

LagunTest: A NLP Based Application to Enhance Reading Comprehension

Itziar Gonzalez-Dios, Kepa Bengoetxea, Amaia Aguirregoitia

University of the Basque Country (UPV/EHU)

Rafael Moreno *Pitxitxi*, 2. 48013 Bilbao (Bizkaia)

{itziar.gonzalezd, kepa.bengoetxea, amaia.aguirregoitia}@ehu.eus

Abstract

The ability to read and understand written texts plays an important role in education, above all in the last years of primary education. This is especially pertinent in language immersion educational programmes, where some students have low linguistic competence in the languages of instruction. In this context, adapting the texts to the individual needs of each student requires a considerable effort by education professionals. However, language technologies can facilitate the laborious adaptation of materials in order to enhance reading comprehension. In this paper, we present *LagunTest*, a NLP based application that takes as input a text in Basque or English, and offers synonyms, definitions, examples of the words in different contexts and presents some linguistic characteristics as well as visualizations. *LagunTest* is based on reusable and open multilingual and multimodal tools, and it is also distributed with an open license. *LagunTest* is intended to ease the burden of education professionals in the task of adapting materials, and the output should always be supervised by them.

Keywords: language technologies for education, reading comprehension, multilingual and multimodal applications

1. Introduction

In the Basque Autonomous Community, Basque and Spanish are used as official languages for teaching, whereas English is introduced at an early age. In this education system, most of the students study in Basque, but for a majority it is not their native language (immersion programme). Furthermore, there is an increasing number of schools teaching some of the subjects through English, which is a foreign language for most of the students.

On the other hand, one of the challenges for nowadays schools is to integrate children from different countries and with different native languages in a trilingual education system. At school, these incoming students are expected to read and understand contents in Basque and Spanish, which are foreign languages for them.

In the last years of primary education and in secondary education, students must deal with a large amount of written information in a second language (L2), and it is not easy for education professionals to offer the individual support that is ideally required. Besides, evaluations have proved that the reading comprehension ability has been decreasing during the last years in the three languages used at school, above all in the fourth degree of primary (ISEI-IVEI, 2016a; ISEI-IVEI, 2016b; ISEI-IVEI, 2016c). Therefore, it is crucial to provide teachers and education professionals with tools that facilitate the adaptation of written texts to maximize learning.

In this context, language technologies and natural language processing (NLP) tools can support personalization and adaptation of written contents to promote understanding. Similarly, automatic text adaptation can ease the burden on the education professionals and boost their efficiency to teach contents.

In this paper we present *LagunTest*¹, a web application based on open source language technologies that aims to help education professionals to enhance the comprehension

of texts written in Basque or English to assist students with different linguistic competence. Specifically, *LagunTest* i) identifies the most frequent words to ease global understanding using a word cloud; ii) offers additional information through images representing the meanings, definitions and synonyms adapted to the student's level; iii) displays the different PoS (part of speech) using colors and iv) represents the dependency tree of the sentence to highlight the morphosyntactic characteristics. *LagunTest* is based on multilingual, multimodal, open source and reusable NLP tools and resources. Its code is available at <https://github.com/kepaxabier/LagunTest> under GNU General Public License v3.0.

This paper is structured as follows: In Section 2 we present the related work; in Section 3, we detail the design criteria and technical resources used to develop *LagunTest* and we explain its functionalities in Section 4. We discuss the limitations of the application in Section 5 and we conclude and outline the future work in Section 6. The text we use as example is shown in Appendix A.

2. Related work

Adapting educational material and building educational applications have directed the attention of many researchers in NLP. Examples of their outcomes are for instance the works presented in main NLP conferences, journals and in the series of BEA workshops organized by ACL SIGEDU². Besides, works on inclusive and adaptive technologies and in automatic text simplification have also been presented in the specialised workshops such as PITER (Predicting and Improving Text Readability for target reader populations) organised in 2012, 2013 and 2014; NLP4ITA (Natural Language Processing for Improving Textual Accessibility) in 2012 and 2013; ATS-MA (Automatic Text Simplification-Methods and Applications in the Multilingual Society) in 2014; ISI-NLP (Improving Social Inclusion using NLP: Tools and resources), QATS (Quality Assessment for Text

¹<http://178.128.198.190:8080/>

²<https://sig-edu.org/>

Simplification) and Computational Linguistics for Linguistic Complexity (CL4LC) in 2016; and ATA (Workshop on Automatic Text Adaptation) in 2018.

