GR4L2
Building CEFR-graded resources for second and foreign language learning

Graded resources: from linguistic engineering to practical applications

Núria Gala

December 7th, 2021 - Louvain-la-Neuve

UMR 7309 • CNRS • AMU
PAROLE ET
LANGAGE

Aix*Marseille
université
Socialement engagée

Overview

Graded resources: roots and context

Back to the roots : empiricism

Grade schools, reading standards

Readability, word-lists and corpus linguistics

Graded resources

Linguistic engineering and practical applications

Using graded reading materials

Analyzing language complexity

Modelling student difficulties

Using the resources in the classrooms or in personalized training

Graded resources: roots and context

Back to the roots : empiricism

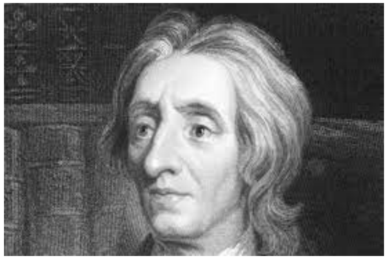
Grade schools, reading standards

Readability, word-lists and corpus linguistics

Graded resources

Empiricism

Derivation from the ancient Greek word *empeiria* ("experience")



John Locke (1632 – 1704)
British empiricist: "truth and knowledge arise out of observation and experience rather than manipulation of accepted or given ideas". Need for children to have concrete experiences to learn.

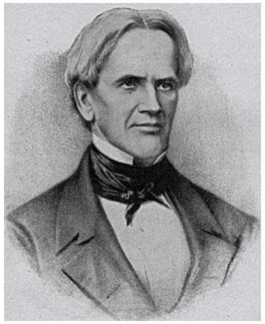
Jean Jacques Rousseau (1712 – 1778)
His philosophy of education: learning through experiencing, "child-centered" education.

Influence on **education**: grade schools, reading tests, adapted reading materials. **Putting the learner at the forefront.**

From one-room schools… to grade schools

Early 19th « Children under the age of five were often mixed in with adults in their twenties. Additionally, classrooms were frequently overcrowded, housing as many as eighty students at a time. Because of the overcrowding, already scarce textbooks and learning materials had to be spread even more thinly amongst students. As a result, class time amounted to a tedious recitation of facts and instructors struggled to devote individual attention to students. »

Ted Brackemyre. *The Rise of Public Education in Early America*. 2021 U.S. History Scene.
https://ushistoryscene.com/article/rise-of-public-education/

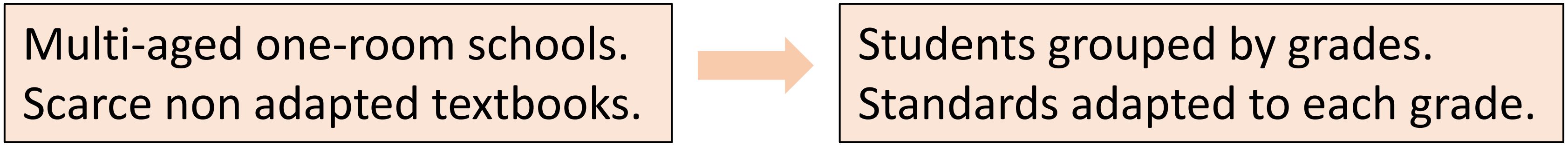


Horace Mann (1796 – 1859), promoter of public education.

1847 first graded school (Boston) with books prepared for each grade.

Students learn best with materials written for their current reading level.

Reading standards were set for each grade.



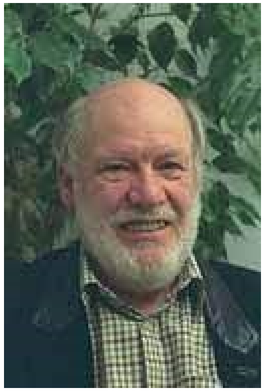
Multi-aged one-room schools.
Scarce non adapted textbooks.

Students grouped by grades.
Standards adapted to each grade.

Empiricism

Derivation from the ancient Greek word *empeiria* (“experience”)

L. Bloomfield and Z. Harris, US distributionalism (1940s-1960s), based on behaviorist psychological theories and on direct observation of environments: *“you shall know a word by the company it keeps”* (Firth, 1957).



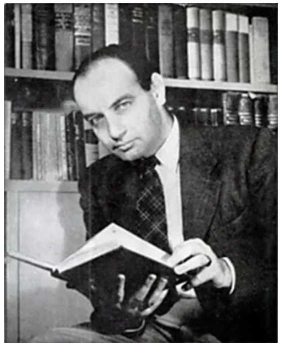
J. Sinclair and G. Leech (1970s-2000s), Corpus linguistics: study of language through its samples, e.g., corpus-driven lexicons for foreign learners of English.

Influence on **language teaching**: readability formulae, word-lists, corpus.
Putting word distributions at the forefront.

Predicting text readability
Vocabulary learning through word lists

First readability formulae (20th)

- B. A. Lively and S. L. Pressey, predicting readability based on word-frequencies (*A method for measuring the vocabulary burden of text-book*, 1923)
- **R. Flesch** (*Marks of Readable Style : A study in Adult Education,* 1943 et 1948)
  Reading Ease score: length (syllables/word, words/sentence)



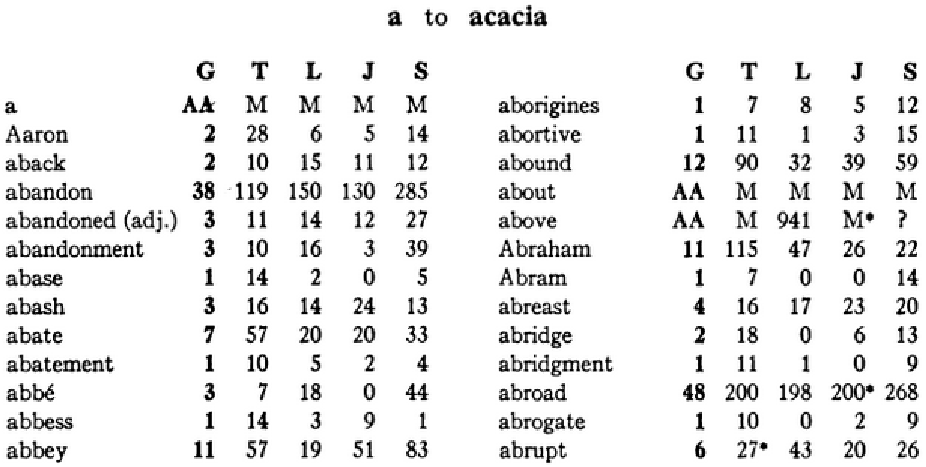
Computational readability (21st)

NLP and machine learning Collins-Thompson & Callan (2005), François (2009)

Neural approaches, deep learning Deutsch, Jasbi & Shieber (2020), Martinc, Pollak & Robnik-Šikonja (2021)

Teachers' Book of Words (Thorndike, 1921)

Basic English (Ogden, 1930)

