*Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*, pages 1142–1151
Marseille, 11–16 May 2020


# Simplifying Coreference Chains for Dyslexic Children

**Rodrigo Wilkens, Amalia Todirascu**
LiLPa – University of Strasbourg
rswilkens@gmail.com, todiras@unistra.fr

**Abstract**

This work aims to generate adapted content for dyslexic children for French, in the context of the ALECTOR project. Thus, we developed a system to adapt texts at the discourse level. This system modifies coreference chains, which are markers of text cohesion, by using rules. These rules were designed following a careful study of coreference chains in both original and simplified text versions. Moreover, aiming to evaluate the reliability of the proposed transformation rules, we analysed several coreference properties as well as the concurrent simplification operations in the aligned texts. This information is coded together with a coreference resolution system and a text rewritten tool. The proposed system comprises a coreference module specialised in written text and seven text transformation operations. Firstly, we evaluate the text simplification results by manual validation of three judges. The identified errors were grouped into five classes that combined can explain 93% of the errors. The second evaluation step consisted of measuring the simplification perception by 23 judges, which allow us to measure the simplification impact of the proposed rules.

**Keywords:** Automatic Text Simplification (ATS), Dyslexic Children, Coreference Resolution, French

## 1. Introduction

Automatic text simplification (ATS) aims to adapt content to a specific audience. Thus, it may be used to support learning activities (foreign or native language acquisition) and to generate appropriate content for people with language disabilities. Several ATS projects target specific audiences (e.g. autism (Yaneva and Evans, 2015) and dyslexia (Rello et al., 2013)), and they have been applied to several languages, such as English (Barbu et al., 2013; Woodsend and Lapata, 2011), Spanish (Saggion et al., 2011), Portuguese (Aluísio et al., 2008; Wilkens et al., 2014), Japanese (Inui et al., 2003), Dutch (Ruiter et al., 2012; Bulté et al., 2018), Italian (Dell'Orletta et al., 2011; Barlacchi and Tonelli, 2013), and French (Seretan, 2012; Brouwers et al., 2014). It might address a specific language level (e.g. lexical or syntactic simplification), or it might be applied at different levels at the same time. The lexical simplification proposes the replacement of complex terms by simpler ones (Rello et al., 2013; François et al., 2016) (Billami et al., 2018), while the syntactic simplification intends to transforms complex syntactic structures into simpler ones (Seretan, 2012; Brouwers et al., 2014). However, these transformations might break the discourse structure by violating cohesion or coherence constraints.

Coherence and cohesion are text properties which help the text understanding (Hobbs, 1979; Schnedecker, 1997b), (Charolles, 2006). In addition, several cohesion devices, such as lexical chains (Hirst and St-Onge, 1995) or coreference chains (Schnedecker, 1997b), are also strongly related to text readability and complexity (Pitler and Nenkova, 2008). Nevertheless, few existing systems (e.g. Siddharthan (2006; Canning (2002)) operate at this level taking into account these discourse constraints after the syntactic simplification process, or replacing anaphoric pronouns by their antecedents (Quiniou and Daille, 2018).

In this line, we study cohesion constraints related to coreference chains and its applicability to discourse simplification for French. Therefore, in this paper, we evaluate the feasibility of a system for automatic simplification acting at the discourse level, aiming to provide adapted content to dyslexic children. This system attempts to preserve the text cohesion based on a detailed manual analysis of French standard and dyslexic simplified corpus. Our work is carried out in the context of the research project ALECTOR, whose goal is to build a French ATS system, aiming to provide fully adapted text content for dyslexic children. It proposes an end-to-end system that addresses the lexical[1], syntactic and discourse levels. In this paper, we focus specifically on the discourse level module, proposing new features, such as the discourse-level transformation, changing the structure of the coreference chains.

This paper is organized as follows. We present the context and motivation (Section 2.) and the architecture of the discourse simplification module (Section 3.), by detailing the automatic coreference detection and text rewrite modules. Later, we present the corpus analysis results (Section 4.), studying coreference chain properties and discourse transformations, that supports the proposed cohesion rules in Section 4.4.. Then, we present the evaluation of the automatic simplification (Section 5.). Finally, the Section 6. presents conclusions and final remarks.

## 2. Text cohesion and ATS

Text cohesion, a crucial feature for text understanding, is reinforced by explicit cohesive devices, such as lexical or coreference chains. Thus, lexical chains contain noun phrases and their synonyms or hyponyms/hypernyms, or noun phrases from the same domain or topic: *music–piano–concert*). Coreference chains contain various expressions referring to the same discourse entity: *Emmanuel Macron – the French president – his* discourse), and anaphoric (an anaphora and its antecedent: *the village – it*) chains. However, coreference and anaphora resolution may pose difficulty to dyslexic people (Vender, 2017; Jaffe et al., 2018). Moreover, when concurrent referents are present in the text, the task of pronoun resolution is harder (Givón, 1993; McMillan et al., 2012; Li et al., 2018), considering that pro-

[1]The lexical simplification proposed for ALECTOR project can be seen in Billami et al. (2018)

nouns may be ambiguous and their resolution depends also on the main topic (Le Bouëdec and Martins, 1998).

The simplification systems frequently ignore the existing cohesive devices, such as anaphora or coreference chains. This aspect is taken into account, for instance, in Siddharthan (2006) and Brouwers et al. (2014). Moreover, Štajner et al. (2012), Todirascu et al. (2016) and Pitler and Nenkova (2008) propose discourse-related features (e.g. entities densities and syntactic transitions) to evaluate text readability in addition to other lexical or morphosyntactic properties. Some existing simplification systems propose rules aiming to preserve discourse cohesion, such as replacing anaphor by their antecedent (Canning, 2002; Quiniou and Daille, 2018), using coreference detection systems to solve coreference issues (Barbu et al., 2013) or using discourse markers to maintain cohesion constraints (Siddharthan, 2006). Siddharthan (2006) proposes a model based on centering theory (Grosz et al., 1995) to recover broken cohesion links, by using a specific pronoun resolution system for English. The model allows the replacement of pronoun by its immediate antecedent. Summing up, only a few ATS approaches propose discourse simplification rules or rules checking discourse constraints, mainly for English.

In this line, we propose a set of discourse simplification rules taking into account more complex rules and contexts, and we build a system to produce simplified French texts addressing dyslexic children. This system is original in the sense of the coreference accessibility literature as well as corpora observations.

