

# The interface between readability and automatic text simplification: identifying difficulties to support simple writing



Thomas François



**UCL**  
Université  
catholique  
de Louvain

Assistant Professor at UCLouvain

Invited talk at ATA-18

8 November, 2018





# Plan

- 1 Introduction : text accessibility
- 2 Is assessing text readability possible ?
- 3 The current challenges of text simplification
- 4 Highlighting difficulties
  - Introduction to the CEFRLex project
  - AMesure
- 5 General conclusion

# Current needs for reading

- 18.9% of adults from 33 countries do not have the minimal functioning skills in reading for professional purposes (OCDE report)  
(<http://www.oecd.org/fr/els/l-importance-des-competences-9789264259492-fr.htm>).
- In parallel, the place of language in professional contexts has substantially increased [Boutet, 2001]
- In everyday life, we are more and more confronted with specialized texts :
  - **medical information** in case of chronic pathology (Grabar et al., 2018)
  - **administrative texts** for various purposes
  - **technological knowledge** is ubiquitous

## Current research efforts

- **Readability** : methods to automatically assess the **global** reading difficulty of a text for a given **reader population**.
- **Automatic text simplification** : methods to automatically adapt the style of a text in order to make it more accessible for reading, while preserving most of the meaning.
- **Simple writing aids** : methods to help writers to express themselves in a more effective and more accessible style.

These fields have reach different points in their development...

# Plan

- 1 Introduction : text accessibility
- 2 Is assessing text readability possible ?
- 3 The current challenges of text simplification
- 4 Highlighting difficulties
  - Introduction to the CEFRLex project
  - AMesure
- 5 General conclusion

## A typical classic readability formula

[Flesch, 1948] :

$$\text{Reading Ease} = 206,835 - 0,846 w/l - 1,015 s/l$$

where :

**Reading Ease (RE)** : a score between 0 and 100 (a text for which a 4th grade schoolchild would get 75% of correct answers to a comprehension test)

**w/l** : number of syllables per 100 words

**s/l** : mean number of words per sentence.

- Classic formulas made use of linear regression and a few linguistic **surface** aspects.
- Vocabulary difficulty is assessed in a simple way (syllables, list, letters, etc.)

# Main periods in readability

## 5 major periods in readability :

- 1 The origins** : first works in the field. A lot of interesting perspectives, often forgotten in the current studies !
- 2 Classic period** : formulas are based on linear regression and mostly use two **indices** (one lexical, one syntactic)
- 3 The cloze test era** : concerns arise about motivated features (= cause of difficulty) and difficulty measurement
- 4 Structuro-cognitivist period** : takes into account newly discovered textual dimensions (cohesion, structure, inference load, etc.).  
→ Period of strong criticisms against the classical formulas
- 5 AI readability** : NLP-enabled features are combined with more complex statistical algorithms.  
→ Main success : integrating both classic and structuro-cognitivist approaches.



# What are the uses for readability formulas ?

Readability formula have been used for :

- Selection of materials for textbooks.
- Calibration of books for children [Kibby, 1981, Stenner, 1996].
- Used in scientific experiments to control the difficulty of textual input data.
- Controlling the difficulty level of publications from various administrations (justice, army, etc..) and newspapers.
- More recently, checking the output of automatic summarization, machine translation, etc.  
[Antoniadis and Grusson, 1996, Aluisio et al., 2010, Kanungo and Orr, 2009].
- Assessing automatic text simplification systems  
[Štajner and Saggion, 2013, Woodsend and Lapata, 2011, Zhu et al., 2010]

Contexts where having a single value indicator is fine !

## Shortcomings of this single score : lost information

- Producing this estimate of the reading difficulty of a text is costly :
  - [Feng et al., 2010] : compute 273 features, 28 retained
  - [François and Fairon, 2012] : compute 406 features, 46 retained
  - [Vajjala and Meurers, 2012] : compute 46 features
  - ...
- Use of a lemmatiser, a tagger and/or a syntactic parser
- The results of these analyses is forgotten (ex. word frequency, complex syntactic structures, etc.)

## Shortcomings of this single score : reliability

- Readability assumes that we know which texts are more difficult than other...  
—> what means “difficult” ? How can we measured it ?
- It is measured through another variable, easier to measure and correlated with difficulty  
—> we call it the **critereion** !
- Several criteria have been used in readability ... none are perfect !  
—> Human judgments, textbooks, comprehension tests, cloze test, reading speed, eye-tracking data, etc.

