" formula for French as a foreign language. *Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning* (pp. 466–477).

Gomes, L., & Pereira Lopes, J. G. (2011). Measuring spelling similarity for cognate identification. *Progress in Artificial Intelligence*, 624–633.

Graesser, A. C., & McNamara, D. (2004). Coh-Metrix: Analysis of Text on Cohesion and Language. *Behavior Research Methods*, *36*(2).

Greenfield, J. (2004). Readability formulas for EFL. *JALT Journal*, *26*(1), 5–24.

Gunning, R. (1969). The Fog Index after twenty years. *Journal of Business Communication*, *6*(2), 3–13.

Haastrup, K. (1991). *Lexical Inferencing Procedures, Or, Talking about Words: Receptive Procedures in Foreign Language Learning with Special Reference to English*. Gunter Narr Verlag.

Hall, P. A. V, & Dowling, G. R. (1980). Approximate string matching. *ACM Computing Surveys (CSUR)*, *12*(4), 381–402.

Heilman, M. J., Collins-Thompson, K., Callan, J., & Eskenazi, M. (2007). Combining lexical and grammatical features to improve readability measures for first and second language texts. *Proceedings of NAACL-HLT* (pp. 460–467).

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition*, *106*(1), 501–11.

Inkpen, D., Frunza, O., & Kondrak, G. (2005). Automatic identification of cognates and false friends in French and English. *Proceedings of the International Conference Recent Advances in Natural Language Processing* (pp. 251–257).

Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics*, *21*(1), 47–77.

Kaushanskaya, M., & Rechtzigel, K. (2012). Concreteness effects in bilingual and monolingual word learning. *Psychonomic Bulletin & Review*, *19*(5), 935–941.

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel*.

Kondrak, G. (2000). A new algorithm for the alignment of phonetic sequences. *Proceedings of the 1st NAACL* (pp. 288–295).

Kondrak, G., & Dorr, B. (2004). Identification of confusable drug names: A new approach and evaluation methodology. *Proceedings of the 20th international conference on Computational Linguistics* (pp. 952–958).

Kroll, J. F., & Stewart, E. (1994). Category interference in translation andpicture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–174.

Larsson, P. (2006). *Classification into Readability Levels Implementation and Evaluation*. Uppsala University, Sweden.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension, *22*(1), 15–30.

Lemhöfer, K., & Dijkstra, T. (2008). Native language influences on word recognition in a second language: A megastudy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 12.

Lervåg, A., & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of child psychology and psychiatry, and allied disciplines*, *51*(5), 612–20.

Lin, D. (1998). Automatic retrieval and clustering of similar words. *Proceedings of the 17th international conference on on Computational linguistics* (Vol. 2, pp. 768–774). Association for Computational Linguistics.

List, J.-M. (2012). LexStat: Automatic detection of cognates in multilingual wordlists. *Proceedings of the EACL 2012 Joint Workshop of LINGVIS\UNCLH* (pp. 117–125).

Lotto, L., & De Groot, A. M. B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning*, *48*(1), 31–69.

McDonald, S., & Ramscar, M. (2001). Testing the distributional hypothesis: The influence of context on judgements of semantic similarity. *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 611–616).

McLaughlin, G. H. (1969). SMOG grading: A new readability formula. *Journal of reading*, *12*(8), 639–646.

Meara, P. (1988). Learning words in an L1 and an L2. *Polyglot*, *9*(3), 1–11.

Melamed, I. D. (1999). Bitext maps and alignment via pattern recognition. *Computational Linguistics*, *25*(1), 107–130.

Mitkov, R., Pekar, V., Blagoev, D., & Mulloni, A. (2008). Methods for extracting and classifying pairs of cognates and false friends. *Machine Translation*, *21*(1), 29–53.

Montalvo, S., Pardo, E. G., Martinez, R., & Fresno, V. (2012). Automatic cognate identification based on a fuzzy combination of string similarity measures. *IEEE International Conference on Fuzzy Systems*, 1–8.

Mulloni, A., & Pekar, V. (2006). Automatic detection of orthographic cues for cognate recognition. *Proceedings of the 5th international conference on Language Resources and Evaluation*, 2387–2390.

Nakov, S. (2009). Automatic identification of false friends in parallel corpora: statistical and semantic approach, *3*, 133–158.