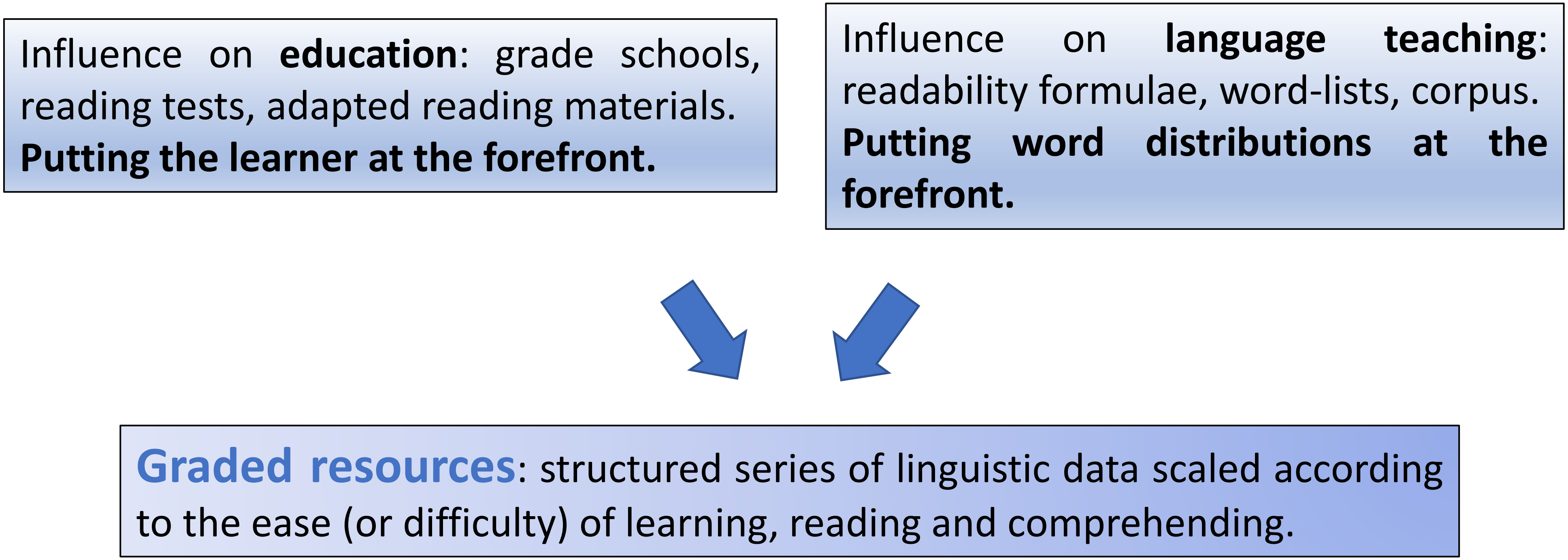


a to **acacia**

	G	T	L	J	S
a	**AA**	M	M	M	M
Aaron	**2**	28	6	5	14
aback	**2**	10	15	11	12
abandon	**38**	·119	150	130	285
abandoned (adj.)	**3**	11	14	12	27
abandonment	**3**	10	16	3	39
abase	**1**	14	2	0	5
abash	**3**	16	14	24	13
abate	**7**	57	20	20	33
abatement	**1**	10	5	2	4
abbé	**3**	7	18	0	44
abbess	**1**	14	3	9	1
abbey	**11**	57	19	51	83
aborigines	**1**	7	8	5	12
abortive	**1**	11	1	3	15
abound	**12**	90	32	39	59
about	**AA**	M	M	M	M
above	**AA**	M	941	M•	?
Abraham	**11**	115	47	26	22
Abram	**1**	7	0	0	14
abreast	**4**	16	17	23	20
abridge	**2**	18	0	6	13
abridgment	**1**	11	1	0	9
abroad	**48**	200	198	200•	268
abrogate	**1**	10	0	2	9
abrupt	**6**	27•	43	20	26

Thorndike & Lorge (1921)
https://catalog.hathitrust.org/Record/000987642

Graded resources



Influence on **education**: grade schools, reading tests, adapted reading materials. **Putting the learner at the forefront.**

Influence on **language teaching**: readability formulae, word-lists, corpus. **Putting word distributions at the forefront.**

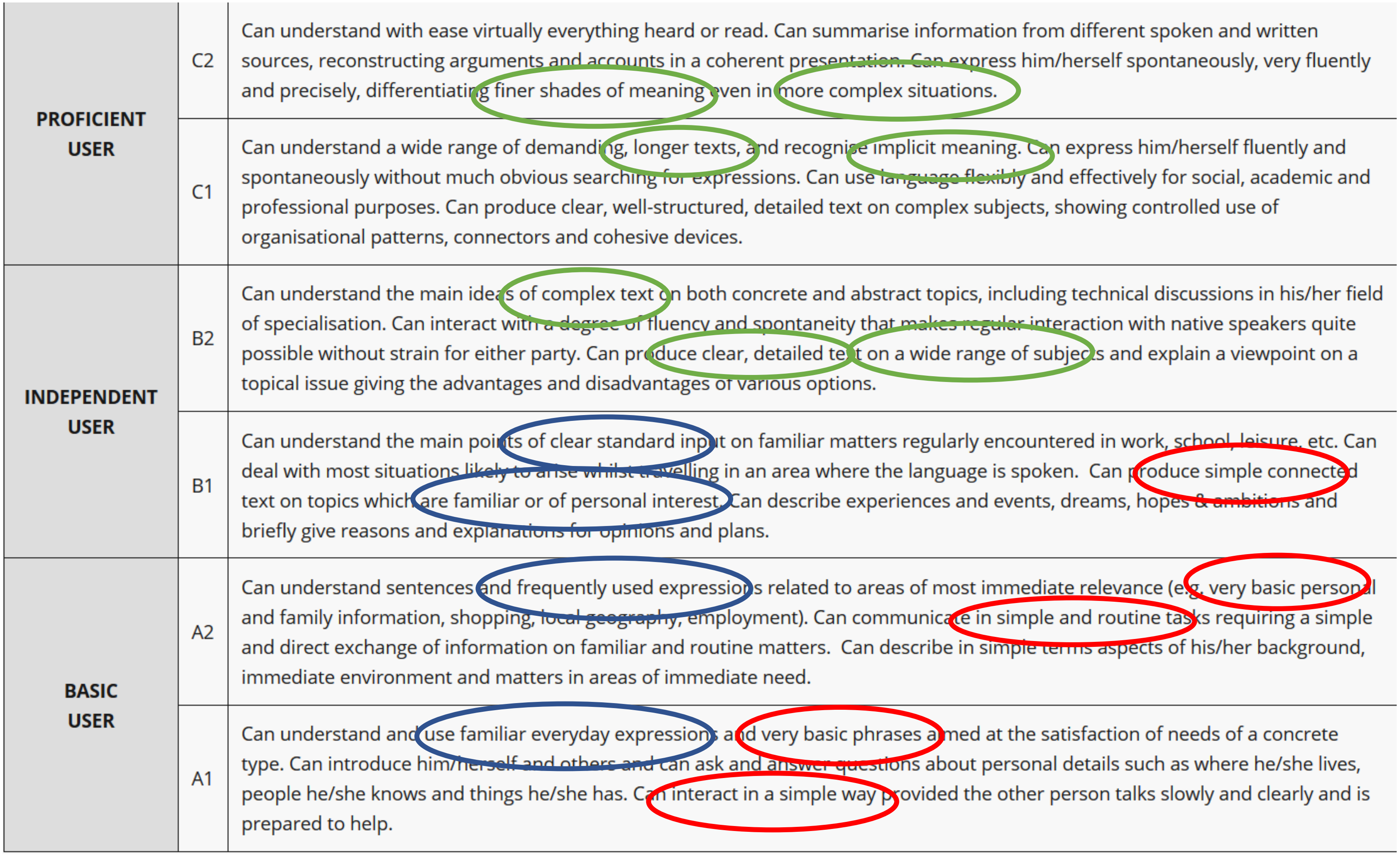
Graded resources: structured series of linguistic data scaled according to the ease (or difficulty) of learning, reading and comprehending.

