University of Pardubice

Faculty of Arts and Philosophy

# The Modification of an Authentic English Text for Non-native Students of English

Bachelor Thesis

2017

**Univerzita Pardubice**
**Fakulta filozofická**
Akademický rok: **2015/2016**

# ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

| | |
|---|---|
| Jméno a příjmení: | **Tomáš Kánský** |
| Osobní číslo: | **H13176** |
| Studijní program: | **B7507 Specializace v pedagogice** |
| Studijní obor: | **Anglický jazyk - specializace v pedagogice** |
| Název tématu: | **Modifikace autentického anglického textu pro studenty, kteří studují angličtinu jako cizí jazyk** |
| Zadávající katedra: | **Katedra anglistiky a amerikanistiky** |

Zásady pro vypracování:

Cílem této práce je prozkoumat možnosti modifikace anglicky psaných textů na úrovni vstupu zprostředkovávajícímu obsah těchto autentických textů studentům anglického jazyka různých úrovní. V teoretické části proto student nejprve definuje koncept autentického textu a poté popíše různé způsoby modifikace zohledňující kontext a specifika příslušného jazyka a další kulturní, sociologické a vzdělávací potřeby čtenáře. V praktické části bude na základě předem vydefinovaných kritérií provedena analýza vybraných modifikovaných textů, student zhodnotí kvalitu těchto modifikací a případně navrhne další způsoby modifikace.

Rozsah grafických prací:

Rozsah pracovní zprávy:

Forma zpracování bakalářské práce: **tištěná/elektronická**

Jazyk zpracování bakalářské práce: **Angličtina**

Seznam odborné literatury: **viz příloha**

Vedoucí bakalářské práce: **Mgr. Irena Reimannová, Ph.D.**
Katedra anglistiky a amerikanistiky

Datum zadání bakalářské práce: **30. dubna 2016**

Termín odevzdání bakalářské práce: **31. března 2017**



L.S.

prof. PhDr. Karel Rýdl, CSc.
děkan

doc. Šárka Bubíková, Ph.D.
vedoucí katedry

V Pardubicích dne 30. listopadu 2016

## Příloha zadání bakalářské práce

Seznam odborné literatury:


ANONYMOUS, no date. Grading of Language. In: Grading of Language [online]. [Accessed 14 March 2016]. Retrieved from: http://www.penguinreaders.com/pr/teachers/grading-of-language.html.

CARTER, Ronald, no date. What Is Advanced-Level Vocabulary? The Case of Chunks and Clusters. In: [online]. [Accessed 14 March 2016]. Retrieved from: http://www.tesol.org/docs/default-source/new-resource-library/symposium-on-vocabulary-3.pdf?sfvrsn=0.

CRIADO1, RAQUEL, 2009. The Distribution of the Lexical Component in ELT Coursebooks and its Suitability for Vocabulary Acquisition from a Cognitive Perspective. A Case Study. In: International Journal of English Studies [online]. 2009. Vol. 9, no. 3, pp. 3960. [Accessed 14 March 2016]. Retrieved from: http://revistas.um.es/ijes/article/view/99511/95091.

DEWEERDT, JEAN PAUL, 2001. A defence of simplification. In: Prospect [online]. December 2001. Vol. 16, no. 3, pp. 5565. Retrieved from: http://www.ameprc.mq.edu.au/docs/prospect_journal/volume_16_no_3/prospect_16,3_

GUARIENTO, William and MORLEY, John, 2001. Text and task authenticity in the EFL classroom. In: ELT Journal [online]. October 2001. Vol. 55, no. 4, pp. 347353. [Accessed 14 March 2016]. Retrieved from: http://eltj.oxfordjournals.org/content/55/4/347.abstract.

KESHAVARZ, Mohammad Hossein and ATAI, Mahmoud Reza, 2007. Content schemata, linguistic simplification, and EFL readers' comprehension and recall. In: Reading in a Foreign Language [online]. April 2007. Vol. 19, pp. 1933. [Accessed 14 March 2016]. Retrieved from: http://nflrc.hawaii.edu/rfl/april2007/keshavarz/keshavarz.pdf.

KHANG, Nguyen Duy, 2010. Teachers' Perceptions about Readability and Modication of Authentic Texts Chosen for Teaching Reading in the Vietnamese Context. In: Journal of NELTA [online]. December 2010. Vol. 15, no. 1-2, pp. 8997. [Accessed 14 March 2016]. Retrieved from: http://www.nepjol.info/nepal/index.php/nelta/article/download/4613/3824.

KRASHEN, Stephen, 1989. We Acquire Vocabulary and Spelling by Reading: Additional Evidence for the Input Hypothesis. In: The Modern Language Journal [online]. 1989. Vol. 73, no. 4, pp. 440464. [Accessed 11 March 2016]. Retrieved from: http://www.jstor.org/stable/326879.

KYTE, George C. and NEEL, Virginia M., 1953. A Core Vocabulary of Spelling Words. In: The Elementary School Journal [online]. September 1953. Vol. 54, no. 1, pp. 2934. [Accessed 11 March 2016]. Retrieved from: http://www.jstor.org/stable/998313.

LYNCH, Tony. Communication in the language classroom. Oxford: Oxford University Press, 1996. Oxford handbooks for language teachers. ISBN 0-19-433522-4.

NATION, Paul and CARTER, Ronald (eds.), no date. VOCABULARY ACQUISITION [online]. Nottingham, England: Department of English Studies, University of Nottingham. [Accessed 14 March 2016]. Retrieved from: http://www.aila.info/download/publications/review/aila06.pdf.

NUNAN, David. Designing tasks for the communicative classroom. Cambridge: Cambridge University Press, 1989. ISBN 0-521-37915-6.

O'KEEFF, Anne, no date. VOCABULARY INSTRUCTION. In: MIRR - Mary Immaculate Research Repository [online]. [Accessed 14 March 2016]. Retrieved from: https://dspace.mic.ul.ie/bitstream/handle/10395/1691/o'keeffe, a., (2012)

'vocabulary instruction' (book chapter) (pre-published version).pdf?sequence=2&isallowed=y.
OLMOS, Carmen, 2009. An Assessment of the Vocabulary Knowledge of Students in the Final Year of Secondary Education. Is Their Vocabulary Extensive Enough? In: International Journal of English Studies [online]. 2009. pp. 7390. [Accessed 14 March 2016]. Retrieved from: http://revistas.um.es/ijes/article/viewfile/99531/95111.
ZIPSER, Katharina, 2013. Effective vocabulary building in the classroom environment from B1 onwards. In: Social and Behavioral Sciences [online]. 2013. Vol. 70, pp. 397405. [Accessed 14 March 2016]. Retrieved from: http://www.sciencedirect.com/science/article/pii/s1877042813000785.

Prohlašuji:

Tuto práci jsem vypracoval samostatně. Veškeré literární prameny a informace, které jsem v práci využil, jsou uvedeny v seznamu použité literatury.

Byl jsem seznámen s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorský zákon, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Souhlasím s prezenčním zpřístupněním své práce v Univerzitní knihovně.

V Pardubicích dne 10. 6. 2017

Tomáš Kánský

## ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisor Mgr. Irena Reimannová Ph.D. for her patience and advice as well as to PhDr. Šárka Ježková, Ph.D., Mgr.Eva Nováková and PaedDr. Monika Černá, Ph.D. for their insightful advice.

## ABSTRACT

This thesis deals with the modification of English texts at the input level, which enables the communication of the contents of authentic texts to students of English at various levels of proficiency. The theoretical part is concerned with the phenomenon of authentic text and reading. It then describes grading of language from different perspectives and subsequent possibilities of simplification. In the practical part, selected texts are analyzed according to defined criteria and their suitability for readers at given levels is evaluated.

## KEY WORDS

modification of language, authentic English text, reading, language input

## NÁZEV

Modifikace autentického anglického textu pro studenty, kteří studují angličtinu jako cizí jazyk

## ABSTRAKT

Tato práce se věnuje modifikaci anglicky psaných textů na úrovni vstupu, který zprostředkovává obsah autentických textů studentům anglického jazyka na různých úrovních. Teoretická část se věnuje problematice autentického textu a čtení. Dále popisuje hodnocení obtížnosti textu z různých hledisek a možnost následného zjednodušení. V praktické části jsou analyzovány vybrané texty dle vydefinovaných kritérií, a je zhodnocena jejich vhodnost pro čtenáře na dané úrovni.

## KLÍČOVÁ SLOVA

modifikace textu, autentický anglický text, čtení, jazykový vstup

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# 1. INTRODUCTION

Authentic texts are an important material for the ELT classroom. They can provide the readers with up-to-date information which are relevant to them. However, especially in the case of elementary students of English, there are not enough authentic texts suitable for the readers. The texts are too difficult for them and cannot provide the learners with a comprehensible input, which is essential for language acquisition through, for example, extensive reading, which is an activity to which specifically written or modified texts lend themselves most easily. When one wants to provide learners with information that interests them through a communication channel which they understand, they can reach for a modified text.

Modified texts, however, have been the target of criticism for being "unauthentic" and "unnatural". This stems from the definition of authentic text as being written by native speakers for native speakers. As is explained in sections 2.1–2.3., text which is modified carefully can retain its features of "authenticity", and the binary division of authenticity is no longer relevant. Furthermore, the concept of authenticity lies not only in the text itself, but has much to do with the notion of task.

The third chapter is concerned with grading of language – grading of vocabulary, grammar and structure. Apart from the section 3.4. on grading of structure, several structural difficulties are also described in the eighth chapter. What makes a lexical item difficult is explained in section 3.2. and strategies, both from the perspective of the writer and the reader, on how to deal with difficult words are dealt with in chapter four. Simplifying text lexically often means substituting of one word for another, so synonymy, hyponymy and meronymy are a major part of the thesis (chapters five to seven).

In the practical part, which is the ninth chapter of the thesis, selected modified texts from the website News in Levels are analyzed and evaluated. Employing the English Vocabulary Profile as a reference, the vocabulary of the texts is analyzed and compared to vocabulary lists of the A1 and A2 level. To determine the structural difficulty of the texts, readability of the news articles is calculated using several formulae.

# 2. TEXT AS INPUT AND COMMUNICATION CHANNEL

## 2.1. The Case of the "Authentic Text"

There have been many definitions presented of what an authentic text is, with e.g. Joy (2011) providing an extensive review.

In the 1960s, Lynch (1966) summarized the idea: "The definition that most teachers accept is that authentic texts are samples of language used by and for native speakers" (p. 124). This idea has opened a door for criticism against modified text, calling it unauthentic and unnatural. However, as ISP Nation and Deweerdt (2001) aptly put it: "Many criticisms of simplified texts apply only to poorly simplified texts and to the poor use of such texts in curriculum planning."

The understanding of authentic text as "written for native speakers" was challenged by Widdowson in 1976 (as cited in ISP Nation & Deweerdt, 2001):

> Widdowson pointed out that "authentic" or "authenticity" can be viewed in a different way. Authenticity is not a characteristic of texts, but is the result of the interaction between a reader and a text. If a learner reads a text, and responds to it in a way that we might expect of someone who comprehends the text, then reading the text is authentic for that learner. This response might involve understanding the text, enjoying its message, seeing the strengths and weaknesses in its content and expression, or seeing its contribution to a wider field. As we shall see later, a text that is authentic, in that it was produced for native speakers, may be too difficult for learners to respond to in an authentic way. This view of authenticity is similar to the modern view of validity. A test is not valid in its own right, but is valid when it is used for the purpose for which it was designed and examined.

A text written with the intended reader in mind, whether it's a native-speaking child or a reader to whom English is a second language, does not become unauthentic. It may be more fitting to call such a text "a specifically written text" (Swift, 2009a). To conclude: "The binary division of authenticity is no longer relevant and efficient; rather, it is to be considered in relation to the context where the material is used" (Shomoossi & Ketabi, 2007).

Nuttall (2005, pp. 177-178) mentions authenticity, not paying much attention to its definition (her definition of an authentic text is as a text written for use by the foreign language community, not for language learners) but focuses on the properties of an authentic text compared to a modified or a specifically written one. She sees authentic texts as an important

material for developing the students' ability to understand discourse and the ability to develop their capacity to infer.

## 2.2. The Modification of an "Authentic" English Text

Swift (2009b) divides modification of texts simply as lexical control, structural control, but also adds information control. These three types of control seem to include all of the aspects which are manipulated when modifying a text. Swift (2009a) also points out that modification does not exclusively mean simplification, providing an example of an enhanced text, one enriched with technical terms in order to provide advanced students of Business English with more vocabulary.

While talking of syntactic and lexical modification seems sensible, adding information control may pose certain difficulties. It is impossible to modify a text both lexically and syntactically without any loss or addition of information. The author of a graded reader should try to modify the form of text while keeping as much information as possible. It is trying to rebuild a structure constructed from one material into a structure of the same function from a material of a different kind. Figure 1 is a visual representation of this kind of effort. The blue bubble illustrates the original text. It communicates large amounts of information, but it is also quite difficult to understand. The green bubble is an example of a perfectly modified text; it conveys the same amount of information as the original text but is much simpler. The red bubble, on the other hand, is a visualization of a poorly modified text; it easier to understand by only one degree but most of the information of the original text was excluded.

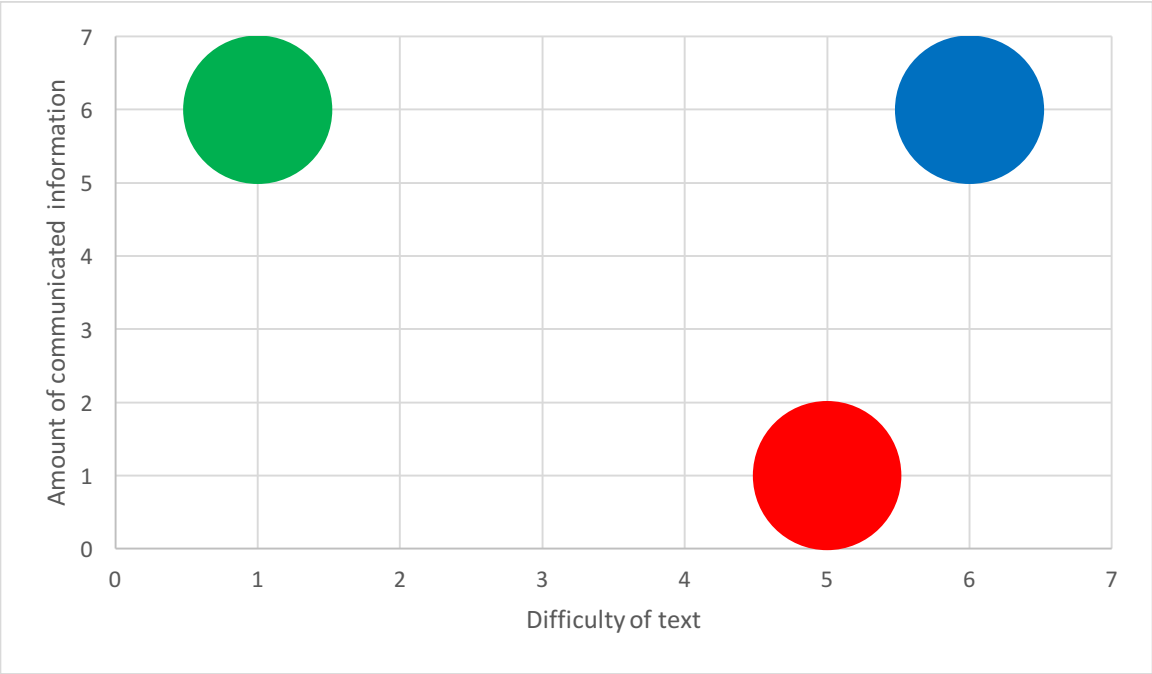


**Figure 1 Visualization of original and modified texts**

It has been agreed that simplification carries risks and must be done carefully and with discretion (Guariento & Morley, 2001; Nuttall, 2005). Difficult words cannot be simply excised, they must be substituted carefully with another word, noun phrase or a clause (Guariento & Morley, 2001). Both Nuttall (2005) and Guariento and Morley (2001) agree modified texts can be useful, with Nuttall (2005) stating that: "Linguistically difficult texts are unlikely to be suitable for developing most reading skills, especially if they result in the use of translation, or any kind of substantial intervention from the teacher" and Guariento and Morley (2001) commenting on the use as well as production of such texts: "While simplification of text, especially for lower levels, is justified, it appears to be difficult to execute seamlessly" (Guariento & Morley, 2001).

## 2.3. Authenticity and Task

When a text is too difficult for a student to read without frustration and subsequent demotivation, teachers have an option to design the task in a way which is feasible to execute by the student. A task like finding all new vocabulary, highlighting the predicate etc. This, however, is not an authentic task. "Some texts lend themselves to this kind of competence training" Guariento and Morley (2001) state: "…but an experienced teacher could argue that learners want to understand a text they are reading" points out Swift (2009a). Indeed, understanding a text being read could arguably be the most authentic task. To conclude in the

words of Guariento and Morley (2001): "'Authenticity' lies not only in the 'genuineness' of text, but has much to do with the notion of task."

## 2.4. The Importance of Content Schemata

Keshavarz and Atai (2007) conducted a study on the effects of syntactic simplification, lexical simplification, syntactic-lexical simplification, and content schemata (content familiarity) on two groups of participants – one at a low degree of proficiency and the other at a high one. As the study shows: "There is a significant effect of the content and EFL proficiency, but not of the linguistic simplification, on reading comprehension and recall" (Keshavarz & Atai, 2007).

Yule (2010) defines schema as a "conventional knowledge structure in memory for specific things" (p. 313). The author of a modified text must take into account the intended reader's background knowledge. He or she, unlike the author, may not have a "supermarket schema" (food displayed on shelves, arranged in aisles, shopping carts and baskets, check-out counter, and other conventional features) as part of their background knowledge (Yule, 2010) and an explicit explanation might be needed.

The author (News in Levels, 2016) of the following paragraph: "People in England love tea. They also dunk biscuits in tea. One man takes this to an extreme" modified from the narration of a video news report (ODN, 2016) seems to have either implicitly or explicitly considered the possibility of the target reader not knowing the phenomenon of "dunking biscuits in tea" and decided to explicitly provide context which is not to be found in the source material.

## 2.5. Shared Assumptions

More general than content schemata is the idea of shared assumptions, e.g. assumptions about the world and the way it works, about moral or political views and so on (Nuttall, 2005). As Nuttall (2005) comments: "Problems arise when there is a mismatch between the presuppositions of the writer and those of the reader."

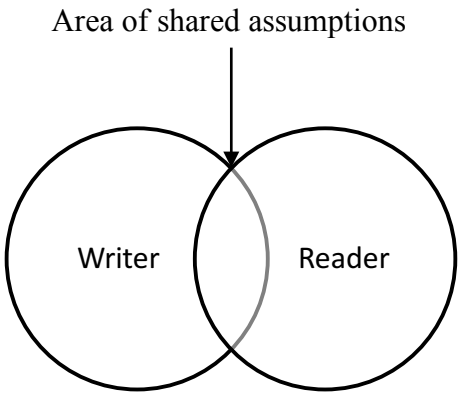


**Figure 2 The writer's and reader's assumptions**

Note. From *Teaching Reading Skills in a Foreign Language*, p. 7, by C. Nuttall, 2005, Oxford: Macmillan.

> Figure 2 is a simple way of showing how, for any two people, some kinds of experience are shared while others are not. The shaded area where the circles overlap represents what the two people have in common. In this area is all the knowledge – including knowledge of language – that they share. It also includes more intangible things like attitudes, beliefs, values and all the unspoken assumptions shared by people brought up in the same society. In the unshaded areas are the things not shared: the experiences and knowledge that are unique to each individual. (Nuttall, 2005)

## 2.6. Reading as a Skill

Reading is one of the four communication skills, which are divided into receptive skills (reading and listening) and productive skills (speaking and writing). Once a commonly-held view that "receptive skills are somehow passive, whereas production skills are in some way more active" (Harmer, 2007) has been questioned and disproved. Obviously, when we speak or write, we are producing language, but receptive skills also require language activation. "We have to think to understand, using any or all of our language knowledge to get meaning from what we are seeing or hearing" (Harmer, 2007). Figure 3 illustrates a fairly widely held view of reading, which is wrong.

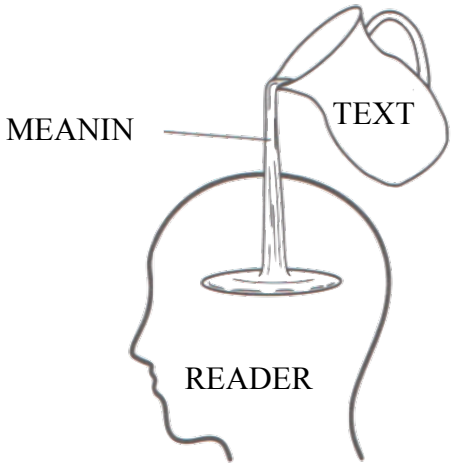


**Figure 3 One view of reading**

Note. From *Teaching Reading Skills in a Foreign Language*, p. 5, by C. Nuttall, 2005, Oxford: Macmillan.

> The text is full of meaning like a jug full of water; the reader's mind soaks it up like a sponge. In this view, the reader's role is passive; all the work has been done by the writer and the reader has only to open his mind and let the meaning pour in. (Nuttall, 2005, p. 5)

In the 1980s and 1990s, there was a shift to an active reader (Wallace, 2001) and then common ground was found:

> More recently the ground has shifted again to talk of reading as "interactive" rather than simply "active". Readers are seen as negotiating meaning; meaning is partial within the

text and writers' intentions may not be privileged over readers' interpretations. (Wallace, 2001)

As writing and reading is a sort of a distance communication between the writer and the reader, seeing reading as interactive may be suitable. Figure 4 adapted from Nuttall (2005, p. 4) illustrates the communication process.

## 2.7. The Writer–Reader Relationship

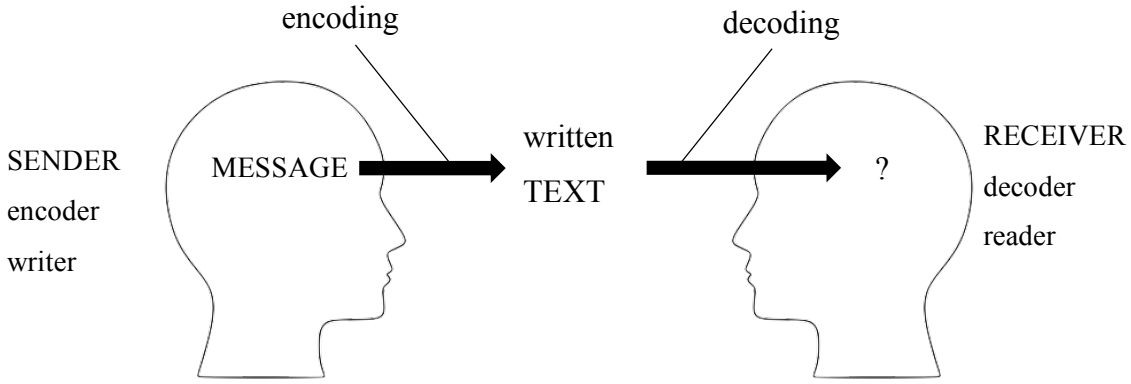


**Figure 4 The communication process**

Note. From *Teaching Reading Skills in a Foreign Language*, p. 4, by C. Nuttall, 2005, Oxford: Macmillan.

As Nuttall (2005) explains, the writer or encoder on the left has a message which they want to share with the receiver. To make this possible, they must encode it – put it into words (Nuttall, 2005). Once encoded, it is available outside their mind as a written text (Nuttall, 2005). The text is then accessible to the mind of another person who reads it, and who may then decode the message it contains: "After being decoded, the message enters the mind of the decoder and communication is achieved" (Nuttall, 2005).

The model is rather simplistic and things can go wrong at any stage, hence the question mark in the decoder's mind (Nuttall, 2005). We cannot be sure whether they have received the message intended (Nuttall, 2005). "However, the process is clear enough for us to say that reading means getting out of the text as nearly as possible the message the writer put into it" (Nuttall, 2005).

To teach reading in a suitable way, a teacher must understand the reading process. According to Wallace (2001), reading, in the most general sense, can be possibly seen as practice, product or process. Reading as a practice "has been the interest of anthropologists and social psychologists whose concern is with reading and writing practices as linked to their uses in everyday life," writes Wallace (2001) and seeing reading as a product means focusing "on the form and meaning of written texts and their constituent parts" Wallace (2001). The third perspective, i.e. reading as a process "pays relatively greater attention to the role of the reader in the ongoing processing of written language and the strategies that she or he draws on in constructing meaning from text" (Wallace, 2001).

### 2.7.1. Bottom-up Reading

For approaches to reading which emphasize text-based features at word and sentence level, the term bottom-up has been used (Wallace, 2001). "In bottom-up processing, the reader builds up a meaning from the black marks on the page: recognizing letters and words, working out sentence structure" (Nuttall, 2005).