Nation, P. (2003). The role of the first language in foreign language learning. *Asian EFL Journal*, *5*(2), 1–8.

Odlin, T. (1989). *Language Transfer: Cross-Linguistic Influence in Language Learning*. Cambridge University Press.

Paribakht, T., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. *Second language vocabulary acquisition: A rationale for pedagogy*, 174–200.

Pitler, E., & Nenkova, A. (2008). Revisiting readability: A unified framework for predicting text quality. *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (pp. 186–195). Association for Computational Linguistics.

Potter, M. C., So, K.-F., Eckardt, B. Von, & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, *23*(1), 23–38.

Ringbom, H., & Jarvis, S. (2009). The importance of cross-linguistic similarity in foreign language learning. *The Handbook of Language Teaching* (pp. 106–118).

Sato, S., Matsuyoshi, S., & Kondoh, Y. (2008). Automatic assessment of Japanese text readability based on a textbook corpus. *6th International Conference on Language Resources and Evaluation (LREC)*. Marrakech, Morocco.

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. *Attention and Awareness in Foreign Language Learning*, 1–63.

Schwarm, S. E., & Ostendorf, M. (2005). Reading level assessment using support vector machines and statistical language models. *Proceedings of the 43rd Annual Meeting of the ACL* (pp. 523–530).

Sepúlveda Torres, L., & Aluisio, S. M. (2011). Using machine learning methods to avoid the pitfall of cognates and false friends in Spanish-Portuguese word pairs. *Proceedings of the 8th Brazilian Symposium in Information and Human Language Technology* (pp. 67–76).

Sigurd, B., Eeg-Olofsson, M., & Van Weijer, J. (2004). Word length, sentence length and frequency - Zipf revisited. *Studia Linguistica*, *58*(1), 37–52.

Simard, M., Foster, G. F., & Isabelle, P. (1992). Using cognates to align sentences in bilingual corpora (pp. 67–81).

Smith, E. A., & Senter, R. J. (1967). *Automated readability index*. Cincinnati University Ohio.

Tharp, J. B. (1939). The measurement of vocabulary difficulty. *The Modern Language Journal*, *24*(3), 169–187.

Tonelli, S., Manh, K. T., & Pianta, E. (2012). Making readability indices readable. *Proceedings of NAACL-HLT: Workshop on Predicting and Improving Text Readability for target reader populations* (pp. 40–48).

Uitdenbogerd, S. (2005). Readability of French as a foreign language and its uses. *Proceedings of the Australian Document Computing Symposium* (pp. 19–25).

Vajjala, S., & Meurers, D. (2012). On improving the accuracy of readability classification using insights from second language acquisition. *Proceedings of the 7th Workshop on Innovative Use of NLP for Building Educational Applications (BEA7)* (pp. 163–173).

Verspoor, M., & Lowie, W. (2003). Making Sense of Polysemous Words. *Language Learning*, *53*(3), 547–586.

Volodina, E., & Pijetlovic, D. (2013). Towards a gold standard for Swedish CEFR-based ICALL. *2nd Workshop on NLP for Computer-Assisted Language Learning* (pp. 48–65). Linköping University Electronic Press, Linköping Universitet.

Vor der Brück, T., Hartrumpf, S., & Helbig, H. (2008). A readability checker with supervised learning using deep syntactic and semantic indicators. *11th International Multiconference: Information Society* (pp. 92–97).

Vygotsky, L. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Webb, S. A., & Chang, A. C.-S. (2012). Second language vocabulary growth. *RELC Journal*, *43*(1), 113–126.

Zesch, T., & Gurevych, I. (2010). Wisdom of crowds versus wisdom of linguists–measuring the semantic relatedness of words. *Natural Language Engineering*, *16*(1), 25–59.

Zhang, D., & Koda, K. (2011). Contribution of morphological awareness and lexical inferencing ability to L2 vocabulary knowledge and reading comprehension among advanced EFL learners: testing direct and indirect effects. *Reading and Writing*, *25*(5), 1195–1216.

Zobel, J., & Dart, P. (1996). Phonetic string matching: lessons from information retrival. *Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 166–172).

Zobl, H. (1980). The formal and developmental selectivity of L1 influence on L2 acquisition. *Language Learning*, *30*(1), 43–57.

View publication stats