Remarks:

- Lexicons *vs* corpora
- Scales often correspond to stablished learning grades, i.e., CEFR
- Teacher judgments of the difficulty *vs l*earner abilities

CEFR (2001)

(Figueras, 2012)



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

« The data in the scaling studies were intuitive teacher judgments rather than samples of performance. » (Fulcher, 2010)

« The lack of systematicity may be indicative of some **incongruities in the way reading materials were graded with CEFR levels**, which may call for a more critical reflection. » (Tack, 2021)

© N. Gala https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions

Linguistic engineering and practical applications

Using graded reading materials

Analyzing language complexity

Modelling student difficulties

Using the resources in the classrooms or in personalized training

Using graded reading materials

Lété, Sprenger-Charolles & Colé (2004)

Scarce tools for studying child language development.
Frequency effect : one of the earliest empirical observations in cognitive psychology.

MANULEX

First grade-level lexical database built from text-books of year 2000.

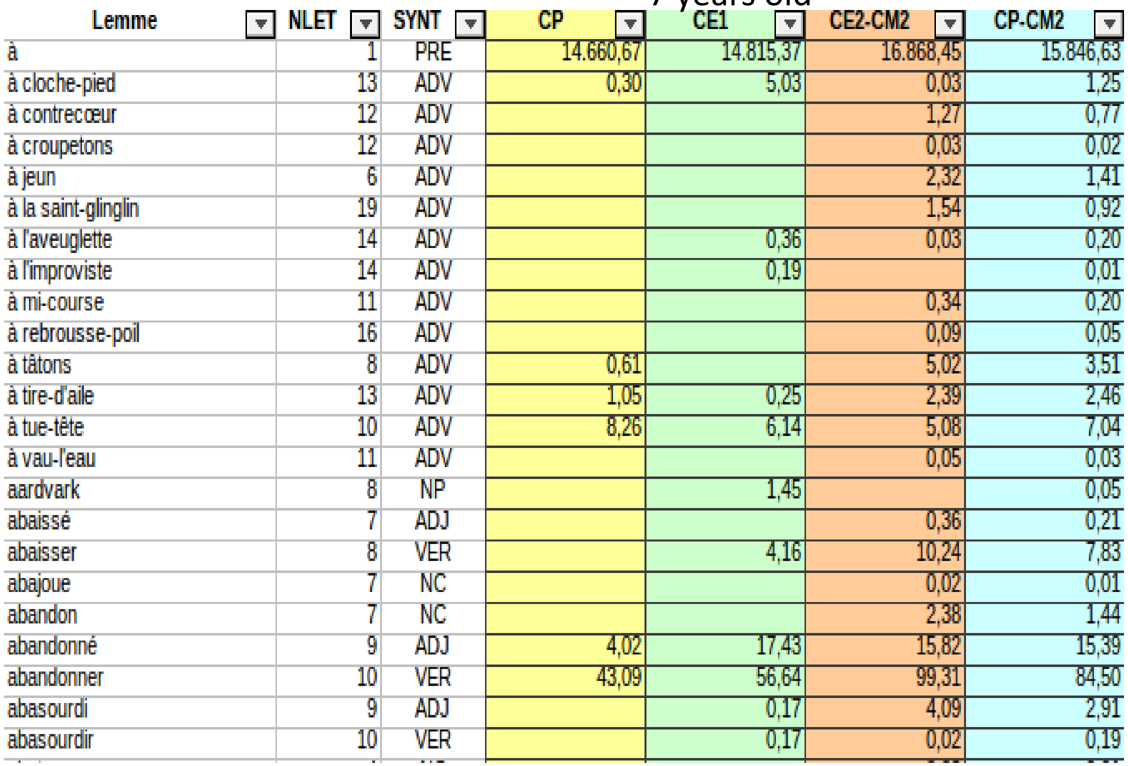
Frequency distributions of words observed across text-books for French L1.

5 grades in primary schools, grouped into 3 (according to the ease of reading): CP 6, CE1 7, CE2 to CM2 8-10 years old.

6 years old

7 years old

8-10 years old



Lemme	NLET	SYNT	CP	CE1	CE2-CM2	CP-CM2
à	1	PRE	14.660,67	14.815,37	16.868,45	15.846,63
à cloche-pied	13	ADV	0,30	5,03	0,03	1,25
à contrecœur	12	ADV			1,27	0,77
à croupetons	12	ADV			0,03	0,02
à jeun	6	ADV			2,32	1,41
à la saint-glinglin	19	ADV			1,54	0,92
à l'aveuglette	14	ADV		0,36	0,03	0,20
à l'improviste	14	ADV		0,19		0,01
à mi-course	11	ADV			0,34	0,20
à rebrousse-poil	16	ADV			0,09	0,05
à tâtons	8	ADV	0,61		5,02	3,51
à tire-d'aile	13	ADV	1,05	0,25	2,39	2,46
à tue-tête	10	ADV	8,26	6,14	5,08	7,04
à vau-l'eau	11	ADV			0,05	0,03
aardvark	8	NP		1,45		0,05
abaissé	7	ADJ			0,36	0,21
abaisser	8	VER		4,16	10,24	7,83
abajoue	7	NC			0,02	0,01
abandon	7	NC			2,38	1,44
abandonné	9	ADJ	4,02	17,43	15,82	15,39
abandonner	10	VER	43,09	56,64	99,31	84,50
abasourdi	9	ADJ		0,17	4,09	2,91
abasourdir	10	VER		0,17	0,02	0,19

http://manulex.org

FLELex: grading French L2 vocabulary

François, Gala, Watrin & Fairon (2014) ; Tack, François & Fairon (2016)