In the educational domain, reading comprehension and reading strategies have been investigated, even from institutional perspectives (National Reading Panel (US), 2000; RAND Reading Study Group, 2002). Moreover, in the last years, the use of technologies at schools in order to assist students with reading difficulties has also been a research line of interest e.g. (Gasparini and Culén, 2012; Haßler et al., 2016; Crossley et al., 2017).

Regarding the educational technologies, most of the works have focused on English and major languages. For example, R-A Reading (resource assisted reading)³ offers additional contexts and definition for the words e but, it is only available for French and English. Multidict⁴ offers a definition adapted to two levels of difficulty, different languages and different dictionaries. Wordlink⁵ makes web pages more accessible by linking the words to Multidict.

But less spoken languages such as Basque have plenty of prototypes and tools in this area e.g. question-answering for education (Aldabe et al., 2006; Aldabe et al., 2013), auto-evaluation of essays (Castro-Castro et al., 2008), readability assessment (Gonzalez-Dios et al., 2014) and automatic text simplification (Gonzalez-Dios, 2016). The creation of multilingual vocabulary exercises by means of NLP tools has also been explored (Agirrezabal et al., 2019).

Regarding levelled materials, Clilstore⁶ offers learning materials in different languages (including Basque, but limited) which are organised in levels according to the CEFR. Moreover, these materials include links to definitions from dictionaries of the presented vocabulary.

3. Design criteria and resources

In this section we present the resources and language technologies that we have been used to build the application. All the resources we use are open sourced or open licensed. Moreover, we have decided to use multilingual and multimodal resources, so that the application can be easily adapted to other languages.

3.1. Determining vocabulary level

In order to obtain an application adaptable to the different language levels and literacy skills of the students, we have included a feature to select among three vocabulary levels: beginner, intermediate and advanced. This level selection allows to adjust the performance and the results to the level of the student and it determines which words will be adapted and which will be displayed. By adjusting the output to the selected level we avoid oversimplification of the text but also the difficulties of too demanding words for the student. To define the levels, we have followed two strategies: a tool-based strategy and a corpus-based strategy.

In the case of English, we have followed a tool-based strategy: we have used the Wordfreq tool (Speer et al., 2018),

which provides estimations on how often a word is used in 36 languages, for example in Spanish and English. This tool returns the word frequency of a word in a corpus of 10^9 words as the logarithm in base 10 of the number of times a word appears per billion words in different sources such as Wikipedia, Subtitles, News, Books, Web Texts, Twitter, Reddit and others. In Table 1 we show examples for each scale of values that Zipf returns. Following van Heuven et al. (2014), a Zipf value equal to 3 means that the word appears 10^3 times for every 10^9 words, that is, 1 per million. So, using wordfreq and based on Begoetxea et al. (2020), where different values have been tested for educational purposes, we have determined the following values for the vocabulary levels:

- **Beginner level:** words with a Zipf value less than or equal to 8
- **Intermediate level:** words with a Zipf value less than or equal to 5
- **Advanced level:** words with a Zipf value less than or equal to 3

In the case of Basque, as Wordfreq does not provide frequencies for this language, we have followed a corpus-based strategy. Specifically, we have performed a corpus analysis to determine the vocabulary levels. The resource we have used is the frequency list of the corpus *Lexikoaren Behatokia*⁷ from *Euskaltzaindia*, the Academy of the Basque Language. This list was created on the 2014 version of the corpus that had 41,773,391 words and it has been used before in automatic text simplification studies for Basque (Gonzalez-Dios, 2016). Based on the values of the list and the distributions of the frequencies, we have stipulated three levels of words:

- **Beginner level:** words whose lemma appears 100,000 or less than 100,000 times in the corpus
- **Intermediate level:** words whose lemma appears 34 or less than 34 times in corpus
- **Advanced level:** words whose lemma appears 6 or less than 6 times in the corpus