## 3. The Discourse Simplification Architecture

Aiming to transform text at discourse level maintaining text cohesion markers, such as coreference chains, we develop a discourse simplification system. First, the system automatically detects the coreference chains in the original texts and then applies cohesion rules. Finally, a text rewritten tool applies all text transformations. Attempting to assess the feasibility of the automatic simplification system, we lack a coreference annotation system and a rewritten tool. Fortunately, we can train the English state of the art system in a French corpus that present similar text types with those used in our study to build an automatic coreference annotator (presented in Section 3.1.). Concerning the rewritten tool, we compared available tools, identifying that none of them performs all required operations. Thus, we propose our own set of text transformation operations used in this paper, and we describe then in Section 3.2.

The discourse transformation rules and the evaluation of its applicability in an automatic system require detailed linguistic information (e.g. PoS, syntactic dependencies and lemma) to be represented as discourse cohesion rules. Thus, we analysed coreference chain properties in a corpus of tales and tales adapted to dyslexic children (Section 4.1.). This analysis focuses on the cohesive elements (Section 4.2.) and other syntactic operations (Section 4.3.) also related to the discourse. On the basis on these analyses, we define the discourse simplification rules proposed in this work (Section 4.4.). In the remain of this section, we present the automatic coreference annotation module and the text rewriting tool.

### 3.1. Automatic Coreference Annotation

Given that the French state of the art in automatic coreference annotation is largely based on the oral register (e.g. (Désoyer et al., 2016; Grobol, 2019)), we trained a new coreference model based on Democrat corpus (Landragin, 2016).[2] This is a large corpus with 689k words and texts written in different centuries. We select only texts written in the 19th century or later[3], which means 10k word documents of newspaper, encyclopedic articles, biographies, excerpts from novels, short stories, and treatises.

In order to obtain a model trained on DEMOCRAT corpus, we use a state-of-the-art end-to-end system for English (Kantor and Globerson, 2019), which can be adapted to French. Moreover, following Grobol (2019), we also split it into two independently trained modules: mention identification and coreference resolution. Another notable modification in the original system is the inclusion of singleton processing (i.e. mention that are not contained by any coreference chain) in the system's output. This modification required us to change the default $\lambda$ value to 0.27, $\lambda$ being the expected proportion of mentions in the corpus. This is a critical step which allows us to process rules that address all mentions no matter the chain size. Note that, due to computational restrictions, we divided all corpus documents into parts of no more than 2k words. The evaluation of the coreference errors are presented in Section 5.1., and a detailed description of the coreference model can be found in (Wilkens et al., 2020).

### 3.2. Text rewrite

Text rewrite applies several text transformation and changes the structure of the sentences: subordinate suppression, sentence split and phrase changing. Since our architecture requires a rewritten tool that allows change the sentence structure without violating the grammar, we compare the the following available tools: Tregex and Tsurgeon (Levy and Andrew, 2006), Semgrex (Chambers et al., 2007), and Semgrex-Plus (Tamburini, 2017).

Levy and Andrew (2006) provide tree query (Tregex) and manipulation (Tsurgeon) tools that can operate on constituent tree data structures. Complementary to Tregex queries, Tsurgeon operates at node and relation level, to change the structure of the trees, allowing the following operations: node rename, delete, insert, move, replace, prune, excise, adjoin, and coindex.

Aiming to speed up the process of graph search and to reduce errors while moving from Tregex constituent formalism to a dependency one, Chambers et al. (2007) proposed Semgrex. For instance, Semgrex allows the queries to be used to identify the direct or indirect governor association, limiting or not the distance between the elements, or even the word positional relation (e.g. immediately precedes, right sibling, right immediate sibling, same nodes).[4]

[2] https://github.com/boberle/cofr

[3] The French 19th-century literature is similar to current one.

[4] These tools are widely accepted in the NLP community, as pointed to their number of citations (more than 334 papers citing Tregex and Tsurgeon, while Semgrex is cited in more than 69.)

Tamburini (2017) developed Semgrex-Plus to convert dependency treebanks into different formats. This tool incorporates Semgrex query language, adding rewrite patterns, which may be applied at both token tags and dependency levels. Semgrex-Plus supports three rewriting operations: replace the tag of a graph node, and inserts or deletes a dependency edge between two graph nodes.
For parsing querying, we select Semgrex, while it fits our needs. But, regarding the sentence rewriting goal, we opt to create a new Semgrex-based sentence processing tool[5], given the parser restrictions and the small set of operations available on Semgrex-Plus. Concerning the operations, we developed the following:

1. *Insert* injects a node (or tree) in another node;
2. *Delete* removes a node and its subtree from the sentence graph;
3. *Split* detaches a node and its subtree;
4. *Move* detaches a node and its subtree from a tree node, attaching it in another node of the same tree;
5. *Replace tag label* replaces the node information (e.g. surface and PoS-tag);
6. *Replace node* substitutes a node by another one; and
7. *Copy subgraph* creates a deep copy of a node or a tree.

The *insert*, *delete*, *move*, and *replace node* methods are directly based on Tsurgeon while the *replace label* is both based on Tsurgeon and Tsurgeon-plus. The *split* operation is also inspired by Tsurgeon *excise* and *adjoin* operations. The *copy* operation was developed due to the need to copy a part of a sentence into different trees.
The development of these operations required us to make choices about the words' positions after an operation. One possible solution is to code the language grammar, in our case French grammar. However, it would demand a lot of effort while resulting in a language-dependent tool. Thus, inspired by Tsurgeon, we opt to code the final position in the operation using the following elements: direction (left or right), reference node (current node or its parent), and dependency scope (current node or its subtree). The direction element allows horizontal control of the operation, while the reference node allows vertical control. The dependency scope element allows controlling the operation distance. For example[6], when targeting the head-word, the insertion of the adjective $beautiful$ in the phrase $these_1 \underset{\text{det}}{\longrightarrow} mangoes_2$ results in $these_1 \underset{\text{det}}{\longrightarrow} mangoes_3 \underset{\text{amod}}{\longleftarrow} beautiful_2$, while the insertion of the adverb $all$, targeting dependency tree for the same phrase, results in $these_2 \underset{\text{det}}{\longrightarrow} mangoes_3 \underset{\text{advmod}}{\longleftarrow} all_1$.
Another issue related to the copy operation is the Semgrex reference. When copying a node or subgraph, it creates a version independent of the original element. While this process allows changing the properties of the new element, this new element will not be associated with any Semgrex query. Due to that, the operations may not be sequentially applied in some cases.
These operations are combined as rules in order to rewrite the text. However, this process is driven by the cohesion changes required to simplify the discourse. In this sense, we detail the process of defining the cohesion rules necessary for our simplification system in the next section.