# Expert judgments

## Pros and cons

**Pros** : supposedly reliable, rather convenient (no subjects)

**Cons** : population is not directly tested

→ we model the experts' view of difficulty for the given population

## Issue of heterogeneity

- [van Oosten et al., 2011] had 105 texts assessed by experts (as pairs) and clustered them by similarity of judgements (train one model per cluster).  
→ this leads to different models, whose intracluster performance > intercluster.
- [François et al., 2014a] had 18 experts annotate 105 administrative texts (with an annotation guide)  
→  $0.10 < \alpha < 0.61$  per batch (average = 0.37).
- High agreement seems difficult to reach in readability (SemEval 2012 :  $\kappa = 0.398$  on the test set).

# Using textbooks

## Pros and cons

**Pros** : very convenient (no subjects and no experts !)

→ most popular criterion in AI readability, due to the large training corpus needed

**Cons** : population is not directly tested, heterogeneity of the annotations

- Very few corpora available : Weekly Reader is mostly used [Schwarm and Ostendorf, 2005, Feng et al., 2010, Vajjala and Meurers, 2012]  
→ **risk** : high dependence towards one training corpus, as McCall and Crabbs lessons in classic period [Stevens, 1980]
- This dependence has consequences :
  - formulas will be specialized towards this corpus (coefficients)
  - always the same population and type of texts considered
- Problem of heterogeneity between textbook series [François, 2014]

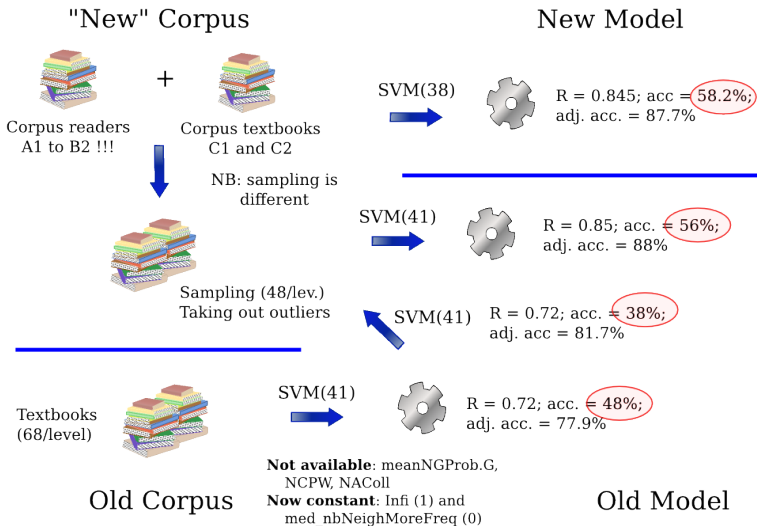
# Example of heterogeneity in a corpus

A1	A2	B1	B2	C1	C2
/	/	-746	-763	-766	-787
-705	-723	/	/	/	/
/	-749	-757	/	/	/
-690	/	/	/	/	/
/	/	/	-758	-766	-777
-694	/	-746	/	/	/
-725	/	/	/	/	/
-696	-730	-753	/	/	/
-731	-742	-733	-766	/	/
/	/	/	/	-787	-778
-664	-712	-756	/	/	/
-711	-740	-752	/	/	/
-683	-740	/	/	/	/
-700.09	-732.9	-750.75	-763.52	-771	-779

# Conclusion about criterion

- No optimal criterion !
- Recent investigation of :
  - eye-tracking [Singh et al., 2016, Bingel et al., 2018a]
  - crowdsourcing [De Clercq et al., 2014]
- Criterion (and corpus) is probably the factor that impact the most readability formulas performance (difficult to compare all work)

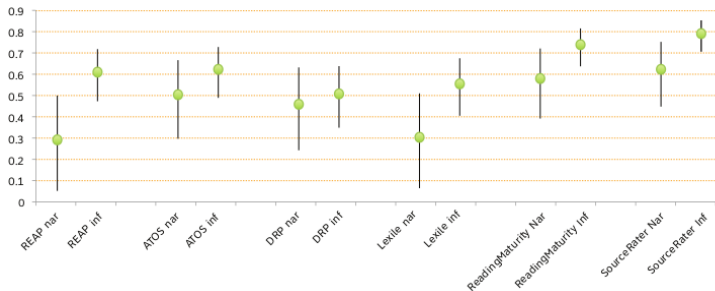
# Other issue : generalization of algorithms





## Effect of the genre

- [Nelson et al., 2012] distinguishes between performance of various famous models on narrative and informative texts



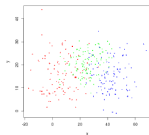
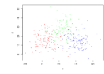
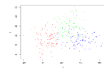
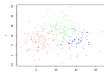
# The curse of readability

- Averaging means more noisy data

Reading test

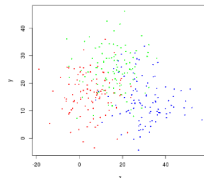


Variance explained = 0.8



Aggregating data

Variance explained = 0.5



## Conclusions on readability

- Readability methods produce a single difficulty value whose reliability is subject to caution
- Probably more efficient to model the reading ability at the individual level (personalised models)
- No much work using neural networks or deep learning so far [Liu et al., 2017]  
→ difficult to combine both !
- Try to better use of the information collected to highlight difficult linguistic phenomena in texts.