> Our image of bottom-up processing might be a scientist with a magnifying glass examining the ecology of a transect – a tiny part of the landscape the eagle surveys. The scientist develops a detailed understanding of that one little area (which might represent a sentence in the text); but full understanding only comes if this is combined with knowledge of adjacent areas and the wider terrain, so that their effects on one another can be recognized. In other words, bottom-up and top-down approaches are used to complement each other. (Nuttall, 2005, p. 17)

A conscious use of bottom-up reading can be made when the initial reading leaves us confused (Nuttall, 2005), and in that case "we must scrutinize the vocabulary and syntax to make sure we have grasped the plain sense correctly" (Nuttall, 2005).

### 2.7.2. Top-Down Reading

In contrast to the bottom-up approach, approaches taking the reader rather than the text as a point of departure have been termed top-down (Wallace, 2001), and they give greater emphasis to what the reader brings to reading.

"In top-down processing, we draw on our own intelligence and experience – the prediction we can make, based on the schemata we have acquired – to understand the text" (Nuttall, 2005).

> We might compare this approach to an eagle's eye view of the landscape. From a great height, the eagle can see a wide area spread out below: it understands the nature of the whole terrain, its general pattern and the relationships between various parts of it, far better than an observer on the ground. (Nuttall, 2005)

### 2.7.3. Interactive Reading

As interactive reading (not to be confused as the interaction between the encoder and decoder mentioned earlier) is understood as shifting from bottom-up to top-down processing continually (Nuttall, 2005). "Now adopting a top-down approach to predict the probable meaning, then moving to the bottom-up approach to check whether that is really what the writer says" (Nuttall, 2005).

## 2.8. Modified English Texts for Extensive Reading

Because modified texts are written with the reader and their level of English in mind, they can serve as a great material for extensive reading (Waring, 2011). Extensive reading is reading for pleasure, enjoyable reading motivating students to read more and learn more, reading which lets them consolidate and strengthen partly known language by encountering it repeatedly (Harmer, 2007, p. 56).

Day agrees with Waring, saying extensive reading (ER) is "i-1" reading. It is reading for fun, it is reading for fluency, reading what the students are interested in, reading as fast as they like, and as much as possible (R. R. Day, 2003, 2007).

The following two figures illustrate the processes which take place in case the reader reads a text which is too hard (Figure 5), or a text at a suitable level (Figure 6).

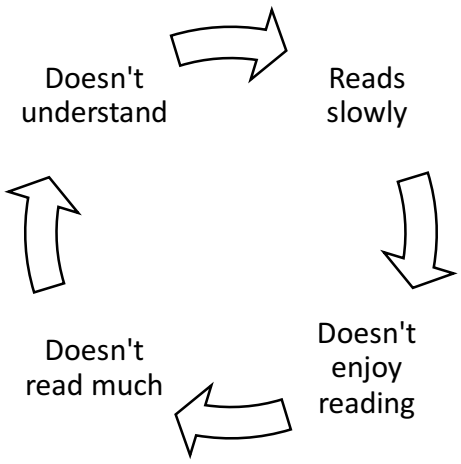


**Figure 5 The cycle of reading frustration**

Note. From *Teaching Reading Skills in a Foreign Language*, p. 127, by C. Nuttall, 2005, Oxford: Macmillan

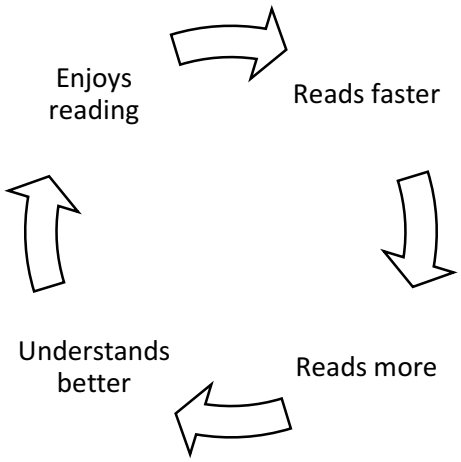


**Figure 6 The cycle of reader's growth**

Note. From *Teaching Reading Skills in a Foreign Language*, p. 127, by C. Nuttall, 2005, Oxford: Macmillan

R. Day and Bamford (2002) also defined top ten principles for teaching reading:

1. The reading material is easy.
2. A variety of reading material on a wide range of topics must be available.
3. Learners choose what they want to read.
4. Learners read as much as possible.

5. The purpose of reading is usually related to pleasure, information and general understanding.
6. Reading is its own reward.
7. Reading speed is usually faster rather than slower.
8. Reading is individual and silent.
9. Teachers orient and guide their students.
10. The teacher is a role model of a reader.

In terms of vocabulary, we can talk about extensive reading when the students know 98% or more of the vocabulary, according to Waring (2011). Figure 7 illustrates this.

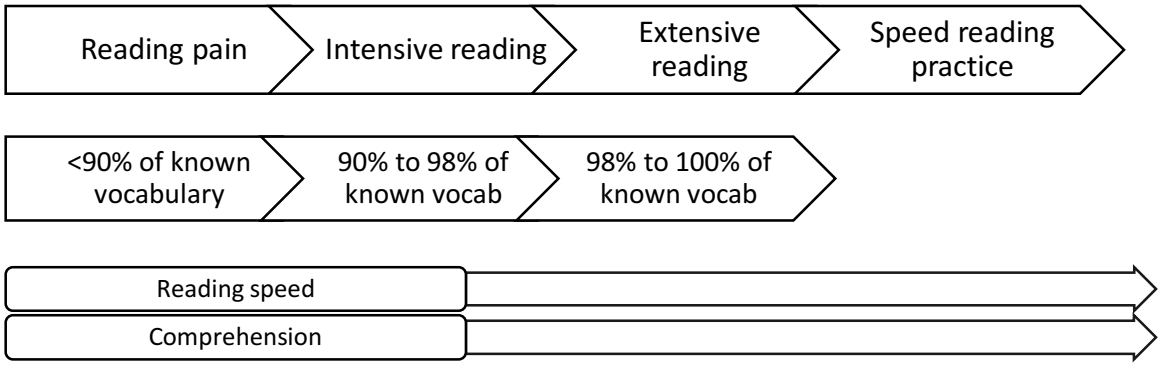


**Figure 7 Lexical density for ER**

The offer of graded readers is such that a reader at any given level may find a text suitable for him or her, starting with phonics and individual words accompanied by visuals, then carrying on to read texts including whatever language has been introduced in the classroom, see Figure 8.

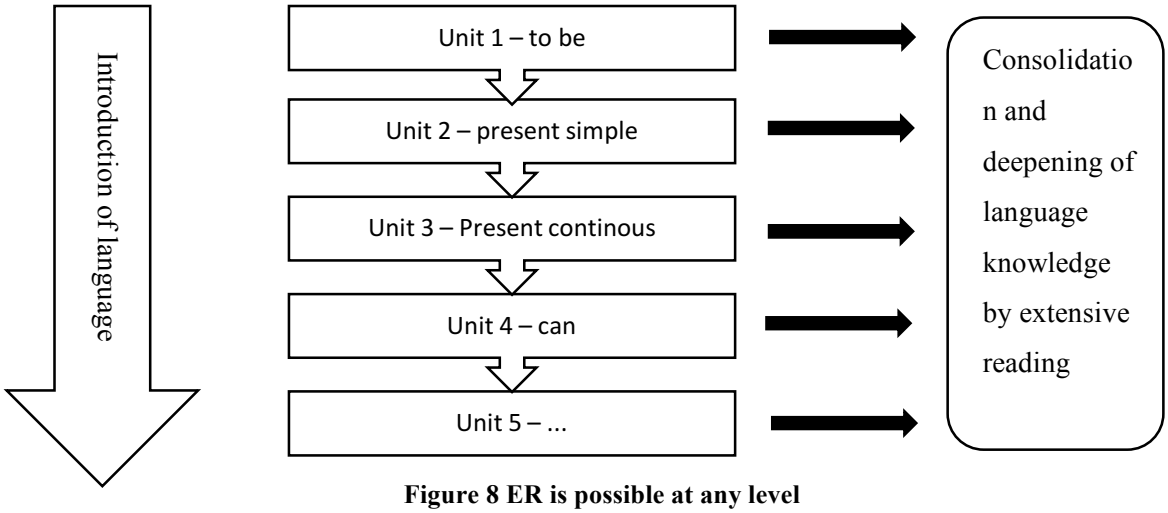


**Figure 8 ER is possible at any level**

Note: Adapted from http://www.robwaring.org/presentations/KAPEE/KAPEE_Vocab_talk_lite.pptx. R. Waring, 2011.

Nuttall (2005, p. 175) agrees with Waring, suggesting that "books for extensive reading should have a smaller proportion of new words – 1 percent perhaps." A high proportion of new words would defeat the aims of a reading program, and "would not serve as an effective way of teaching vocabulary, either, because we learn words best when they occur in a well-understood context," believes Nuttall (2005, p. 175). Two to three percent of new words is "a lot" according to Nuttall (2005) and such texts may be acceptable for intensive reading in her opinion.

### 2.8.1. Vocabulary Acquisition

Waring (2011) presents data from *Sequences* by Heinle Cengage concerning the acquisition of vocabulary in the context of learning from a 5-level "typical" course book, in this thesis adapted into Table 1 and Table 2. Having set the "acquisition" at 20 occurrences – which is arguably quite high – Waring expects students to know 883 words (456 + 202 + 225) by the end of the three years receptively, and 200 words productively. This does not include the learning of collocations, colligations, idioms, phrases, multiple meaning, lexical chunks, sentence heads, etc.

Table 1 presents Cengage's data on the recycling rate of words in a typical 5-level course. The total number of words is 225,000, and the number of different words is 3,239. It is worth mentioning that 40 function words (in, of, the, by, etc.) accounted for 41.2 percent of the total word count in the series.

**Table 1 Typical learning from course books**

| Occurrences | 50 + | 30-49 | 20-29 | 10-19 | 5-9 | 1-4 | Total |
|---|---|---|---|---|---|---|---|
| Different words | 456 | 202 | 225 | 466 | 575 | 1315 | 3,239 |
| | 15.31% | 6.24% | 6.95% | 14.39% | 17.75% | 40.60% | 100% |

Note: Adapted from http://www.robwaring.org/presentations/KAPEE/KAPEE_Vocab_talk_lite.pptx. R. Waring, 2011.

**Table 2 Typical learning from course books plus extensive reading**

| Occurrences | 50 + | 30-49 | 20-29 | 10-19 | 5-9 | 1-4 | Total |
|---|---|---|---|---|---|---|---|
| Different words | 1,023 | 283 | 250 | 539 | 570 | 1,325 | 3,990 |
| | 25.64% | 7.09% | 6.27% | 13.51% | 14.29% | 33.21% | 100% |

Note: Adapted from http://www.robwaring.org/presentations/KAPEE/KAPEE_Vocab_talk_lite.pptx. R. Waring, 2011.

With the addition of one graded a week, a 76 percent improvement in "learnt" vocabulary is seen – from 880 words to 1556 words (1,023 + 283 + 250). This the result of the increased number of words reaching the "acquisition level" – from 27 to 39 percent (25.64% + 7.09% + 6.27%). Waring (2011) also claims that "The students will have a better sense of how the vocabulary and grammar fit together, and have a better sense of collocation, and other deeper aspects of vocabulary acquisition as well as picking up phrases." As it is apparent from the the data, adding a graded reader per week can more than double learners' vocabulary.

Nuttall (2005, p. 62) does not present any data but agrees with Waring:

> The L1 reader did not learn his fifty thousand words by being taught them; most were learnt by meeting them in context. Usually this involved assimilating the meaning gradually, after frequent encounters. In the classroom, students simply do not get enough exposure for this natural assimilation to be possible. Therefore, solutions outside the classroom must be found. An extensive reading programme is the single most effective way of improving vocabulary.

### 2.8.2. Semantic Features

Modifying a text lexically means substituting words and phrases deemed too difficult by words or phrases suitable for the intended reader. This is no easy task. The writer modifying a text must first know how certain vocabulary is graded, and only then can he or she find an easier equivalent. It is a search for a word, phrase, clause or a whole sentence with which the reader is most familiar and which expresses as much of the original idea as possible. Absolute synonymy is rare, but in search for the perfect word in a word-for-word substitution, one could hypothetically compare the semantic features, which could be treated as the "basic elements involved in differentiating the meaning of each word in a language from every other word" according to Yule (2010, p. 114), of the concerned words. Should one for some reason wish to substitute the word *girl* for the word *woman*, they would change the semantic feature +*adult* to –*adult*. Table 3 gives an overview of some of the semantic features in selected words.

**Table 3 Semantic features of *table*, *horse*, *boy*, *man*, *girl* and *woman***

| | **table** | **horse** | **boy** | **man** | **girl** | **woman** |
|---|---|---|---|---|---|---|
| **animate** | – | + | + | + | + | + |
| **human** | – | – | + | + | + | + |
| **female** | – | – | – | – | + | + |
| **adult** | – | + | – | + | – | + |

Note. From *The Study of Language*, p. 114, by G. Yule, 2010, Cambridge: Cambridge University Press.

It remains to be seen how authors of modified texts substitute one word for another, but we can presume it is done implicitly rather than analytically, as that would take a considerable amount of time, although software like AntWordProfiler (Laurence, 2014) for grading and simplifying texts exists.

# 3. GRADING OF LANGUAGE

The Common European Framework of Reference is commonly used in the Czech context to specify the language level of student's books and other materials. The English Profile Programme "has set out to provide the definitive guide to what learners of English know at each CEFR level" (Cambridge ESOL & Cambridge University Press, 2011, p. 58). "It has done this by combining leading educational organizations (Cambridge ESOL, Cambridge University Press, The British Council, English UK), world-leading research institutions (University of Cambridge, University of Bedfordshire, University of Nottingham and other EP Network Partners) and extensive data about real English language use (Cambridge English Corpus, Cambridge Learner Corpus, a range of other corpora from around the world, a wide range of coursebooks, exam specifications and wordlists)" (Cambridge ESOL & Cambridge University Press, 2011, p. 58).

It is mentioned in the *English Profile: Introducing the CEFR for English* (Cambridge ESOL & Cambridge University Press, 2011, p. 58) that while there have been previous attempts to describe English at different CEFR levels, they "have been produced by language specialists largely using their insight as expert users and teachers of the language. However, English Profile's results are based on observed data, providing concrete evidence of what learners throughout the world can do at each CEFR level" (Cambridge ESOL & Cambridge University Press, 2011, p. 58). While the English Profile can serve as a guide to both students and teachers alike, it would seem it was created primarily as a descriptive rather than a prescriptive set (see Appendix S for global scale descriptors for CEFR levels).

## 3.1. Grading of Vocabulary

Publishers of graded readers can specify the vocabulary and grammar used, refer to the CEFR, or do both like for example the Penguin Readers, which provides a comprehensible description of their grading (Penguin Readers). News in Levels, whose texts were adapted for analysis in the practical part of the thesis, on the other hand, only vaguely defines their three levels as each having the 1,000 "most important words" (News in Levels).

### 3.1.1. Vocabulary Levels According to Paul Nation

Paul Nation is Emeritus Professor in Applied Linguistics at the School of Linguistics and Applied Language Studies (LALS) at Victoria University of Wellington, New Zealand. His

specialist interests are the teaching and learning of vocabulary and language teaching methodology. Nation (2012) created the BNC/COCA word family lists, which consist of 29-word family lists – unlike the vocabulary lists adapted for the practical part of the thesis from the English Vocabulary Profile (Cambridge English Language Assessment et al., 2012), which consist of individual words; each treated as its own word family – and can be used for profiling texts using the Range software or AntWordProfiler (Laurence, 2014).

### 3.1.2. Vocabulary Levels According to English Vocabulary Profile

The English Vocabulary Profile (Cambridge English Language Assessment et al., 2012) is a full six-level English profile covering levels A1–C2 for British and American English. It is based on extensive research using the Cambridge Learner Corpus (CLC):

> This is a collection of several hundred thousand examination scripts written by learners from all over the world that is added to every year. Combined with solid evidence of use in many other sources related to general English, such as examination vocabulary lists and classroom materials, the CLC confirms what learners can and cannot do at each level.

It is impossible to grade all the vocabulary and one can easily search for a word and receive no results (see Appendix G for a list of vocabulary not included in the English Vocabulary Profile). Some words may be arguably graded easily as C1–C2 without the help of the profile due to their field of discourse or scarce usage among the general public (e.g. mare, stallion, filly). Others are fairly common (shin, pup, trout, drone, beige, etc.). Certain words may be important to learners of particular nationalities and even though it is a script written by "learners from all over the world" (Cambridge English Language Assessment et al., 2012) and "will not solely be concerned with English as it is spoken in the UK, or in other English-speaking countries" (Cambridge ESOL & Cambridge University Press, 2011, p. 58), some words (*soy, palm, date*) which could arguably be important to a student from a particular region, are missing from the profile or are assigned higher level of difficulty (see Appendix I). The English Vocabulary Profile can very well provide a general sense for grading of vocabulary, which a beginning English teacher may have not yet developed. By no means, however, can it be taken as an absolute for reasons demonstrated. One of the missing features is a list of topics and vocabulary which learners at each level know (something which is often required in the context of the Czech "maturita" exam).

The English Vocabulary Profile also mentions what is meant when students "know a word":

> For us, "knowing" a word is a cumulative process, which implies lifelong learning, as further meanings and uses are acquired. Take for example the word *know*. The English Vocabulary Profile entry for this word stretches from A1 to C2 level, with figurative and idiomatic uses coming in at the higher CEFR levels, suggesting that there are additional 53 meanings and phrases containing *know* that are not acquired until the C levels. (Cambridge ESOL & Cambridge University Press, 2011, pp. 53-54)

## 3.2. Difficult Lexical Items

Apart from words which are higher up the CEFR scale, there are certain types of lexical items which may present a difficulty to an English learner. One of those lexical items are idioms. They are lexical items which consist of several words whose individual meanings are different from the one of the idiom. The English Vocabulary Profile (Cambridge English Language Assessment et al., 2012) lists common idioms but they do not occur in levels lower than B2.

Another group of words which may pose a difficulty are polysemous words and homonymous words. Although a line between polysemy and homonymy cannot be always easily drawn (Peprník, 1992, p. 105), we talk about polysemy in context of a word having two or more related meanings, e.g. *foot* of a person, of a bed or of a mountain (Yule, 2010). Homonyms are words with the same form that are unrelated in meaning, e.g. *mole* on skin and *mole* the animal (Yule, 2010). Intermediate students of English are likely to encounter both polysemy and homonymy, as it is fairly common in English. It could be recommended for teachers to explain their students the nature of polysemy and give a few examples of homonymy.

Hyponymy in regards to difficulty of words is dealt with in chapter 6, but Nuttall (2005) points out another problem which may arise when the student fails to recognize a superordinate or a hyponym referring to one entity. This is especially true when the superordinate and its hyponym are different parts of speech, or when the hyponym is expressed by a sentence or an even longer stretch of text, believes Nuttall (2005, p. 67), providing the following two examples:

1. The diseased limb was cruelly deformed; leprosy has no mercy.
2. A six-year-old boy recently gave a performance of the Beethoven violin concerto. This feat was reported in the press.

The problem of not recognizing that two or more different expressions co-refer may often occur, as there is a preference in English for "elegant variation" as Nuttall (2005, p. 91) observes, and such problems can arise not only in the case of hyponymy but also in the case of synonymy and metaphor.

## 3.3. Grading of Grammar

The English Profile (Cambridge English Language Assessment et al., 2015b) lists not only vocabulary but also grammar in its English Grammar Profile:

> The English Grammar Profile (EGP) is a sister resource to the English Vocabulary Profile, and has been put together by Anne O'Keeffe (Limerick University) and Geraldine Mark, the co-authors, along with Ron Carter and Mike McCarthy, of English Grammar Today (Cambridge University Press). Mark and O'Keeffe investigated the extensive data in the Cambridge Learner Corpus to establish when learners begin to get to grips with different linguistic structures.

Criteria features for A1 level are not in the booklet *English Profile: Introducing the CEFR for English* (Cambridge ESOL & Cambridge University Press, 2011), as "investigation of this level is underway" (Cambridge ESOL & Cambridge University Press, 2011, p. 12), however, a list of features can be found in the online English Grammar Profile (Cambridge English Language Assessment et al., 2015a). The grammar known to an A1 learner is quite limited. However, the learner is (according to the online Grammar Profile) able to:

- Use future with "will" in affirmative form.
- Use past simple in affirmative form.
- Use present continuous in affirmative form.
- Use present simple in affirmative and negative.
- Make questions with "can" and "to be".

On the other hand, the learner is according to the online Grammar Profile unable to use common irregular plural forms, and make questions in present simple. It is worth noting that the English Grammar Profile is written in "can-use" statements and does not mention ability to understand certain grammar. See Appendix J for key distinguishing features of learner English by CEFR level.

## 3.4. Grading of Structure

The *English Profile: Introducing the CEFR for English* (Cambridge ESOL & Cambridge University Press, 2011) also presents structural features that are significant for each level (apart from A1 level). Utterances of an A2-level learner consist mainly of simple sentences (Cambridge ESOL & Cambridge University Press, 2011, p. 12) with a more complex sentence structure consisting of, for example, two clauses joined by *that* (see Appendix K for all of the structural features for levels A2–C2).

> To sum up, as learners progress from level A1 through to B2, they gradually acquire new structures which can be identified as characteristic of each level. Once they reach C levels, learners' progress is characterized by increased structural accuracy and by greater lexical accuracy and range rather than by the addition of new structures to their repertoire. (Cambridge ESOL & Cambridge University Press, 2011, p. 15)

It is also worth mentioning the average length of utterances at each level, illustrated in Table 4.

**Table 4 Average length of utterances at each level**

| **CEFR** | A2 | B1 | B2 | C1 | C2 |
|---|---|---|---|---|---|
| **Avg. length of utterance** | 7.9 | 10.8 | 14.2 | 17.3 | 19 |

Nuttall (2005) deals with difficult structures in chapters 6–7 of her book *Teaching Reading Skills in a Foreign Language*. Some of the structures can be found in chapter 8 of this thesis.

### 3.4.1. Readability Indexes

Readability indexes of a text, "which can be useful but not completely reliable" thinks (Nuttall, 2005, pp. 175-176), can be calculated. Nuttall (2005) mentioned software which counts the number of syllables of each word and the number of sentences of the text. She explains that longer words tend to be more difficult and translate to a more difficult text. Similarly, the fewer sentences there are in a text, the more difficult the text is because fewer sentences means longer, more complex sentences (Nuttall, 2005). To establish the length of words, syllables or individual characters can be counted. Coleman Liau Index and Automated Readability Index (ARI) rely on counting characters, words and sentences. Flesch Kincaid Grade Level, Gunning Fog Score and SMOG Index consider number of syllables as well as complex words (words with three or more syllables). Opinions on which formula is the most accurate vary. Counting

of syllables is most difficult to automate according to The Readability Test Tool (WebpageFX, 2017), as "the English language does not comply to strict standards" (WebpageFX, 2017). The Readability Test Tool, which is used to establish the readability of texts in the practical part of the thesis, calculates the overall readability using all of the mentioned formulae, making it the most comprehensible tool available. See Appendices L–P for more information on each readability index.

### 3.4.2. Discourse Analysis

When somebody begins writing a text, they have something in mind – a body of facts, an emotion, an argument – which they want to communicate and it is up to the writer's decision how he or she organizes the text (Nuttall, 2005, p. 20). The writer can begin with the earliest possible incident, or a later episode and then go backwards to show how the story unfolded (Nuttall, 2005, p. 20). He or she can begin with an example, then drawing a generalization or vice versa (Nuttall, 2005, p. 20). "These choices – the way the meaning in the text are organized to convey the message – are what we mean by *discourse*" (Nuttall, 2005, p. 20). The author of a modified text must consider how to organize the text to make it comprehensible for the intended reader. Narrating a story chronologically may be easier to understand than reporting it through one of the story's characters (as is common in detective stories, for example).

As is shown in the practical chapter 9, section 9.6, the author of a modified text often expands implicit propositions. This may sometimes result in what some call unnaturalness or lengthiness of modified texts, as texts modified to an elementary level of English simply lack the means to express large amounts of information or complex ideas in an elegant way. However, a good author must know which information is essential and which is not. If no information is omitted, the expansion of what is implied or inferred in the original text may result in the modified text actually being longer than the original.

# 4. DEALING WITH DIFFICULT WORDS

## 4.1. Ignoring Difficult Words

Not all words are equally important. Both readers and authors of modified text should be aware of this. Nuttall (2005, pp. 64-66) insightfully mentions how important it is to know when a word may be safely ignored, which is something that many students have never contemplated, as it may seem wrong because it is not done in class (Nuttall, 2005, pp. 64-66). Convincing students that ignoring new words is both acceptable and necessary, sees Nuttall (2005) as the biggest hurdle. To avoid unnecessary substitution or omission of some words, the author modifying a text must be able to tell which words are crucial to the reader's understanding of the text, which words are not, and which can be inferred from the context.