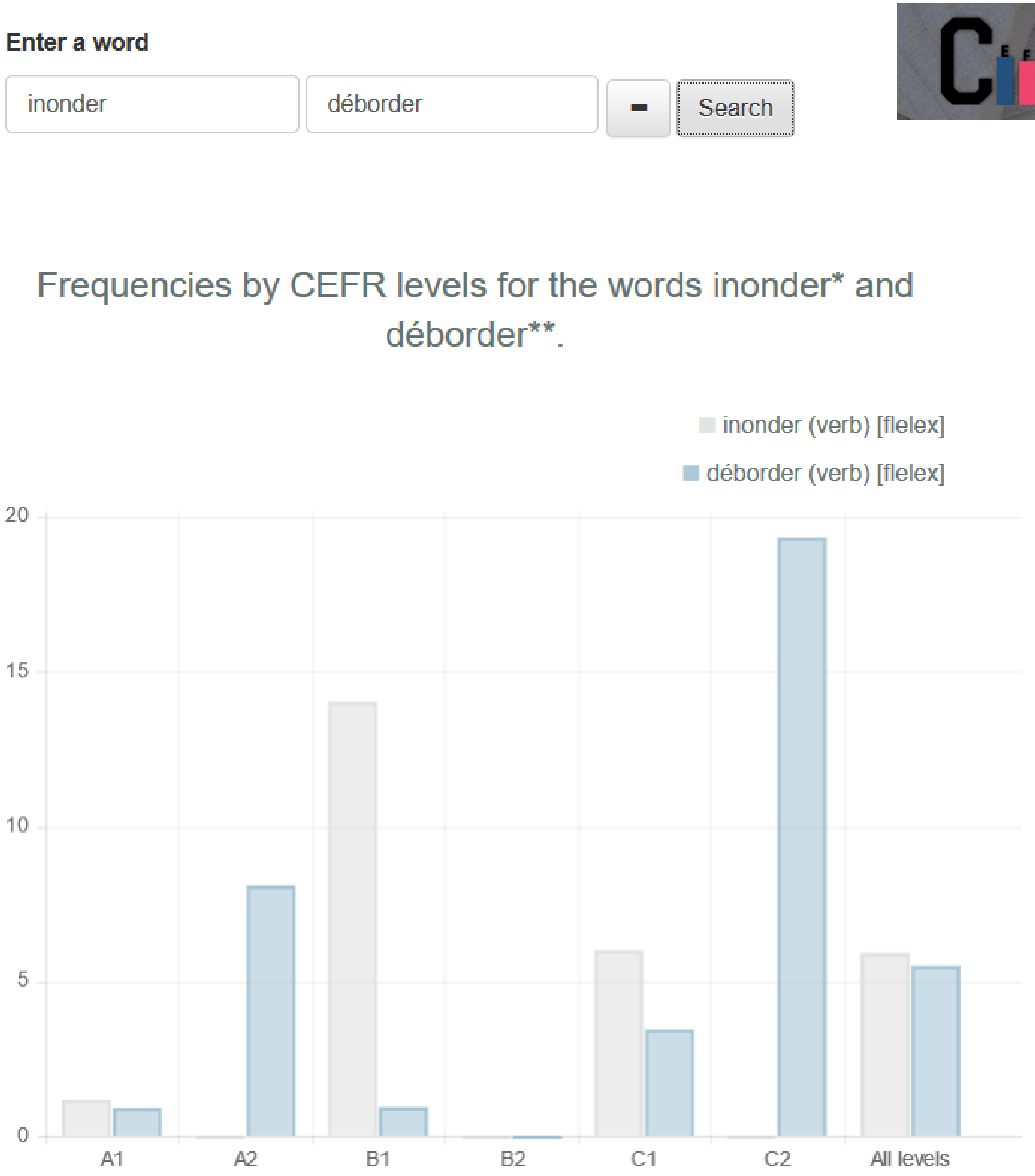
Word frequencies by difficulty level according the CEFR scale, first resource of the CEFRLex project.

777,000 words distributed across several textual genres.

Available online, possibility of comparison between 2 words.

Possibilities of using the FLELex resource: evaluating lexical simplification and vocabulary knowledge for learners of French

http://cental.uclouvain.be/flelex/



Enter a word

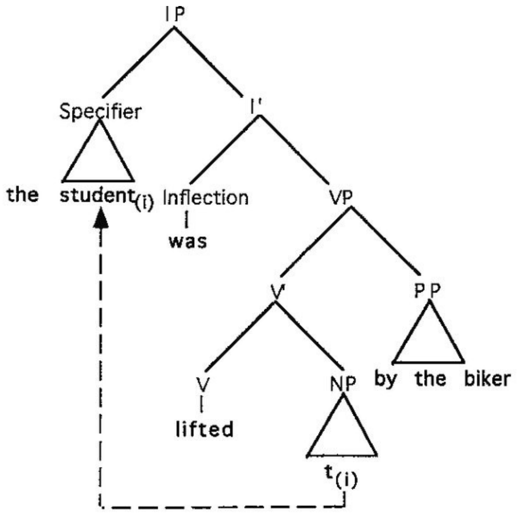
inonder

déborder

Frequencies by CEFR levels for the words inonder* and déborder**.

Analyzing language complexity

Gala, François, Bernhard & Fairon (2014)



How complex is a complex word ?
Complexity (objective) / difficulty (subjective)

Orthography

- Length (phonemes, letters, syllables)
- Orthographical neighbourhood
- Grapheme-phoneme coherence
- Syllable structure

Morphology

- Length (morphemes)
- Frequency of morphemes
- Size of the morphological family

Semantics

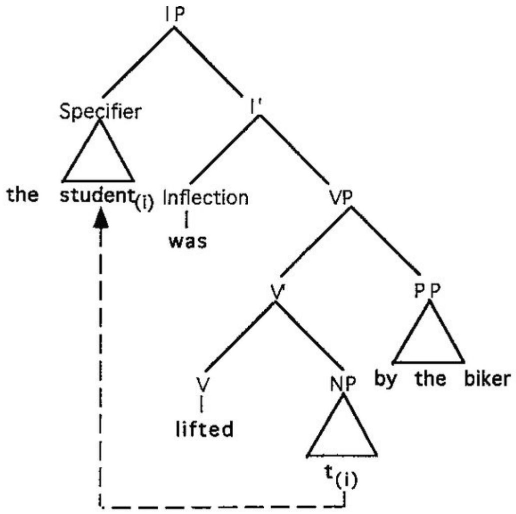
- Polysemy

Example in French (theft … burglary… robbery) :

vol – fuite – attaque – effraction – cambriolage – chapardage – acte de brigandage - maraudage

Analyzing language complexity

Gala, François, Bernhard & Fairon (2014)



How complex is a complex word ?
Complexity (objective) / difficulty (subjective)

Orthography

- Length (phonemes, letters, syllables)
- Orthographical neighbourhood
- Grapheme-phoneme coherence
- Syllable structure

Morphology

- Length (morphemes)
- Frequency of morphemes
- Size of the morphological family

Semantics

- Polysemy

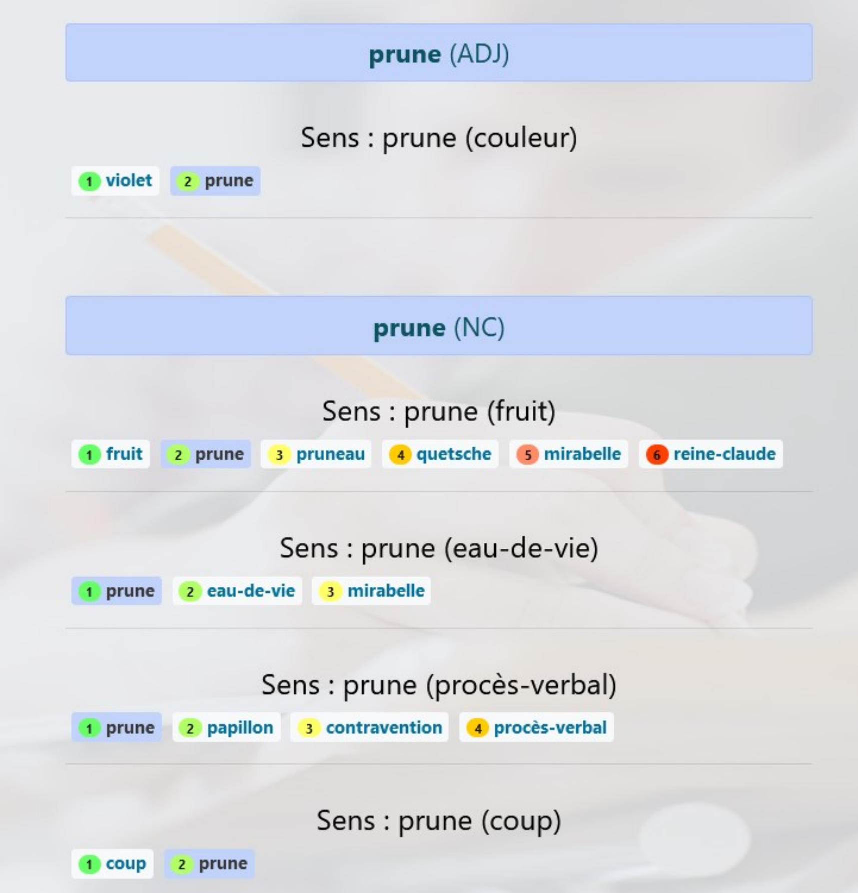
Example in French (theft … burglary… robbery) :

vol – fuite – attaque – effraction – cambriolage – chapardage – acte de brigandage - maraudage