# Plan

- 1 Introduction : text accessibility
- 2 Is assessing text readability possible ?
- 3 The current challenges of text simplification**
- 4 Highlighting difficulties
  - Introduction to the CEFRLex project
  - AMesure
- 5 General conclusion

## Is simplifying useful ?

We take for granted that simplifying helps the reading process. What are the evidence ?

- Manual simplification seems to help.
- [Rello et al., 2013] : using shorter words causes readers with dyslexia to read significantly faster and significantly increased their comprehension.  
→ no effect on control group, though !
- [Gala and Ziegler, 2016] : simplified texts are read significantly faster and with fewer (speech) errors by dyslexic children.  
→ no significant effect on comprehension

Still true : “reading comprehension can be improved for **reader with poor literacy** by” manually simplifying texts [Siddharthan, 2014]

# Is simplifying useful ?

## What about automatic simplification ?

- “a summary helps, but the information highlighted in texts do not”  
[Margarido et al., 2008]  
—> only percentages ; no effect computed
- Use of Lickaert scales  
[Wubben et al., 2012, Woodsend and Lapata, 2011]  
—> no effect computed

Still too much evidence for ATS !

# The lack of theoretical grounds

- Not clear what should be simplified !
- Good synthesis by [Siddharthan, 2014] :
  - lexical and simplification revision (L'Allier, 1980)
  - making discourse relations explicit (Beck et al., 1991)
  - reformulating causal relations (Linderholm et al., 2000)
- No full-fledged psycholinguistic model about complexity to rank words/syntactic structures/... [Gala et al., 2018]
- Current approaches are based on data... and are as good as their data (Wikipedia, Newsela)

# Current challenges

- Goes beyond the lexical and syntactic levels !
- Drawing from readability, investigate adaptive approach to take into account the variety of texts and readers [Bingel et al., 2018b]
- Rely more on psycholinguistics evidences to motivate simplifications.
- Develop more parallel data, especially for languages other than English





# Plan

- 1** Introduction : text accessibility
- 2** Is assessing text readability possible ?
- 3** The current challenges of text simplification
- 4** Highlighting difficulties
  - Introduction to the CEFRLex project
  - AMesure
- 5** General conclusion



# Plan

- 1 Introduction : text accessibility
- 2 Is assessing text readability possible ?
- 3 The current challenges of text simplification
- 4 Highlighting difficulties**
  - Introduction to the CEFRLex project
  - AMesure
- 5 General conclusion



# Objectives of the CEFRLex project

- To offer lexical resources describing word distributions in textbooks across the 6 CEFR levels.



- Possible uses :

- Targeted vocabulary learning (which word to learn at which level)
- Comparing the frequency of usage of synonyms
- Using it within a language model for various iCALL tasks (readability, etc.)
- Apply it for automatic text simplification (ATS)

# The CEFRLex project : current projects

## FLELex (French L2)

- Available at <http://cental.uclouvain.be/flelex/>
- Publication : [François et al., 2014b]
- Team : Thomas François, Núria Gala, Anaïs Tack, Patrick Watrin, Cédrick Fairon

## EFLLex (English L2)

- Available at <http://cental.uclouvain.be/cefrlex/>
- Publication : [Dürlich and François, 2018]
- Team : Thomas François, Luise Dürlich

# The CEFRLex project : current projects

## SVALex (Swedish L2 - reception)

- Available at <http://cental.uclouvain.be/svalex/>
- Publication : [François et al., 2016]
- Team : Thomas François, Elena Volodina, Ildikó Pilán, Anaïs Tack

## SweLLex (Swedish L2 - production)

- Available at <http://cental.uclouvain.be/svalex/>
- Publication : [Volodina et al., 2016]
- Team : Elena Volodina, Ildikó Pilán, Lorena Llozhi, Baptiste Degryse and Thomas François

# The CEFRLex project : current projects

## NT2Lex - Dutch L2

- Available at <http://cental.uclouvain.be/cefrlex/>
- Publication : [Tack et al., 2018]
- Team : Anaïs Tack, Thomas François, Piet Desmet, Cédric Fairon

## ELELex - Spanish L2

- Not yet available (publication in preparation)
- Publication : scheduled for 2019
- Team : Thomas François, Barbara De Cock, Irwing Palacios