## 4.2. Structural Clues

Ignoring inessential words is the first step but there are other strategies for dealing with lexical items that block comprehension. Another techniques according to Nuttall (2005, pp. 69-77) are the interpretation of structural clues, inference from context, and the use of the dictionary. With the technique of structural clues, students can be taught how to establish at least the grammatical category of a word by looking at its position in the sentence, which can then tell them the kind of meaning to look for (Nuttall, 2005).

Another clue, suggests Nuttall (2005), is the internal structure of a word. She points out that teachers can start making students aware of affixes "as soon as they meet the simple ones such as *UNhappy*, *teachER*, *DISagree*, e*xaminATION,*" but admits serious work can begin when "students have a big enough vocabulary to provide more than a few examples of one affix, bases, and co-occurrence of affixes." Nuttall (2005) concludes that as there is a finite set of affixes and many of them are highly productive: "The analytical approach to morphology pays big dividends in enabling students to work out the meaning of new words." Apart from understanding the structure of complex words, students also need to understand the patterns of compound words, with Nuttall (2005) advising teachers to draw attention to them when encountered in texts.

## 4.3. Inference from Context

While the structural approach is valuable, it may not deliver enough meaning for the reader's purpose, so he or she must make use of the context (Nuttall, 2005, pp. 72-75). Inference is a

skill we all have in our first language and most of the learnt vocabulary is learnt using it; we encounter spoken and written words frequently and in situations we understand, so we are able to infer their meaning, explains Nuttall (2005, pp. 72-75). Nuttall (2005) also stresses the importance of showing students how to infer meaning from context by presenting a sequence of sentences with each specifying an unknown word to the point of rough understanding, as well as demonstrating a limited amount of words, which can fit into certain context. Authors modifying a text must make sure the final text offers enough clues for inference of a certain word, and that there are not too many new words, i.e. the text is not too lexically dense.

## 4.4. Using a Glossary or a Dictionary

If a word cannot be substituted with a suitable option, the writer may choose to highlight the word and define it in a glossary. Deciding when to substitute a word and when to define may be difficult at times, but authors of graded readers should avoid defining too many words, as the reader may be tempted to see the glossary unnecessarily. Generally, it can be said that it is suitable to define a word in the following four cases.

(1) The word is crucial for the understanding of the clause, sentence, paragraph, or the overall message of the reading text. As an example can be given the word *agreement* (B2) in the news about Donald Trump's intention to withdraw from the Paris Agreement written by News in Levels (2017c). As the news is about the Paris Agreement, the word *agreement* is important, and the author wanted everyone to understand it. Describing the agreement and calling it a plan (A2), for example, would be an option, but the reader would be presented with unspecific information and would not be able to search for more information as easily (see Appendix F).

(2) The word cannot be substituted by another one due to a great of loss in meaning. (3) A different word or phrase would sound odd in the given context. The word *thief* found in another text from News in Levels (2017b) is at the B1 level, and the author chose to define it. Had they wished to substitute it instead, some of the most suitable options according to the *Oxford Thesaurus of English* (Waite, 2009) would be (a) *criminal* – B1, (b) *robber* – (rob B1), (c) *burglar* – B2, (d) *pickpocket*, (e) *sneak thief*, (f) *mugger* and (g) *stealer* – (steal A2). The English Vocabulary Profile (Cambridge English Language Assessment et al., 2012) does not list the agent nouns of the options b and g, only the verbs themselves. Even despite the English

Grammar Profile (Cambridge English Language Assessment et al., 2015a) not listing the suffix *-er*, we can presume the words to be at the same level as the verbs.

Options d–f would be too difficult for the intended reader. *Criminal*, *robber* and *stealer* would be possible but *robber* and *stealer* are rarely used and sound odd. The word *criminal* does not really fit into the context either, as the purpose of the sentence "The monkey is a thief" is to highlight an unusual characteristic of the monkey due to what it has done, rather than labelling it as a criminal – a rather vague term which does not tell us anything about the crimes committed, so it would not function as an exclamation in the sentence. However, the whole sentence could be perhaps replaced by a similar exclamation expressing surprise, e.g. "How shocking!" or "What a bad monkey!". See Appendix H for the whole article.

The last reason for defining a word (4) could be the author's wish to draw attention to the word for instructional reasons, for example. Teachers may also recommend students to use dictionaries, but there is a usual tendency to use dictionaries "far too often" believes Nuttall (2005, p. 76), so students need to know when to look up a new word.

## 4.5. How to Define a Word

When a word is defined for one or more of the reasons presented above, it should be defined in a language at a level the intended readers know well. It should also be defined in regards of its context. However, it should not be a placeholder for a paraphrase of the surrounding portion of the text such as the definition of the word *mark* reading "The witch's kiss turns the girl's skin a different color" found in a retold version of *The Wizard of Oz* (Baum, 2007): "It is a long way. But my kiss will protect you. No one will hurt you when they see this *mark*."

# 5. SYNONYMY

## 5.1. Absolute Synonymy

To substitute a word for another, one must often engage in a search for its synonym in a thesaurus. Synonymy in the English language is well developed, as English has adopted much vocabulary from other languages. As Peprník (1992) states, for some words, one can find even more than one synonym, for example: *end – finish – complete – terminate*, or *help – aid – assistance*. Modifying texts would be ease if most synonymous words were equal in their meaning. This, however, is not the case, as absolute synonymy is rare and most synonyms are partial. Peprník (1992) lists the following absolute synonyms: *greenhouse – hothouse*, *kind – sort*, *doorman – doorkeeper*. Cruse (2000) suggests as candidates for absolute synonymy the pairs *sofa – settee*, and *pullover – sweater*.

## 5.2. Synonymy According to Peprník

Apart from absolute synonyms, Peprník and Cruse differ in their classifications. Peprník (1992, p. 101) distinguishes absolute and partial synonyms. Partial synonyms which differ in scope of denotation, connotation and distribution (Peprník, 1992, p. 101). He also recognizes several subgroups of partial synonyms – ideographic synonyms, stylistic synonyms, geographic synonyms and expressive synonyms. Figure 9 illustrates Peprník's classification. As it is apparent, while synonyms may share the central semantical meaning, most synonyms also add or exclude some part of denotation or connotation.

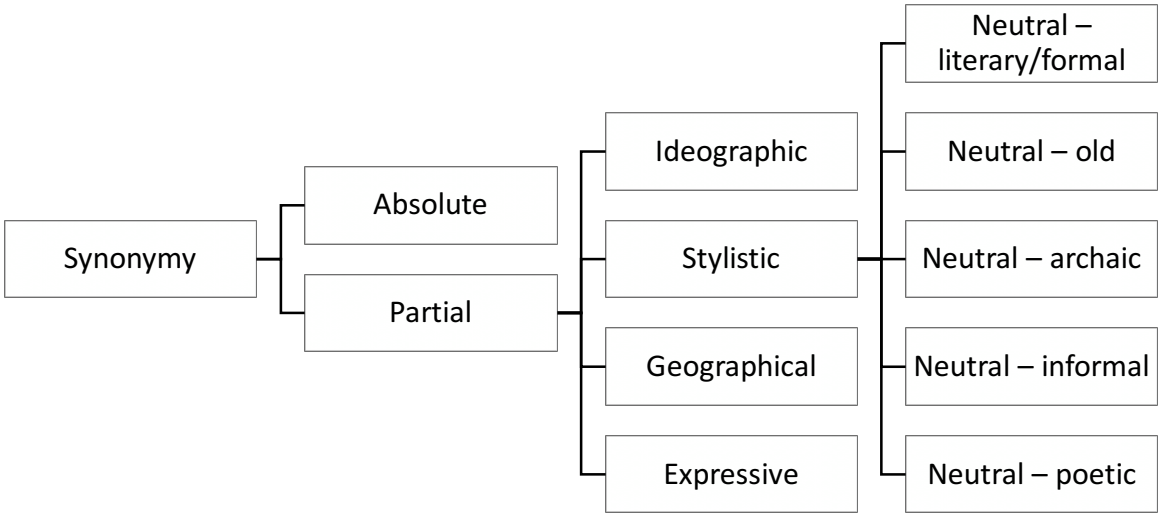


**Figure 9 Synonymy according to Peprník**

### 5.2.1. Ideographic Synonyms

Ideographic synonyms differ in distribution and scope of denotation, i.e. they distinguish nuances in meaning (Peprník, 1992, p. 101). In other words, ideographic synonym is a word that is very similar to another word, but really means a different idea, e.g. the words *skinny* and *thin*. The word *skinny* has a negative connotation, whereas *thin* has a more positive meaning (litteacher8, 2012). Should one then wish to substitute the word *skinny* (which is not included in the English Vocabulary Profile but is arguably more difficult than *thin*) for the word *thin* they would lose the connotation in which these two words differ. The attitude towards the thin person expressed by the word *skinny* may then be expressed in another sentence should the word *skinny* be substituted by simply *thin* or the word may be highlighted and defined as something as "thin but not in a nice way," "too thin," or "unhealthily thin." Which words to substitute and which ones define depends on how much meaning would be changed in case of the substitution and on the importance of the word in terms of its context – see section 4.4.

Table 5 lists some of the examples of ideographic synonyms according to Peprník (1992, pp. 101-102). The following series of tables include the classification of each word and its synonym according to the English Vocabulary Profile (Cambridge English Language Assessment et al., 2012). If a word has not been assigned a classification, it is not included in the English Vocabulary Profile.

**Table 5 Ideographic synonyms and their grading**

| Word | CEFR class. | Synonym | CEFR class. |
|---|---|---|---|
| Nouns | | | |
| blossom | | bloom | |
| expectation | B2 | anticipation | C2 |
| mistake | A2 | error | B2 |
| Adjectives | | | |
| high | A2 | tall | A1–A2 |
| happy | A1 | lucky | B1 |
| rich | A2 | wealthy | B2 |
| Verbs | | | |
| buy | A1 | get | A1 |
| expect | B1 | wait | A1 |

| | | | |
|---|---|---|---|
| pull | A2 | draw | B2 |
| reject | B2 | refuse | B1 |
| seem | B1 | appear | B1 |
| shut | A2 | close | A1 |
| remember | A1 | recall | B2 |

The difference in denotation can be also related to intensity, and greater intensity is sometimes accompanied with expressive or stylistic connotation, see Table 6 (Peprník, 1992, p. 102).

**Table 6 Difference in denotation in terms of intensity in ideographic synonyms**

| Word | CEFR | Syn. | CEFR | Syn. | CEFR | Syn. | CEFR | Syn. | CEFR | Syn. | CEFR |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Nouns | | | | | | | | | | | |
| work | A2 | drudgery | | toil | | labour | C1 | | | | |
| poverty | B2 | penury | | destitution | | | | | | | |
| Adjectives | | | | | | | | | | | |
| large | A2 | big | A1 | enormous | B1 | huge | B1 | vast | B2 | | |
| thrifty | | parsimonious | | niggardly | | miserly | | tight-fisted | | | |
| Verbs | | | | | | | | | | | |
| break | A2 | smash | B2 | | | | | | | | |
| cry | B1 | shout | A2 | scream | B1 | shriek | | yell | B2 | | |
| laugh | A2 | chuckle | | titter | | giggle | C2 | guffaw | | | |
| long | C2 | yearn | C2 | hanker | | pine | | | | | |
| look | A1 | peep | C2 | glance | B2 | gaze | B2 | watch | A1 | view | C2 |

### 5.2.2. Stylistic Synonyms

Peprník (1992) states that stylistic synonyms are based on oppositeness of stylistically neutral or stylistically marked words. In case of literary and formal words, it is usually the case that words of the Romance origin are formal or literary, and words of the Anglo-Saxon origin are neutral. Words of Latin or Greek origin are also more formal than those of French origin, Peprník adds. He also comments on stylistically informal synonyms, saying that creation of phrasal verbs is a great source of informal verbs as well as clipping, e.g. *microphone – mike*.

**Table 7 Stylistic synonyms according to Peprník**

| | Word | CEFR class. | Synonym | CEFR class. |
|---|---|---|---|---|
| Neutral – literary or formal words | | | | |

| | | | | |
|---|---|---|---|---|
| Nouns | | | | |
| 1. | choice | B1 | selection | B2 |
| 2. | people | A1 | folk | B1 |
| 3. | largeness | B1–B2* | magnitude | |
| 4. | marriage | B1 | matrimony | |
| 5. | movement | B2 | motion | |
| 6. | runaway | | fugitive | |
| 7. | half | A2 | semi | B2 |
| 8. | friendship | B1 | amity | |
| 9. | freedom | B2 | liberty | B2 |
| 10. | faithfulness | B1–B2* | fidelity | |
| 11. | happiness | B1 | felicity | |
| 12. | killing | B1 | homicide | |
| 13. | speed | B1 | velocity | |
| Adjectives | | | | |
| 14. | beginning | B1–B2* | initial | B2 |
| 15. | cheeky | B2 | insolent | |
| 16. | deep | B2 | profound | C2 |
| 17. | precise | B2 | accurate | B1 |
| 18. | tearful | | lachrymose | |
| 19. | wild | A2 | ferocious | |
| 20. | wretched | C2 | miserable | B1 |
| 21. | lonely | B1 | solitary | C2 |
| 22. | quick | A1 | rapid | B2 |
| 23. | funny | A1 | ridiculous | B2 |
| Verbs | | | | |
| 24. | admit | B1 | concede | C2 |
| 25. | begin, start | both A1 | commence | C2 |
| 26. | consider | B2 | deem | C2 |
| 27. | leave | A1 | depart | B1 |
| 28. | leave | B1 | abandon | B2 |

Asterisked words (No. 3, 10, 14) were not to be found in the English Vocabulary Profile. However, in row 74 and 76 of the C1 English Grammar Profile (Cambridge English Language Assessment et al., 2015a), it is stated that students "can use nominalized forms in academic or business context to make something more formal" and "can use the '–ing' form of verbs as abstract nouns." The cited two rows state usage of nominalization; understanding the meaning of nominalized nouns (and their possible use as pre-modifiers as is the case with word No. 14 in Table 8) could not be found in the English Grammar Profile. One can only guess the approximate level, and the words were graded as B1–B2.

Should we assume all the literal, formal and poetic words synonymous to the neutral ones are not in the English Vocabulary Profile due to their being "very advanced" or "above C2", 26 out of the 28 synonymous pairs in Table 7 see an increase in difficulty in the neutral–literary relation, and neutral–formal relation. This comes perhaps as no surprise, as one could assume students learn the meaning of words such as *people*, *marriage* and *happiness* first, instead of words such as *folk*, *matrimony* or *felicity*. Similar trend can be clearly seen in the case of neutral–poetic pairs in Table 8. In contrast, informal words are not considered easier by the English Vocabulary Profile (Cambridge English Language Assessment et al., 2012) when compared to their neutral counterparts. Four of the 10 pairs saw an increase in difficulty in the neutral–informal relation, and only one pair saw a decrease in difficulty. The neutral and informal words in the remaining 5 pairs were of the same difficulty according to (Cambridge English Language Assessment et al., 2015b). To conclude, it is evident that stylistically marked words are more difficult than their neutral counterparts.

**Table 8 Stylistic synonyms**

| | Word | CEFR | Synonym | CEFR |
|---|---|---|---|---|
| Neutral – old fashioned | | | | |
| 1. | anorak | | windcheater | |
| 2. | mirror | A2 | looking glass | |
| 3. | record player | | gramophone | |
| Neutral – archaic | | | | |
| 4. | I think | | methinks | |
| Neutral – informal | | | | |
| 5. | abandon | B2 | give up | B1 |

| | | | | |
|---|---|---|---|---|
| 6. | appear | B1 | turn up | B2 |
| 7. | choose | A1 | pick out | B1 |
| 8. | continue | B1 | go on | B1 |
| 9. | extinguish | | put out | B1 |
| 10. | suggest | B1 | put forward | C1 |
| 11. | overtake | B2 | get ahead | |
| 12. | sink | B1 | get down | B1 |
| 13. | enter | A2 | come in | A2 |
| 14. | escape | B1 | get away | B2 |
| Neutral – poetic | | | | |
| Nouns | | | | |
| 15. | anger | B2 | ire | |
| 16. | clothes | A1 | array | |
| 17. | forehead | B1 | brow | |
| 18. | girl | A1 | maiden | |
| 19. | horse | A1 | steed | |
| 20. | lawn | C1 | green sward | |
| 21. | meadow | | mead | |
| 22. | morning | A1 | morn | |
| 23. | sea water | A1 | brine | |
| 24. | valley | B1 | vale | |
| Adjectives | | | | |
| 25. | beautiful | A1 | fair | |
| 26. | brave | B1 | doughty | |
| 27. | strange | A2 | uncouth | |
| 28. | unhappy | A2 | hapless | |
| Verbs | | | | |
| 29. | hear | A1 | hearken | |
| 30. | kill | | slay | |
| 31. | see | A1 | behold | |
| 32. | stay | A1 | dwell | C1 |
| 33. | think | A1 | deem | C2 |

### 5.2.3. Other Types of Synonymy

Peprník (1992) also mentions geographic synonyms – words of British, American, Scottish and Australian English, e.g. *tap – faucet*. Expressive synonyms are based on the oppositeness neutral–expressive, e.g. *hard work – drudgery* (Peprník, 1992).

## 5.3. Synonymy According to Cruse

Cruse (2000) categorizes synonyms as absolute, propositional and near. A pair of words are propositional synonyms if they can be used in any truth functional expressions without the change of the value of the sentence:

| | |
|---|---|
| Mary took the *can* from the pantry. | John loves his *violin*. |
| Mary took the *tin* from the pantry. | John loves his *fiddle*. |

The most important differences, two or more of which usually come into play at one time, are the differences of expressive meaning, differences of stylistic level (on the colloquial-formal dimension), and the differences of presupposed field of discourse (Cruse, 2000). In the sentences "He broke his *shin*" and "He broke his *fibula*" the difference comes from the field of discourse.

Near synonyms, also called notional synonyms, lie on the border between synonymy and non-synonymy and establishing the line between propositional synonymy and near-synonymy, and near-synonymy and non-synonymy may not be straightforward (Cruse, 1987, pp. 84-195; 2000, pp. 156-161).

# 6. HYPONYMY

More often than not, it is possible to substitute a word for its hyponym or its superordinate. In some cases, the hyponyms tend to be the easier words. The difficulty of words is according to the English Vocabulary Profile (Cambridge English Language Assessment et al., 2012).

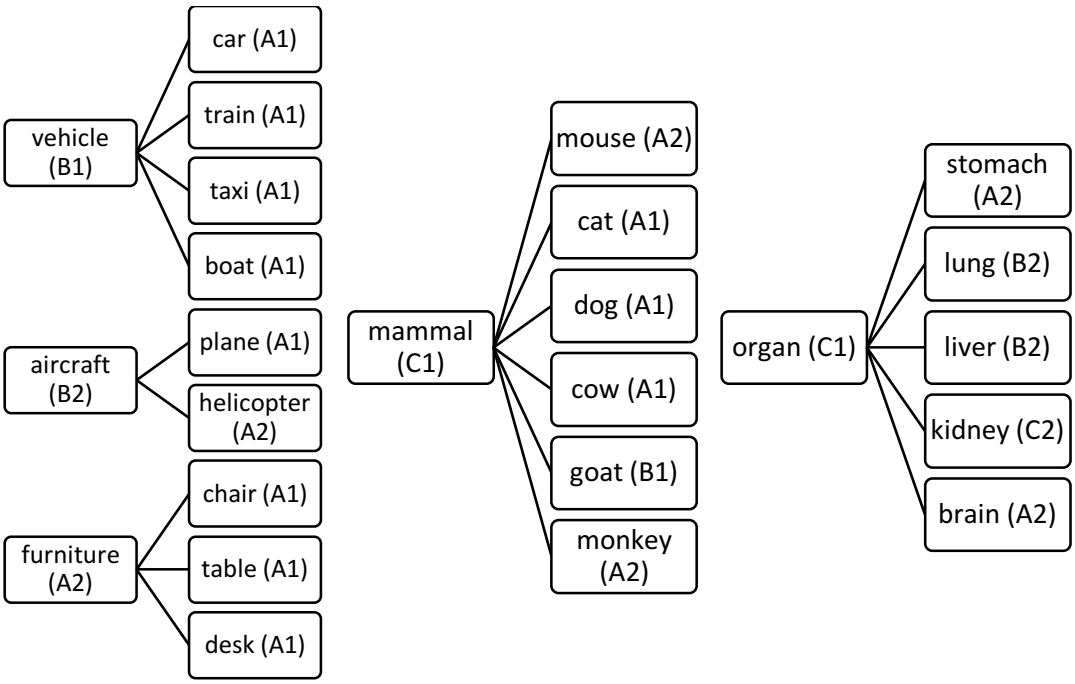


**Figure 10 Superordinates can be more difficult**

In the case of relation of some words to their superordinates (certain verbs, specific names of flowers, trees, dog breeds, names of young animals or males and females), it is usually the other way around; the superordinate is the easier word. The words lacking a classification are not listed in the English Vocabulary Profile (Cambridge English Language Assessment et al., 2012), but could be considered C1–C2 due to their specific meaning and arguably scarce usage among the general public.

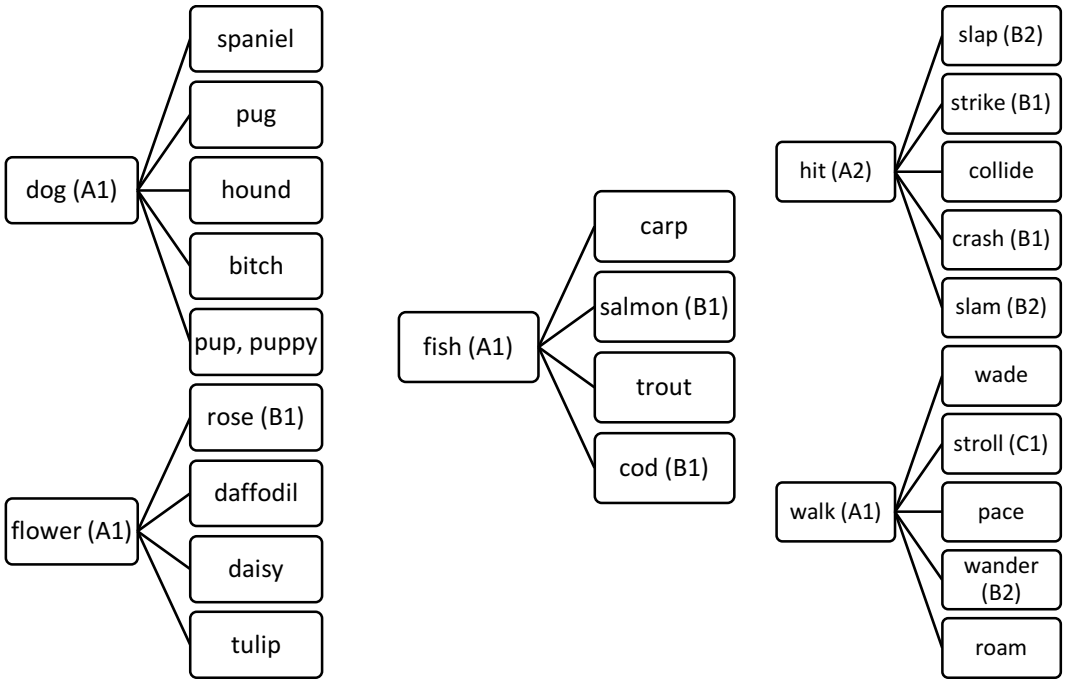


**Figure 11 Hyponyms can be more difficult**

Hyponymy is a relation of inclusion, however, as Cruse (2000) says: "What includes what depends on whether we look at meaning extensionally or intensionally."

> From the extensional point of view, the class denoted by the superordinate term includes the class denoted by the hyponym as a subclass; thus, the class of fruit includes the class of apples as one of its subclasses. Looking at the meanings intensionally, we may say that the meaning (sense) of apple is richer than that of fruit and includes, or contains within it, the meaning of fruit. (Cruse, 2000)

For this reason, one must be careful when substituting one word for another; mind not to lose too much meaning in given context. It is usually more suitable to be specific rather than vague.

# 7. MERONYMY

Meronymy is another relation of inclusion, which is the lexical reflex of the part-whole relation, e.g. *hand – finger*, *teapot – spout*, *car – engine*, *tree – branch* etc. (Cruse, 2000). As Cruse (2000) explains, *finger* is the meronym (or partonym) and *hand* the holonym in the case of *hand – finger* relation. As with hyponymy, the idea of inclusion in different directions depends on whether one takes an extensional or an intensional view (Cruse, 2000). Concerning grading of vocabulary, the situation is the same as in the case of hyponymy; either a meronym or its holonym may be the easier word, or be assigned the same level of difficulty.