## 4. Corpus Analysis and Cohesion Rules

The first step in our methodology consists of examining text addressed to dyslexic children and to compare them with texts addressed to regular reader. Thus we compile a parallel corpus (original and manually simplified text for dyslexic children). This corpus consists of five paired tales (1,143 words for the dyslexic texts and 1,969 words for the original texts) adapted by an association helping dyslexic children `https://methodolodys.ch/`. In the next subsection, we define the coreference chain properties, used to compare the original and simplified texts. We analyse the changes in the coreference chains due to text simplifications in order to define the cohesion rules.

### 4.1. Coreference chain properties

We manually annotate the corpus with coreference chains, using SACR (Oberle, 2018). Next, we compare the coreference chains in both simplified and original texts to detect discourse simplification rules. For this comparison, we compute the following properties of the coreference chains, proposed by Todirascu et al. (2017) who, inspired by Schnedecker (1997a), studied properties in different text genres. These properties influence text readability (Todirascu et al., 2016) and the usage preference by adults and children (Todirascu et al., 2017):

- The length of the chain and the average distance between the mentions (L2L);
- The distribution of various types of referring expressions included in coreference chains;
- The lexical stability ratio (Perret, 2000), helping to evaluate the lexical variety of referring expressions included in a chain[7]; and
- The density of mentions (between two consecutive referring expressions in the same chain) and of the annotations.

We compare the chain's properties in the original and simplified texts. Additionally, the referring expressions might be sorted according to its accessibility (Ariel, 1990), from low accessible to high accessible (e.g. personal and reflexive pronouns). Thus, the first mention of a discourse entity might be a low accessible one (e.g. proper noun, indefinite noun phrases with a reference function), while

[5] `https://github.com/rswilkens/text-rewrite`

[6] In the example, the arrows indicate the dependency relation and the subscripted numbers indicate the word index.

[7] The stability coefficient is computed as the number of referring expressions from one chain divided by the number of the different referring expressions included in that chain. A high coefficient means the variety of the referring expressions is higher.

the other mentions should be highly accessible expressions (e.g. demonstrative noun phrases, personal pronouns, or relative pronouns). These properties will be exploited to define discourse simplification rules. Complementary, we also manually identify the transformations operated on align text pairs using the MEDITE tool (Fenoglio and Ganascia, 2006). We use this alignment to distinguish different transformation types. Following, we compare the influence of the simplification operations on the chain's structure by performing quantitative and qualitative analysis. Then, the corpus analysis is compiled into simplification rules. The design of these rules aims to change the text in a way that the final version presents similar coreference chain properties. It includes the properties listed above as well as the accessibility level (Ariel, 2001; Ariel, 1990) and the centering theory (Grosz et al., 1995).

## 4.2. Comparison of Coreference Chains

We start our study of the cohesive elements by comparing the properties and transformations of five text pairs. Each of those was manually annotated with coreference chains and anaphoric links. Due to the lack of available data containing original and simplified texts for dyslexic people, our corpus is relatively small when comparing with others simplification corpus, such as Paetzold and Specia (2016) that contains 929 sentences.[8] Moreover, manual coreference annotation is a time-consuming and challenging task, in terms of referring expression identification (delimiting the expression and finding its type) and of chain identification (linking the referring expression with a given chain).

The corpus annotation targets the referring expressions composing the reference chains: named entities (person, organisation, places), indefinite and definite noun phrases, demonstrative noun phrases, personal pronouns, relative pronouns, reflexive pronouns, and other referring expressions, such as possessive determiners.[9] Besides the number of the coreference or anaphoric chains, each referring expression has been annotated with its syntactic function and the expression's type.

Regarding the coreference properties, as can be seen in Table 1, there is a reduction in the adapted texts when compared to the originals. Furthermore, we observe a significant difference in the text pairs in the following properties: link count (p=0.01), stability coefficient (p=0.01), chain density (p=0.04), link density (p=0.008) and annotation density (p=0.02). In addition, we also identify correlations between the different versions for most of the properties (0.74 for link count, 0.81 for stability coefficient, 0.72 for chain density, and 0.74 for link density). These results indicate that there exists a difference between original and adapted texts in coreference level, but despite this, the relationship of the properties are kept. We also observe a correlation (-0.717) between the length of the chains and the number of chains. Longer chains are correlated with less chains (in average 10.618 against 7.0) in dyslexic texts.

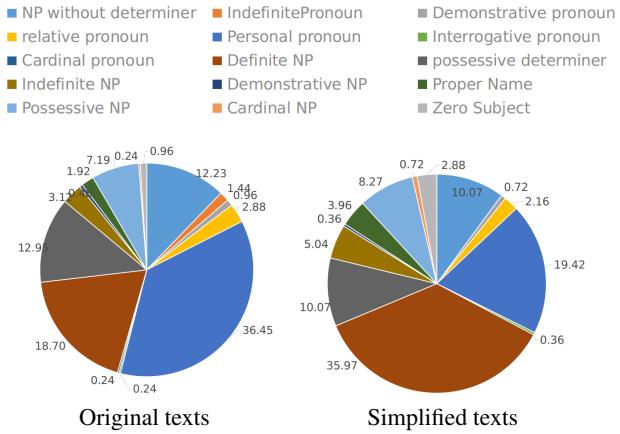
[8]Note that different simplification approaches, such as discourse simplification, requires more text in order to find simplification examples and their contexts.

[9]The annotation guide is presented by Todirascu et al. (2017).