# Common methodology

- 1** Collect a corpus of texts intended for L2 learners (from textbooks or simplified readers)  
→ The texts must be labelled with a CEFR level
- 2** Find the lemma and the part-of-speech tag of each word in the corpus  
→ Issue : what is a word ? MWE !
- 3** Estimate the frequency distribution of each lemma using a robust estimator  
→ dispersion index [Carroll et al., 1971] to normalize frequencies
- 4** Iterative process : manual postprocessing of the resource to correct NLP errors precedes a new frequency estimation step

# Example : entries from EFLLex and ELELex

lemma	tag	A1	A2	B1	B2	C1	total
cat	NN	2.940	202.796	31.681	33.339	28.9847	65.019
empty	JJ	86.492	150.888	65.947	194.801	123.405	156.021
explore	VB	20.578	54.677	73.625	46.070	56.961	69.590
obviously	RB	0	11.034	2.589	68.463	36.665	30.689
tiresome	JJ	0	0	0	0.315	0.815	0.611
video	NN	2.467	0.556	34.825	23.802	13.248	18.431
write	VB	934.708	378.337	760.734	536.380	713.326	549.909
shopping centre	NN	0	5.040	2.589	0	0.815	1.946
sign up	VB	0	0.887	10.789	2.499	6.216	5.302

lemma	tag	A1	A2	B1	B2	C1	Total
incluir	VM	5.4	60.35	31.3	90.4	258.04	74.1
llamada	NCF	9.6	45.3	56	40.8	9.1	44.6
monumental	AQ0	0	0	1.3	2	0	1.14
malentendido	NCM	0	0	0.4	0	0	0.04
acerca de	SP	0	0	4.9	21.4	18.3	8.2
al fin y al cabo	RG	0	0	0.4	15.5	10.2	4.1



# Demonstration

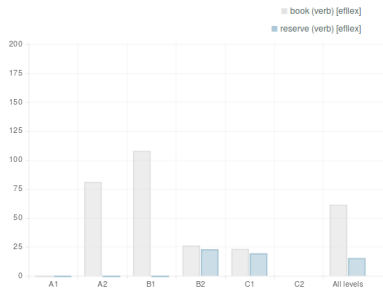
## Make a query in EFLLex

---

Enter a word

Frequencies by CEFR levels for the words book and reserve.



## A few figures about the ressources

Resource	# entries	# Hapax	# MWE	<i>r</i> with list
FLELex-TT	14,236	4,697	2,038	Lexique3 (0, 84)
EFLLex	15,280	/	/	BNC (0.97)
SVALex	15,681	/	1,450	/
ELELex	24,804	/	5,456	/
NT2Lex-CGN	15,227	/	459	Subtlex-NL (0.69)

NB : in NT2Lex-CGN, 4,431 (31%) of the single-word entries are compounds

# CEFRLex project : Assets

- Allow to discriminate words within a level :

Word	Pos	A1	A2	B1	B2	C1
about	RB	2014,8	1210,2	984,4	483,9	238,2
(to) camp	VB	38,7	26,4	11,7	12,3	0

- Frequency estimation from authentic pedagogical documents
- Disambiguation as regards POS-tags and/or senses  
[Tack et al., 2018]

# FLELex relates words to a given CEFR level

FLELex FLELex Search FLELex Download FLELex Analyse a text with FLELex

## Analyse a text with FLELex

With FLELex, it is possible to analyse the lexical complexity of a French text for a specific CEFR proficiency level. All you need to do is introduce a text of your choice and we'll do the analysis for you. For additional tips and tricks on how to interpret the analysis, please consult the "How-to" tab below.

New text Analysis How-to ▾

Lexical complexity for level **A2**

La présidente **nouvellement** **étue** demande l'abolition de la taxe sur le **capital**.



Webmaster: CENTAL (Centre de traitement automatique du langage)  
Collège Erasme, 1 place Blaise Pascal, B-1348 Louvain-la-Neuve (Belgique)

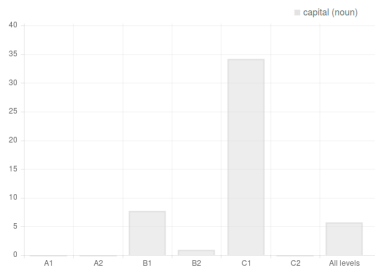




# From distribution to level

**Problem** : How to transform a distribution into a single level ?

**Example** : the distribution of *capital*



\* Frequencies of capital in FLELex-TT.