# 8. STRUCTURAL DIFFICULTIES

The previous section was concerned with lexical items, their properties and their control. The following sections deal with syntax, discourse and structural control, mainly according to Nuttal's chapters 6–7; grading of structure is also dealt with in section 3.4. It is important that structure of text corresponds with the level of the intended reader as well because comprehension can be blocked even when all the vocabulary is familiar (Nuttall, 2005, p. 175). Although the reading lesson is not the place for extensive grammar teaching, reading does require grammatical skills, so it may be suitable for the teacher to explain whatever structures his or her students will encounter.

Nuttall (2005, p. 175) claims new grammatical forms such as tenses and structural words often cause no problems, but what is most likely to cause problems is structural difficulty, sentence length and complexity. In case of longer sentences, more needs to be remembered or pinned-down by the reader (Nuttall, 2005, p. 92).

Generally speaking, to simplify a text syntactically, Nuttall (2005, p. 82) suggests removing all the optional parts until only the essentials remain and then one can restore the optional elements one by one, fitting them into the structure carefully so students can make sense of them. The following is a list of textual structures which may pose a difficulty.

## 8.1. The Concealed

### 8.1.1. Recognizing and Interpreting Pro-forms

Pro-forms such as *it*, *our*, *this*, *those*, *one*, *such*, *he*, *she*, *they*, *their* may pose problems when the reader cannot easily identify the referent from the text. The author uses pro-forms to avoid unnecessary repetition and while the referents of most pro-forms may seem obvious to an experienced reader, it is not always the case for beginning readers of English. Repetition of specific words instead of using pro-forms may be also useful in terms of vocabulary acquisition (see subsection 2.8.1). The following two sentences from the modified news (News in Levels, 2017a) about the London Bridge Attack, a terrorist attack which took place in June of 2017, illustrate a correct use of referencing. If *the suspects* and *these* were both replaced with *they*, the sentence would still be comprehensible to an experienced reader but would pose a difficulty otherwise.

Original:
Eight minutes after the attack began, policemen shot the suspects dead. *The suspects* were wearing what looked like explosive vests, but *these* were later established to be hoaxes.

Changed for the sake of bad example:
Eight minutes after the attack began, policemen shot the suspects dead. *They* were wearing what looked like explosive vests but *they* were later established to be hoaxes.

The first *they* refers to the suspects but a reader may need to stop to think to make sure it was not the policemen. He or she would not find the answer in the text because the author would presume the reader to have certain schemata and assumptions. The second *they* may pose another problem, especially if the reader does not understand the word *hoaxes*. The difficulties of recognizing pro-forms is also dealt with in section 3.2. To conclude, it could be said that if a text is to be easy, it must always be apparent who does what in a sentence.

### 8.1.2. Elliptical Expressions

The English language has economizing tendencies, and people prefer "to omit rather than repeat information that the reader's common sense can readily supply" (Nuttall, 2005, p. 89). Ellipsis is "the omission from speech or writing of a word or words that are superfluous or able to be understood from contextual clues" (Stevenson, 2010). There is little difficulty in short simple sentences like the ones found in Nuttall (2005, p. 89) "The days are hot and the night cool" and "They came although they were asked not to" but "I carried the bag and my friend the suitcase" and "He told us where it was hidden and despite the disapproving glances of the others promised to show us the way" can be more challenging.

## 8.2. The Condensed

This section lists and explains syntactical structures which incorporate large amounts of information into a relatively short piece of text and are therefore difficult to interpret. The general trend of the author is to identify and expand such structures so they are presented to the reader in a clear manner.

### 8.2.1. Nominalizations

Nominalization is the process of making a noun out of a verb, adjective or a whole clause. While nominalization is a powerful means for expressing complex abstract ideas, overusing it means blocking comprehension not only for students of English. Sword (2012) likes to call nominalizations "zombie nouns" because when there are too many, they block comprehension:

"The proliferation of nominalizations in a discursive formation may be an indication of a tendency towards pomposity and abstraction." The sentence contains seven nominalizations but because the reader cannot easily identify the subject and the verb, the theme and rheme, the sentence fails to tell us who is doing what. Luckily, Sword (2012) shows us how the at first seemingly complicated sentence can be modified: "Writers who overload their sentences with nominalization tend to sound pompous and abstract."

Nuttall (2005, pp. 83-84) focuses more on noun groups created from a clause. In such cases, the noun phrases conceal "unstated propositions." Nuttall's example can be used: "The implementation of the recommendation that child allowances should be restricted to the first three children was delayed for several years." *Implementation* conceals: "Someone recommended that child allowances should be restricted" and *recommendation* "Someone implemented the recommendation."

It is apparent that nominalized clauses pack large amounts of information. In order to make the text more comprehensible, it may be suitable to phrase the unstated propositions: *It was recommended that child allowances should be restricted to the first three children*. *The implementation of this recommendation was delayed for several years.* Advanced students of English can learn to recognize nominalizations and establish their underlying propositions – a job which the author of a modified text intended for elementary learner must do.

### 8.2.2. Modifying structures

Anything which contributes to complexity and length of sentences, such as relative clauses, nominal clauses, participle, infinitives, post-modification and pre-modification, etc., can potentially be a problem.

## 8.3. The Misplaced

The line between what is condensed and what is misplaced is not clear, but it can usually be seen which is predominant. Nominalized clauses would represent almost the perfect case of the condensed and using coordinating conjunctions in a confusing manner would be the misplaced. The location of participle, infinitive or preposition clauses and phrases may seem confusing to an elementary reader but a native reader or a linguist would naturally classify them as

condensations – indeed, subordinate clauses, both finite or non-finite, can be condensed into participles or infinitives.

### 8.3.1. Confusing Usage of Coordinating Conjunctions

Conjunctions such as *and*, *but* and *or* coordinating noun groups are not likely to cause difficulty, but they can also join clauses. Problems arise when there is a text where it is hard to determine which items are joined, as Nuttall (2005, p. 82) illustrates: "He looked at the child with surprise that he should know such words at his age and indignation that he should be permitted to use them." The item that parallels *indignation* is not *age* nor *words*, but *surprise* (Nuttall, 2005, p. 82). Rewriting may help the reader understand.

# 9. ANALYSES OF TEXTS

## 9.1. Quantitative Analysis of Vocabulary

For the quantitative analysis of vocabulary, three sets of modified texts were adapted from the website News in Levels, which modifies authentic material for the purposes of students of English. The website offers the news articles in three different levels (level 1, level 2 and level 3). All of the three levels are modified texts. The three corpora were adapted from the three levels of news articles published from February 20th to June 5th of 2017. As there are two articles published every workday, each corpus consists of 201 news articles.

For the purposes of the vocabulary analysis, all the articles were reduced to only the body of the articles – definitions of the "difficult words" were omitted as well as the reference to the original source of the text, and the titles of the articles. The definitions of the articles are not analyzed in this thesis. The complete corpora are available for download at https://uloz.to/!epWaDP1Q1eKp/corpora-zip, protected by the password *corpora*. For an idea of the nature of the texts, first six articles are appended and referenced in subsections 9.1.1–9.1.3.

Each level is titled the same (see Appendix E) so it would make sense for the titles to be comprehensible for readers at each level. One title is on average 3.1 words long, and the average percentage of words known to a learner, according to the English Vocabulary Profile, at the A1–A2 level is 55.9 percent.

Word (Microsoft, 2017) and AntWordProfiler (Laurence, 2014) differ in the word count of texts. However, the difference is negligible (approximately 0.6 percent). In the following analyses, data provided by the AntWordProfiler is presented.

There is not much specification regarding the vocabulary used for certain levels at News in Levels, apart from stating that: "Level 1 has the 1000 most important words. Level 2 has the 2000 most important words, Level 3 has the 3000 most important words" (News in Levels).

The corpora are compared to 1,465 tokens of the A1 and A2 English Vocabulary Profile; 594 words were adapted from the A1 English Vocabulary Profile and 871 from the A2 level (see Appendix T for the complete list of A1–A2 vocabulary list used for the analysis). These do not

include all the lexical items of the profile, as phrases had to be omitted due to the inability of the software to search text for phrases (for an idea about the nature of the phrases, see Appendix B). The phrases provide more information in terms of the context in which the vocabulary can be used by the learner at the particular level, and their absence should not therefore be of much influence on the analyses.

### 9.1.1. Quantitative Analysis of Level-1 News Articles

The whole corpus consists of 14,170 tokens, meaning one news article is on average 71 words long (see Appendix C for the first six articles). Out of the 14,170 tokens, 9,876 are included in the A1–A2 vocabulary profile, which makes it 69.70 percent. When compared with A1 words only, there is a 61 percent concordance (8,644 words), meaning A2 words account for 8.7 percent of the analyzed vocabulary (1,232 words). That means that 30.30 percent of the words (4,294 words) are more difficult than the A level.

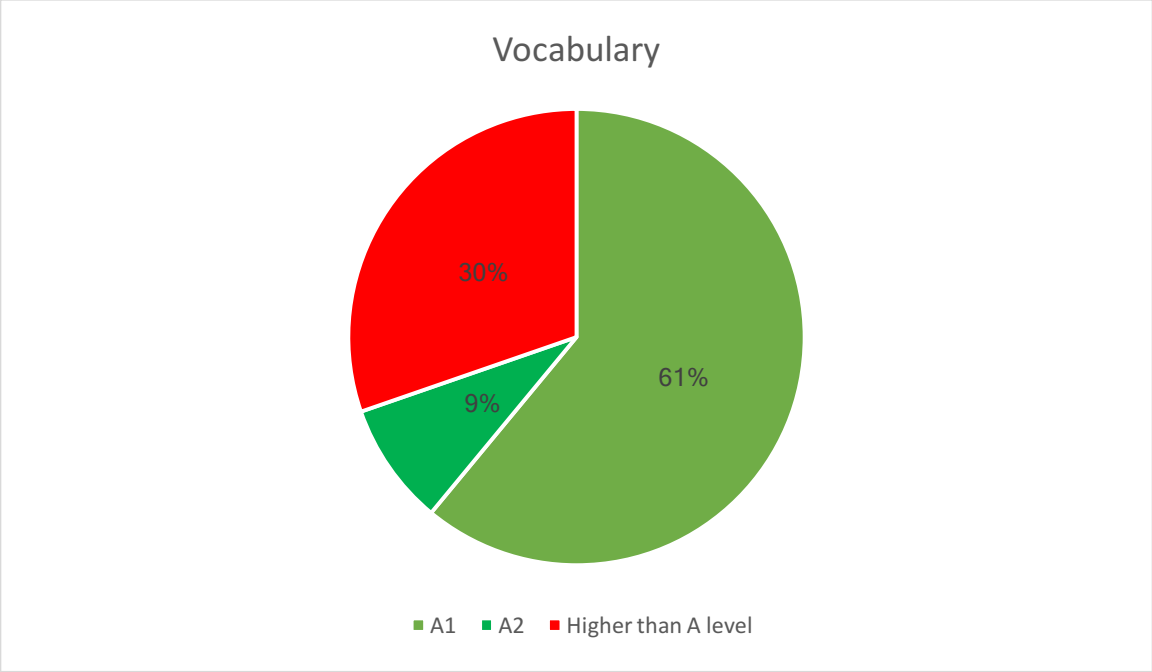


**Figure 12 Proportion of A1, A2 and difficult vocabulary**

Most of the level-1 articles include three (occasionally four) words which are highlighted and then defined at the end of the text. It remains unknown how the authors of the articles decide whether to define a word, but we can presume that the three words per article are at the B or C levels and therefore worth defining. As there are 201 articles, 603 of the difficult words were

subtracted from the 4,294 words which are in the "higher than A level" section of the pie chart, resulting in 3,691 of defined words. They account for around 4 percent of all the words.

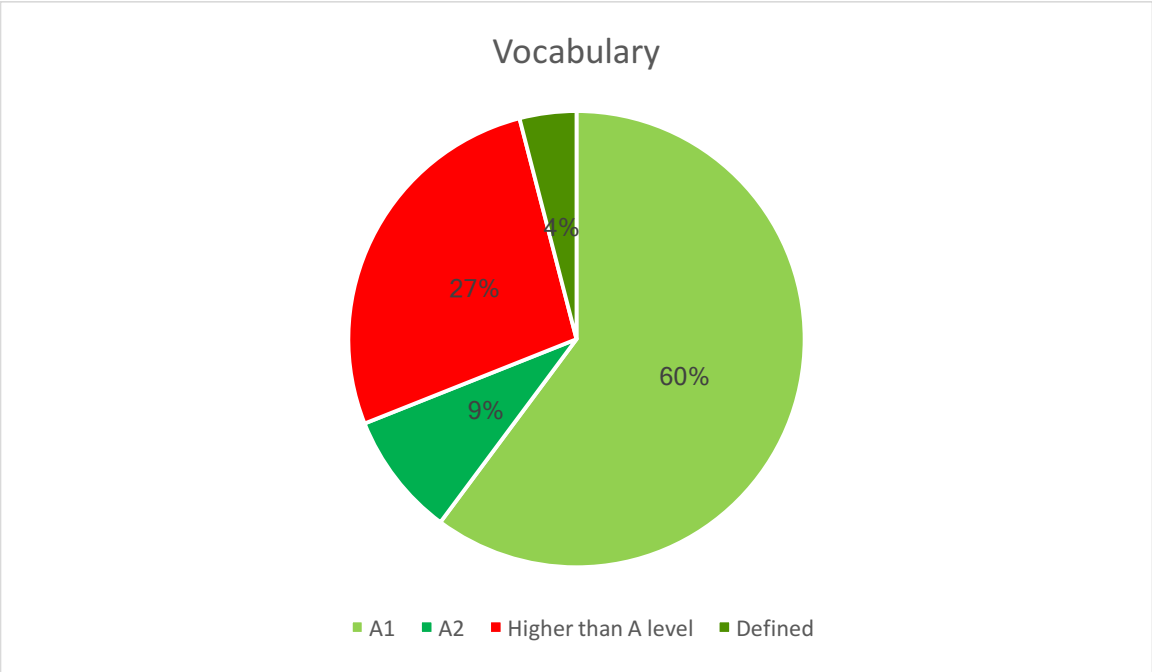


**Figure 13 Proportion of A1, A2, difficult and defined vocabulary**

### 9.1.2. Quantitative Analysis of Level-2 News Articles

There are 12,404 tokens in the whole corpus so the average article is 62 words long (see Appendix Q for the first six articles). Out of the 12,404 tokens, 6,983 tokens, i.e. 56.30 percent, are at the A1 level; 1,173 or 9.5 percent are at the A2 level, and 4,283 tokens or 34.5 percent are more difficult than the A2 level – are at the B1 level or higher. As is the case in level-1 articles, there are typically three words defined per article.

### 9.1.3. Quantitative Analysis of Level-3 News Articles

There are 14,638 tokens in the whole corpus (see Appendix R for the first six articles). The average length of one article is 73 words, and 7,659 tokens of the corpus (52.22 percent) are at the A1 level; 1,387 tokens are at the A2 level (9.5 percent), and 5,592 tokens are more difficult than the A2 level (38 percent). The number of defined words in level-3 articles varies.

### 9.1.4. Comparison of the Quantitatively Analyzed Texts

While the average length of a level-1 article is only 2 words shorter than its level-3 counterpart, the text conveys much less information. This is due to the compactness of the level 3 (see section 9.6) as well as omission of some information.

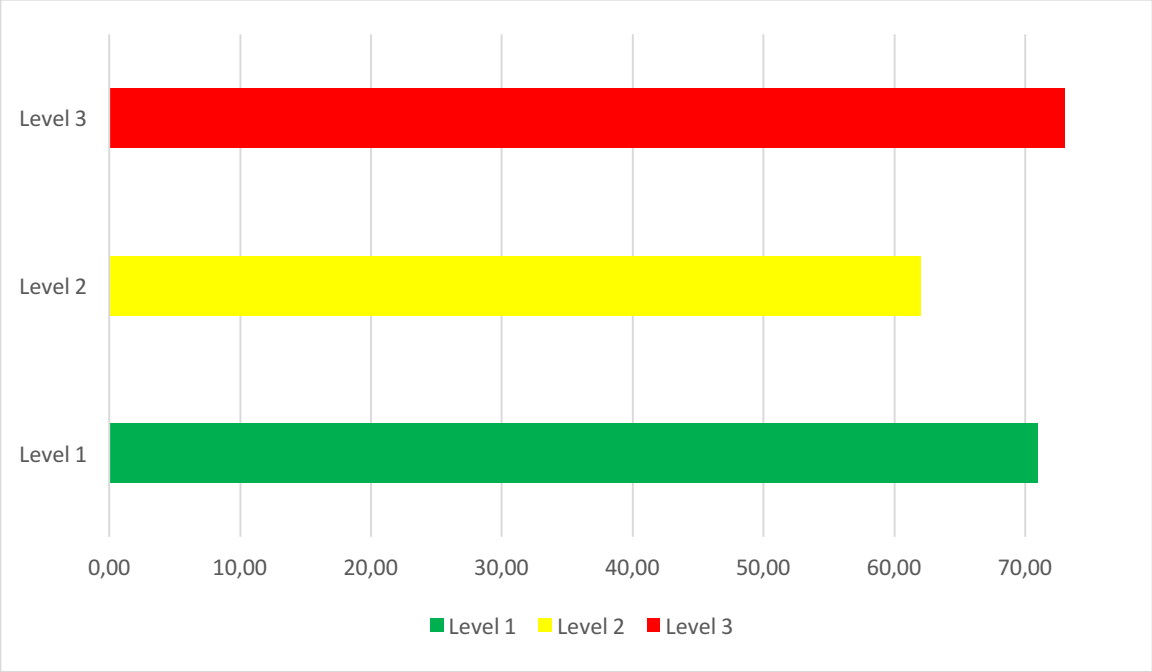


**Figure 14 Comparison of the average article length**

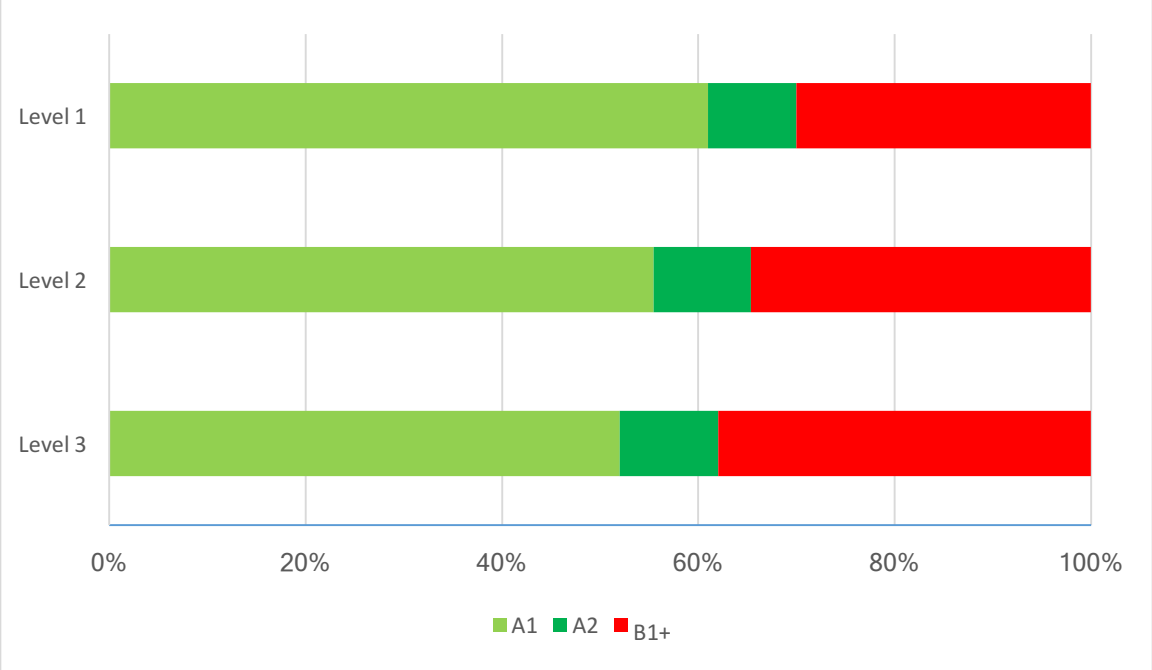


**Figure 15 Proportion of graded vocabulary at the given levels of text**

**Table 9 Percentage of graded vocabulary at different levels**

| | A1 | A2 | B1+ |
|---|---|---|---|
| **Level 3** | 52% | 10% | 38% |
| **Level 2** | 56% | 10% | 35% |
| **Level 1** | 61% | 9% | 30% |

As Table 9 shows, there is a 5 percent increase in the amount of B1+ vocabulary between level 1 and level 2 news, and a 3 percent increase in the amount of B1+ vocabulary between level 2 and 3. Table 10 shows the average amount of words which are at the B or higher levels in an average level 1–3 articles.

**Table 10 Average number of B1+ words**

| | Number of B1+ vocabulary | Average length of article |
|---|---|---|
| **Level 1** | 21 | 71 |
| **Level 2** | 21 | 62 |
| **Level 3** | 28 | 73 |

## 9.2. Qualitative Analysis of Vocabulary

A set of eight triplets of articles was searched (see Appendix D) for vocabulary which was substituted in one of the levels, and each word was assigned a grade according to the English Vocabulary Profile (Cambridge English Language Assessment et al., 2012). The articles were published at News in Levels between May 18th and 22nd 2015. Level-3 articles are rewritten narrations of authentic (intended for native speakers) video news reports. Level-2 and level-1 articles are modified texts.

To determine the difference in the level of vocabulary (the difference between level-3 vocabulary and level-2 vocabulary is in column x.3–x.2, and the difference between level-2 vocabulary and level-1 vocabulary in column x.2–x.1), the CEFR levels were assigned values, see Table 11.

**Table 11 Value assigned to CEFR levels**

| CEFR | A1 | A2 | B1 | B2 | C1 | C2 |
|---|---|---|---|---|---|---|
| value | 1 | 2 | 3 | 4 | 5 | 6 |

When there is a pair of vocabulary, with one word at the B2 level (value 4) and the other at the A1 level (value 1), the difference is 3.

**Table 12 Qualitative analysis of vocabulary**

| | Original text | CEFR | Mod. text | CEFR | x.3–x.2 | Mod. text | CEFR | x.2–x.1 |
|---|---|---|---|---|---|---|---|---|
| | Text 1.3 | | Text 1.2 | | | Text 1.1 | | |
| 1. | python | C2? | snake | A2 | 4 | snake | A2 | 0 |
| 2. | undergo surgery | C1+B2 | operation | B1 | 1.5 | operate | B1 | 0 |
| 3. | remove | B1 | remove | B1 | 0 | take out | A1 | 2 |
| 4. | swallow | B2 | eat | A1 | 3 | eat | A1 | 0 |
| | Text 2.3 | | Text 2.2 | | | Text 2.1 | | |
| 5. | receive reports | A2+B1 | call | A2 | 0.5 | call | A2 | 0 |
| 6. | inspection | C2 | find out | A2 | 4 | look | A1 | 1 |
| 7. | fake | C1 | not real | A1–A2 | 3.5 | not real | A1–A2 | 0 |
| 8. | officer | B1 | police | A2 | 1 | police | A2 | 0 |
| 9. | spot | B2 | discover | B1 | 1 | find | A1 | 2 |
| 10. | diver | B1 | diver | A2 | 1 | somebody | A2 | 0 |
| 11. | place (verb) | C1 | place (verb) | C1 | 0 | put | A1 | 4 |
| 12. | amusement | B2 | fun | A1 | 3 | fun | A1 | 0 |
| | Text 3.3 | | Text 3.2 | | | Text 3.1 | | |
| 13. | petrified | B2 | scared | B1 | 1 | scared | B1 | 0 |
| 14. | protective | B2 | protective | B2 | 0 | protect | B1 | 1 |
| 15. | threatened | C1 | threatened | C1 | 0 | danger | B1 | 2 |
| | Text 4.3 | | Text 4.2 | | | Text 4.1 | | |
| 16. | enter | A2 | enter | A2 | 0 | want to come | A1+A1 | 1 |
| 17. | narrow | C2 | narrow | C2 | 0 | choose | A1 | 5 |
| 18. | contenders | C2? | – | – | – | people | A1 | |
| | Text 6.3 | | Text 6.2 | | | Text 6.1 | | |
| 19. | capture | B2 | film (verb) | B1 | 1 | film (verb) | B1 | 0 |
| 20. | batter | C2? | batter | C2? | 0 | come | A1 | 5 |
| 21. | *several* central US states | A2 | central US states | | – | central US states | | 0 |
| 22. | damage | B1 | damage | B1 | 0 | destroy | B1 | 0 |
| 23. | structure | C2 | structure | C2 | 0 | building | A2 | 4 |

| | Text 7.3 | | Text 7.2 | | | Text 7.1 | | |
|---|---|---|---|---|---|---|---|---|
| 24. | essential | B1 | important | A1 | 2 | – | – | |
| 25. | operate | B2 | – | – | – | control, drive | B1+A1 | |
| | Text 8.3 | | Text 8.2 | | | Text 8.1 | | |
| 26. | structure | C2 | building | A2 | 4 | building | A2 | 0 |
| 27. | bring* | C2? | will be | A1 | 1 | is | A1 | 0 |

Grades accompanied by a question mark were guessed by the author of this thesis, as they were not listed in neither of the English Profiles.