ReSyf: a lexicon with graded synonyms

Gala, François & Fairon (2013), Billami, François & Gala (2018)

https://cental.uclouvain.be/resyf/



prune (ADJ)

Sens : prune (couleur)

1 violet · 2 prune

prune (NC)

Sens : prune (fruit)

1 fruit · 2 prune · 3 pruneau · 4 quetsche · 5 mirabelle · 6 reine-claude

Sens : prune (eau-de-vie)

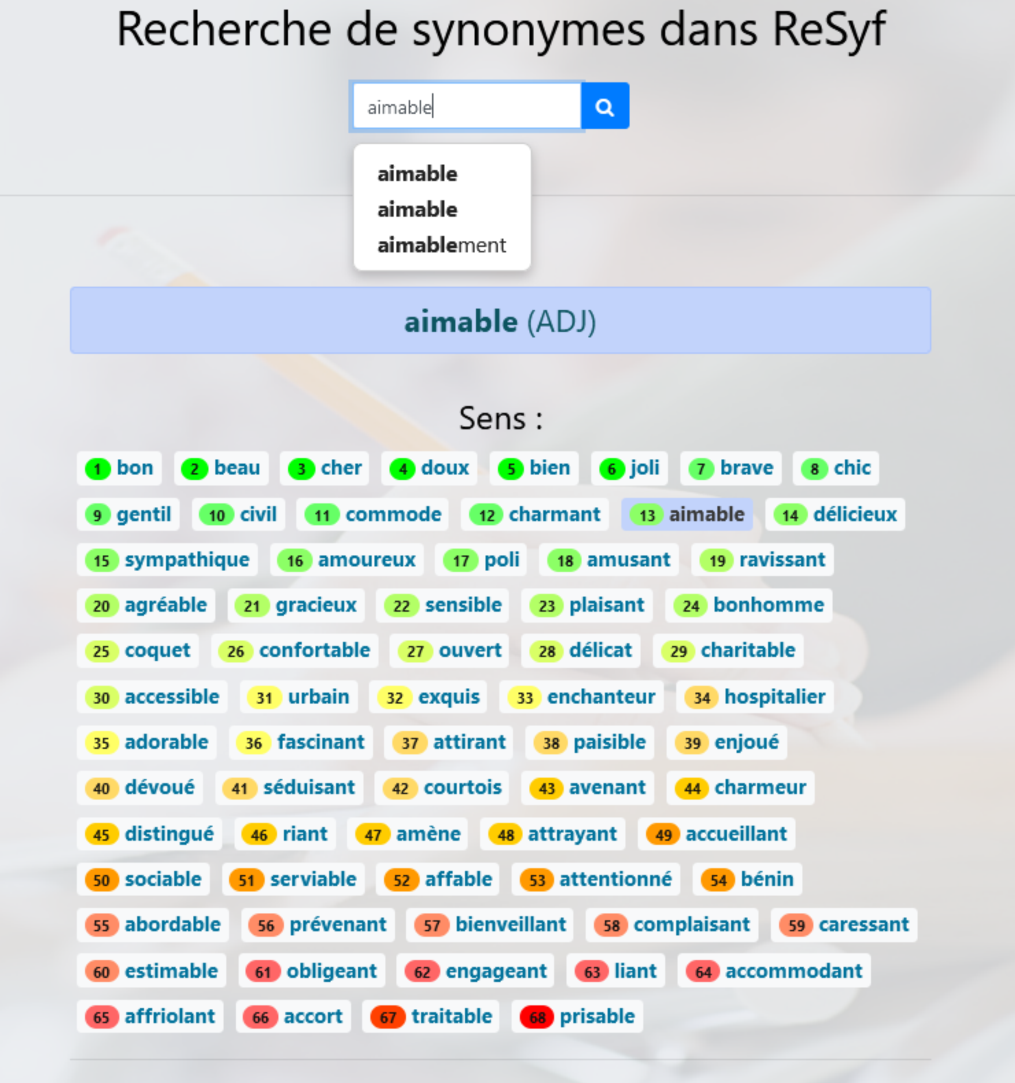
1 prune · 2 eau-de-vie · 3 mirabelle

Sens : prune (procès-verbal)

1 prune · 2 papillon · 3 contravention · 4 procès-verbal

Sens : prune (coup)

1 coup · 2 prune



Recherche de synonymes dans ReSyf

aimable
aimable
aimablement

aimable (ADJ)

Sens :

1 bon · 2 beau · 3 cher · 4 doux · 5 bien · 6 joli · 7 brave · 8 chic
9 gentil · 10 civil · 11 commode · 12 charmant · 13 aimable · 14 délicieux
15 sympathique · 16 amoureux · 17 poli · 18 amusant · 19 ravissant
20 agréable · 21 gracieux · 22 sensible · 23 plaisant · 24 bonhomme
25 coquet · 26 confortable · 27 ouvert · 28 délicat · 29 charitable
30 accessible · 31 urbain · 32 exquis · 33 enchanteur · 34 hospitalier
35 adorable · 36 fascinant · 37 attirant · 38 paisible · 39 enjoué
40 dévoué · 41 séduisant · 42 courtois · 43 avenant · 44 charmeur
45 distingué · 46 riant · 47 amène · 48 attrayant · 49 accueillant
50 sociable · 51 serviable · 52 affable · 53 attentionné · 54 bénin
55 abordable · 56 prévenant · 57 bienveillant · 58 complaisant · 59 caressant
60 estimable · 61 obligeant · 62 engageant · 63 liant · 64 accommodant
65 affriolant · 66 accort · 67 traitable · 68 prisable

Interface: **D. Ricci & B. Delmée** -(2017-2018)
supervised by T. François (CENTAL) & N. Gala (LPL)

Modelling individual difficulties

Larmuseau, Cornelis, Lancieri, Desmet & Depaepe (2020); Tack (2021)

Aims:
- gauging individual overall cognitive load to process (read, understand) a word
- accounting for individual differences between readers

Implicit / indirect measures:
- reading times, eye fixations, physiological data (brain signal, heart rate, skin temperature)