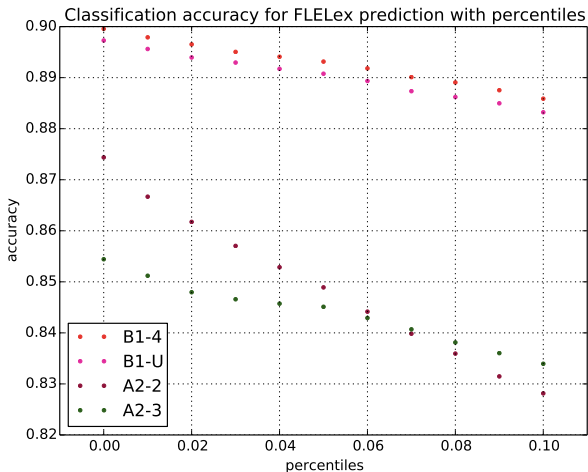
... is transformed into B1

# From distribution to level

## Experiment from [Tack et al., 2016a]

- Collect annotations from 4 Flemish learners of FFL (A2 and B1) on 51 short texts  
→ learners report unknown words via a web interface.
- Then, various thresholds (frequency value, quantile) were tested in order to maximize the prediction of unknown words for the 4 subjects.
- Surprisingly, the best discretization function is “first occurrence” !

# From distribution to level



## Using FLELex as a way to predict lexical knowledge

Rule to decide between known (0) and unknown (1) words :

$$c = \begin{cases} 1, & \text{if } l > L \vee \exists ! l \\ 0, & \text{otherwise} \end{cases}$$

	Lexical words	Grammatical words	Total
Learner A2-2	86.6%	99.2%	89.7%
Learner A2-3	81.1%	99.2%	87.4%
Learner B1-4	91.3%	99.7%	92.3%
Learner B1-U	90.8%	99.8%	92.0%

**TABLE :** Accuracy for the prediction of the lexical knowledge of the 4 learners using FLELex.



# Discussion

- In the interface, predictions appears as overoptimistic (too much A1 words)
- The evaluation seems good, but...
  - The model behave better on known words than on unknown ones (less numerous)
- This is a consequence of the rule “first occurrence”, which appears as too simple !

	Connu	Inconnu
apprenant A2-2	95.7% (0.92)	4.3% (0.42)
apprenant A2-3	88.1% (0.94)	11.9% (0.38)
apprenant B1-4	97.0% (0.94)	3.0% (0.40)
apprenant B1-U	96.7% (0.94)	3.3% (0.37)

**TABLE :** Percentage of known and unknown words + recall for predictions based on FLELex [Tack et al., 2016a].

# Plan

- 1 Introduction : text accessibility
- 2 Is assessing text readability possible ?
- 3 The current challenges of text simplification
- 4 **Highlighting difficulties**
  - Introduction to the CEFRLex project
  - **AMesure**
- 5 General conclusion

# The AMesure project

Administrative texts can be hard to understand :

## AMesure is...

- A platform for simple writing of administrative texts, supported by the FWB (Federation Wallonia-Brussels).
- **Objectives :**
  - 1 Draw the attention of writers to complex linguistic phenomena ;
  - 2 Provide pieces of advice for manual simplification of administrative texts (based on simple writing guides).

# AMesure : current situation

- A global readability score (readability formula, in A) [François et al., 2014a]
- Assessment of several linguistics dimensions of the text (B)
- Highlighting complex phenomena in the text (C)
- Suggestions for simple writing for each sentence (D)

## Difficulté du texte



## Analyse détaillée: TextField

### PaLETTE d'analyse

- Analyse des phrases
- Subordonnées
- Toutes
- Relatives
- Complétives
- Autres
- Voix passive
- Toutes

### Texte annoté :

En introduisant le formulaire C 131 A travailleur sur lequel vous avez complété la partie I, auprès de votre organisme de paiement. Votre employeur peut envoyer par voie électronique une Déclaration de début de travail à temps partiel avec maintien des droits à l'extranet de la sécurité sociale (cette déclaration sera envoyée auprès de votre organisme de paiement) ou il peut compléter un formulaire papier C 131 A employeur (que vous devez introduire vous-même auprès de votre organisme de paiement).

Phrase  
La phrase contient 6 difficultés syntaxiques. Essayez de simplifier la phrase.  
La phrase fait 59 mots. Le maximum conseillé est de 15.  
Évitez la multiplication des parenthèses dans une phrase.  
Évitez la multiplication des subordonnées.

# Detecting complex phenomena in administrative texts

Currently detected :

- Subordinated clauses :
  - relative clauses
  - object clause (fr. *complétive*)
  - adverbial clause
- Passive sentence
- Brackets
- Abbreviations (list-based and rules-based)
- Complex words (frequency-based)

# Detecting syntactic structures

Implementation based on [Brouwers et al., 2014]

Sentences to  
analyse



Syntactic parsing  
(Berkeley Parser)



Defining rules  
(based on a corpus)

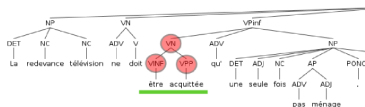


Applying regular  
expressions  
(via Tregex)

En Région wallonne, une taxe annuelle d'un montant de 100 €  
doit être payée lorsque l'on détient un appareil de télévision,  
quel que soit l'usage qui en est fait.