3.2. Choosing the NLP tools

For the automatic analysis of the text, LagunTest can be easily adapted to any model which is the state-of-the-art in segmentation (tokenization and sentence-splitting), lemmatization, POS tagging and dependency parsing task for over 50 languages. In this paper, we have tested NLP-Cube (0.1.0.7) and StanfordNLP (0.2.0) (Qi et al., 2019), that were the best systems in English on CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies (Zeman and Hajič, 2018) but we have decided to use the StanfordNLP 0.2.0 tool⁸ that was trained

³www.lexitutor.ca

⁴<https://multidict.net/multidict>

⁵<https://multidict.net/wordlink/>

⁶<https://multidict.net/clilstore/>

⁷<http://lexikoarenbehatokia.euskaltzaindia.net/aurkezpena.htm>

⁸<https://stanfordnlp.github.io/stanfordnlp/>

Zipf value	fpmw	Examples
1	0.01	antifungal, bioengineering, farsighted, harelip, proofread
2	0.1	airstream, doorkeeper, neckwear, outsized, sunshade
3	1	beanstalk, cornerstone, dumpling, insatiable, perpetrator
4	10	dirt, fantasy, muffin, offensive, transition, widespread
5	100	basically, bedroom, drive, issues, period, spot, worse
6	1,000	day, great, other, should, something, work, years
7	10,000	and, for, have, I, on, the, this, that, you

Table 1: Examples based on the SUBTLEX-UK word frequencies. fpmw= frequency per million words.

on 70 languages using the Universal Dependencies framework (Nivre et al., 2016). We have limited the analysis to the following processes: tokenisation, sentence splitting, PoS tagging, NER and syntactic parsing.

As a semantic resource, we have used WordNet (Miller, 1995). Exactly, for the texts in English we have used the version included in the NLTK (Bird et al., 2009) and for the texts in Basque the version of the MCR (Gonzalez-Agirre et al., 2012).

In order to perform Word Sense Disambiguation (WSD), we have used UKB, which is available for several languages. UKB applies the so-called Personalized PageRank on a Lexical Knowledge Base (LKB) such as WordNet or Basque WordNet (Pociello et al., 2011) to rank the vertices of the LKB and, thus, disambiguate the words (Agirre et al., 2014).

3.3. Obtaining the images

To make the information in written texts as visual as possible, we have decided to provide images of the words. To obtain the images, we have used three resources: ImageNet (Deng et al., 2009), Wikidata and Wikipedia. ImageNet is a collection of images mapped to WordNet. There are images for 21,841 synsets. Wikidata is a free, collaborative and multilingual database which currently contains 475 images mapped to WordNet. And, finally, Wikipedia is a collaborative encyclopedia from which you can directly obtain images of the names or entities that appear in the raw text, but, in this case, in an ambiguous manner.

To display the images, we have determined that the application selects the image first by means of the WordNet sense identifier (the output of the WSD tool). If there is no image in ImageNet or Wikidata for that sense, the application selects image from Wikipedia with the wordform in the text (without disambiguation).

4. Functionalities

In this section we describe the functionalities provided by *LagunTest* (Figure 1) to assist teachers on adapting the content. Guided by usability principles, *LagunTest*:

- allows the user to insert text manually or import the text in a document by uploading files in any of the following formats: .txt, .doc, .docx and .odt.
- allows texts in two languages: English or Basque. The user can select the language from a dropdown list on the initial screen.

- computes and displays the results according to the selected level of vocabulary (beginner, intermediate or advanced) in the selected language.
- offers an overall visual representation of the content of the texts by organizing the most important content words of the text as a wordcloud.
- integrates visual and textual information to assist in grasping the meaning of words of less frequent words.

The available options are accessible and visible at the top on the initial screen. The user only needs to set the language and level from the options at the bottom of the initial screen and submit the text to get the results.

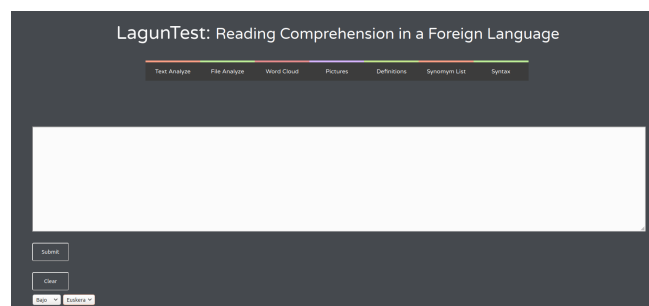


Figure 1: Main Page of LagunTest Web Tool

LagunTest is organized into seven tabs which are described in detail in the following subsections. The first two tabs (*Analyze Text* and *Analyze File*) are used to insert the text and the following five to access the visualizations and the rest of lexico-semantic and syntactic information available after submitting the text. To illustrate the functionalities, we will use as example throughout the following subsections the first paragraph of the article *Milk* in Simple Wikipedia,⁹ which is also presented in Appendix 1, targeted at a learner with basic level of English.