Table 13 shows the number of instances of every degree of simplification in a pair of words. Noteworthy is the high amount of no change, i.e. value 0. There are nine instances of no change in difficulty in the level 3–2 relation due to no substitution. In the level 2–1 relation, there are ten instances of no substitution and four instances of substitution for a word at the same level. The total change value in the level 3–2 relation is 31.5, meaning there is an average 1.3 degree of simplification. In the level 2–1 relation, the total change equals 27, meaning there is an average simplification by one level.

**Table 13 Number of instances of change in vocabulary**

| Δ | No. of instance in x.3–x.2 relation | No. of instance in x.2–x.1 relation |
|---|---|---|
| 0 | 9 | 14 |
| 0.5 | 1 | 0 |
| 1 | 6 | 3 |
| 1.5 | 4 | 0 |
| 2 | 1 | 3 |
| 3 | 6 | 0 |
| 3.5 | 1 | 0 |
| 4 | 3 | 2 |
| 5 | 0 | 2 |

The three corpora were additionally compared to A1 and A2 English Vocabulary Profile (Cambridge English Language Assessment et al., 2012) using the AntWordProfiler (Laurence,

2014) just as the larger corpora analyzed in section 9.1. Figure 16 illustrates the proportion of vocabulary at each level, a picture not of much difference to Figure 15. The data can be also compared between Table 14 and Table 9.

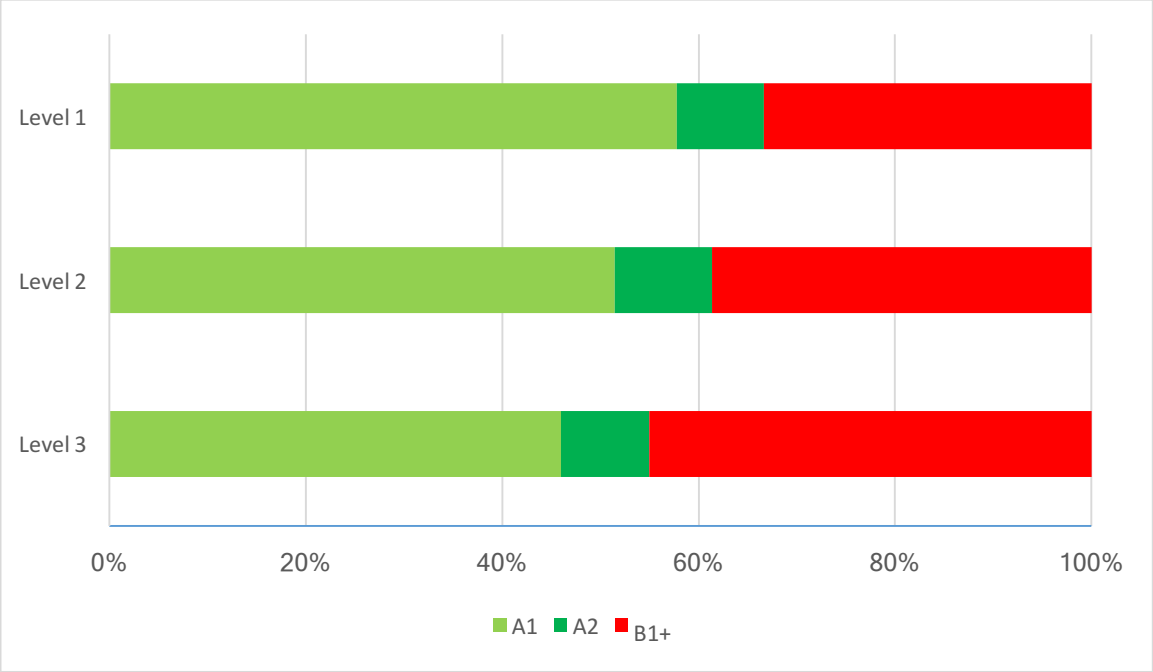


**Figure 16 Proportion of graded vocabulary at the given levels – qualitative analysis**

**Table 14 Percentage of graded vocabulary at different levels – qualitative analysis**

| | A1 | A2 | B1+ |
|---|---|---|---|
| **Level 3** | 46% | 9% | 45% |
| **Level 2** | 52% | 10% | 39% |
| **Level 1** | 52% | 8% | 40% |

## 9.3. Gradience in Classification

The 27th triplet of vocabulary in Table 12 is worth mentioning, as it illustrates the inability to perfectly divide lexical and syntactical properties of text (or rather morpho-syntactic properties in this case). Not only is the verb *bring* in the sense *cause (someone or something) to be in a particular state or condition* (Stevenson, 2010) not listed in the English Vocabulary Profile (Cambridge English Language Assessment et al., 2012), it is in the form of present participle, functioning as a condensed relative clause. Both C levels of the English Grammar Profile

(Cambridge English Language Assessment et al., 2015a) mention usage of non-finite structures, though only in a fixed set of phrases.

Just as interpreting some vocabulary borders on grammatical, structural or syntactical control, it may also border with information control. As explained in the theoretical chapter 5, absolute synonymy is rare, and much meaning can be lost when a word is substituted with a word other than its absolute synonym.

## 9.4. Conclusion of the Vocabulary Analyses

The data acquired by the analyses of vocabulary seem to suggest that the main difference between the difficulty of articles lies not in the vocabulary used but mainly in the grammar and structure. The proportion of B1+ words at each level is roughly the same, and while the qualitative analysis of vocabulary in section 9.2 suggest a simplification by one CEFR level, the degree of simplification is not considerable.

## 9.5. Analysis of Readability

As explained in subsection 3.4.1, there are several available formulas for calculating readability indexes. The Readability Test Tool (WebpageFX, 2017) calculates several readability indexes and interprets the data into grades corresponding with the grades of the American education system and the age of the pupils (grade 1, 6–7 Yo; grade 2 7–8 Yo, etc.). Table 15 lists all the readability indexes as well as the overall grade. Table 16 lists the statistics of texts.

**Table 15 Readability indices of the texts**

| Readability indices | Level 1 | Level 2 | Level 3 |
|---|---|---|---|
| Test results | Grade 4, 9–10 Yo | Grade 7, 12–13 Yo | Grade 9, 15–16 Yo |
| Flesch Kincaid Reading Ease | 85 | 74.3 | 64 |
| Flesch Kincaid Grade Level | 3.1 | 5.9 | 8.6 |
| Gunning Fog Score | 5.1 | 8 | 10.8 |
| SMOG Index | 4.5 | 6.5 | 8.3 |
| Coleman Liau Index | 8.4 | 9.7 | 10.8 |
| Automated Readability Index | 1.3 | 4.9 | 8.5 |

**Table 16 Text statistics**

| Text statistics | Level 1 | Level 2 | Level 3 |
|---|---|---|---|
| No. of sentences | 2,137 | 1,054 | 856 |
| No. of words | 14,348 | 12,576 | 14,837 |
| No. of complex words | 1,088 | 1,237 | 1,704 |
| Percent of complex words | 7.58% | 9.84% | 11.48% |
| Average words per sentence | 6.71 | 11.93 | 17.33 |
| Average syllables per word | 1.36 | 1.42 | 1.48 |

The data in the tables suggests that the difference in difficulty of texts lies in the complexity of structure. Level-1 articles can be understood by a nine or ten-year-old, level-2 articles by a twelve or thirteen-year-old, and level-3 articles can be understood by teenagers older than fifteen. The difference in complexity is also apparent from the average amount of words per sentence, also illustrated in Figure 17. As the average length of utterance of a A2 learner is 7.9 words, it can be said that level-1 news articles are suitable for such learner, arguably even for an A1 learner. Level-2 news articles are in terms of structure suitable for learners at the B levels, and level-3 news articles for C learners. While there is a difference in the percentage of complex words at each level, it is not as significant as the difference in average sentence lengths.

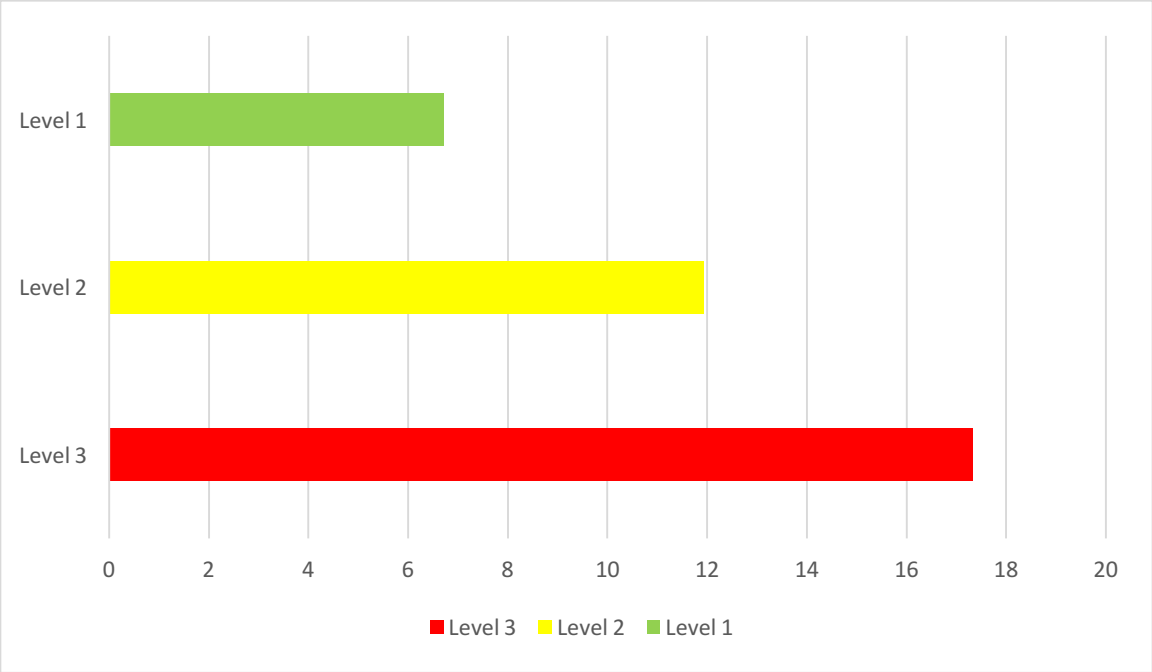


**Figure 17 Average words per sentence at each level**

## 9.6. Discourse Analysis

Eight pairs of texts (see Appendix A) – the original text and its modified counterpart – were analyzed in terms of distribution of information. The modified texts progress chronologically because perfect tenses are not used in level-1 articles. Text 1a states one of the events first, jumping in medias res (*A pet python called Winston has undergone surgery to remove a pair of barbecue tongs it had swallowed)* and explains what had preceded the surgery later: *The snake drama unfolded when…*

On the other hand, the modified text 1b starts with the fact that there is a man who has a snake which he feeds a rat, etc.: *A man from Australia has a snake. He feeds it a rat. He holds the rat with tongs. The snake eats the tongs, too*. This "chronologicalization" is a trend seen among all the modified texts. Although there are no discourse markers that would signal the sequence of the events, the sequence in which the events occurred is clear thanks to iconicity in grammar – the principle of sequential order.

It is worth noting that what is expressed in the modified texts in simple sentences must be extracted from the original texts from a whole variety of forms – sentences, clauses, participles, modifications, or even single words. As illustrated visually in the Appendix A, *A pet python* implies that somebody owns the python and has it at home, most likely. The author of the modified text must recognize this, and decide when to introduce this information. In the case of text 1b, it is at the very beginning: *A man from Australia has a snake*.

## 9.7. Conclusion of the Analyses

The analyses show that the modified texts, unlike their original counterparts, describe events in their chronological order and reduce the amount of the implied and inferred to a minimum by explicitly stating each proposition, with simple sentences being used in the level-1 articles and moderately complex sentences in level-2 articles. Structurally, articles at the level-1 are suitable for A-level learners, level-2 articles for level B learners, and level-3 articles for level C learners. This was found out by calculating the readability indexes of the texts as well as by discourse analysis of selected texts.

In terms of the lexis used at each level, the data shows there is not much difference in difficulty, with even the level-1 articles containing on average 30 percent vocabulary at level B1 or higher.

This, however, is understandable as the analyzed texts inform on a large variety of topics, and the English Vocabulary Profile may or may not list them. Unlike grammar and structure, each student may know different sets of vocabulary based on his or her personal interests. Depending on the reader's vocabulary, on how many words he or she knows, can the texts serve as materials for intensive or extensive reading. While there is a tendency for lower article levels to be easier in terms of vocabulary, the main differentiating features are the structural ones. The statements: "Level 1 has the 1000 most important words. Level 2 has the 2000 most important words, Level 3 has the 3000 most important words" (News in Levels) cannot, however, be true.

# 10. CONCLUSION

Apart from briefly elaborating on the concept of authenticity, reading, and uses of modified English text in the beginning of the thesis, the thesis concerned itself mainly with grading of language and its modification. Texts can be modified lexically, grammatically or structurally, and certain pieces of information can be omitted or added.

While absolute synonymy is rare in the English language, words can be often substituted for one another to achieve lexical simplification, though there is always some loss of meaning. Structural modification to achieve better readability often means the expansion of unstated proposition, i.e. stating explicitly what is inferred or implicit.

The analyses of the news articles indicate a great difference in the readability of each level, a great difference in structural difficulty. The lexical difficulty of the texts does not see such a difference. While the description of the news articles on the website News in Levels was determined to be untrue, the news articles are suitable for readers at any of CEFR's six levels.

# 11. RESUMÉ

Tato práce se věnuje modifikaci anglicky psaných textů na úrovni vstupu, který zprostředkovává obsah autentických textů studentům anglického jazyka na různých úrovních. Cílem práce je na základě teoretických poznatků a analýzy textu zjistit vhodnost analyzovaných textů, zpráv pro studenty anglického jazyka, jenž jsou publikovány na stránkách News in Levels. Prvních osm kapitol jsou kapitolami teoretickými. Devátá kapitola je kapitola věnovaná analýze textů.

První kapitola je zaměřená na koncept autentického textu. Autentickým textem se tradičně myslí text, který je psán rodilými mluvčími pro rodilé mluvčí. Modifikovaný text, tedy text, který je určen pro čtenáře jejichž mateřštinou není angličtina, byl nebo je vnímán jako text neautentický, a byl často kritizován pro svoji neautentičnost a nepřirozenost. Toto vnímání textu bylo ovšem v roce 1976 zpochybněno H. G. Widdowsonem, který tvrdí, že autentičnost textu se nenachází v textu samotném, ale v interakci mezi čtenářem a textem. Čte-li někdo text, rozumí mu a reaguje na něj, je pro onoho čtenáře čtení autentickým zážitkem. W. Guairento a J. Morley s Widdowsonem nepřímo souhlasí – tvrdí, že autentičnost textu nespočívá v textu samotném, ale v textové úloze, tedy čtení.

Modifikace textu v nejširším slova smyslu, je hledání jazykových forem, které jsou zamýšlenému čtenáři nebližší, a které zároveň vyjadřují stejné informace jako text originální. V případě, že je text trochu lehčí, ale hodně informací je vypuštěno, jedná se o špatně modifikovaný text. Perfektním textem by byl text, který vyjadřuje vše, co je vyjádřeno v textu originálním, ale využívá k tomu čtenářovi známé textové formy.

Sekce 2.4. a 2.5. nahlíží na text jako na komunikační kanál. Médium, pomocí kterého je dosaženo komunikace mezi autorem a čtenářem. Autor informace kóduje do textu, čtenář je následně dekóduje, čímž dochází ke komunikaci. V této komunikaci může dojít k chybě, jak už ze strany kodéra či dekodéra. Ač je čtení jako jedna z jazykových kompetencí nazýváno kompetencí pasivní, dochází při něm k aktivaci jazykových znalostí dekodéra. V první kapitole jsou také zmíněny typy čtení – čtení ze spodu (bottom-up reading) a čtení z vrchu (top-down reading). Při čtení ze spodu čtenář konstruuje informace z černých značek na papíře – znaků, slov, vět. Při čtení z vrchu dochází k projekci čtenářových schémat do textu, odhadování a předvídání informací, které jsou pomocí textu komunikovány. Interaktivním čtení

(nezaměňovat s interakcí, komunikací mezi autorem textu a čtenářem) je myšleno průběžné uplatňování jak čtení ze spodu, tak čtení z vrchu.

Poslední sekce první kapitoly se zabývá využitím modifikovaných textů pro výuku čtení, hlavně čtení extenzívního. Extenzivní čtení (EČ) je čtení velkého množství textů za účelem pobavení či získání informací, nikoliv za účelem setkání se s novými gramatickými jevy či slovíčky. Extenzivní čtení rozvíjí čtecí dovednosti, dochází k nárůstu rychlosti čtení, plynulosti a upevnění jazykových prostředků, které čtenář již zná. Materiály pro EČ musí být tedy pro čtenáře srozumitelné a musí jej zajímat. Studenti anglického jazyka na vyšších úrovních mohou číst autentické (v tradiční slova smyslu) texty, ovšem studenti začátečníci, či mírně pokročilí, autentickým textům nerozumí, jsou pro ně příliš těžké. Pro tyto studenty jsou vhodné texty modifikované, které obsahují stejné informace jako texty autentické, ovšem využívají snazšího jazyka.

Druhá kapitola se věnuje stupňování jazyka dle English Profile, projektu, který v souladu s evropským referenčním rámcem pro jazyky popisuje jazykové schopnosti studentů anglického jazyka na daných úrovních (A1–C2). Jazyk lze odstupňovat ze tří hledisek – z hlediska slovní zásoby, hlediska gramatického a hlediska strukturálního. English Vocabulary Profile ohodnocuje slovní zásobu, která je známá uživateli dané úrovně. English Grammar Profile ohodnocuje gramatické jevy. Hodnocení jazyka z hlediska strukturálního je uvedeno v brožuře *English Profile Introducing the CEFR for English*. V druhé kapitole bakalářské práce je také zmíněno možné stupňování jazyka dle profesora Paula Nationa. Kromě slovní zásoby, která je ohodnocena jako obtížnější, jsou idiomy, což jsou několikaslovné lexikální jednotky nesoucí význam, jenž nelze vyvodit z významů jednotlivých slov, uvedeny jako těžší lexikální jednotky. K problémům v porozumění také dochází, když čtenář mylně identifikuje dvě lexikální jednotky jako jednotky odkazující na dvě entity, nikoliv jako jednotky odkazující na entitu jednu.

Třetí kapitola uvádí způsoby, kterými se lze vyrovnat s těžkými lexikálními jednotkami. Těžké lexikální jednotky mohou být v některých případech ignorovány – studenti by neměli míti tendence vyhledávat každé nové slovíčko ve slovníku. Měli by být instruování, jak poznat, která slova jsou v textu důležitá a která nikoliv, a nedůležitá slova ignorovat. Díky poměrně rigidnímu slovosledu anglické věty je možno určit slovní druh slova. Studenty lze také naučit interpretovat morfologické vlastnosti slov za účelem odhadnutí významu nového slova. Dalším,

v pořadím čtvrtým způsobem vypořádání se s neznámými lexikálními jednotkami je vyvozování z kontextu. Autor modifikovaného textu by se měl snažit, aby slova, která mohou předpokládanému cílovému čtenáři činit potíže byla obklopena kontextem, kterému čtenář rozumí a dokáže z něj odhadnout význam slova těžkého. Učitel by měl čtenáře poučit, jak lze z kontextu význam slov dedukovat.

Čtenář má samozřejmě možnost slovo vyhledat ve slovníku. Slovník je vhodnou pomůckou, nicméně by se neměl nadužívat. Autor má možnost obtížná slova v textu zvýraznit a definovat. Množství definovaných slov by mělo býti střídmé, a můžeme ho považovat až jako poslední možnost, když nelze slovo vhodně nahradit nebo ho opsat.

Synonymie, hyponymie a meronymie jsou přemětem dalších kapitol. Synonymie je v anglickém jazyce velmi rozvinutá, neboť angličtina v průběhu let přijala mnoho cizích slov. K některým slovům je možné nalézt i několik synonym. Absolutní synonymie je ovšem vzácná a většina synonym jsou synonyma částečná – jejich významy jsou shodné pouze částečně. Čtvrtá kapitola se zabývá synonymy poměrně podrobně. Mezi velkou skupinu částečných synonym patří synonyma ideografická a stylistická. Ideografická synonyma se liší v distribuci a rozsahu denotace. Ideografickými synonymy jsou například slova *mistake* (chyba) a *error* (omyl). Stylistická synonyma jsou založena na protikladech neutrální – příznakové. Příznaky se někdy kombinují, např. jedno slovo může být současně poetické i zastaralé. Příkladem stylistických synonym by byla slova *choice – selection* (výběr). Slovo *choice* je stylisticky neutrální. Slovo *selection* je příznakové – knižní. V práci je zmíněna problematika synonymie i dle A. Cruse, nicméně nejvíce je v práci citována práce J. Peprníka.

Hyponymum je slovo generického významu, jednotkou vyšší abstrakce. Např. slovo *pes* je hyponymum ke slovu *štěně*. V textu práce je ukázáno, že slova, jež jsou hodnocena jako snazší, mohou být jak slovy nadřazenými – *dog* (pes) je lehčí slovo než *pup* (štěně) – tak i slovy podřadnými – *car* (auto) je snazší než *vehicle* (vozidlo).

Meronymie je dalším vztahem mezi dvěma nebo více slovy jenž reflektuje vztah na úrovni část–celek, např. *zápěstí – ruka.* Tak jako u hyponymie, může být slovo označující část (meronymum) či slovo označující celek (holonymum) v daném případě oním snazším či těžším slovem páru.

Sedmá kapitola obsahuje větné konstrukce, jež mohou čtenářům působit potíže. Kapitola je rozdělena na tři sekce – zájmenné a eliptické konstrukce, informace zhušťující konstrukce a konstrukce jenž svým umístěním přispívají k nejasnostem. Jak již bylo zmíněno, k problémům v porozumění může docházet, když čtenář mylně identifikuje referenci dvou slov. Ke stejným nedorozuměním dochází i u odkazování na entity pomocí zájmenných frází. Užitím elipsy neboli výpustky, může také dojít k zastření významu. Konstrukce zhušťující informace mohou být také přítěží – hutnému textu je těžší porozumět. Poslední sekce pojednává o tom případě, kdy je užito spojek tak, že čtenáři nemusí být jasné, zdali daná spojka koordinuje výčet jednotlivých slov či podřadných vět.

Předposlední kapitolou je kapitola praktická. Je rozdělena do sedmi sekcí, z nichž tři jsou stěžejní. První sekce analyzuje texty z webové stránky News in Levels. Jak již bylo zmíněno, na stránkách News in Levels jsou publikovány malé modifikované články, jenž čtenáře informují o dění ve světě. Stránka nabízí ony modifikované články ve třech úrovních. Úroveň 3 je nejtěžší, úroveň 2 je středně těžká a úroveň 1 je nejlehčí. Dle těchto úrovní byly vytvořeny tři korpusy. Každý korpus obsahuje 201 článků.

Pro zjištění obtížnosti slovní zásoby byly ony korpusy porovnány se slovní zásobou úrovně A1 a A2 dle English Vocabulary Profile. Analýzou bylo zjištěno, že rozdíl v poměru slov, jež jsou na úrovni B1 nebo více, je mezi jednotlivými úrovněmi minimální. Osm trojic textů bylo vybráno pro kvalitativní analýzu slovní zásoby. Bylo zjištěno, že rozdíl v použité slovní zásobě není velký. Výsledek této kvalitativní analýzy byl tudíž ve shodě s analýzou kvantitativní.

Aby byla zjištěna čtivost (readability) textů, která koreluje se strukturální složitostí, byly jednotlivé texty zadány do The Readability Test Tool, jež uplatňuje a kombinuje několik vzorců pro výpočet čtivosti. Celkový výsledek je prezentován v kontextu průměrného žáka základní školy v USA. Analýzou textů byly zjištěny velké rozdíly mezi jednotlivými úrovněmi textů. Texty na úrovni 1 jsou srozumitelné žákům čtvrté třídy a průměrná věta se skládá z 6,71 slov. Texty na úrovni 2 by byly srozumitelné žákům sedmé třídy a průměrná věta se skládá z 11,93 slov. Textům na úrovni 3 by rozuměli až žáci třídy deváté a průměrná věta se skládá z 17,33 slov. Analýzy slovní zásoby a analýzy čtivosti – větné konstrukce poukazují na to, že jednotlivé úrovně textů se liší především ve větné konstrukci, nikoliv v použité slovní zásobě.