"VN < (V | VINf | VPP "+etre()+ " \$. (VPP "+notVIntransitif()+)");





# Evaluation

- Results of the previous system (François et al., 2018)
- Test data = 13 administrative texts (319 clauses, 134 passives, 57 abbreviations)

Phenomena	Precision	Recall
all subordinated clauses	0.89	0.88
relative clauses	0.98	0.85
object clauses	0.6	0.67
adverbial clauses	0.83	0.84
passives	0.9	0.92
abbreviations	0.57	0.4

# Generating the advices

- Theoretical reference = simple writing guides by the administrations...
- 7 cases have been implemented

Problem	Condition
number of nested syntactic structures	$\geq 3$
total number of clauses	$> 3$
length of the sentence	$> 15$ words
length of the longest nested clause	$> 10$ words
length of text between brackets	$> 10$ words
number of subordinated clauses	$\geq 3$



# Example

AMESURE Home Aide

## Analyse détaillée: Exemple

### Palette d'analyse

- Analyse des phrases
- Subordonnées
  - Toutes
  - Relatives
  - Complétives
  - Autres
- Voix passive
  - Toutes
- Information secondaire
  - Parenthèses
- Complexité lexicale
  - Mots complexes
  - Abréviations

### Texte annoté :

Le rôle est un document fiscal global qui reprend le nom de l'ensemble des redevables ainsi que le montant de l'impôt ou de la taxe dont ils sont redevables. En matière de redevance télévision, vous êtes tenu par la loi, en tant que redevable, de payer la redevance dans le délai fixé par l'invitation à payer qui vous est adressée par l'administration fiscale wallonne (Direction générale opérationnelle de la Fiscalité). Si vous n'effectuez pas le paiement réclamé par cette invitation, le montant dû est alors enrôlé et un avertissement-extrait de rôle (qui est donc un extrait individuel du rôle vous concernant personnellement) vous sera alors envoyé. Le montant de redevance repris sur cet avertissement-extrait de rôle est considéré, au regard de la loi, comme une dette certaine au profit du trésor public, qui doit être payée immédiatement. Attention: L'Administration fiscale n'adresse pas de rappel après envoi de cet avertissement-extrait de rôle. En cas de non paiement, votre dossier sera donc transmis chez un huissier de justice et votre dette initiale sera donc augmentée de frais supplémentaires. C'est pourquoi, en cas de réception d'un tel document, il vous est instamment conseillé de verser sans délai la somme réclamée. Si vous contestez cette taxation, vous pouvez introduire une réclamation par écrit dans les 6 mois selon les formes précisées sur l'avertissement-extrait de rôle et explicitées dans le feuillet explicatif qui y est joint. à noter que l'introduction d'une éventuelle demande d'exonération n'est pas suspensive ni de l'obligation du paiement, ni de l'introduction d'une réclamation administrative

— Extrait pris du site: <http://www.wallonie.be/fr/redevance-impayee-et-facilites-de-paiement>

### Phrase

Pas plus de deux imbrications de phénomènes complexes.

La phrase contient 6 difficultés syntaxiques. Essayez de simplifier la phrase.

La phrase fait 43 mots. Le maximum conseillé est de 15.

Eviter la multiplication des subordonnées.



## Next version in preparation

- Diagnosis organised in three main dimensions (lexical, syntactic, discursive)
- Further investigation of abbreviations :
  - density of abbreviations
  - detecting abbreviations without an extended forms in the text
  - detecting technical terms
- Global ratio of complex structures for the text
- Suggest simpler synonyms using ReSyf [Billami et al., 2018]

Soon out !

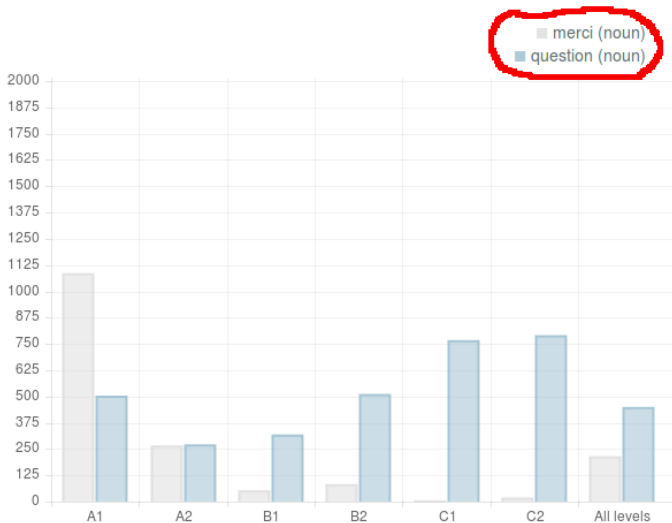
# Plan

- 1 Introduction : text accessibility
- 2 Is assessing text readability possible ?
- 3 The current challenges of text simplification
- 4 Highlighting difficulties
  - Introduction to the CEFRLex project
  - AMesure
- 5 General conclusion