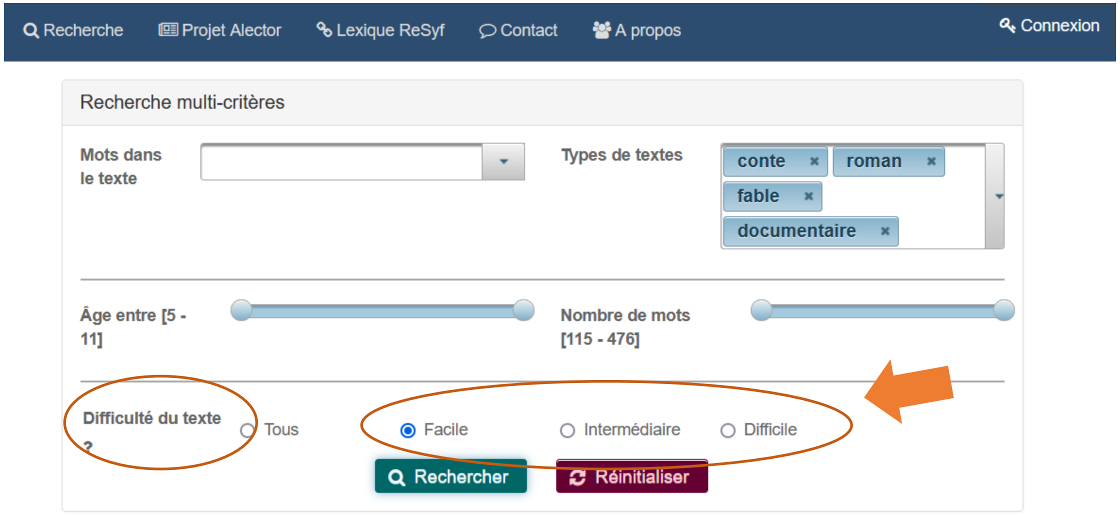
Explicit / direct measures:
- vocalization (read-aloud), verbalization (think-aloud), self-assessment

➡ Building graded resources which

- include individual indirect measures for grading vocabulary
- propose texts / exercises according to personalized needs

ALECTOR: a parallel corpus

Gala, Tack, Javourey-Drevet & François (2020)



Recherche multi-critères

Mots dans le texte

Types de textes: conte, roman, fable, documentaire

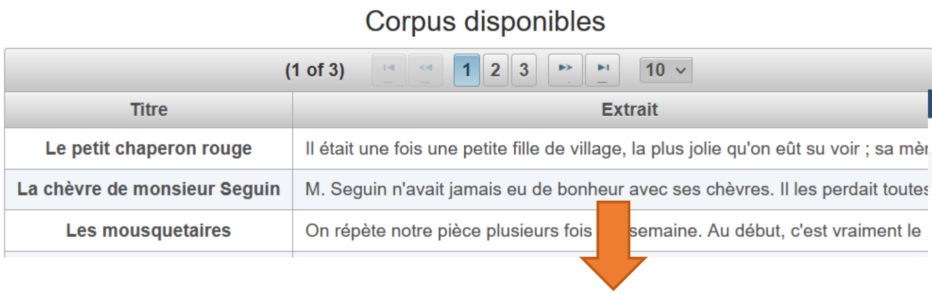
Âge entre [5 - 11]

Nombre de mots [115 - 476]

Difficulté du texte ? ○ Tous ● Facile ○ Intermédiaire ○ Difficile

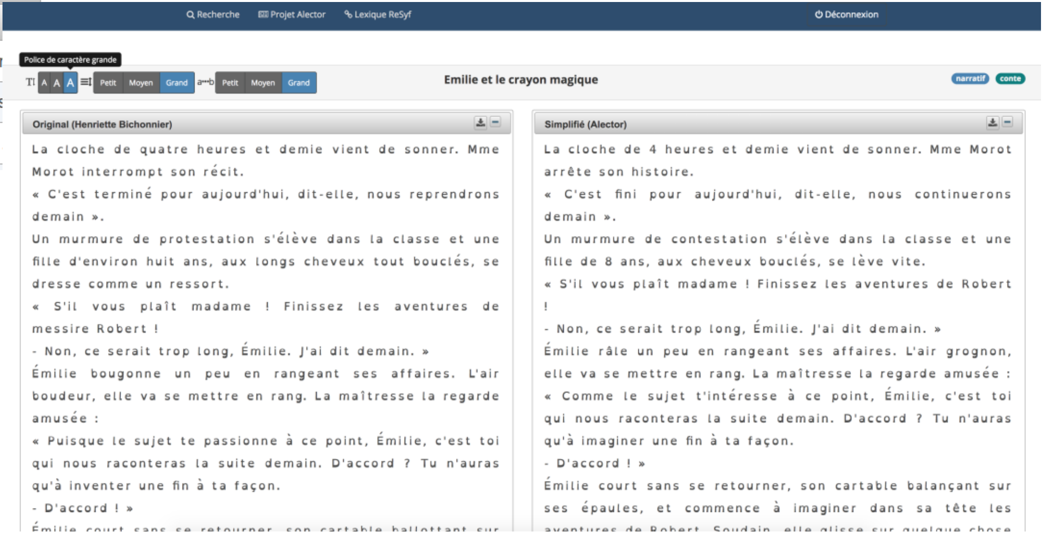
79 original and simplified French texts for reading training online.
Lexical simplifications: Manulex and ReSyf.

Grades according to the difficulty of reading (reading times gathered in 6 schools, 970 children, 2017 to 2019) Javourey-Drevet (2021)



Corpus disponibles

Titre	Extrait
Le petit chaperon rouge	Il était une fois une petite fille de village, la plus jolie qu'on eût su voir ; sa mèr
La chèvre de monsieur Seguin	M. Seguin n'avait jamais eu de bonheur avec ses chèvres. Il les perdait toutes
Les mousquetaires	On répète notre pièce plusieurs fois semaine. Au début, c'est vraiment le



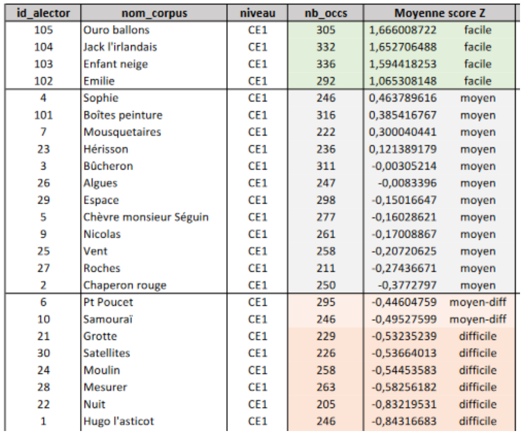
Emilie et le crayon magique

Original (Henriette Bichonnier)

La cloche de quatre heures et demie vient de sonner. Mme Morot interrompt son récit.
« C'est terminé pour aujourd'hui, dit-elle, nous reprendrons demain ».
Un murmure de protestation s'élève dans la classe et une fille d'environ huit ans, aux longs cheveux tout bouclés, se dresse comme un ressort.
« S'il vous plaît madame ! Finissez les aventures de messire Robert !
- Non, ce serait trop long, Émilie. J'ai dit demain. »
Émilie bougonne un peu en rangeant ses affaires. L'air boudeur, elle va se mettre en rang. La maîtresse la regarde amusée :
« Puisque le sujet te passionne à ce point, Émilie, c'est toi qui nous raconteras la suite demain. D'accord ? Tu n'auras qu'à inventer une fin à ta façon.
- D'accord ! »
Émilie court sans se retourner, son cartable ballottant sur

Simplifié (Alector)