# 12. BIBLIOGRAPHY


Baum, L. F. (2007). *The Wizard of Oz.* Retold by E. Walker. Oxford: Macmillan.

Cambridge English Language Assessment, Cambridge University Press, British Council, University of Cambridge, University of Bedfordshire, & EnglishUK. (2012). English Vocabulary Profile. Retrieved from http://vocabulary.englishprofile.org/staticfiles/about.html

Cambridge English Language Assessment, Cambridge University Press, British Council, University of Cambridge, University of Bedfordshire, & EnglishUK. (2015a). English Grammar Profile Online. Retrieved from http://www.englishprofile.org/english-grammar-profile/egp-online

Cambridge English Language Assessment, Cambridge University Press, British Council, University of Cambridge, University of Bedfordshire, & EnglishUK. (2015b). EnglishProfile. Retrieved from http://www.englishprofile.org

Cambridge ESOL, & Cambridge University Press. (2011). *English Profile: Introducing the CEFR for English* Retrieved from http://www.englishprofile.org/images/pdf/theenglishprofilebooklet.pdf

Cruse, A. (1987). *Lexical Semantics*. Cambridge: Cambridge University Press.

Cruse, A. (2000). *Meaning in Language An Introduction to Semantics and Pragmatics*. Oxford: Oxford University Press.

Day, R., & Bamford, J. (2002). Top Ten Principles for Teaching Extensive Reading. *Reading in a Foreign Language, 14*(2).

Day, R. R. (2003). What is Extensive Reading? *Teacher Talk*(21).

Day, R. R. (Producer). (2007, 27.12.2016). From Intensive to Extensive Reading: Building on Success. Retrieved from https://www.youtube.com/watch?v=f7Fq5fU5eQ8

Guariento, W., & Morley, J. (2001). Text and task authenticity in the EFL classroom. *ELT Journal Volume, 55*(4), 347-353.

Harmer, J. (2007). *The Practice of English Language Teaching* (Fourth Edition ed.). Harlow: Pearson Education Limited.

ISP Nation, & Deweerdt, J. P. (2001). A defence of simplification. *Prospect, 16*(3), 55-65.

Joy, J. J. L. (2011). The Duality of Authenticity in ELT. *The Journal of Language and Linguistic Studies, 7*(2), 7-23.

Keshavarz, M. H., & Atai, M. R. (2007). Content schemata, linguistic simplification, and EFL readers' comprehension and recall. *Reading in a Foreign Language, 19*(1), 16.

Laurence, A. (2014). AntWordProfiler (Version 1.4.1). Tokyo, Japan: Waseda University.

litteacher8. (2012, 8.12.2012). Citation of discussion forum references using EndNote [What are the difference between ideographic, stylistic, absolute and contextual synonyms?]. Retrieved from https://www.enotes.com/homework-help/synonyms-408799

Lynch, T. (1966). *Communication in the Language Classroom*. Oxford: Oxford University Press.

Microsoft. (2017). Word pro Mac (Version 15.34).

Nation, P. (2012). The BNC/COCA word family lists. Retrieved from http://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/Information-on-the-BNC_COCA-word-family-lists.pdf

News in Levels. News in Levels. Retrieved from https://www.newsinlevels.com

News in Levels. (2015a). Bad weather in the USA – level 1. Retrieved from https://www.newsinlevels.com/products/bad-weather-in-the-usa-level-1/

News in Levels. (2015b). Bad weather in the USA – level 2. Retrieved from https://www.newsinlevels.com/products/bad-weather-in-the-usa-level-2/
News in Levels. (2015c). Bad weather in the USA – level 3. Retrieved from https://www.newsinlevels.com/products/bad-weather-in-the-usa-level-3/
News in Levels. (2015d). BB King is dead – level 1. Retrieved from https://www.newsinlevels.com/products/bb-king-is-dead-level-1/
News in Levels. (2015e). BB King is dead – level 2. Retrieved from https://www.newsinlevels.com/products/bb-king-is-dead-level-2/
News in Levels. (2015f). BB King is dead – level 3. Retrieved from https://www.newsinlevels.com/products/bb-king-is-dead-level-3/
News in Levels. (2015g). Black bears scare tourists – level 1. Retrieved from https://www.newsinlevels.com/products/black-bears-scare-tourists-level-1/
News in Levels. (2015h). Black bears scare tourists – level 2. Retrieved from https://www.newsinlevels.com/products/black-bears-scare-tourists-level-2/
News in Levels. (2015i). Black bears scare tourists – level 3. Retrieved from https://www.newsinlevels.com/products/black-bears-scare-tourists-level-3/
News in Levels. (2015j). Paper plane championship – level 1. Retrieved from https://www.newsinlevels.com/products/paper-plane-championship-level-1/
News in Levels. (2015k). Paper plane championship – level 2. Retrieved from https://www.newsinlevels.com/products/paper-plane-championship-level-2/
News in Levels. (2015l). Paper plane championship – level 3. Retrieved from https://www.newsinlevels.com/products/paper-plane-championship-level-3/
News in Levels. (2015m). Self-driving cars in the streets – level 1. Retrieved from https://www.newsinlevels.com/products/self-driving-cars-in-the-streets-level-1/
News in Levels. (2015n). Self-driving cars in the streets – level 2. Retrieved from https://www.newsinlevels.com/products/self-driving-cars-in-the-streets-level-2/
News in Levels. (2015o). Self-driving cars in the streets – level 3. Retrieved from https://www.newsinlevels.com/products/self-driving-cars-in-the-streets-level-3/
News in Levels. (2015p). The Shanghai Tower – level 1. Retrieved from https://www.newsinlevels.com/products/the-shanghai-tower-level-1/
News in Levels. (2015q). The Shanghai Tower – level 2. Retrieved from https://www.newsinlevels.com/products/the-shanghai-tower-level-2/
News in Levels. (2015r). The Shanghai Tower – level 3. Retrieved from https://www.newsinlevels.com/products/the-shanghai-tower-level-3/
News in Levels. (2015s). Snake eats tongs – level 1. Retrieved from https://www.newsinlevels.com/products/snake-eats-tongs-level-1/
News in Levels. (2015t). Snake eats tongs – level 2. Retrieved from https://www.newsinlevels.com/products/snake-eats-tongs-level-2/
News in Levels. (2015u). Snake eats tongs – level 3. Retrieved from https://www.newsinlevels.com/products/snake-eats-tongs-level-3/
News in Levels. (2015v). An underwater surprise – level 1. Retrieved from https://www.newsinlevels.com/products/an-underwater-surprise-level-1/
News in Levels. (2015w). An underwater surprise – level 2. Retrieved from https://www.newsinlevels.com/products/an-underwater-surprise-level-2/
News in Levels. (2015x). An underwater surprise – level 3. Retrieved from https://www.newsinlevels.com/products/an-underwater-surprise-level-3/

News in Levels. (2016). Seventy-metre Biscuit Dunk. Retrieved from http://www.newsinlevels.com/products/seventy-metre-biscuit-dunk-level-1/

News in Levels. (2017a). London Bridge Attack – level 3. Retrieved from https://www.newsinlevels.com/products/london-bridge-attack-level-3/

News in Levels. (2017b). Monkey steals money – level 1. Retrieved from https://www.newsinlevels.com/products/monkey-steals-money-level-1/

News in Levels. (2017c). Paris Agreement and Trump – level 1. Retrieved from https://www.newsinlevels.com/products/paris-agreement-and-trump-level-1/

Nuttall, C. (2005). *Teaching Reading Skills in a Foreign Language*. Oxford: Macmillan.

ODN (Producer). (2016, 2016). Guinness World Record won for bungee 'dunk' into cup of tea. Retrieved from https://www.youtube.com/watch?v=UBf7WC19lpw

Penguin Readers. Grading of Language. Retrieved from http://www.penguinreaders.com/pr/teachers/grading-of-language.html

Peprník, J. (1992). *Anglická lexikologie*. Olomouc: Univerzita Palackého.

Shomoossi, N., & Ketabi, S. (2007). A Critical Look at the Concept of Authenticity. *Electronic Journal of Foreign Language Teaching, 4*(1), 149–155.

Stevenson, A. (2010). *Oxford Dictionary of English* (3rd ed.). Oxford: Oxford University Press.

Swift, S. (Producer). (2009a, 2016). Authentic and simplified texts Pt 2. Retrieved from https://www.youtube.com/watch?v=1aAJjcaDU34

Swift, S. (Producer). (2009b, 2016). Authentic texts or simplified texts Part One. Retrieved from https://www.youtube.com/watch?v=AMYQK8vMxQ0&t=18s

Sword, H. (Producer). (2012, 7.6.2017). Beware of nominalizations (AKA zombie nouns) - Helen Sword. Retrieved from https://www.youtube.com/watch?v=dNlkHtMgcPQ

Waite, M. (2009). *Oxford Thesaurus of English*. Oxford: Oxford University Press.

Wallace, C. (2001). Reading. In R. Carter & D. Nunan (Eds.), *The Cambridge Guide to Teaching English to Speakers of Other Languages* (pp. 21-27). Cambridge: Cambridge University Press.

Waring, R. (Producer). (2011, 9.1.2017). Why Extensive Reading doubles your students' vocabulary. Retrieved from http://www.robwaring.org/presentations/KAPEE/KAPEE_Vocab_talk_lite.pptx

WebpageFX. (2017). READABILITY TEST TOOL Quick and easy way to test the readability of your work. Retrieved from https://www.webpagefx.com/tools/read-able/

Widdowson, H. G. (1979). *Explorations in Applied Linguistics*. Oxford: Oxford University Press.

Yule, G. (2010). *The Study of Language* (4th ed.). Cambridge: Cambridge University Press.

# 13. APPENDICES

Appendix A *Discourse analysis*

Text 1a
Snake eats tongs – level 3
18-05-2015 07:00
A pet python called Winston has undergone surgery to remove a pair of barbecue tongs it had swallowed. Vets at Adelaide University in Australia operated on the man's python because they weren't sure he'd be able to regurgitate the tongs without causing itself internal injuries.
The outline of the tongs inside Winston could be seen lodged in his stomach and these X-ray images later revealed the item wedged inside. The snake drama unfolded when Winston's owner Aaron used the tongs to feed a dead rat to him which he then gripped on to. Aaron was unable to free the tongs so left them with Winston only to return later to find he'd swallowed them whole.
Snakes do have an ability to regurgitate food if they change their mind, but vets weren't convinced Winston was going to even if he tried.
He is reported to now be making a good recovery. (News in Levels, 2015u)

Text 1b
Snake eats tongs – level 1
18-05-2015 07:00
A man from Australia has a snake. He feeds it a rat. He holds the rat with tongs. The snake eats the tongs, too.
This is dangerous for the snake. Doctors operate on it. They take out the tongs.
The snake's name is Winston. He is getting better. (News in Levels, 2015s)

Text 2a
An underwater surprise – level 3
18-05-2015 15:00
It's enough to chill you to the bone at first glance, but this skeleton discovery is not what it seems. Police in Arizona thought they were about to embark on a grim investigation when they received reports of human remains at the bottom of the Colorado River but on closer inspection, and much to the relief of emergency crews, the so-called remains were just two fake skeletons enjoying a tea party.
The discovery of the underwater bash came after officers received reports that a snorkeler had spotted human bones but they soon found the pair donning sunglasses and lounging in lawn chairs on the riverbed. One of the skeletons bore a sign reading, "Bernie living the dream in the river."
Two divers have come forward to claim responsibility for creating the skeleton tea party. The pair apparently placed the skeletons named Bernie and Bernadette in the river in August 2014 for their own amusement. The county lieutenant is now hoping that the dive site will become a tourist attraction for the area. (News in Levels, 2015x)

Text 2b
An underwater surprise – level 1
18-05-2015 15:00
This news is from Arizona, USA. Somebody snorkels in the Colorado River. The person finds something. It is two human skeletons!
The person calls the police. The police come. They look at the skeletons. They see that they are not real.

Somebody put the skeletons into the water for fun. The skeletons even have names – Bernie and Bernadette. They may bring tourists to the area. (News in Levels, 2015v)

Text 3a
Black bears scare tourists – level 3
19-05-2015 15:00

This black bear and her three cubs chased a group of petrified tourists on a bridge in Yellowstone National Park. The video shows the bear bounding along the bridge with her cubs keeping up as the tourists tried to keep a safe distance. A man can be heard yelling, "Go, go!" while some people on the bridge kept taking photographs, even as the bear veered in their direction.

Parent bears are highly protective of their cubs and will attack if they feel threatened. People weren't always as cautious of bears in the park – from 1910 to the 1960s, Yellowstone Park managers allowed visitors to feed black bears along the roads, although the National Park Service officially frowned on this activity.

Black bears soon became the symbol of Yellowstone. (News in Levels, 2015i)

Text 3b
Black bears scare tourists – level 1
19-05-2015 15:00

Yellowstone National Park is a national park in the USA. The black bear is the park's symbol. An incident happens in the park. A group of bears chase tourists. One of the bears is a mother. The other bears are her cubs. Bear mothers protect their cubs. If the cubs are in danger, they can attack.

The tourists are scared. They get inside their cars. Others take photographs of the bears. (News in Levels, 2015g)

Text 4a
Paper plane championship – level 3
19-05-2015 07:00

Finalists of the fourth edition up the Red Bull Paper Wings world paper aeroplane championships descended on Salzburg, Austria on Saturday to settle the 2015 titles for the furthest throw, most aerobatic flight and the aircraft that flew for the longest time.

Organisers said that 46,000 people around the world entered the competitions, with over 500 qualifying contests narrowing the contenders down to just 200. And this has got to be one of the only sports where it's okay to fold under pressure.

The furthest throw was a whopping 53.22 metres for Veselin Ivanov of Bulgaria who designed the pencil-slim aircraft that flew straight as an arrow. In the aerobatics category, the winner Avedis Tchamitchian of Lebanon, scored a maximum of 50 points from the judges.

And there was a nail-biting finale for the longest air time title, the narrowest seen at the Red Bull Paper Wings World Finals, as Armenia's Karen Hambardzumyan launched his aircraft high up into the hangar's roof and it descended in an elegant spiralling flight, lasting 14.36 seconds. (News in Levels, 2015l)

Text 4b
Paper plane championship – level 1
19-05-2015 07:00

People hold a championship in Austria. It is a paper plane championship.

Many people want to come to the event, but they choose only 200 people. You can win in three categories.

One man throws his plane 53.22 metres. He wins in the category. (News in Levels, 2015j)

Text 5a
BB King is dead – level 3
21-05-2015 07:00
The King of Blues, guitarist and singer BB King, has died at the age of 89. BB King took the blues from rural juke joints to the mainstream and influenced a generation of rock guitarists from Eric Clapton to Stevie Ray Vaughan.
He was best known for his hits "My Lucille," "Sweet Little Angel" and "Rock Me, Baby" and will be forever linked with the Gibson Guitars he named Lucille. Rolling Stone Magazine placed him behind only Jimi Hendrix and Duane Allman in its list of the 100 greatest guitarists of all-time, and until recently, King performed in at least a hundred concerts a year.
He was recently taken to hospital with a diabetes-related illness and earlier this month posted on Facebook that he was in hospice care at his home. He died in his sleep in Las Vegas. (News in Levels, 2015f)

Text 5b
BB King is dead – level 1
21-05-2015 07:00
This news is about BB King. He was the King of Blues. He played the guitar. He was perfect at it.
Recently, he went to hospital. He was not well. Then he went home. He was in hospice care. He died. He was 89. He died in his sleep. He died in Las Vegas. (News in Levels, 2015d)

Text 6a
Bad weather in the USA – level 3
21-05-2015 15:00
Hailstones big enough to shatter the windscreen of a car.
"Oh! There it goes! The windshield!"
The moment the baseball-sized stones start falling from the sky is captured on camera near Tipton in Oklahoma. Thunderstorms and floods battered several central US states on May 17th. They followed on from a tornado series that hit large parts of the area a day earlier, cutting power lines and damaging structures.
Parts of Texas, Oklahoma, Arkansas, Tennessee and Kansas were either under a flash flood watch or flash flood warning. There were reportedly 29 tornadoes on Saturday that hit states ranging from Louisiana to Wyoming.
Texas and Oklahoma were the hardest hit areas. The damage in Oklahoma was mostly in the southwestern and northeastern parts of the state.
"Oh! There it goes! The windshield!" (News in Levels, 2015c)

Text 6b
Bad weather in the USA – level 1
21-05-2015 15:00
This happens in central US states. There is bad weather. There are many tornados. They hit large parts of the area. They destroy buildings. After the tornados, floods and storms come. Hailstones fall from the sky. A person films this on a camera. The camera is inside a car. The hailstones fall on the car. (News in Levels, 2015a)

Text 7a
Self-driving cars in the streets – level 3
22-05-2015 07:00

Google's new self-driving cars are to finally hit the roads this summer with steering wheels. The company will begin testing the vehicles on public roads in the coming months, but it turns out they will have steering wheels and brakes, something they hadn't envisioned a year ago.
Engineers will operate 25 prototypes which use the same software as Google's Lexus sport-utility vehicles that have already self-driven about 10,000 miles (16,000 kilometres) a week in recent months.
The Internet search company announced a year ago that it planned to build a fleet of self-driving cars, saying the prototypes wouldn't have steering wheels, accelerator pedals or brake pedals because they wouldn't need them.
It turns out, however, that the critical control devices are still rather essential. Built in Detroit, the cars will be equipped with removable steering wheels and pedals so test engineers can take over driving if needed.
The company also said it would test new passenger and pedestrian protection technologies, including a firm front end and flexible windshield.
It also said the speed of the prototype will be limited to 25 miles per hour (40 kilometres per hour) to decrease the likelihood of severe injury in a collision. (News in Levels, 2015o)

Text 7b
Self-driving cars in the streets – level 1
22-05-2015 07:00
Google builds cars. These cars are special. They don't need a driver. They are self-driving cars. Google starts to test the cars. They will drive in the streets. For the testing, however, the cars will have a driver. He or she will control the car. If something goes wrong, the driver will drive it. The maximum speed of the cars is 40 kilometres per hour.
People build these cars in Detroit, USA. (News in Levels, 2015m)

Text 8a
The Shanghai Tower – level 3
22-05-2015 15:00
The Shanghai Tower is set to be the world's second tallest building when it's completed in the summer. The 120 floors of the 632-metre structure will make the building the tallest in China but bringing it just under 200 metres lower than the Burj Khalifa in Dubai.
"So, when Sears Tower was built, there was a recession. When the Petronas Towers were built, there was a recession. When the Burj Khalifa was built, there was a recession. This is not the world's tallest building so there's no recession. So, that's the world's tallest building curse, not the skyscraper curse.
So, it's actually better to be number two than number one. I think one of the most profitable skyscrapers in the world is the Empire State Building. But for the first 10 years, it didn't make any money because it was opening during a recession.
I think that goes back to the skyscraper—the world's tallest building curse. There was a recession right after that. But now, it's one of the most successfully rented buildings in the world because it's iconic, people know it and people want to be in it."
The Shanghai Tower is set to house mainly financial institutions and government agencies. (News in Levels, 2015r)

Text 8b
The Shanghai Tower – level 1
22-05-2015 15:00
People are building a skyscraper. It is called the Shanghai Tower. It is in Shanghai, China. It has 120 floors. It is 632 metres tall.

It is the tallest building in China, but it is not the tallest building in the world. It is the second tallest building.
Burj Khalifa is the tallest building in the world. Burj Khalifa is in Dubai, Saudi Arabia. It is 200 metres taller than the Shanghai Tower. (News in Levels, 2015p)

Appendix B *Phrases of the A1 level omitted from the analysis of the vocabulary*

all right
a.m.
as well
I am Spanish/a teacher
there is/there are/there was, etc.
be called sth
CD player noun
come from swh/sth
of course
credit card
dining room
do the cleaning/cooking, etc.
what does sb do?
for example
Excuse me
(that's) fine
Don't forget...
get here/there/home/to work, etc.
get a bus/train/taxi, etc.
get (sb) up
go shopping
go out
good for you
good afternoon
good evening
good morning
good night
half past one/two/three, etc.
Happy Birthday/New Year, etc.
How are you?
ice cream
last week/year/Monday, etc.
would like sth/to do sth
Would you like...?
live in/at, etc.
how many
mobile phone
more beautiful/difficult/easily, etc.
(Good) morning
in the morning
next week/year/Monday, etc.
one/two/three, etc. o'clock
Yes, please
p.m.
put sth down/in/on, etc.
it rains/it is raining
see you later/soon/tomorrow, etc.
see you soon
take a picture/photo(graph)
thank you
There is/are/was, etc.
too small/hard/much, etc.
T-shirt
not very good/tall/happy, etc.
wake (sb) up or wake up (sb)
Well done!
Don't worry (about sth)
Would you like ...?
be ... years old

Appendix C *Corpus of level-1 news articles*

A lorry drives into a department store in the centre of Stockholm. At least two people die. Police look for the attacker. They send people away from the city centre and the main central station. The government wants to find out more. It also tells the public to be careful and to listen to the police.
The attack comes after other terrorist attacks in France, Germany and the United Kingdom.

This happens in the USA. A man leaves his home early in the morning. He sees a bear. The bear walks on its hind legs. It walks like a human.
The man uses his phone to film the bear. He uploads the video on the Internet. People watch the video. They say that they know the bear. They saw it, too!

There is a prison on Sumatra. The prison is built for 300 people. There are 1,900 people in it. On Friday, 200 prisoners escape. They are out of their cells during prayers.
Police and soldiers come to the prison. They catch 50% of the prisoners on that same day.

A baby otter is in a canal in Arizona. It is not well. It is hungry, dehydrated, and full of fleas. Luckily, people see it.
They take it to a centre which helps animals. There they give it milk and fish. It gets better quickly. The otter starts a new life in a wildlife park.

A bus is driving on a highway. It is a church bus. It is in Arkansas, USA. A girl is on the bus. She sits at a door of the bus.
The door opens. She falls out. It is a shocking moment. A man films it on camera.
A man is driving behind the bus. He sees all of this. He quickly goes to the girl. Traffic is busy. He picks up the girl.
The girl's family talks. They think that she will be fine.

This story is about a man. He is from the USA. He is a university professor. He goes to Nepal. He climbs a mountain. The mountain is covered in ice.
There is a hole in the ice. It is 22 metres deep. The man falls in it. He doesn't go all the way down. He stops somewhere in the hole. He cannot move. His arm and five ribs are broken.
He does not want to die. He climbs out of the hole. People find him the next day.