The original texts contain several referents, while some referents disappear in simpler texts (secondary characters or objects) which explains the variation in the number of chains. Additionally, the average distance between two consecutive links is higher in original rather dyslexic texts, as a consequence of text deletions.

| Properties | Adapted | Original |
|---|---|---|
| Avg chain size | 10.376 | 10.86 |
| Avg L2L distance | 14.550 | 11.920 |
| Avg Link length | 1.500 | 1.450 |
| Avg chain count | 6.200 | 7.800 |
| Avg link count | 55.600 | 83.4 |
| Avg chain density | 0.012 | 0.009 |
| Avg coeff stab | 0.607 | 0.471 |
| Avg link density | 0.113 | 0.093 |
| Avg annotation density | 0.162 | 0.139 |

Table 1: Coreference chains properties




Original texts Simplified texts

Figure 1: The distribution of types of referring expressions in the chains for original and simplified texts

The composition of the chains varies with the complexity of the texts as shown in Figure 1. The percentage of personal pronouns included in coreference chains is larger in the original text (36.5%), rather than in simplified texts (19.4%). The simplified texts present significantly more definite nouns and possessive determiners (respectively, 36.0% in the simplified texts, and 18.7% in the original texts). Indeed, the pronouns have been deleted or replaced by their referent. In original texts, possessive determiners (12.9% vs 10.1 % in simple texts) or relative pronouns are frequent, while the percent of proper names is more important than in simpler texts. This observation is in line with studies in cognitive psychology and neuroscience, such as Li et al. (2018), Jaffe et al. (2018), McMillan et al. (2012) and Heine et al. (2006), showing that the pronouns require more reasoning to find its antecedent.

If the definite noun phrases is frequent, the reference chains in original texts contains synonyms or hypernyms, while in simpler texts, it contains several repetitions of the same referent. This is measured with the help of the stability coefficient (Perret, 2000), and a high coefficient indicates a variety of referring expressions. We observed more sta-

ble chains in dyslexic corpora (0.47) than original corpora (0.60). In other words, the dyslexic texts are characterized by less variation in the forms used in the reference chains than in the original texts.

The coreference chains properties comparison in the original and simplified texts represents the first step to define the cohesion rule. The next step is the identification of changes in the structure of the coreference chains induced by simplifications, before defining the cohesion rules.

### 4.3. Simplifications and Coreference Chains

We use the alignment tool MEDITE (Fenoglio and Ganascia, 2006) to detect the changes between the original and the simplified texts, studying insertions, deletions, and modifications of character strings. Additionally, this analysis allows us to focus on specific parts of the text. So, we add two extra pairs of texts to our corpus. One of the main challenge in comparing the both text versions is that several discourse entities disappear in the simplified texts due to lexical or syntactic simplifications. The text simplification is a complex task and most of the transformations combines several levels (i.e., lexical, syntactic or discourse).

We manually align 98 pairs of sentences with similar information and we identify the changes modifying the structure of the coreference chains. The sentences were selected to contain referring expressions included into coreference chains. Then, we compare the changes operated by simplification: adding extra information and deleting secondary information. Additionally, we measure the percentage of each simplification type: lexical (18.94%), syntactic (46.32%), and discourse (34.74%).

The discourse transformations involve pronoun suppression, determiner changing, and replacing the pronouns with some antecedents. In the following example, the relative clause *qui se dirigent vers eux* is deleted. As a consequence, the relative pronoun *qui/who* and the personal pronoun *eux/them* disappear:

*Original version:* En chemin, ils aperçoivent, au loin, des bandits qui se dirigent vers eux. *In English:* In their way, they saw, far away, the bandits who went to them.

*Simplified version:* En chemin, ils aperçoivent au loin des bandits. *In English:* In their way, they saw, faw away, the bandits.

In the next example, several syntactic transformations are applied. First, the complex NP *une partie du lait de la seule vache qu'ils possèdent/ a part of the milk of the only cow they own* is deleted. In this case, the personal pronoun *ils/they* disappears. Then the NP *le reste/the rest of the milk* is replaced by a new NP containing some information from the suppressed NP (*le lait de leur unique vache/the milk of their unique cow*). In this transformation the relative clause is replaced by the possessive determiner *leur/their* (*la seule vache qu'ils possèdent* becomes *leur unique vache*). The NP *au marché/at the market* is replaced by an explanation (*pour survivre/to survive*). Finally, the pronominal subject of the main clause *ils* is replaced by the antecedent *Jack et sa maman*.

*Original version:* $\text{Ils}_1$ boivent une partie du lait de la seule vache qu'ils possèdent et vendent le reste au marché. *In English:* $\text{They}_1$ drink a part of the milk produced of the only cow they have and sell the rest of milk at the market. *Simplified version:* Jack et sa $\text{maman}_1$ boivent et vendent le lait de leur unique vache pour survivre. *In English:* Jack and his $\text{mother}_1$ drink and sell the milk of their unique cow to survive.

### 4.4. Cohesion Rules

We compiled the patterns observed in the corpus analysis into transformation rules applied at discourse level. These rules change the properties of coreference chains. For our purpose, we use accessibility theory (Ariel, 1990), which proposes a hierarchy of referring expressions from those with low accessibility (newly introduced expressions, such as proper nouns or definite noun phrases) to high accessible expressions (personal pronouns, demonstrative pronouns or determiners). Moreover, we use the centering theory (Grosz et al., 1995). This theory predicts the situations when the attention centre shifts to a new one, which results in a change of the syntactic function of the discourse center. We exploit these observations when proposing the following rules for discourse simplification (four addressing substitution and one for deletion):

**R1. Replacing Pronouns by Its Antecedent:** *reduce coreference ambiguity*

The personal pronoun should be replaced if it is potentially ambiguous, i.e. several referents might be selected. This strategy reduces the amount of processing inferences done by the reader to link one referring expression to its antecedents. To avoid ambiguities, we replace referring expressions with high accessibility by those expressions with low accessibility (e.g. the personal pronoun is replaced with a referent, which might be a definite noun phrase or a proper noun). If possible, the antecedents with the same syntactic function should be selected as the referring expression. For example, the pronoun *elle* is replaced by the subject of the previous sentence (*Madame Dupont*):

*Original version:* La deuxième amie dit que la soupe a une odeur agréable. Madame Dupont est en colère contre elle. $\underline{\text{Elle}}_1$ la trouve hypocrite. *In English:* The second friend says that the soup smells good. Mrs Dupont is angry with her. $\underline{\text{She}}_1$ considers her as hypocrite.

*Simplified version:* La deuxième amie dit que la soupe a une odeur agréable. Madame Dupont est en colère contre elle. $\underline{\text{Madame Dupont}}_1$ la trouve hypocrite. *In English:* The second friend says that the soup smells good. Mrs Dupont is angry with her. $\underline{\text{Mrs Dupont}}_1$ considers her as hypocrite.