# Conclusions

- Highlighting reading difficulties is a simpler task than ATS and a more meaningful one than readability  
→ It is also easier to test on humans
- Suffer from the same lack of theoretical definition of complexity  
→ Allows to test independently each category of phenomena
- Adaptation is also required (Yimam and Biemann, 2018)  
→ Personalising the models seems to account for a larger part of the variance [Tack et al., 2016b]

# Thank you for your attention



# References I



Aluisio, S., Specia, L., Gasperin, C., and Scarton, C. (2010).

Readability assessment for text simplification.

In *Fifth Workshop on Innovative Use of NLP for Building Educational Applications*, pages 1–9, Los Angeles.



Antoniadis, G. and Grusson, Y. (1996).

Modélisation et génération automatique de la lisibilité de textes.

In *ILN 96 : Informatique et Langue Naturelle*.



Billami, M., François, T., and Gala, N. (2018).

Resyf : a french lexicon with ranked synonyms.

In *27th International Conference on Computational Linguistics (COLING 2018)*.



Bingel, J., Barrett, M., and Klerke, S. (2018a).

Predicting misreadings from gaze in children with reading difficulties.

In *Proceedings of the Thirteenth Workshop on Innovative Use of NLP for Building Educational Applications*, pages 24–34.

## References II



Bingel, J., Paetzold, G., and Søgaard, A. (2018b).

Lexi : A tool for adaptive, personalized text simplification.

In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 245–258.



Boutet, J. (2001).

La part langagière du travail : bilan et évolution.

*Langage et société*, (4) :17–42.



Brouwers, L., Bernhard, D., Ligozat, A.-L., and François, T. (2014).

Syntactic sentence simplification for french.

In *Proceedings of the 3rd International Workshop on Predicting and Improving Text Readability for Target Reader Populations (PITR 2014)*.



Carroll, J., Davies, P., and Richman, B. (1971).

*The American Heritage word frequency book*.

Houghton Mifflin Boston.



De Clercq, O., Hoste, V., Desmet, B., Van Oosten, P., De Cock, M., and Macken, L. (2014).

Using the crowd for readability prediction.

*Natural Language Engineering*, 20(3) :293–325.

## References III



Dürlich, L. and François, T. (2018).

EFLLex : A Graded Lexical Resource for Learners of English as a Foreign Language.  
In *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018)*, pages 1–7.



Feng, L., Jansche, M., Huenerfauth, M., and Elhadad, N. (2010).

A Comparison of Features for Automatic Readability Assessment.  
In *COLING 2010 : Poster Volume*, pages 276–284.



Flesch, R. (1948).

A new readability yardstick.  
*Journal of Applied Psychology*, 32(3) :221–233.



François, T., Brouwers, L., Naets, H., and Fairon, C. (2014a).

AMesure : une formule de lisibilité pour les textes administratifs.  
In *Actes de la 21e Conférence sur le Traitement automatique des Langues Naturelles (TALN 2014)*.



## References IV



François, T. and Fairon, C. (2012).

An "AI readability" formula for French as a foreign language.

In *Proceedings of the 2012 Conference on Empirical Methods in Natural Language Processing (EMNLP 2012)*, pages 466–477.



François, T., Gala, N., Watrin, P., and Fairon, C. (2014b).

FLELex : a graded lexical resource for French foreign learners.

In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014)*.



François, T., Volodina, E., Ildikó, P., and Tack, A. (2016).

SVALex : a CEFR-graded lexical resource for Swedish foreign and second language learners.

In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016)*, pages 213–219.

# References V



François, T. (2014).

An analysis of a french as a foreign language corpus for readability assessment.

*In Proceedings of the 3rd workshop on NLP for Computer-assisted Language Learning, NEALT Proceedings Series Vol. 22, Linköping Electronic Conference Proceedings 107, pages 13–32.*



Gala, N., François, T., Javourey-Drevet, L., and Ziegler, J. C. (2018).

La simplification de textes, une aide à l'apprentissage de la lecture.

*Langue française, (3) :123–131.*



Gala, N. and Ziegler, J. (2016).

Reducing lexical complexity as a tool to increase text accessibility for children with dyslexia.

*In Proceedings of the Workshop on Computational Linguistics for Linguistic Complexity (CL4LC), pages 59–66.*



Kanungo, T. and Orr, D. (2009).

Predicting the readability of short web summaries.

*In Proceedings of the Second ACM International Conference on Web Search and Data Mining, pages 202–211.*



## References VI



Kibby, M. (1981).  
Test Review : The Degrees of Reading Power.  
*Journal of Reading*, 24(5) :416–427.