Appendix D *Analyzed texts with highlights of words which were substituted*

Text 1.3
Snake eats tongs – level 3
18-05-2015 07:00
A pet python called Winston has undergone surgery to remove a pair of barbecue tongs it had swallowed. Vets at Adelaide University in Australia operated on the man's python because they weren't sure he'd be able to regurgitate the tongs without causing itself internal injuries.
The outline of the tongs inside Winston could be seen lodged in his stomach and these X-ray images later revealed the item wedged inside. The snake drama unfolded when Winston's owner Aaron used the tongs to feed a dead rat to him which he then gripped on to. Aaron was unable to free the tongs so left them with Winston only to return later to find he'd swallowed them whole.
Snakes do have an ability to regurgitate food if they change their mind, but vets weren't convinced Winston was going to even if he tried.
He is reported to now be making a good recovery. (News in Levels, 2015u)

Text 1.2
Snake eats tongs – level 2
**18-05-2015 07:00**
A man from Australia has a pet snake called Winston. One time, the **owner** used **barbecue tongs** to feed a rat to the snake.
The snake **gripped** on to the tongs as well, and the man could not take the tongs from him. The owner left the tongs with Winston. When he returned, the tongs were gone – Winston ate them. Snakes can bring back food which they don't like, but tongs could hurt the snake. The pet snake had to have an operation – doctors removed the tongs and Winston is getting better. (News in Levels, 2015t)

Text 1.1
Snake eats tongs – level 1
18-05-2015 07:00
A man from Australia has a snake. He feeds it a rat. He holds the rat with tongs. The snake eats the tongs, too.
This is dangerous for the snake. Doctors operate on it. They take out the tongs.
The snake's name is Winston. He is getting better. (News in Levels, 2015s)

Text 2.3
An underwater surprise – level 3
18-05-2015 15:00
It's enough to chill you to the bone at first glance, but this skeleton discovery is not what it seems. Police in Arizona thought they were about to embark on a grim investigation when they received reports of human remains at the bottom of the Colorado River but on closer inspection, and much to the relief of emergency crews, the so-called remains were just two fake skeletons enjoying a tea party.
The discovery of the underwater bash came after officers received reports that a snorkeler had spotted human bones but they soon found the pair donning sunglasses and lounging in lawn chairs on the riverbed. One of the skeletons bore a sign reading, "Bernie living the dream in the river."
Two divers have come forward to claim responsibility for creating the skeleton tea party. The pair apparently placed the skeletons named Bernie and Bernadette in the river in August 2014

for their own amusement. The county lieutenant is now hoping that the dive site will become a tourist attraction for the area. (News in Levels, 2015x)

Text 2.2
An underwater surprise – level 2
18-05-2015 15:00
A person in Arizona was snorkelling in the Colorado River when he discovered something shocking – two human skeletons!
He called police to the site and found out that the skeletons were not real. What's more, they were relaxing and enjoying a tea party.
Two divers placed the skeletons Bernie and Bernadette in the river in August 2014. They did it for fun.
The skeletons may bring more tourists to the area. (News in Levels, 2015w)

Text 2.1
An underwater surprise – level 1
18-05-2015 15:00
This news is from Arizona, USA. Somebody snorkels in the Colorado River. The person finds something. It is two human skeletons!
The person calls the police. The police come. They look at the skeletons. They see that they are not real.
Somebody put the skeletons into the water for fun. The skeletons even have names – Bernie and Bernadette. They may bring tourists to the area. (News in Levels, 2015v)

Text 3.3
Black bears scare tourists – level 3
19-05-2015 15:00
This black bear and her three cubs chased a group of petrified tourists on a bridge in Yellowstone National Park. The video shows the bear bounding along the bridge with her cubs keeping up as the tourists tried to keep a safe distance. A man can be heard yelling, "Go, go!" while some people on the bridge kept taking photographs, even as the bear veered in their direction.
Parent bears are highly protective of their cubs and will attack if they feel threatened. People weren't always as cautious of bears in the park – from 1910 to the 1960s, Yellowstone Park managers allowed visitors to feed black bears along the roads, although the National Park Service officially frowned on this activity.
Black bears soon became the symbol of Yellowstone. (News in Levels, 2015i)

Text 3.2
Black bears scare tourists – level 2
19-05-2015 15:00
Yellowstone National Park is a national park in the USA. Its symbol is the black bear.
A black bear and her cubs chased a group of tourists on a bridge in the park. People filmed the moment and it showed the scared tourists, as well as those who just kept taking photographs.
Bear mothers are highly protective of their cubs, and they will attack if they feel threatened.
People weren't always as careful around the park's bears as they are now. From 1910 to the 1960s, people were allowed to feed the bears along the roads. (News in Levels, 2015h)

Text 3.1
Black bears scare tourists – level 1

19-05-2015 15:00
Yellowstone National Park is a national park in the USA. The black bear is the park's symbol. An incident happens in the park. A group of bears chase tourists. One of the bears is a mother. The other bears are her cubs. Bear mothers protect their cubs. If the cubs are in danger, they can attack.
The tourists are scared. They get inside their cars. Others take photographs of the bears. (News in Levels, 2015g)

Text 4.3
Paper plane championship – level 3
19-05-2015 07:00
Finalists of the fourth edition of the Red Bull Paper Wings world paper aeroplane championships descended on Salzburg, Austria on Saturday to settle the 2015 titles for the furthest throw, most aerobatic flight and the aircraft that flew for the longest time.
Organisers said that 46,000 people around the world entered the competitions, with over 500 qualifying contests narrowing the contenders down to just 200. And this has got to be one of the only sports where it's okay to fold under pressure.
The furthest throw was a whopping 53.22 metres for Veselin Ivanov of Bulgaria who designed the pencil-slim aircraft that flew straight as an arrow. In the aerobatics category, the winner Avedis Tchamitchian of Lebanon, scored a maximum of 50 points from the judges.
And there was a nail-biting finale for the longest air time title, the narrowest seen at the Red Bull Paper Wings World Finals, as Armenia's Karen Hambardzumyan launched his aircraft high up into the hangar's roof and it descended in an elegant spiralling flight, lasting 14.36 seconds. (News in Levels, 2015l)

Text 4.2
Paper plane championship – level 2
19-05-2015 07:00
People held the world paper aeroplane championships in Austria. The event settled three 2015 titles. The titles were for the furthest throw, most aerobatic flight, and the aircraft which flew for the longest time.
Around 46,000 people around the world entered the championships, but five hundred qualifying contests narrowed them to just 200.
The furthest throw was an amazing 53.22 metres. The longest flight was 14.36 seconds long. (News in Levels, 2015k)

Text 4.1
Paper plane championship – level 1
19-05-2015 07:00
People hold a championship in Austria. It is a paper plane championship.
Many people want to come to the event, but they choose only 200 people. You can win in three categories.
One man throws his plane 53.22 metres. He wins in the category. (News in Levels, 2015j)

Text 5.3
BB King is dead – level 3
21-05-2015 07:00
The King of Blues, guitarist and singer BB King, has died at the age of 89. BB King took the blues from rural juke joints to the mainstream and influenced a generation of rock guitarists from Eric Clapton to Stevie Ray Vaughan.

He was best known for his hits “My Lucille,” “Sweet Little Angel” and “Rock Me, Baby” and will be forever linked with the Gibson Guitars he named Lucille. Rolling Stone Magazine placed him behind only Jimi Hendrix and Duane Allman in its list of the 100 greatest guitarists of all-time, and until recently, King performed in at least a hundred concerts a year.
He was recently taken to hospital with a diabetes-related illness and earlier this month posted on Facebook that he was in hospice care at his home. He died in his sleep in Las Vegas. (News in Levels, 2015f)

Text 5.2
BB King is dead – level 2
21-05-2015 07:00
The King of Blues, guitarist and singer BB King, died. He was 89 years old.
He influenced a generation of rock guitarists and was one of the world’s best guitarists. King performed in at least a hundred concerts a year.
Recently, he had an illness and people took him to hospital. Earlier this month, he said on Facebook that he was in hospice care at home. He died in his sleep in Las Vegas. (News in Levels, 2015e)

Text 5.1
BB King is dead – level 1
21-05-2015 07:00
This news is about BB King. He was the King of Blues. He played the guitar. He was perfect at it.
Recently, he went to hospital. He was not well. Then he went home. He was in hospice care. He died. He was 89. He died in his sleep. He died in Las Vegas. (News in Levels, 2015d)

Text 6.3
Bad weather in the USA – level 3
21-05-2015 15:00
Hailstones big enough to shatter the windscreen of a car.
“Oh! There it goes! The windshield!”
The moment the baseball-sized stones start falling from the sky is captured on camera near Tipton in Oklahoma. Thunderstorms and floods battered several central US states on May 17th. They followed on from a tornado series that hit large parts of the area a day earlier, cutting power lines and damaging structures.
Parts of Texas, Oklahoma, Arkansas, Tennessee and Kansas were either under a flash flood watch or flash flood warning. There were reportedly 29 tornadoes on Saturday that hit states ranging from Louisiana to Wyoming.
Texas and Oklahoma were the hardest hit areas. The damage in Oklahoma was mostly in the southwestern and northeastern parts of the state.
“Oh! There it goes! The windshield!” (News in Levels, 2015c)

Text 6.2
Bad weather in the USA – level 2
21-05-2015 15:00
Thunderstorms and floods battered central US states on May 17th. They came after a tornado series which hit large parts of the area one day earlier. They cut power lines and damaged structures.
From the inside of a car, a camera filmed one interesting moment. Baseball-sized hailstones fell on the car’s windshield and broke it. This happened in Oklahoma. (News in Levels, 2015b)

Text 6.1
Bad weather in the USA – level 1
21-05-2015 15:00
This happens in central US states. There is bad weather. There are many tornados. They hit large parts of the area. They destroy buildings. After the tornados, floods and storms come. Hailstones fall from the sky. A person films this on a camera. The camera is inside a car. The hailstones fall on the car. (News in Levels, 2015a)

Text 7.3
Self-driving cars in the streets – level 3
22-05-2015 07:00
Google's new self-driving cars are to finally hit the roads this summer with steering wheels. The company will begin testing the vehicles on public roads in the coming months, but it turns out they will have steering wheels and brakes, something they hadn't envisioned a year ago.
Engineers will operate 25 prototypes which use the same software as Google's Lexus sport-utility vehicles that have already self-driven about 10,000 miles (16,000 kilometres) a week in recent months.
The Internet search company announced a year ago that it planned to build a fleet of self-driving cars, saying the prototypes wouldn't have steering wheels, accelerator pedals or brake pedals because they wouldn't need them.
It turns out, however, that the critical control devices are still rather essential. Built in Detroit, the cars will be equipped with removable steering wheels and pedals so test engineers can take over driving if needed.
The company also said it would test new passenger and pedestrian protection technologies, including a firm front end and flexible windshield.
It also said the speed of the prototype will be limited to 25 miles per hour (40 kilometres per hour) to decrease the likelihood of severe injury in a collision. (News in Levels, 2015o)

Text 7.2
Self-driving cars in the streets – level 2
22-05-2015 07:00
Google announced a year ago that it wanted to build self-driving cars. It said that the cars would not have steering wheels, accelerator pedals or brake pedals. They said that the cars would not need them.
This summer, Google is going to finally test the cars in the streets. However, the cars will still have steering wheels, accelerator pedals and brake pedals. These things are still important. Also, the speed of the cars will be limited to 40 kilometres per hour. People build the cars in Detroit, USA. (News in Levels, 2015n)

Text 7.1
Self-driving cars in the streets – level 1
22-05-2015 07:00
Google builds cars. These cars are special. They don't need a driver. They are self-driving cars. Google starts to test the cars. They will drive in the streets. For the testing, however, the cars will have a driver. He or she will control the car. If something goes wrong, the driver will drive it. The maximum speed of the cars is 40 kilometres per hour.
People build these cars in Detroit, USA. (News in Levels, 2015m)

Text 8.3

The Shanghai Tower – level 3
22-05-2015 15:00
The Shanghai Tower is set to be the world's second tallest building when it's completed in the summer. The 120 floors of the 632-metre structure will make the building the tallest in China but bringing it just under 200 metres lower than the Burj Khalifa in Dubai.
"So, when Sears Tower was built, there was a recession. When the Petronas Towers were built, there was a recession. When the Burj Khalifa was built, there was a recession. This is not the world's tallest building so there's no recession. So, that's the world's tallest building curse, not the skyscraper curse.
So, it's actually better to be number two than number one. I think one of the most profitable skyscrapers in the world is the Empire State Building. But for the first 10 years, it didn't make any money because it was opening during a recession.
I think that goes back to the skyscraper—the world's tallest building curse. There was a recession right after that. But now, it's one of the most successfully rented buildings in the world because it's iconic, people know it and people want to be in it."
The Shanghai Tower is set to house mainly financial institutions and government agencies. (News in Levels, 2015r)

Text 8.2
The Shanghai Tower – level 2
**22-05-2015 15:00**
People are building China's tallest building. Its name is the Shanghai Tower and they will finish it this summer. With 120 floors, it will be 632 metres tall.
It won't be the world's tallest building – it will be 200 metres lower than the Burj Khalifa in Dubai. However, a chief architect explained that this is not a bad thing. He says that when a building which is the tallest in the world is built, there is a recession.
There was a recession when the Empire State Building was built. For the first ten years, it didn't make any money! But now, it is one of the most successfully rented buildings. It is iconic, people know it and people want to be in it.
Shanghai Tower will house mainly financial institutions and government agencies. (News in Levels, 2015q)

Text 8.1
The Shanghai Tower – level 1
22-05-2015 15:00
People are building a skyscraper. It is called the Shanghai Tower. It is in Shanghai, China. It has 120 floors. It is 632 metres tall.
It is the tallest building in China, but it is not the tallest building in the world. It is the second tallest building.
Burj Khalifa is the tallest building in the world. Burj Khalifa is in Dubai, Saudi Arabia. It is 200 metres taller than the Shanghai Tower. (News in Levels, 2015p)

Appendix E *The titles of the news articles are identical across the levels*

## Man Saves a Dog – level 3

16-06-2017 15:00 Level 1 Level 2 Level 3

A passer-by **spotted** a dog **struggling** in the River Thames, so he jumped in to help. He saved the dog by attaching it to rescue chains near Westminster Bridge, but then he found himself in need of saving.

The **RNLI** launched a lifeboat which pulled the **Good Samaritan** and the dog from the water, which then re-united with its owner.

Difficult words: **spot** (see), **struggling** (having problems), **RNLI** (Royal National Lifeboat Institution), **Good Samaritan** (a helpful person who helps because he wants to).

You can watch the video news lower on this page.

## Man Saves a Dog – level 1

16-06-2017 15:00 Level 1 Level 2 Level 3

A dog falls into the **River Thames**. It **needs** help. A man sees it. He jumps into the water. He saves the dog.

However, now the man is in trouble. He needs help. People call for a **lifeboat**. It comes and saves him.

Difficult words: **River Thames** (a large river in London), **need** (if you need something, you must have it or something bad happens), **lifeboat** (a boat which saves people).

You can watch the original video in the Level 3 section.

Appendix F *The importance of the word "agreement"*

Paris Agreement and Trump – level 1
05-06-2017 15:00
The Paris Agreement is an **agreement** on **emitting** less greenhouse gases. Trump wants to pull back from the agreement.
This makes a lot of people angry. Leonardo DiCaprio, Justin Trudeau and Hillary Clinton tweet about this. Emmanuel Macron, Australia Greens Party Leader, and the Japanese Environment Minister talk about this. The Japanese Minister says that he is very angry.
However, some Americans are happy. They are mostly **miners**. They want to keep their jobs.
Difficult words: **agreement** (something that you promise to do), **emit** (put into the air), **greenhouse gas** (a gas which makes the planet warmer), **miner** (a person who mines – gets something like coal out of the ground).
You can watch the original video in the Level 3 section.

Appendix G *Vocabulary not found in the English Vocabulary Profile*

nuclear family
foster family
step family
godfather
godmother
beige
turquoise
submarine
hovercraft
drone
paper clip
hole punch
sharpener
binder

mare
stallion
filly
cub
bitch
pup
trout
carp
ape
orangutan
chimpanzee
gorilla
creek
spaniel

pug
hound
daffodil
daisy
beech
willow
spruce
fir
boletus
trolleybus
artery
shin
nipple
freckle

Appendix H *News article used to demonstrate the importance of context*

Monkey steals money – level 1
09-06-2016 15:00
A shop has a visitor. The visitor is a monkey. A man in the shop gives a piece of fruit to the monkey.
Then the monkey goes into the shop. He sits on a chair. Then he takes some money and he runs away. The monkey is a **thief**!
The man wants the money back. He tries to give the monkey a banana. The monkey does not want it. He runs away with 10,000 rupees. This is about 150 US dollars.
The man knows this monkey. The monkey was at the shop before. Now the man is more **careful**.
Difficult words: **thief** (somebody who steals), **careful** (when you do not do dangerous things).

Appendix I *Vocabulary missing from the English Vocabulary Profile*

palm (C1)
cactus
bamboo
soy (C1)
yam
coconut (B1)
date – fruit (C1)
passion fruit
mango (A2)
yak
chopstick
sesame
tuktuk
rickshaw
koala
panda

Appendix J *Key features of learner English by CEFR level*

| Level | Some key features | Examples from the CLC at the appropriate level |
|---|---|---|
| **A2** | • Simple sentences<br>• Sentences with clauses joined by *that*<br>• Descriptive phrases introduced by a past participle<br>• Simple direct *wh-* questions<br>• Simple sentences using infinitives<br>• Other infinitives<br>• Some modals | • We came back and went to bed<br>• I know that you have a new house too<br>• There are beautiful paintings painted by famous Iranian painters<br>• What are you going to wear?<br>• I want to buy a coat<br>• … something to eat<br>• We must be there at 7 o'clock in the morning. |
| **B1** | • *-ing* clauses<br>• *Whose* relative clauses<br>• Indirect questions<br>• Clauses with *what* as subject/object<br>Verb + object + infinitive<br>• easy + infinitive<br>• Some complex auxiliaries<br>• Additional modal uses | • Maria saw him taking a taxi<br>• … this famous painter whose pictures I like so much<br>• Guess where it is?<br>• This is what I think<br>• I ordered him to gather my men.<br>• The train station is easy to find.<br>• would rather, had better<br>• I have invited all his friends, so we should be 28 people. |
| **B2** | • *-ing* clause before the main clause<br>• It + verb + infinitive phrase<br>• *Wh*-clause as subject of main clause<br>• Reported speech<br>• Lexically-specific verbs/adjectives + infinitive | • Talking about spare time, I think we could go to the Art Museum<br>• It would be helpful to work in your group as well.<br>• What came after was what really changed my summer!<br>• I told him I loved his songs.<br>• … proved to be wrong, turned out to be …, expected to … |
| **C1** | • Lexically-specific verbs + object + infinitive<br>• Might for permission<br>• Fewer grammatical errors with agreement, countability or word formation | • I believe her to be this country's best representative.<br>• Might I tell you what we [should/will] discuss? |
| **C2** | • Some new lexically-specific verbs + object + infinitive<br>• Longer utterances with greater accuracy | • They declare some products to be the hits of the season |

Note: Adapted from *English Profile: Introducing the CEFR for English*, pp 11-12, by Cambridge ESOL, & Cambridge University Press, 2011.

Appendix K *Structural features that are significant for each level (from A2 C2 levels)*

**A2**

By the time the learner reaches A2, certain language features are emerging as being criterial, distinct characteristics. The average length of utterance is 7.9 words (based on learner data from the Cambridge Learner Corpus). These consist usually of simple sentences such as *We came back and went to bed* (see examples in A2.1 in the table below), *I met a lot of interesting people* (A2.2), or *I can give you my guitar* (A2.3). More complex sentences produced by A2 learners are of the type *I knew that you have a new house* or *I think the zoo is an interesting place*. As these examples show, the *that* which may be used to join clauses in sentences of this type may be included or omitted by learners at this level (A2.4). Users at this level are also able to write sentences such as *There are beautiful paintings painted by famous Iranian painters* where *there* is a descriptive phrase introduced by a past participle (A2.8). However, similar descriptive phrases introduced by a present participle are a characteristic of a B1 rather than an A2 user (B1.3).

A2 learners are able to produce simple direct *wh-* questions as well as statements (A2.6). For example, *What are you going to wear? How did you know I liked skateboards?* They are also able to produce simple sentences using infinitives – *I want to buy a coat* or *I would like to sell a book*, for example (A2.5). Infinitives are also found to be used correctly by A2 users in phrases like *something to eat* (A2.7).

A2 users also have gained the ability to use some modals in some of their basic senses. They can for instance use may, can and might for possibility (A2.10), must for obligation (A2.11) and should for advice (A2.12). In other words, an A2 level of English is characterised by the use of basic and relatively simple structures.

**B1**

As far as B1 is concerned, it can first of all be noted that learners typically continue to write steadily longer sentences as they go up the levels. The mean length of a B1 utterance is 10.8 words in comparison with 7.9 at A2 level and 14.2 at B2 level. One reason for the increasing length is that learners are able to cope with more complex sentences. For example, they produce sentences containing *-ing* clauses, which follow and complement the direct object of a main clause: *Maria saw him taking a taxi* (B1.2) or which function as adverbs and follow the main clause they modify: *He was sitting there, drinking a coffee and writing something* (B1.11). Interestingly, placing the *-ing* clause before the main clause as in *Talking about spare time, I think we could go to the Art Museum*, is a feature that characterises B2 users of the language (B2.1).
There are a number of other types of complex sentence that appear to be criterial at B1 level. While learners may cope well with simple relative clauses using *who* and *which* at lower levels, the use of *whose* as a relative pronoun is typical of this level e.g. *I met a very nice boy*

*whose name's John* (B1.6). Indirect questions similarly appear to be a marker of B1 language e.g. *Guess where it is, I don't know what to do* (B1.8 9). Simple clauses beginning with *what (…what I think, What I saw…)* acting as either the object or subject of a main sentence are also successfully used by learners at this level, for example, *This is what I think, What I saw was so amazing* (B1.7).

B1 learners are getting more confident in using the infinitive in more complex structures. While A2 level students cope well with verb + infinitive (e.g. *I would like to buy…* or *I want to sell*… A2.5), at B1 level students make use of structures consisting of verb + object + infinitive, with or without to e.g. *I would like you to come, I want you to do it, I helped him bake the cake* (B1.1). In addition, there are a number of other uses of infinitives linked with specific vocabulary items that seem to feature as typical features of a learner having reached B1 (B1.12 14). The level significant word in each case is bolded in the following examples: *Monica **seems** to be good. I was **supposed** to go to the English class* (B1.12). *Your friends **expect** you to spend an amazing holiday with them. I would **like** you to spend a weekend at my house. I **want** you to say hi [to] everybody* (B1.13). *The train station is **easy** to find* (B1.14).

Different auxiliary and modal uses help to map a learner's progress from one level to another. As far as B1 is concerned, the complex auxiliaries would rather and had better are typical features of a B1 learner (B1.10). So also are certain modal uses such as may for permission (*May I borrow ….*), *must* for logical necessity (*He is having a great time and must be really happy there*) and *should* for probability (*I have invited all his friends, so we should be 28 people*) (B1.16 18). The number of error types that significantly improve from A2 to B1 is fairly modest, and improvements focus on the use of quantifiers (e.g. *some, a lot of, any, several, other, every* etc.).

To sum up, the B1 language user has mastered the basic structures of the language and is beginning to attempt to produce more complex language.

**B2**

The mean length of a B2 user's utterance is 14.2 words, an increase of 3.4 words on Bl. This can be explained in part by the fact that increasingly complex sentences are used as the learner progresses through the levels.

B2 language use typically, for example, may include sentences beginning with an adverbial clause introduced by an *-ing* word (B2.1) e.g. *Talking about spare time, I could go to the Art Museum*.

Another example of a more complex structure which characterises a B2 user is a sentence introduced by *It* and followed by an infinitive phrase e.g. *It would be helpful to work in your group as well* (B2.2). Yet another B2 structure is a sentence beginning with a *Wh-* word introducing a clause acting as the subject of the main clause of the sentence (B2.3) e.g. *What attracted me the most was the possibility of meeting people of the same interests*. A third

example of a structure characterising a B2 level user is a main sentence with a direct object followed by a subordinate complement clause with or without that (B2.4), for example, *I told him (that) I loved his songs.*

As at B1 level, there are certain verbs and adjectives which, when associated with infinitive structures, can be seen as characteristic of the level (B2.6-9); examples of some of these are given below with the significant words highlighted in bold:

*My worries **proved** to be wrong.*
*Unfortunately for me the situation **turned out** to be opposite to what I thought it was.*
*And whenever money is involved, some problems are **likely** to happen.*
*You are **sure** to arrive at work on time.*
*I would **prefer** my accommodation to be in log cabins.*
*How many hours a day should I be **expected** to work?*
*Your theatre is **known** to present excellent spectacles.*
*So zoos could be the only place where people could spend their time avoiding the pollution we are **obliged** to live with every day.*
*Woods is **thought** to stand for all of white people and this book could have an influence on them.*
*The grammar and vocabulary are a bit **hard** to learn.*

There are no instances of modal use which can be seen as being criterial for a B2 level of language. Like in B1, the number of error types that appear to improve substantially at this level remains small.

**C1**

At C1 level the average length of utterance continues to increase, rising from B2's 14.2 to 17.3 words.

It is noticeable once learners move into C levels that the criterial, distinguishing features of their language are a combination of lexis and structure. Learners use structures mastered at earlier levels but with a much wider range of vocabulary and in more accurate ways. Verbs such as *chance*, *believe*, *find*, *suppose*, *take*, *assume*, *discover*; *feel* and *prove* used in complex structures which include an infinitival clause are characteristic of the language of a C1 level learner (C1.1-3). Here are some examples (see Table 6 for more):

*• I chanced to know about your competition from an international magazine.*
*• Being born and raised in Mexico, I believe her to be the country's best representative in the world.*
*• I can assure you that the strike isn't as worrying as you suppose it to be.*
*• Secondly, the low cost of membership and entry was assumed to be an advantage as well.*
*• The internet is a valuable tool, which can be proved to be the most important aspect in the learning process.*

The only strictly grammatical features which might be seen as criterial, new features of language as the learner moves from B2 to C1, are the use of a double *'s* genitive structure (e.g. *the bride's family's house*) (C1.4) and the use of *might* to convey the idea of permission (*Might I tell you what we discussed*?) (C1.5).

In general, however, it is not mainly the case that the C1 learner is mastering new grammatical features, it is more that he or she is using those that were already available at B2 level in a much more accurate way. Far fewer grammatical errors with, say, agreement, countability or word form occur at C1 than at B2 level.

**C2**

The trends noted at C1 continue into C2 in that what marks out C2 users of the language is the fact that they have a greater grammatical accuracy and a wider lexical range than C1 users; in other words there are no specific new structures appearing at this level. C2 users also continue the trend of producing a longer average utterance than at lower levels; the mean length of utterance for C2 is 19 words as opposed to 17.3 at C1.