4.1. Analyze Text and Analyze File tabs: Inserting text

The *LagunTest* application allows to insert the text in two ways: 1) in the *Analyze Text* tab, the user can enter the text by typing it into a text box; 2) in the *Analyze File* tab the user can upload a document in one of the following formats:

⁹https://simple.wikipedia.org/wiki/Main_Page

.txt, .doc, .docx and .odt. This avoids the need of converting the text to a specific format.

Once the text has been inserted, the user needs to select the language of the text (Basque or English). Additionally, the user needs to choose a level (basic, intermediate or advanced). Based on the language level, the application adapts the output as presented in Section 3. Finally, the user must click on the submit button, and automatically, the user can check the available information using one of the following tabs: *WordCloud*, *Pictures*, *Definitions*, *Synonym List* and *Syntax*.

4.2. WordCloud tab: overall representation of the text as a wordcloud

Wordclouds are visualization tools that highlight the relative frequency of words in a text. Wordclouds are very useful to quickly identify the more frequent words in the text, since they will appear bigger and bolder. This is a way to pull out the most pertinent parts of textual data. If a student can easily identify the words that appear most frequently in the text, he probably knows effortlessly which are the most pertinent parts, and therefore, the ones he should focus on because they are critical to understand the ideas. Additionally, this visual representation offers a different perspective on the text. Some students will probably have a more visual learning style, and may benefit from observing this global representation. Teachers can use the resulting wordcloud to stimulate reflection on the contents, and open a dialogue with the students intended to relate the words and link the concepts represented by the words.

In Figure 2 we show the wordcloud created by *LagunTest* for the example, where *milk*, *mammals* and *babies* appear in a bigger font.



Figure 2: Wordcloud of the example sentence

4.3. Pictures tab: Images of some of the most difficult words from the text

The *Pictures* tab shows the images of the words in the text. In educative contexts, images can be useful to evoke the words. In the current example, the application has returned 9 Wikipedia images for the words *liquid*, *glands*, *breasts*, *udders*, *teats*, *babies*, *teeth*, *calcium* and *bones* and 6 ImageNet image for the words *Milk*, *mammal*, *cow*, *dog*, *humans* and *food*.

In Figure 3 we show the image returned for *udders*, where that part of the body of the cow is shown. The rest of the returned images can be accessed through the urls of Table 2. We also show where they come from.



Figure 3: Image returned for the word *udders*

4.4. Definitions tab: definitions and examples

The *Definitions* tab shows the definitions of the nouns and verbs if they are in the respective wordnets. Definitions and examples can be useful to help understanding the concepts in the texts.

In the case of the beginner level, the application offers the definitions and examples for the following words: *milk*, *liquid*, *made*, *mammals*, *cows*, *dogs*, *humans*, *glands*, *breasts*, *udders*, *teats*, *babies*, *have*, *teeth*, *given*, *eat*, *food*, *nutrients*, *help*, *grow*, *source*, *calcium* and *bones*. As an example, we show the definition “any warm-blooded vertebrate having the skin more or less covered with hair” and the example “young are born alive except for the small subclass of monotremes and nourished with milk” for the word *mammals*.

In addition, both in the *Definitions* and *Synonyms* tabs, the words of the text entered are colored according to their PoS: nouns in light green, verbs in red, adjectives in dark blue, proper names in pink, pronouns in dark green, particles in yellow and finally adverbs in orange.

4.5. Synonym List tab: Synonyms

The *Synonym List* tab shows the synonyms of the nouns and verbs that appear in the respective wordnets for the beginner level. Synonyms can be suitable when introducing new vocabulary.

Below, we show the synonyms that were obtained from the example for some of the words (for brevity, we do not include the whole list).