**R2. Changing Determiner:** *clearly mark new entities introduction*

The determiner is an important marker of accessibility of a NP. Indefinite noun phrases are usually useful to introduce a new entity in the discourse, while definite nouns phrases (formed with a definite article or demonstrative determiner) have high accessibility and refer to known entities. Demonstrative NPs are more accessible than definite NPs (Ariel, 1990). The determiners might change the position in the accessibility scale, from high accessibility to less accessible referring expressions (e.g. *the dog* $\Rightarrow$ *a dog*), or the

other way around (for some contexts). Thus the composition of the reference chain changes. For example: *Original version:* Le$_1$ renard; Cette$_2$ hyène. *In English:* The$_1$ fox; This$_2$ hyena
*Simplified version:* Un$_1$ renard; La$_2$ hyène. *In English:* A$_1$ fox; The$_2$ hyena.

**R3. Determining Repeated Pronouns:** *reduce working memory demands*

A sequence of repeated pronoun should be replaced by its antecedent when a pronoun (e.g. *il*) is the subject of the main phrase and repeated several times in a different sentences (e.g. *le renard/the fox* ). For example:
*Original version:* Le renard$_1$ avait très soif. Il$_2$ aperçut un puits. Sur la poulie, il y avait une corde, et, à chaque bout de la corde, il y avait un seau. Il$_3$ s'assit dans un des seaux et fut entraîné au fond. Heureux, il$_4$ but pendant de longues minutes. *In English:* The fox$_1$ was very thirsty. It$_2$ saw a well. On the pulley, there was a rope, and at each end of the rope, there was a bucket. It$_3$ sat in one of the buckets and was dragged to the bottom. Happy, it$_4$ drank for long minutes.
*Simplified version:* Le renard$_1$ avait très soif. Le renard$_2$ aperçut un puits. Sur la poulie, il y avait une corde, et, à chaque bout de la corde, il y avait un seau. Le renard$_3$ s'assit dans un des seaux et fut entraîné au fond. Heureux, le renard$_4$ but pendant de longues minutes. *In English:* The fox$_1$ was very thirsty. The fox$_2$ saw a well. On the pulley there was a rope, and at each end of the rope there was a bucket. The fox$_3$ sat in one of the buckets and was dragged to the bottom. Happy, the fox$_4$ drank for long minutes.

**R4. Deleting Pronouns:** *remove inefficient information*

As a side effect of the syntactic simplification, some referring expressions disappear (e.g. indefinite pronouns, such as *chacun*, *quelqu'un*, and possessive determiners, such as *son*), when parts of the main sentences are removed. For example:
*Original version:* Ils avaient chacun$_1$ leur particularité : un était plutôt naïf, l'autre plutôt peureux et le dernier plutôt bavard. *In English:* They had each one$_1$ his particularity: one was quite naive, the other quite timid, and the last one quite talkative.
*Simplified version:* Ils avaient *leurs* particularité*s* : le naïf, le peureux, le bavard. *In English:* They boys had *their* particularities: the naive, the fearful, and the talkative.

**R5. Mention Paraphrasing:** *make noun phrases more readable*

A higher referring expression might be replaced by a less accessible referring expression (Ariel, 1990). In other words, a demonstrative NP might be replaced by a more generic definite NP when a less accessible element replace both the determiners and the noun is replaced by a generic representation of the concept (e.g. hypernym). For example, *ces renards* ⇒ *les animaux*. The possessive determiners are a harder case since the sentence containing the determiner is transformed into a more explicit structure (e.g. *Les lions... leur pays* ⇒ *le pays de lions*; in English, *The lions... their country* ⇒ *the country of lions*). An alternative is to replace the possessive NP (e.g. *son mari*) by its referent (e.g. *M. Dupont*). For example:
*Original version:* Mme Dupont a préparé sa soupe. Son mari$_1$ dit, pour la première fois, qu'il n'aime pas sa soupe. *In English:* Mrs Dupont had prepared her soup. Her husband$_1$ says, for the first time, that he does not like her soup.
*Simplified version:* Mme Dupont a fait sa soupe. M. Dupont$_1$ dit, pour la première fois, qu'il n'aime pas sa soupe. *In English:* Mrs Dupont cooked her soup. Mr. Dupont$_1$ said for the first time that he does not like her soup.
The rules proposed in this section feed the simplification system. They are coded using the operations present in Section 3.2., and their evaluation is presented in the next Section 5..

## 5. Evaluation

In this paper, we propose simplification guidelines, explaining the rules acting at the discourse level and the properties changes in the coreference chains. In this section, we evaluate the finds of this work in two different levels: text transformation errors (Section 5.1.) and simplification perception (Section 5.2.).
This evaluation objectives to measure the simplification in terms of readability, simplicity and grammatically. Hence, we compared the original text against both automatically and manually simplified. One of the difficulties in this evaluation is to ensure that the texts are the most comparable as possible since some parts of the manually simplified are very different from the original one. Therefore, we select text extracts of our corpus, by aligning them with CATS-Align (Stajner et al., 2018), a text alignment tools designed to texts in the same language that takes into account both n-gram and word similarity. This process resulted in a subcorpus (1,339 words) used for the rest of this section.

### 5.1. System evaluation

Given the context of the rewrite tool, the corpus analysis presented in Section 4., and the cohesion rules presented in Section 4.4., we code simplification rules aiming to perform sentence split, subordination suppression, determiner changing, phrase changing and anaphoric resolution. The sentence split rules break coordinate conjunctions that link sentence to sentence or phrase to phrase structures. At the same step, the subordinate structures are removed. In these two rule sets, we opted to reduce the recall, by applying rigorous grammar structures, in order to avoid ungrammatical sentences.
The replacing of the determiners addresses the mention determination proposed by R2 (e.g. *a* replacing *the*). It replaces all determiners in mentions by others with a high accessibility in the hierarchy proposed by Ariel. This process is not applied to possessive determiner, and if a mention is the first element of its chain. It is replaced by an indefinite determiner; otherwise, it is replaced by a definite determiner. Regarding the mentions with possessive determiners, they are replaced by the mention referenced by the

determiner or paraphrased (e.g. *son mari* by *M. Dupont*) following R5.

The anaphoric resolution is processed in two steps. At first, following R3, we resolve all subject or object pronouns with at least one ambiguous pronoun between it and its reference. Later, applying R1, we resolve all remaining pronouns in the subject position.