Lexical range in relation to specific structures continues to expand. For example *declare*, *presume*, *remember* and the adjective *tough* are used with the following complex infinitive complements at C2 level (C1.1-3):

*They declare some products to be the hits of the season, thus creating fashion and few of us want to be unfashionable.*
• *He presumed work to be the way to live.*
• *Not only meetings with new people are presumed to give new experiences.*
*She remembered her father to be a lively, tall and broad shouldered man with a beard that tickled when he bent down to kiss her goodnight.*

• *What she knew would be really tough to live with was the reason of his death.*

At the same time, the C2 level sees the highest number of significant grammatical error improvements of all CEFR levels. Like in C1, there are significant reductions of error rates for most of the error types identified in the Cambridge Learner Corpus. Overall, it can be said that it is at the C levels where learners appear to be mastering and accurately using the majority of grammatical features in English.

Note: Adapted from *English Profile: Introducing the CEFR for English*, pp. 12-15, by Cambridge ESOL, & Cambridge University Press, 2011.

Appendix L *Flesch-Kincaid readability*

Developed by Rudolf Flesch and J. Peter Kincaid, the Flesch-Kincaid readability scores are the most widely used measures of readability. And they are used by the United States military to evaluate the readability of their manuals.

The first number, Flesch-Kincaid reading ease, is based on a ranking scale of 0-100, and the higher your score, the better. Low scores indicate text that is complicated to understand. So if your website receives a low Flesch-Kincaid reading ease score, you will likely need to simplify your text.

For most business writing, a score of 65 is a good target, and scores between 60 and 80 should generally be understood by 12 to 15 year olds.

Flesch-Kincaid reading ease formula: 206.835 - 1.015 x (words/sentences) - 84.6 x (syllables/words).

The second number, Flesch-Kincaid grade level, tells you the American school grade you would need to be in to comprehend the material on the page.

As a measure, most of your writing should be able to be understood by students in seventh grade.

For example, The Huffington Post's website has an average grad level of about 7, meaning that it should be easily understood by 12 to 13 year olds.

Flesch-Kincaid grade level formula: 0.39 x (words/sentences) + 11.8 x (syllables/words) - 15.59.

Both Flesch-Kincaid reading ease and grade level use the same core metrics: word length and sentence length. But they correlate inversely. If you receive a high score on the reading ease test, you should receive a lower grade level score.

Note. Adapted from https://www.webpagefx.com/tools/read-able/flesch-kincaid.html

Appendix M Gunning Fog Index

Developed by American businessman, Robert Gunning, the Gunning Fog Index estimates the years of formal education needed to comprehend a passage of text on the first reading.
Gunning noticed that many high school students were unable to read, and he attributed this to a writing problem, believing that newspapers and business documents were full of complicated text.

He went on to become the first person to take readability research into the workplace, and he developed a consulting firm dedicated to readability. Gunning then worked with newspapers and popular magazines, helping writers and editors understand how to write to their target audience.

To calculate the Gunning Fog Index, you need to take a passage of text at least 100 words and count the number of exact words and syllables. Then, divide the total number of words in the sample by the total number of sentences. This will give you the Average Sentence Length (ASL).

Next, you will need to count the number of words that contain three or more syllables that are not proper nouns, combinations of easy or hyphenated words, or two-syllable verbs made into three by adding -es and -ed endings. Then, you will need to divide that number by the total number of words in the sample passage. This will give you the Percent Hard Words (PHW).

Finally, you will need to add the ASL and PHW and multiply the result by 0.4.

Gunning Fog Index formula: Grade level= 0.4 (ASL + PHW).

The principle behind the Gunning Fog Index formula is that short sentences written in plain English receive a better score than longer sentences written in complex language.

A Gunning Fox Index score of 7 or 8 is ideal, and anything higher than 12 is too complex for most people to read.

Popular magazines, such as TIME and the Wall Street Journal average Gunning Fog scores of 11, while Shakespeare has a Gunning Fog Index of about 6.

Though considered a fairly accurate measure of readability, the Gunning Fog Index has some limitations. For instance, it doesn't take into account that not all multi-syllable words are difficult.

Note. Adapted from https://www.webpagefx.com/tools/read-able/gunning-fog.html

Appendix N *Coleman Liau Index*

Developed by linguists Meri Coleman and T.L. Liau, the Coleman Liau Index is designed to evaluate the U.S. grade level necessary to understand the text.

Instead of syllables per word and sentence lengths, the Coleman Liau Index relies on characters and uses computerized assessments to understand characters more easily and accurately.

Coleman and Liau developed the formula to automatically calculate writing samples instead of manually coding the text. Unlike syllable-based readability indicators, the Coleman Liau Index does not require users to take into account the syllable-counts in the text. Therefore, passages can easily be scanned into a word processor to calculate the Coleman Liau Index.

According to Coleman and Liau, word length in letters is a better predictor of readability than word length in syllables.

Coleman Liau Index formula: 5.89 x (characters/words) - 0.3 x (sentences/words) – 15.8.

The score you receive indicates the U.S. school level a person needs to understand the text. So if you receive a rating of 9.5, your text should be easily understood by a student in 9th or 10th grade.

If you’re writing to elementary school students and your Coleman Liau score indicates the necessity of a high school reading level, you’re probably going to need to re-evaluate your content. On the other hand, you probably shouldn’t score a third-grade level if you’re writing to college students.

In order to effectively communicate with your target audience, your writing needs to reflect the grade level of the people you are trying to reach.

Note. Adapted from https://www.webpagefx.com/tools/read-able/coleman-liau-index.html

## Appendix O Automated Readability Index

The automated readability index is a readability test designed to measure the how easy your text is to understand. Similar to other popular readability tools, the ARI gives you an estimate of the U.S. grade level necessary to comprehend a passage of text.

For example, if your ARI score is 10, then your text should be understood by high school students.

ARI is derived from ratios that represent word difficulty (number of letters per word) and sentence difficulty (number of words per sentence).

Most readability scores consist of two factors. The first factor relates to sentence structure, and typically measures the average number of words per sentence. Readability indices also take into account word structure, and are generally based on the average number of syllables per word or the proportion of easy words determined by referencing a word list.

Automated Readability Index formula: 4.71 x (characters/words) + 0.5 x (words/sentences) - 21.43.

Your Automated Readability Index score approximates the age needed to understand the text. For a breakdown of grade levels in the U.S., you can reference this chart:

- 5-6 yrs. old — Kindergarten
- 6-7 yrs. old — First Grade
- 7-8 yrs. old — Second Grade
- 8-9 yrs. old — Third Grade
- 9-10 yrs. old — Fourth Grade
- 10-11 yrs. old — Fifth Grade
- 11-12 yrs. old — Sixth Grade
- 12-13 yrs. old — Seventh Grade
- 13-14 yrs. old — Eighth Grade
- 14-15 yrs. old — Ninth Grade
- 15-16 yrs. old — Tenth Grade
- 16-17 yrs. old — Eleventh grade
- 17-18 yrs. old — Twelfth grade
- 18-22 yrs. old – College

Unlike most readability measures, ARI takes into account characters per word instead of syllables per word. This is because the number of characters is more accurately interpreted by computer programs.

If your ARI doesn't accurately reflect the age or grade level of your target audience, you may need to revisit your text and make the necessary edits to ensure your writing resonates with the intended target.

Note. Adapted from https://www.webpagefx.com/tools/read-able/automated-readability-index.html

Appendix P SMOG Index

Created by G Harry McLaughlin, the SMOG Index estimates the years of education a person needs to comprehend a piece of writing, and it was created as an improvement of other readability formulas.

Though some believe SMOG is an acronym for Simple Measure of Gobbledygook, it is widely recognized to be a nod to Robert Gunning's Fog Index.

To calculate the SMOG Index, you will need to take the entire passage of text being evaluated and select 10 sentences in a row near the beginning, 10 sentences in a row in the middle, and 10 sentences in a row near the end of the text.

Then, you will count every word with three or more syllables in each group of 10 sentences—even if the same word appears multiple times.

Next, you need to calculate the square root of the number derived in the previous step, and round it off to the nearest 10.

That number is your SMOG Index score.

The SMOG formula is generally accepted as appropriate for secondary age (4th grade to college) readers.

Unlike many other readability measures, which test for around 50-75% comprehension, McLaughlin used a 100% correct-score criterion.

Your SMOG Index score approximates the age needed to understand the text. And you can reference this chart for a breakdown of grade levels in the U.S.

- 5-6 yrs. old — Kindergarten
- 6-7 yrs. old — First Grade
- 7-8 yrs. old — Second Grade
- 8-9 yrs. old — Third Grade
- 9-10 yrs. old — Fourth Grade
- 10-11 yrs. old — Fifth Grade
- 11-12 yrs. old — Sixth Grade
- 12-13 yrs. old — Seventh Grade
- 13-14 yrs. old — Eighth Grade
- 14-15 yrs. old — Ninth Grade
- 15-16 yrs. old — Tenth Grade
- 16-17 yrs. old — Eleventh grade
- 17-18 yrs. old — Twelfth grade
- 18-22 yrs. old – College

Note. Adapted from https://www.webpagefx.com/tools/read-able/smog-index.html

Appendix Q *Corpus of level-2 news articles*

A lorry drove into a department store in the centre of Stockholm. According to the police, there are at least two dead in this awful terrorist attack.
Police evacuated the city centre and the main central station. They did not arrest anyone yet, but they are working on it.
The Swedish Prime Minister said that the government is doing everything in its power to find out what happened. He also recommended that the public be careful and listen to police broadcasts.
The attack came after other terrorist incidents in France, Germany and the United Kingdom.

An American man was leaving home early in the morning when he saw a bear. He took out his phone and filmed it, because it walked on its hind legs!
The bear didn't know that the man filmed it. It seemed like it went for a walk. After the man uploaded the video on the Internet, several people said that they saw the bear, too.

About 200 inmates escaped from a prison on Sumatra. It happened on Friday when the prisoners were out of their cells during prayers. The capacity of the prison is for 200 people, but it is holding almost 1,900 people.
Hundreds of police and soldiers came to the prison and set up roadblocks around the city. They caught over half of the prisoners on the same day.

Utility workers found a baby otter alone and saved it. It was struggling to survive in a canal in Arizona. The workers took it to a wildlife centre. It was dehydrated, hungry, and full of fleas.

People gave it cat's milk and fish, and it got better soon. It will now start a new life in a wildlife park.

A four-year-old girl from Arkansas was on a church bus when the bus door opened. She fell out onto the highway. A man filmed the shocking moment on camera.

A volunteer firefighter was driving behind the bus. He quickly came to the girl and picked her up. He would not normally do this, but there was a lot of traffic. The girl's family says that they think that she will be fine.

An American university professor was climbing in Nepal when he fell down a hole. The crack in the ice was 22 metres deep, and the man got trapped in it.

The man thought that he was going to die. His arm and five ribs were broken, but he didn't give up and he moved out of the gap. It took him six hours. He reached his tent in another three hours. Help came the next day.

Appendix R *Corpus of level-3 news articles*

A lorry drove into a department store in the centre of Stockholm. Someone hijacked the lorry before the attack. According to the police, there are at least two dead in this awful terrorist attack.
Police evacuated the city centre and main central station. They have not arrested anyone yet, but they are working on it.
The Swedish Prime Minister said that the government is doing everything in its power to find out what happened. He also advised the public to be careful and to listen to police broadcasts.
The attack came after other terrorist incidents in France, Germany and the United Kingdom.

Early one morning, a bear was walking around a neighbourhood like a human in New Jersey, USA. Ian Bowman from Oak Ridge spotted the animal and quickly took out his phone to record the unusual sight.
The bear was blissfully unaware of being filmed as he walked on his hind legs and took a casual stroll. After Ian uploaded the video online, other people in the neighbourhood reported that they had seen the same bear performing that same stunt.

About 200 inmates escaped from an overcrowded prison on Sumatra which is holding about 1,900 people despite its capacity of only 300 people. The inmates escaped on Friday when they were let out of their cells for Friday prayers. They overwhelmed the six guards on duty who the inmates accused of being violent.
Hundreds of police and soldiers were deployed to the prison, and they set up roadblocks around the city. According to the Justice and Human Rights Ministry, they recaptured over half of the prisoners on that same day.

People found a baby otter alone and struggling to survive in a canal in Arizona. Luckily, utility workers spotted and rescued it.
When they took it to a wildlife centre, it was dehydrated, hungry and infested with fleas, but it was not long before its health recovered. The people there fed it cat's milk and mashed trout.
The pup will now start a new life in one of the state's wildlife parks.

A dashboard camera recorded the harrowing moment a four-year-old girl fell out of the back of a bus. The door of the church bus swung open and the young girl fell out onto the Arkansas state highway.
A volunteer firefighter quickly rescued her, as he happened to be driving behind the bus. He would not normally move the girl, but he had to because of the busy traffic. The girl's family told reporters that they believe that she will make a full recovery.

John, who is an associate geography professor at Western Kentucky University, was climbing a mountain in Nepal as part of a research project and fell into a crevasse.
He filmed his struggle inside the 22-metre hole, saying that he thought that he was going to die after breaking his arm and five ribs, but he managed to crawl out the gap. It took him six hours, using only an ice axe, and another three to reach his tent, where he spent the night before rescuers reached him the next day.

Appendix S *Global Scale descriptors for CEFR levels*

| | | |
|---|---|---|
| Proficient User | C2 | Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations. |
| | C1 | Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices. |
| Independent User | B2 | Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options. |
| | B1 | Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics, which are familiar, or of personal interest. Can describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans. |
| Basic User | A2 | Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need. |
| | A1 | Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help. |

Note: Adapted from *English Profile: Introducing the CEFR for English*, p. 4, by Cambridge ESOL, & Cambridge University Press, 2011.

Appendix T *List of A1–A2 vocabulary*

a.m.
able
about
above
accident
across
activity
actor
actually
add
address
adjective
adult
advanced
adventure
adverb
advertisement
advice
aeroplane
afraid
after
afternoon
afterwards
again
against
age
aged
ago
agree
air
airport
alarm
album
alcohol
all
almost
alone
along
already
alright
also
always
amazing
ambulance
among
an
and
angry
animal
another
answer
any
anybody
anymore
anyone
anything
anyway
anywhere
apartment
apple
appointment
April
area
arm
armchair
around
arrive
art
artist
as
ask
assistant
at
attractive
August
aunt
autumn
available
away
awesome
baby
back
bad
badly
badminton
bag
bake
ball
balloon
banana
band
bank
bar
barbecue
baseball
basketball
bat
bath
bathroom
battery
be
beach
bean
bear
beard
beautiful
because
become
bed
bedroom
beer
before
begin
beginner
beginning
behind
believe
belong
below
belt
beside
best
better
between
bicycle
big
bike
bill
biology
bird
birthday
biscuit
bit
black
blackboard
blanket
blonde
blood
blue
board
boat
body
boil
boiled
book
bookcase
bookshelf
bookshop
boot
bored
boring
borrow
boss
both
bother
bottle
bottom
bowl
box
boy
boyfriend
brain
bread
break
breakfast
bridge
bright
brilliant
bring
broken
brother
brown
brush
build
building
burger
bus
business
businessman
businesswoman
busy
but
butter
buy
by
bye
café
cake
calendar
call
called
camera
camp
camping
can't
cannot
cap
capital
car
card
careful
carefully
carpet
carrot
carry
cartoon
case
cash
castle
cat
catch
cathedral
CD
ceiling
cent
centimetre
centre
century
cereal
certainly
chain
chair
champagne
change
channel
chat
cheap
check
cheese
chef
chemist
chemistry

cheque
chess
chicken
child
chilli
chip
chocolate
choose
church
cigarette
cinema
circle
city
class
classical
classmate
classroom
clean
cleaner
clear
clearly
clever
click
climb
climbing
clock
close
closed
clothes
cloud
cloudy
clown
club
coach
coat
coffee
cola
cold
colleague
collect
college
colour
comb
come
comfortable
comic
company
comparative
competition
complete
computer
concert
congratulations
contact
conversation
cook
cooker
cooking
cool
copy
corner
correct
cost
could
country
countryside
course
cousin
cover
cow
crazy
cream
credit
cricket
crisp
cross
crowd
crowded
cry
cup
cupboard
curry
curtain
customer
cut
cycling
dad
daily
dance
dancer
dancing
danger
dangerous
dark
date
daughter
day
dead
dear
December
decide
deep
degree
delay
dentist
department
describe
desert
desk
dessert
details
diary
dictionary
die
difference
different
difficult
digital
dining
dinner
dinosaur
diploma
directions
dirty
disco
discount
discuss
dish
dishes
do
doctor
document
dog
doll
dollar
don't
door
double
down
download
downstairs
Dr
draw
drawer
drawing
dream
dress
dressed
drink
drive
driver
driving
drum
dry
duck
during
DVD
each
ear
early
earn
earring
easily
east
easy
eat
egg
eight
eighteen
eighth
eighty
electric
electricity
elephant
eleven
else
email
empty
end
engine
engineer
enjoy
enough
enter
entrance
envelope
especially
euro
even
evening
ever
every
everybody
everyone
everything
everywhere
exactly
exam
examination
example
excellent
except
excited
exciting
excuse
exercise
exit
expensive
explain
extra
eye
face
fact
factory
fail
fair
fall
family
famous
fan
fantastic
far
farm
farmer
fashion
fast
fat
father
favourite
February
feel
field
fifteen
fifth
fifty
file
film
final
finally
find

fine
finger
finish
fire
first
fish
fishing
fit
five
flat
flight
floor
flower
fly
fog
foggy
folder
follow
food
foot
football
footballer
for
foreign
forest
forget
fork
form
forty
four
fourteen
fourth
free
fresh
Friday
fridge
fried
friend
friendly
from
front
fruit
full
fun
funny
furniture
further
future
game
garage
garden
garlic
gas
gate
geography
get
gift
girl
girlfriend
give

glad
glass
glasses
glove
go
goal
God
gold
golden
golf
good-looking
goodbye
gram
grammar
grandchild
granddad
granddaughter
grandfather
grandma
grandmother
grandpa
grandparent
grandson
granny
grape
grass
great
green
grey
grilled
group
grow
guess
guest
guide
guidebook
guitar
guy
hair
half-price
hall
ham
hand
handbag
happen
happy
hard
hat
hate
have
he
head
headache
health
healthy
hear
heart
heating
heavy
helicopter

hello
help
her
here
hers
herself
hey
hi
high
hill
him
himself
hip-hop
his
history
hit
hobby
hockey
hold
holiday
home
homework
honey
hope
horrible
horse
hospital
hot
hotel
hour
house
housewife
how
however
hundred
hungry
hurt
husband
I
ice
ID
idea
identification
if
ill
immediately
important
improve
in
include
including
indoor
information
insect
inside
instead
instructions
instrument
interested
interesting

international
Internet
into
invitation
invite
island
it
its
itself
jacket
jam
January
jazz
jeans
jewellery
job
join
journey
juice
July
jump
jumper
June
just
keep
key
keyboard
kick
kid
kill
kilo
kilogram
kilometre
kind
king
kiss
kit
kitchen
kite
knife
know
lake
lamp
language
laptop
large
last
late
later
latest
laugh
lazy
learn
leather
leave
left-hand
leg
lemon
lemonade
lend

less
lesson
let's
letter
level
library
licence
lie
life
lift
light
like
line
lion
list
listen
litre
little
live
living
long
look
lose
lost
lot
lots
loud
love
lovely
low
luck
lucky
luggage
lunch
lunchtime
machine
mad
magazine
magic
mail
main
make-up
man
manager
mango
many
map
March
mark
market
married
match
maths
matter
May
maybe
me
meal
mean
meat
mechanic
medicine
meet
meeting
melon
member
memory
menu
message
metre
midday
middle
midnight
might
milk
million
mine
mineral
minus
minute
mirror
Miss
missing
mistake
mix
mobile
model
modern
moment
Monday
money
monkey
month
moon
more
morning
mosque
most
mother
motorbike
motorway
mountain
mouse
mouth
move
movie
MP3
Mr
Mrs
Ms
much
mug
mum
museum
mushroom
music
musical
must
my
myself
name
national
nationality
nature
near
nearly
neck
necklace
need
negative
neighbour
net
never
new
news
newspaper
next
nice
night
nine
nineteen
ninety
ninth
no
nobody
noise
noisy
noon
normal
north
nose
not
note
notebook
notes
nothing
notice
noun
November
now
number
nurse
occupation
October
of
off
offer
office
often
oh
oil
OK
old
omelette
on
once
one
onion
online
only
open
opera
opposite
or
orange
order
other
our
ours
ourselves
out
outside
over
own
p.m.
pack
page
pain
paint
painter
painting
pair
pale
paper
pardon
parent
park
part
partner
party
pass
passenger
passport
past
pasta
path
pay
PC
pear
pen
pence
pencil
penfriend
people
pepper
per
perfect
perfume
perhaps
person
pet
petrol
phone
photo
photograph
photographer
photography
physics
piano
pick

picnic
picture
piece
pig
pillow
pilot
pink
pity
pizza
place
plan
plane
plant
plastic
plate
platform
play
player
playground
pleasant
please
pleased
plural
plus
pocket
point
police
policeman
policewoman
polite
pool
poor
pop
popular
possible
possibly
post
postcard
poster
potato
pound
practice
practise
pray
prefer
prepare
present
pretty
price
print
printer
prize
probably
problem
program
programme
project
pub
pull
pupil
purple
purse
push
put
puzzle
quarter
queen
question
quick
quiet
quiz
rabbit
race
racket
radio
railway
rain
raincoat
rap
rat
read
reading
ready
real
really
reason
receipt
receive
receptionist
record
red
remember
rent
repair
repeat
rest
restaurant
return
rice
rich
ride
right-hand
ring
river
road
roast
rock
roof
room
round
roundabout
rubber
rugby
ruler
run
runner
running
sad
safe
sailing
salad
sales
salesperson
salt
same
sandwich
Saturday
sauce
sausage
save
say
scarf
school
schoolchild
science
scissors
scooter
screen
sea
seat
second
secretary
see
sell
send
sentence
September
serve
set
seven
seventeen
seventh
seventy
several
shall
shame
shampoo
share
she
sheep
sheet
shelf
ship
shirt
shoe
shop
shopping
short
shorts
should
shout
show
shower
shut
sick
side
sightseeing
sign
silver
simple
since
sing
singer
singing
single
singular
sink
sister
sit
site
sitting
six
sixteen
sixth
sixty
size
skate
skateboard
skateboarding
skating
skiing
skirt
sky
sleep
slice
slim
slow
slowly
small
smoke
smoking
snack
snow
snowboarding
so
soap
sock
sofa
soft
software
some
somebody
someone
something
sometimes
somewhere
son
song
soon
sorry
sort
soul
sound
soup
south
space
spare
speak
speaker
special

spell
spelling
spend
spoon
sport
sports
spring
square
stadium
staff
stage
stairs
stamp
stand
star
start
station
stay
steak
steal
still
stomach
stop
storm
story
straight
strange
street
strong
student
studies
study
subject
such
sugar
suit
suitcase
summer
sun
Sunday
sunglasses
sunny
superlative
supermarket
supper
sure
surf
surfing
surname
surprise
surprised
sweater
sweet
sweets
swim
swimming
T-shirt
table
take
talk
tall
taxi
tea
teach
teacher
team
teenager
telephone
television
tell
temperature
ten
tennis
term
terrible
test
text
textbook
than
thank
thanks
that
the
theatre
their
theirs
them
themselves
then
there
these
they
thin
thing
things
think
third
thirsty
thirteen
thirty
this
those
thousand
three
through
throw
thunderstorm
Thursday
ticket
tidy
tie
tights
till
time
timetable
tired
to
toast
today
toe
together
toilet
toilets
tomato
tomorrow
tonight
too
tooth
toothache
toothbrush
top
tour
tourist
towel
town
toy
traffic
train
trainer
tram
travel
tree
trip
trousers
true
try
Tuesday
tune
turn
TV
twelve
twenty
twice
two
type
tyre
umbrella
uncle
under
underground
understand
unfortunately
unhappy
uniform
university
until
unusual
up
upset
upstairs
us
use
useful
usually
variety
various
vegetable
verb
very
video
view
village
violin
visit
visitor
vocabulary
volleyball
wait
waiter
waitress
wake
walk
walking
wall
wallet
want
war
warm
wash
washing-up
watch
water
way
we
wear
weather
Web
website
Wednesday
week
weekday
weekend
weekly
welcome
well
west
wet
what
wheel
when
where
which
while
white
who
whole
why
wide
wife
wild
will
win
wind
window
windy
wine
winner
winter
with
without

| | | | |
|---|---|---|---|
| woman | worried | yeah | young |
| wonderful | worry | year | your |
| wood | worse | years | yours |
| wooden | worst | yellow | yourself |
| wool | would | yes | zero |
| word | wow | yesterday | zoo |
| work | write | yet | |
| worker | writing | yogurt | |
| world | wrong | you're | |

Note. Adapted from http://vocabulary.englishprofile.org/staticfiles/about.html