- **made:** ‘get’
- **cows:** ‘moo-cow’
- **humans:** ‘homo’, ‘man’, ‘human_being’
- **breasts:** ‘titty’, ‘boob’, ‘knocker’, ‘tit’, ‘bosom’
- **babies:** ‘infant’, ‘babe’
- **have:** ‘hold’, ‘have_got’
- **food:** ‘nutrient’
- **help:** ‘aid’, ‘assist’
- **source:** ‘origin’, ‘beginning’, ‘rootage’, ‘root’
- **calcium:** ‘ca’, ‘atomic_number_20’

Word	Source	Url
milk	ImageNet	http://farm4.static.flickr.com/3618/3479565830_54ec716835.jpg
liquid	Wikipedia	https://upload.wikimedia.org/wikipedia/commons/4/43/2006-01-14_Surface_waves.jpg
mammal	ImageNet	http://static.flickr.com/1250/1487336553_f99ee15b0a.jpg
cow	ImageNet	http://farm3.static.flickr.com/2145/2150559343_8de1c310e3.jpg
dog	ImageNet	http://farm3.static.flickr.com/2325/1891011832_cb5d5098c2.jpg
humans	ImageNet	http://farm3.static.flickr.com/2358/1797858275_bbff0e1b73.jpg
glands	Wikipedia	https://upload.wikimedia.org/wikipedia/commons/3/37/405_Modes_of_Secretion_by_Glands_updated.svg
breasts	Wikipedia	https://upload.wikimedia.org/wikipedia/commons/b/bd/Bare_breasts_are_our_weapons_crop.jpg
udders	Wikipedia	https://upload.wikimedia.org/wikipedia/commons/6/61/Cow_udders02.jpg
teats	Wikipedia	https://upload.wikimedia.org/wikipedia/commons/9/91/Bundesarchiv_Bild_183-17369-0004%2C_Barby%2C_Bauer_eine_Melkmaschine_pr%C3%A4sentierend.jpg
babies	Wikipedia	https://upload.wikimedia.org/wikipedia/commons/8/84/Baby_%28126372492%29.jpg
teeth	Wikipedia	https://upload.wikimedia.org/wikipedia/commons/c/cb/Close_up_-_chimpanzee_teeth.png
food	ImageNet	http://farm4.static.flickr.com/3079/2852414223_0d0fa765a0.jpg
calcium	Wikipedia	https://upload.wikimedia.org/wikipedia/commons/5/51/Ca%28aq%296_improved_image.tif
bones	Wikipedia	https://upload.wikimedia.org/wikipedia/commons/2/23/603_Anatomy_of_Long_Bone.jpg

Table 2: Url of the images returned by the application

4.6. Syntax tab: dependency tree

Finally, in the *Syntax* tab, the application shows the syntactic dependencies of sentence, based on the framework of Universal Dependencies. This can be practical when teachers aim at practicing or focusing on a particular syntactic structure.

In Figure 4 we can see the generated syntax tree of the fourth sentence: “Milk has many nutrients to help babies grow and be healthy”.

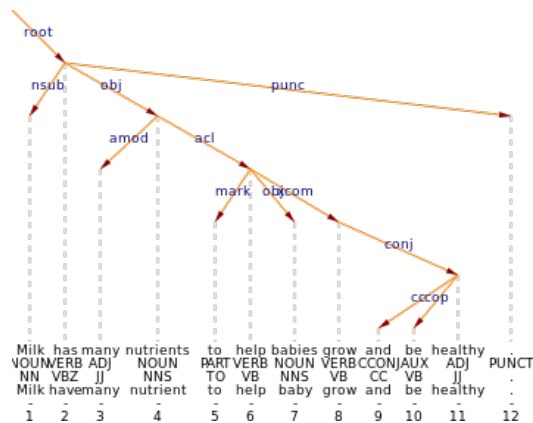


Figure 4: The syntax tree obtained using the StanfordNLP parser

5. Discussion

According to Stanovich (1986), children who read slowly and do not enjoy it, develop the vocabulary knowledge slower, while children who read well learn more words and, therefore read better. This is known as the Matthew effects in reading, which was first introduced in education by Walberg and Tsai (1983).

That is why we claim that using this tool to assist in acquiring vocabulary, particularly at early stages, can promote comprehension and can be used as a facilitator of further

reading. However, since the process used by *LagunTest* is fully automatic, there are some limitations to the results.