La cloche de 4 heures et demie vient de sonner. Mme Morot arrête son histoire.
« C'est fini pour aujourd'hui, dit-elle, nous continuerons demain ».
Un murmure de contestation s'élève dans la classe et une fille de 8 ans, aux cheveux bouclés, se lève vite.
« S'il vous plaît madame ! Finissez les aventures de Robert !
- Non, ce serait trop long, Émilie. J'ai dit demain. »
Émilie râle un peu en rangeant ses affaires. L'air grognon, elle va se mettre en rang. La maîtresse la regarde amusée :
« Comme le sujet t'intéresse à ce point, Émilie, c'est toi qui nous raconteras la suite demain. D'accord ? Tu n'auras qu'à imaginer une fin à ta façon.
- D'accord ! »
Émilie court sans se retourner, son cartable balançant sur ses épaules, et commence à imaginer dans sa tête les aventures de Robert. Soudain, elle glisse sur quelque chose



id_alector	nom_corpus	niveau	nb_occs	Moyenne score Z	
105	Ouro ballons	CE1	305	1,666008722	facile
104	Jack l'irlandais	CE1	332	1,652706488	facile
103	Enfant neige	CE1	336	1,594418253	facile
102	Emilie	CE1	292	1,065308148	facile
4	Sophie	CE1	246	0,463789616	moyen
101	Boîtes peinture	CE1	316	0,385416767	moyen
7	Mousquetaires	CE1	222	0,300040441	moyen
23	Hérisson	CE1	236	0,121389179	moyen
3	Bûcheron	CE1	311	-0,00305214	moyen
26	Algues	CE1	247	-0,0083396	moyen
29	Espace	CE1	298	-0,15016647	moyen
5	Chèvre monsieur Séguin	CE1	277	-0,16028621	moyen
9	Nicolas	CE1	261	-0,17008867	moyen
25	Vent	CE1	258	-0,20720625	moyen
27	Roches	CE1	211	-0,27436671	moyen
2	Chaperon rouge	CE1	250	-0,3772797	moyen
6	Pt Poucet	CE1	295	-0,44604759	moyen-diff
10	Samouraï	CE1	246	-0,49527599	moyen-diff
21	Grotte	CE1	229	-0,53235239	difficile
30	Satellites	CE1	226	-0,53664013	difficile
24	Moulin	CE1	258	-0,54453583	difficile
28	Mesurer	CE1	263	-0,58256182	difficile
22	Nuit	CE1	205	-0,83219531	difficile
1	Hugo l'asticot	CE1	246	-0,84316683	difficile

Interface: **S. Lâm** (2019) supervised by
C. Ramisch (LIS) and N. Gala (LPL)

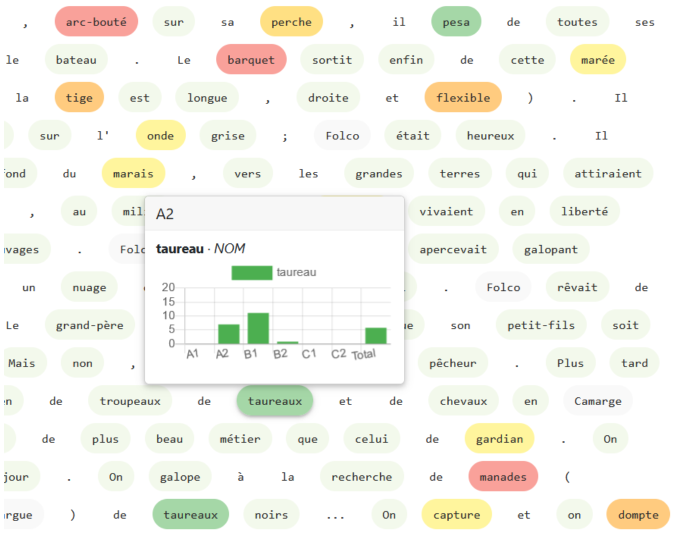
Z score: mean of different readers

http://corpusalector.huma-num.fr/

Using graded resources in the classrooms… or in personalized trainings

For the teacher, in addition to other activities for vocabulary learning:

- Analysing texts before using them in the classrooms, identify complex words for a given grade (choosing or discarding a text)



, arc-bouté sur sa perche , il pesa de toutes ses
le bateau . Le barquet sortit enfin de cette marée
la tige est longue , droite et flexible) . Il
sur l' onde grise ; Folco était heureux . Il
ond du marais , vers les grandes terres qui attiraient
, au mil vivaient en liberté
vages . Folc apercevait galopant
un nuage . Folco rêvait de
Le grand-père e son petit-fils soit
Mais non , pêcheur . Plus tard
n de troupeaux de taureaux et de chevaux en Camarge
de plus beau métier que celui de gardian . On
jour . On galope à la recherche de manades (
rgue) de taureaux noirs ... On capture et on dompte

A2

taureau · *NOM*

During the class:

- Discussing about the knowledge of a word within a grade (whether the word is understood, a synonym can be proposed by the group, the word can be re-used in another context, etc.)

- Studying the morphology, the syntactic properties and the semantics of the word (POS category, cooccurrences, synonyms or thematic links –if possible browse through the semantic links)

In total autonomy:

- Working with texts adapted to the student profile, adaptive learning (Kerr, 2016)

Conclusions and future work

New field with high potential for educational applications.

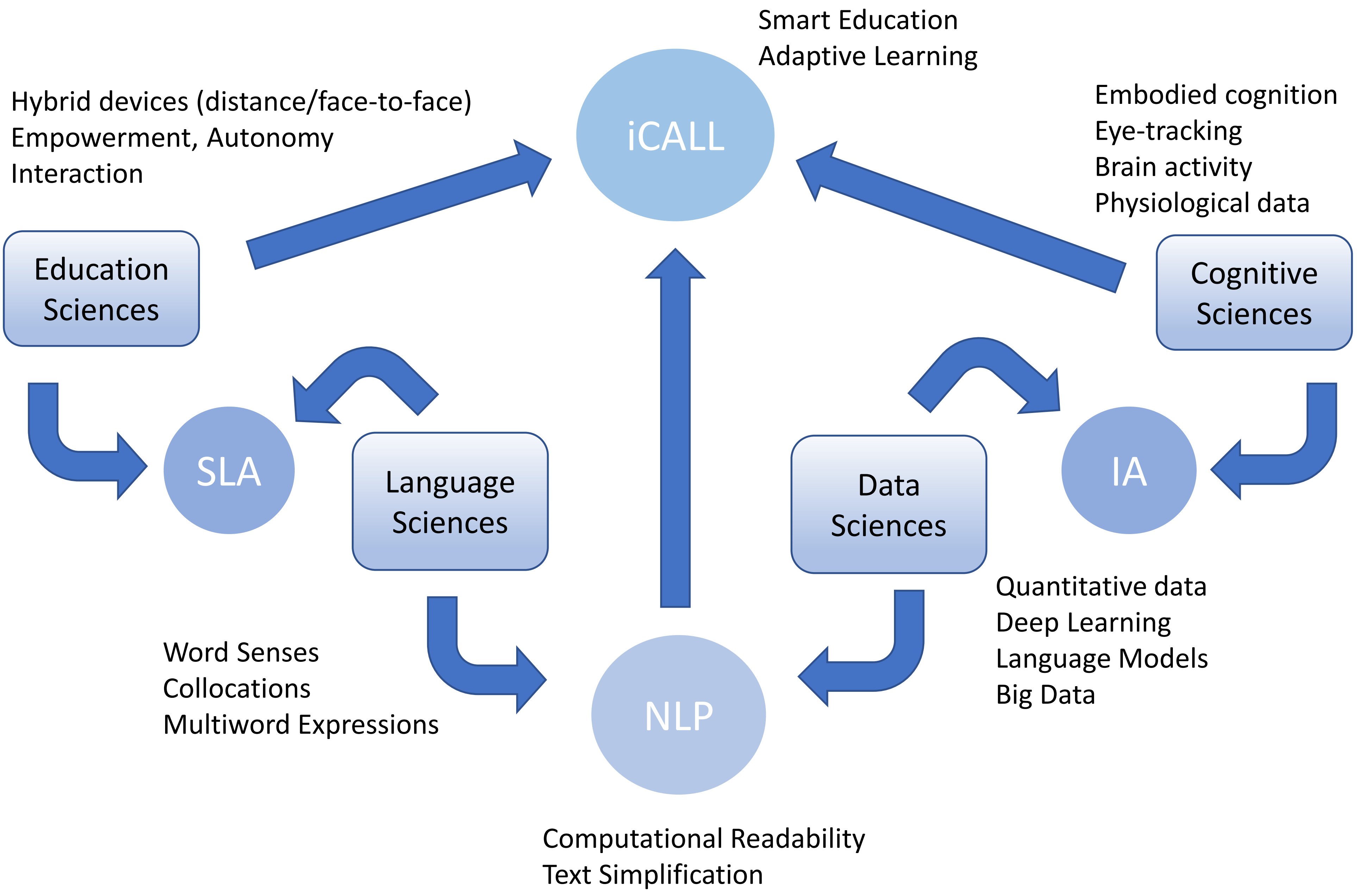
CEFRLex project is a pioneer in graded resources development.