Liu, H., Li, S., Zhao, J., Bao, Z., and Bai, X. (2017).  
Chinese teaching material readability assessment with contextual information.  
In *Asian Language Processing (IALP), 2017 International Conference on*, pages 66–69.  
IEEE.



Margarido, P. R., Pardo, T. A., Antonio, G. M., Fuentes, V. B., Aires, R., Aluísio, S. M., and Fortes, R. P. (2008).  
Automatic summarization for text simplification : Evaluating text understanding by poor readers.  
In *Companion Proceedings of the XIV Brazilian Symposium on Multimedia and the Web*, pages 310–315. ACM.



Nelson, J., Perfetti, C., Liben, D., and Liben, M. (2012).  
Measures of text difficulty : Testing their predictive value for grade levels and student performance.  
*Student Achievement Partners*.

## References VII



Rello, L., Baeza-Yates, R., Dempere-Marco, L., and Saggion, H. (2013).  
Frequent words improve readability and short words improve understandability for people with dyslexia.  
*In IFIP Conference on Human-Computer Interaction*, pages 203–219. Springer.



Schwarm, S. and Ostendorf, M. (2005).  
Reading level assessment using support vector machines and statistical language models.  
*Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*,  
pages 523–530.



Siddharthan, A. (2014).  
A survey of research on text simplification.  
*ITL-International Journal of Applied Linguistics*, 165(2) :259–298.



Singh, A. D., Mehta, P., Husain, S., and Rajakrishnan, R. (2016).  
Quantifying sentence complexity based on eye-tracking measures.  
*In Proceedings of the workshop on computational linguistics for linguistic complexity (cl4lc)*,  
pages 202–212.

## References VIII



Štajner, S. and Saggion, H. (2013).

Readability indices for automatic evaluation of text simplification systems : A feasibility study for spanish.

*In Proceedings of the 6th International Joint Conference on Natural Language Processing (IJCNLP 2013)*, pages 374–382, Nagoya, Japan.



Stenner, A. (1996).

Measuring reading comprehension with the lexile framework.

*In Fourth North American Conference on Adolescent/Adult Literacy.*



Stevens, K. (1980).

Readability formulae and McCall-Crabbs standard test lessons in reading.

*The Reading Teacher*, 33(4) :413–415.



Tack, A., François, T., Desmet, P., and Fairon, C. (2018).

NT2Lex : A CEFR-Graded Lexical Resource for Dutch as a Foreign Language Linked to Open Dutch WordNet.

*In Proceedings of the 13th Workshop on Innovative Use of NLP for Building Educational Applications (NAACL 2018).*



# References IX



Tack, A., François, T., Ligozat, A.-L., and Fairon, C. (2016a).

Evaluating lexical simplification and vocabulary knowledge for learners of french : possibilities of using the flelex resource.

*In Proceedings of the Tenth conference on International Language Resources and Evaluation (LREC'16)*, pages 230–236.



Tack, A., François, T., Ligozat, A.-L., and Fairon, C. (2016b).

Modèles adaptatifs pour prédire automatiquement la compétence lexicale d'un apprenant de français langue étrangère.

*In Actes de la 23e Conférence sur le Traitement Automatique des Langues Naturelles (TALN 2016)*, pages 1–14.



Vajjala, S. and Meurers, D. (2012).

On improving the accuracy of readability classification using insights from second language acquisition.

*In Proceedings of the Seventh Workshop on Building Educational Applications Using NLP*, pages 163–173.

# References X



van Oosten, P., Hoste, V., and Tanghe, D. (2011).

A posteriori agreement as a quality measure for readability prediction systems.

In Gelbukh, A., editor, *Computational Linguistics and Intelligent Text Processing*, volume 6609 of *Lecture Notes in Computer Science*, pages 424–435. Springer, Berlin / Heidelberg.



Volodina, E., Pilán, I., Llozhi, L., Degryse, B., and François, T. (2016).

Swellex : second language learners' productive vocabulary.

In *Proceedings of the joint workshop on NLP for Computer Assisted Language Learning and NLP for Language Acquisition at SLTC, Umeå*, number 130, pages 76–84. Linköping University Electronic Press.



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.

# References XI



Wubben, S., Van Den Bosch, A., and Krahmer, E. (2012).  
Sentence simplification by monolingual machine translation.  
In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics : Long Papers-Volume 1*, pages 1015–1024. Association for Computational Linguistics.



Zhu, Z., Bernhard, D., and Gurevych, I. (2010).  
A monolingual tree-based translation model for sentence simplification.  
In *Proceedings of the 23rd International Conference on Computational Linguistics*, pages 1353–1361. Association for Computational Linguistics.