To evaluate the developed system, we start asking judges to evaluate the system's output for five excerpts. We observe that while some errors are quite obvious (e.g. wrong grammar) others can be difficult to pick. Our group of judges is composed of one French native speaker and two non-natives (the observed disagreements cannot be addressed by the nativeness), and we classify as an error when two of them agree on that. We measured a Fleiss Kappa agreement of 0.744. This evaluation resulted in 42.42% of correct simplifications.

After the system evaluation, we established the error sources, which resulted in 5 error classes (coreference, expression, reference, grammar and enumeration). As can be seen in Table 2, this inspection can explain 93.42% of the errors.

| Error class | % |
|---|---|
| Coreference | 14.47% |
| Expression | 15.79% |
| Reference issues | 52.63% |
| Grammar | 6.58% |
| Enumeration | 3.95% |
| *Total* | 93.42% |

Table 2: Automatic simplification system errors

The coreference class indicates a problem in the referent or link. It has a strong impact in the simplification process because a wrong attachment of a pronoun into a chain poses errors during the coreference resolution step. This error class is directly related to the coreference system performance (see Table 3) that is in line with those published Kantor and Globerson (2019), for English, and Grobol (2019), for French, even that the results cannot be directly compared due to the different language and corpus. Focusing on the coreference system evaluation presented in Table 3, we present the evaluation split according to the use of singletons, due to some proposed rules that process all mentions while others require only those in a chain. The evaluation without singletons indicates the system performance is deeply impacted by them, which is expected since about 80% of the mentions are singletons. The coreference system errors deeply impacted the mention paraphrasing rule, because it tends to link both possessive and possessor to the same entity. Summing up, the coreference errors indicate that the most critical errors in discourse simplification are bounded by the performance of the coreference system, which, even if it is compatible with state of the art, caused 14.5% of errors.

The expression error class groups all error related to language idiosyncrasies, such as idioms and collocations. These elements are not processed by the system now. However, the inclusion of a dictionary could address most of the observed errors.

| | Singleton | Non singleton |
|---|---|---|
| *MUC* | 79.03 | 79.03 |
| $B^3$ | 71.66 | 59.26 |
| $CEAF_e$ | 74.35 | 59.55 |
| *CoNLL* | 75.10 | 65.95 |
| *BLANC* | 65.79 | 59.07 |
| $CEAF_m$ | 69.94 | 61.46 |

Table 3: F1 score of the coreference evaluation, reporting full evaluation and evaluation without singletons

The reference issues class concerning the determiner changes, and the errors in it are mostly related to R2. This class groups several category types, and the most frequent errors are the following: modification of non-accountable nouns that have the determiner defined by French grammar; unique nouns (e.g. *le soleil*/the sun) that usually are preceded by a definite determiner, but depending on the text content, it may be indefinite (e.g. a text about astronomy might address different suns); and noun referenced character introductions (e.g. *le chef de la bande*/the leader of the gang) that, when modified, indicates the presence of other similar characters in another text part.

For the grammar class, the errors indicate some unauthorized grammar modification in the sentence context, such as agreement and superlative. These errors could be solved by identifying specific grammar structures and ignoring them during the simplification since their changing implies to modify the text subject.

Finally, the enumeration class indicates a coreference error related to a sequential characters presentation. These errors are also related to grammar, but they require more than a sentence as context. This class may present a challenge to the system since it mixes different cohesion elements.

At this point, we observe that most of the errors are related to preprocessing, and they could be solved using simple language resources (e.g. dictionary). Interestingly, only two of the text deletions were incorrect. These results make us confident that the system performance can be improved by including more language information. However, our primary goal here is to evaluate the simplification rules. So, we used the results of the evaluation to suppress the wrong modifications, and this was used in simplification judgement step (Section 5.2.)

## 5.2. Simplification judgment

For the simplification perception evaluation, we ask a group of 23 students (Master level in Linguistics and Computational linguistics) to opine about the original and simplified texts. This group is composed of native and non-native french speaks living in France. Aiming to avoid a (non-)nativeness bias, we arrange them similarly in each group.

For each text, a student should pass through a pipeline of four steps:[10] (1) read the text, (2) provide their opinion about readability, simplicity and grammatically, and (3) answer questions about the text. After that, the student should

[10]The evaluation pipeline is quite straightforward even if long (about 8 minutes per text), and no annotator reported issues.

(4) compare the original and simplified texts indicating the meaning preservation, the fluency and grammatically, and the simplification effect, even that the grammaticality and fluidity are already checked. In the second and fourth steps, we asked the students to provide their answers on a Likert scale (0 strong agreement and 5 strong disagreement).

Regarding the questions about the text, we assessed literal (an understanding of the straightforward meaning of the text), reorganization (combination of information from various parts of the text), and evaluation (a global or comprehensive judgment about some aspects of the text) types comprehension, from the taxonomy of the types of comprehension and the forms of questions proposed by Day and Park (2005). The literal assessment questions were outlined as multiple-choice, aiming to test if the judge read the text or just skipped it.[11] The reorganization questions were designed as a true/false question exploring potential misunderstandings of the text. Moreover, we assess the previous knowledge about the text, and all students presented a consistent knowledge for all texts.

We divided the group of 20 students into three groups. It means that on average, we set 2 non-native and 3.8 native speakers in each group.[12] The first group was exposed to the original version of the text, and then to the manually simplified one, aiming to identify the maximum simplification effect to the selected texts. The second group was exposed to the same texts as the first group but in the opposite order. Given that the texts are already simple, no matter their version, our goal with this group is the measure the impact of text simplicity in judges annotation, i.e. how a text can be more difficult. With these two groups, we can measure the simplification range to the selected texts. Finally, the third group was exposed to the automatically simplified and original text versions. Group 3 allows us to measure the simplification impact of the cohesion rules presented in Section 4.4.

Regarding the collected data, we do not observe a statistically significant difference in term of native language for text perception analysis, except for the reading time ($p < 0.001$)[13]. Both native and non-native speakers spend less time reading the automatically simplified (native 72.63 and non-native 100.30 seconds) text than the original one (native 171.39 and non-native 157.98 seconds) $p < 0.001$. All annotators agree on text readability (average of 3.48), simplicity (2.25), and grammatically (2.11) no matter the group language. In the text understanding question, we observe a higher accuracy and time spend to answer in Group 1 (79.6%; 130 seconds) than in Group 2 (75.7%; 115s) and Group 3 (75.0%; 115s).