Regarding the recall, *LagunTest* is limited by the words that are covered in the referenced resources. For example, in the case of the Basque version, not all the nouns have definitions in Basque WordNet. Moreover, regarding the word frequencies, the tools and corpora are general purpose tools containing above all journalist texts for adults and words that can be usual for children such as *milk* can be detected as low frequency words.

Regarding synonyms, we may encounter register issues. For example, for the word *calcium* the application shows synonyms *ca* and *'atomic_number_20*. As a general resource has been used, these words may be counterproductive for schoolchildren and they may not be very suitable for users with a low level of English. In order to overcome this problem, a possibility is to get frequencies by crawling a corpus from websites for children, or to use the Oxford Children’s Corpus (Wild et al., 2012) and filter the results. Moreover, in relation, to the definitions, some of them can be difficult for children.

Regarding the images obtained by the tool, in the image returned for *breast*, a woman is protesting by showing her bare breast. This image may not be the most adequate for children. Besides, in the case of the images obtained from Wikipedia, the tool may offer incorrect images since it has not been disambiguated. On the other hand, web images can also have biases (Crawford and Paglen, 2019). Obtaining suitable images is an open problem in machine learning and Artificial Intelligence.

Summing up, we certainly recommend that the teachers supervise the output of the tool and validate the results for each particular case before using it in the classroom. They know best their students and should decide, for instance, if it is better to work with definitions or with synonyms depending on the task. *LagunTest* is a resource that may offer valuable information to the professionals, but it is worth noting that this type of automatic assistance can never take

the role of a professional educator.

6. Conclusion and future work

In this article we have presented the application *LagunTest*, its design criteria and possible shortcomings. *LagunTest* aims at enhancing reading comprehension by showing images, definitions and synonyms according to the level of knowledge. Through this application we attempt to assist education professionals in adapting materials.

At the moment, *LagunTest* is available to work on vocabulary of English and Basque, but we are working to adapt it to other languages such as Spanish, Galician and Catalan. We also plan to improve its output by adapting it as possible to children by including computational resources for children. Furthermore, the evaluation of the usability of the application and the appropriateness of the contents by education professionals is our main future goals.

7. Acknowledgments

We acknowledge following projects: DL4NLP (KK-2019/00045), DeepReading RTI2018-096846-B-C21 (MCIU/AEI/FEDER, UE), Ixa Group: consolidated group type A by the Basque Government (IT1343-19) and BigKnowledge for Text Mining, BBVA.