Methodological challenges:

- Model more fine-grained gradings (e.g. for multiword expressions and collocations, for domain-specific texts)
- Include cognitive data in the gradings (to go beyond frequency distributions)
- Train personalized models
- Include graded resources in language learning platforms (track learner’s activities and propose adapted contents)
- Extend to a different languages and varieties (e.g. oral)
- …

Interdisciplinarity: linguistics, education, NLP, cognitive sciences…

Graded resources of tomorrow



Smart Education
Adaptive Learning

iCALL

Hybrid devices (distance/face-to-face)
Empowerment, Autonomy
Interaction

Education Sciences

Embodied cognition
Eye-tracking
Brain activity
Physiological data

Cognitive Sciences

SLA

Language Sciences

Data Sciences

IA

Word Senses
Collocations
Multiword Expressions

NLP

Quantitative data
Deep Learning
Language Models
Big Data

Computational Readability
Text Simplification

© N. Gala

Take home message

Graded resources are new resources with high potential for educational applications.

Beyond frequency distributions, there are important methodological challenges requiring interdisciplinary expertise.

nuria.gala@univ-amu.fr

UMR 7309 • CNRS • AMU
PAROLE ET
LANGAGE

Aix Marseille
université
Socialement engagée

ILCB
Institute of
Language, Communication
and the Brain

Thank you

References

Billami, M. B., François, T. & Gala, N. (2018) **ReSyf** a lexical resource with graded synonyms. Proceedings of 27th International Conference on COmputational LINGuistics *COLING* Sta. Fe, US, 2570 - 2581.

Collins-Thompson, K. & Callan, J. (2005) Predicting reading difficulty with statistical language models. *Journal of the American Society for Information Science and Technology* 56(13), 1148-1463.

Deutsch, T., Jasbi, M., & Shieber, S. (2020). Linguistic features for readability assessment. *arXiv preprint arXiv:2006.00377*.

Dubay, W. (2004) The Principles of Readability. *Impact Information*.

Figueras, N. (2012) The impact of the CEFR. *ELT Journal* 66(4). Special issue October 2012. Cambridge University Press.

François, T. (2009) Combining a statistical language model with logistic regression to predict the lexical and syntactic difficulty of texts for FFL. In *Proceedings of the 12th Conference of the EACL: student research workshop*, 19 – 27.

François, T., Gala, N., Watrin, P. & Fairon, C. (2014) **FLELex**: a graded lexical resource for French foreign learners. *International conference on Language Resources and Evaluation (LREC 2014)*, May 2014, Reykjavik, Iceland.

Fulcher, G. (2010) The reification of the Common European Framework of Reference (CEFR) and effect-driven testing. *Advances in Research on Language Acquisition and Teaching*. Greek Association of Applied Linguistics. Thessaloniki, 15-26.

Gala, N., François, T., Bernhard, D. & Fairon, C. (2014) Un modèle pour prédire la complexité lexicale et graduer les mots. In proceedings of the conference *Traitement Automatique des Langues Naturelles TALN'2014*, Marseille, France, 91-102.

Gala, N., François, T. & Fairon, C. (2013) Towards a French lexicon with difficulty measures: NLP helping to bridge the gap between traditional dictionaries and specialized lexicons. *Proceedings from eLex – Electronic Lexicography*, Tallin Estonia, 132 – 151.

Gala, N., François, T., Javourey-Drevet, L. et Ziegler, J.-C. (2018) La simplification de textes, une aide à l'apprentissage de la lecture. Dans *Langue Française « Lire – écrire : Des savoirs scientifiques aux savoirs pratiques »*, 199 (3). Éds. Liliane Sprenger-Charolles et Alain Desrochers. Armand Colin, 123-131.

Gala, N., Tack, Javourey-Drevet, L., François, T. & Ziegler, J. C. (2020) **Alector**: A Parallel Corpus of Simplified French Texts with Alignments of Misreadings by Poor and Dyslexic Readers. *Language Resources and Evaluation for Language Technologies (LREC)*, Marseille, France.

Javourey-Drevet, L. (2021) La simplification de textes comme outil pour améliorer la fluidité et la compréhension de lecture chez les enfants à l'école primaire. Une étude en longitudinal avec des textes littéraires et scientifiques chez des enfants entre 7 et 9 ans. Thèse de doctorat, école doctorale Cognition, Langage, Education. Aix Marseille Université.

Kerr, P. (2016) Adaptive learning. Technology for the language teacher.

Larmuseau, C., Cornelis, J, Lancieri, J., Desmet, P. & Depaepe, F. (2020) Multimodal learning analytics to investigate cognitive load during online problem solving. *British Journal of Educational Technology* vol 51 Issue 5, 1548-1562.

Lété, B., Sprenger-Charolles, L. & Colé, P. (2004) **Manulex**: A grade-level lexical database from French elementary-school readers. *Behavior Research Methods, Instruments & Computers*, 36, 156-166.

Martinc, M., Pollak, S. & Robnik-Šikonja, M. (2021) Supervised and unsupervised neural approaches to text readability. *Computational Linguisticc Journal* volume 47 Issue 1 available online at https://doi.org/10.1162/coli_a_a00398

Tack, A. (2021) *Mark my words. On the automated prediction of lexical difficulty for foreign language readers*. PhD dissertation. KU Leuven & Université catholique de Louvain.

Tack, A., François, T. & Fairon, C. (2016) Evaluating lexical simplification and vocabulary knowledge for learners of French: possibilities of using the **FLELex** resource. In *Proceedings of the 10th International Conference on Language and Resources Evaluation (LREC 2016), Portorož, Slovenia: European Language Resources Association*, 230–236.