On different text versions comparison, there is no significant difference in fluency (1.72) and adequacy (2.28) judgements, as expected. However, there is a significant difference in the simplification evaluation. Group 1 scored the text difference as 1.49, which indicates that even the text was classified as simple, there is room to simplify even more. Contrarily to the 4.51 expected, the Group 2[14] scored 3.29 indicating an intermediate disagreement with the simplification. This result is interesting because it can indicate a possible impact from the text order presentation. Group 3 scored 2.69, indicating a slight agreement that the simplified text is most accessible.

## 6. Conclusion and further work

We presented a study of discourse-level transformations to simplify French texts. We analysed the coreference chains properties and their changes induced by manual simplification. This resulted in a set of simplification guidelines that can be automatically applied. Aiming to evaluate the feasibility of the automatic application of these guidelines, we developed a rule-based system changing the structure of the coreference chains. The rules acting at discourse level were manually defined after a study of narrative texts for children (regular readers) and simplified texts (dyslexic children). In this process, we were required to develop a novel text rewritten tool and to train a new coreference model for French.

During the simplification system evaluation, we identified that most of the miss-transformations are resulting of a lack of language resources in the system's pipeline. That indicates that the proposed rules are appropriate, but they require extra linguistic knowledge. However, in a purely rule-based system, it will require significant development time to tune these rules.

The evaluation of the performed simplifications shows that readability and grammatically did not impact the perception of simplification in this corpus. Furthermore, the discourse simplification provided a weak simplification perception when compared to a full simplification pipeline (i.e. lexical, syntactic, and discourse simplification).

As future work, we intend to improve system performance. We will start including more language resources, but we also intend to explore other approaches than rule-based. We also plan to validate the simplification with both a larger group of students and of dyslexic children. Moreover, we would like to include feedback from the simplification target-group.

## 7. Acknowledgments

This work was supported by the Alector (*Aide à la lecture pour enfants dyslexiques et faibles lecteurs*) project from the French National Research Agency (ANR) (ANR-16-CE28-0005). We thank Bruno Oberle (University of Strasbourg) for his help. We thank the students from the Computational Linguistics and Linguistics MA programmes from University of Strasbourg which participate to the evaluation.

[11] We remove the judgements provided by 3 annotators due to they provided wrong answers for 2 literal assessment questions in different texts, and we also remove the text judgement when the literal assessment question is wrong. This resulted in the suppression of 12 judgements.

[12] According to Schwarzer and Kauchak (2018), two annotators are adequate to simplification evaluation.

[13] The statistical significance presented in this paper is based on ANOVA, and the p-values are adjusted using the Tukey's Honest Significant Difference" method.

[14] Group 2 judged the same texts as Group 1, but text in the inverted way.

## 8. Bibliographical References


Aluísio, S. M., Specia, L., Pardo, T. A., Maziero, E. G., and Fortes, R. P. (2008). Towards brazilian portuguese automatic text simplification systems. In *Proceedings of the eighth ACM symposium on Document engineering*, pages 240–248. ACM.

Ariel, M. (1990). *Accessing noun-phrase antecedents*. Routledge.

Ariel, M. (2001). Accessibility theory: An overview. *Text representation: Linguistic and psycholinguistic aspects*, 8:29–87.

Barbu, E., Martín-Valdivia, M. T., and Ureña-López, L. A. (2013). Open book: a tool for helping asd users' semantic comprehension. In *Proceedings of the Workshop on Natural Language Processing for Improving Textual Accessibility*, pages 11–19.

Barlacchi, G. and Tonelli, S. (2013). Ernesta: A sentence simplification tool for children's stories in italian. In *International Conference on Intelligent Text Processing and Computational Linguistics*, pages 476–487. Springer.

Billami, M. B., François, T., and Gala, N. (2018). ReSyf: a French lexicon with ranked synonyms. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 2570–2581, Santa Fe, New Mexico, USA, August. Association for Computational Linguistics.

Brouwers, L., Bernhard, D., Ligozat, A.-L., and François, T. (2014). Syntactic sentence simplification for french. In *3rd International Workshop on Predicting and Improving Text Readability for Target Reader Populations*.

Bulté, B., Sevens, L., and Vandeghinste, V. (2018). Automating lexical simplification in dutch. *Computational Linguistics in the Netherlands Journal*, 8:24–48.

Canning, Y. M. (2002). *Syntactic simplification of Text*. Ph.D. thesis, University of Sunderland.

Chambers, N., Cer, D., Grenager, T., Hall, D., Kiddon, C., MacCartney, B., De Marneffe, M.-C., Ramage, D., Yeh, E., and Manning, C. D. (2007). Learning alignments and leveraging natural logic. In *Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing*, pages 165–170. Association for Computational Linguistics.

Charolles, M. (2006). De la cohérence à la cohésion du discours. In F. Calas, editor, *Cohérence et discours*, pages 25–38. Presses de l'Université Paris Sorbonne.

Day, R. R. and Park, J.-s. (2005). Developing reading comprehension questions. *Reading in a foreign language*, 17(1):60–73.

Dell'Orletta, F., Montemagni, S., and Venturi, G. (2011). Read-it: Assessing readability of italian texts with a view to text simplification. In *Proceedings of the second workshop on speech and language processing for assistive technologies*, pages 73–83. Association for Computational Linguistics.

Désoyer, A., Landragin, F., Tellier, I., Lefeuvre, A., Antoine, J.-Y., and Dinarelli, M. (2016). Coreference resolution for french oral data: Machine learning experiments with ANCOR. In *International Conference on Intelligent Text Processing and Computational Linguistics (CICLing'2016)*.

Fenoglio, I. and Ganascia, J.-G. (2006). Edite, un programme pour l'approche comparative de documents de genèse. *Genesis (Manuscrits-Recherche-Invention)*, 27(1):166–168.

François, T., Billami, M., Gala, N., and Bernhard, D. (2016). Bleu, contusion, ecchymose: tri automatique de synonymes en fonction de leur difficulté de lecture et compréhension. In *JEP-TALN-RECITAL 2016*, volume 2, pages 15–28.

Givón, T. (1993). Coherence in text, coherence in mind. *Pragmatics & Cognition*, 1(2):171–227.

Grobol, L. (2019). Neural coreference resolution with limited lexical context and explicit mention detection for oral french. In *Computational Models of Reference, Anaphora and Coreference (CRAC)*.

Grosz, B. J., Weinstein, S., and Joshi, A. K. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational linguistics*, 21(2):203–225.