8. Bibliographical References

- Agirre, E., López de Lacalle, O., and Soroa, A. (2014). Random walks for knowledge-based word sense disambiguation. *Computational Linguistics*, 40(1):57–84.
- Agirrezabal, M., Altuna, B., Gil-Vallejo, L., Goikoetxea, J., and Gonzalez-Dios, I. (2019). Creating vocabulary exercises through NLP. *Proceedings of the Digital Humanities in the Nordic Countries 4th Conference, CEUR-WS*, pages 18–32.
- Aldabe, I., Lopez De Lacalle, M., Maritxalar, M., Martinez, E., and Uria, L. (2006). Arikiturri: an automatic question generator based on corpora and nlp techniques. In *International Conference on Intelligent Tutoring Systems*, pages 584–594. Springer.
- Aldabe, I., Gonzalez-Dios, I., Lopez-Gazpio, I., Madrazo, I., and Maritxalar, M. (2013). Two approaches to generate questions in basque. *Procesamiento del lenguaje natural*, 51:101–108.
- Bengoetxea, K., Gonzalez-Dios, I., and Aguirregoitia, A. (2020). AzterTest: Open Source Linguistic and Stylistic Analysis Tool. *Procesamiento del Lenguaje Natural*, 64:61–68.
- Bird, S., Klein, E., and Loper, E. (2009). *Natural language processing with Python: analyzing text with the natural language toolkit*. ” O’Reilly Media, Inc.”.
- Castro-Castro, D., Lannes-Losada, R., Maritxalar, M., Niebla, I., Pérez-Marqués, C., Álamo-Suárez, N., and Pons-Porrata, A. (2008). A multilingual application for automated essay scoring. In *Ibero-American Conference on Artificial Intelligence*, pages 243–251. Springer.
- Crawford, K. and Paglen, T. (2019). Excavating AI: The Politics of Images in Machine Learning Training Sets. <https://www.excavating.ai/>.
- Crossley, S. A., McNamara, D. S., et al. (2017). *Adaptive educational technologies for literacy instruction*. Routledge.
- Deng, J., Dong, W., Socher, R., Li, L., Kai Li, and Li Fei-Fei. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, June.
- Gasparini, A. A. and Culén, A. L. (2012). Tablet pcs— an assistive technology for students with reading difficulties. In *ACHI 2012: The Fifth International Conference on Advances in Computer-Human Interactions*, pages 28–34.
- Gonzalez-Agirre, A., Laparra, E., and Rigau, G. (2012). Multilingual central repository version 3.0. In *LREC*, pages 2525–2529.
- Gonzalez-Dios, I., Aranzabe, M. J., Díaz de Ilarraza, A., and Salaberri, H. (2014). Simple or complex? assessing the readability of basque texts. In *Proceedings of COLING 2014, the 25th international conference on computational linguistics: Technical papers*, pages 334–344.
- Gonzalez-Dios, I. (2016). *Euskarazko egitura sintaktiko konplexuen analisisirako eta testuen sinplifikazio automatikorako proposamena / Readability Assessment and Automatic Text Simplification. The Analysis of Basque Complex Structures*. Ph.D. thesis, University of the Basque Country (UPV/EHU).
- Haßler, B., Major, L., and Hennessy, S. (2016). Tablet use in schools: A critical review of the evidence for learning outcomes. *Journal of Computer Assisted Learning*, 32(2):139–156.
- ISEI-IVEI. (2016a). *Rs1.2.1 Curriculum kompetentzia LHko 4. mailan: Euskara*. Instituto Vasco de Evaluación e Investigación Educativa.
- ISEI-IVEI. (2016b). *Rs1.2.2 Curriculum kompetentzia LHko 4. mailan: Gaztelania*. Instituto Vasco de Evaluación e Investigación Educativa.
- ISEI-IVEI. (2016c). *Rs1.2.4. Nivel de competencia curricular en 4 curso de Educación Primaria: Inglés*. Instituto Vasco de Evaluación e Investigación Educativa.
- Miller, G. A. (1995). Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41.
- National Reading Panel (US). (2000). Teaching children to read: Reports of the subgroups. The Panel.
- Nivre, J., De Marneffe, M.-C., Ginter, F., Goldberg, Y., Hajic, J., Manning, C. D., McDonald, R., Petrov, S., Pyysalo, S., Silveira, N., et al. (2016). Universal dependencies v1: A multilingual treebank collection. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC’16)*, pages 1659–1666.
- Pociello, E., Agirre, E., and Aldezabal, I. (2011). Methodology and Construction of the Basque WordNet. *Language resources and evaluation*, 45(2):121–142.
- Qi, P., Dozat, T., Zhang, Y., and Manning, C. D. (2019). Universal dependency parsing from scratch. *arXiv preprint arXiv:1901.10457*.
- RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: Rand.

- Speer, R., Chin, J., Lin, A., Jewett, S., and Nathan, L. (2018). Luminosinsight/wordfreq: v2.2, October.
- Stanovich, K. E. (1986). Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy. *Reading Research Quarterly*, 21(4):360–407.
- van Heuven, W., Mandera, P., Keuleers, E., and Brysbaert, M. (2014). SUBTLEX-UK: a new and improved word frequency database for British English. *Quarterly journal of experimental psychology*, 67:1176–1190.
- Walberg, H. J. and Tsai, S.-L. (1983). Matthew Effects in Education. *American Educational Research Journal*, 20(3):359–373.
- Wild, K., Kilgarriff, A., and Tugwell, D. (2012). The Oxford Children’s Corpus: Using a Children’s Corpus in Lexicography. *International Journal of Lexicography*, 26(2):190–218, 09.
- Daniel Zeman et al., editors. (2018). *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. ACL, Brussels, Belgium, October.

Appendix A: Example text

Milk is a white liquid made by mammals, like cows, dogs, and humans. It is made in the mammary glands (breasts, udders, or teats) of female mammals. Because newborn babies have no teeth, they must be given milk before they can eat solid food. Milk has many nutrients to help babies grow and be healthy. It is also a rich source of calcium which is good for your bones and teeth.