Heine, A., Tamm, S., Hofmann, M., Hutzler, F., and Jacobs, A. M. (2006). Does the frequency of the antecedent noun affect the resolution of pronominal anaphors?: An erp study. *Neuroscience letters*, 400(1-2):7–12.

Hirst, G. and St-Onge, D. (1995). Lexical chains as representations of context for the detection and correction of malapropisms. *WordNet: An Electronic Lexical Database*, 305, 10.

Hobbs, J. R. (1979). Coherence and coreference*. *Cognitive Science*, 3(1):67–90.

Inui, K., Fujita, A., Takahashi, T., Iida, R., and Iwakura, T. (2003). Text simplification for reading assistance: a project note. In *Proceedings of the second international workshop on Paraphrasing-Volume 16*, pages 9–16. Association for Computational Linguistics.

Jaffe, E., Shain, C., and Schuler, W. (2018). Coreference and focus in reading times. In *Proceedings of the 8th Workshop on Cognitive Modeling and Computational Linguistics (CMCL 2018)*, pages 1–9.

Kantor, B. and Globerson, A. (2019). Coreference resolution with entity equalization. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 673–677.

Landragin, F. (2016). Description, modélisation et détection automatique des chaînes de référence (DEMOCRAT). *Bulletin de l'Association Française pour l'Intelligence Artificielle*, (92):11–15.

Le Bouëdec, B. and Martins, D. (1998). La production d'inférences lors de la compréhension de textes chez des adultes : une analyse de la littérature.

Levy, R. and Andrew, G. (2006). Tregex and tsurgeon: tools for querying and manipulating tree data structures. In *LREC*, pages 2231–2234. Citeseer.

Li, J., Fabre, M., Luh, W.-M., and Hale, J. (2018). The role of syntax during pronoun resolution: Evidence from fmri. In *Proceedings of the Eight Workshop on Cognitive Aspects of Computational Language Learning and Processing*, pages 56–64.

McMillan, C. T., Clark, R., Gunawardena, D., Ryant, N.,

and Grossman, M. (2012). fmri evidence for strategic decision-making during resolution of pronoun reference. *Neuropsychologia*, 50(5):674–687.

Oberle, B. (2018). SACR: A Drag-and-Drop Based Tool for Coreference Annotation. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan, May 7-12, 2018. European Language Resources Association (ELRA).

Paetzold, G. H. and Specia, L. (2016). Unsupervised lexical simplification for non-native speakers. In *Thirtieth AAAI Conference on Artificial Intelligence*.

Perret, M. (2000). Quelques remarques sur l'anaphore nominale aux xive et xve siècles. *L'Information grammaticale*, 87(1):17–23.

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

Quiniou, S. and Daille, B. (2018). Towards a diagnosis of textual difficulties for children with dyslexia. In *11th International Conference on Language Resources and Evaluation (LREC)*.

Rello, L., Baeza-Yates, R., Bott, S., and Saggion, H. (2013). Simplify or help?: text simplification strategies for people with dyslexia. In *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility*, page 15. ACM.

Ruiter, M. B., Beijer, L. J., Cucchiarini, C., Krahmer, E. J., Rietveld, T. C., Strik, H., et al. (2012). Human language technology and communicative disabilities: requirements and possibilities for the future. *Language Resources and Evaluation*, 46(1):143–151.

Saggion, H., Gómez-Martínez, E., Etayo, E., Anula, A., and Bourg, L. (2011). Text simplification in simplext: Making texts more accessible. *Procesamiento del lenguaje natural*, (47):341–342.

Schnedecker, C. (1997a). *Nom propre et chaînes de référence*. Paris.

Schnedecker, C. (1997b). *Nom propre et chaînes de référence*. Centre d'Études Linguistiques des Textes et Discours de l'Université de Metz.

Schwarzer, M. and Kauchak, D. (2018). Human evaluation for text simplification: The simplicity-adequacy tradeoff. In *SoCal NLP Symposium*.

Seretan, V. (2012). Acquisition of syntactic simplification rules for french. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC-2012)*, pages 4019–4026.

Siddharthan, A. (2006). Syntactic simplification and text cohesion. *Research on Language and Computation*, 4(1):77–109.

Štajner, S., Evans, R., Orasan, C., and Mitkov, R. (2012). What can readability measures really tell us about text complexity. In *Proceedings of workshop on natural language processing for improving textual accessibility*, pages 14–22. Citeseer.

Stajner, S., Franco-Salvador, M., Ponzetto, S. P., and Rosso, P. (2018). Cats: A tool for customised alignment of text simplification corpora. In *Proceedings of the 11th Language Resources and Evaluation Conference, LREC 2018, Miyazaki, Japan, May 7-12*.

Tamburini, F. (2017). Semgrex-plus: a tool for automatic dependency-graph rewriting. In *Proceedings of the Fourth International Conference on Dependency Linguistics (Depling 2017)*, pages 248–254.

Todirascu, A., François, T., Bernhard, D., Gala, N., and Ligozat, A.-L. (2016). Are cohesive features relevant for text readability evaluation? In *26th International Conference on Computational Linguistics (COLING 2016)*, pages 987–997.

Todirascu, A., François, T., Bernhard, D., Gala, N., Ligozat, A.-L., and Khobzi, R. (2017). Chaînes de référence et lisibilité des textes : Le projet ALLuSIF. *Langue française*, 195(3):35–52, September.

Vender, M. (2017). *Disentangling Dyslexia: Phonological and Processing Impairment in Developmental Dyslexia*. Frankfurt: Peter Lang.

Wilkens, R., Dalla Vecchia, A., Boito, M. Z., Padró, M., and Villavicencio, A. (2014). Size does not matter. frequency does. a study of features for measuring lexical complexity. In *Ibero-American Conference on Artificial Intelligence*, pages 129–140. Springer.

Wilkens, R., Oberle, B., Landragin, F., and Todirascu, A. (2020). French coreference for spoken and written language. In *Proceedings of the 12th Language Resources and Evaluation Conference, LREC 2020, Marseille, FRance, May 11-16*.

Woodsend, K. and Lapata, M. (2011). Learning to simplify sentences with quasi-synchronous grammar and integer programming. In *Proceedings of the conference on empirical methods in natural language processing*, pages 409–420. Association for Computational Linguistics.

Yaneva, V. and Evans, R. (2015). Six good predictors of autistic text comprehension. In *Proceedings of the International Conference Recent Advances in Natural Language Processing*, pages 697–706.