

Towards the effective integration of NLP in CALL

Piet Desmet, Dirk De Hertog

0. Overview

1. Introduction
2. Technological Readiness of NLP
 1. Native Language
 2. Learner Language
 3. Measure linguistic complexity and proficiency
3. Applications: 7 use cases
4. Future Possibilities

1. Introduction

1.1 Current Situation

Most of the CALL environments do not use NLP

-> use of NLP in classrooms is hardly mainstream practice

“The development of systems using NLP technology is not on the agenda of most CALL experts, and interdisciplinary research projects integrating computational linguists and foreign language teachers remain very rare” (Amaral & Meurers 2011: 6).

cf. also report by the Dutch Language Union:

Onderzoek taal- en spraaktechnologie en onderwijs (Van den Heuvel, T'Sas & Verberne 2012)

1.2 Concerns

Why not? Skepticism -> technological and pedagogical concerns

(1) Native language :

Can NLP account for the full **complexity** of natural human languages?

“I am pessimistic about the possibility of ICALL” (Nyns 1989: 46)

But is this still the case?

(2) Learner language

a) (Erroneous) application of NLP developed for native language

-> But: can NLP be developed to work for learner language?

b) Current binary scoring of errors and automatic correction

-> How do we handle the notion of error and correction?

(3) Tutor: “One size fits all” due to technological restrictions

But: can NLP help us create language and learner models

1.2 Reasonable Targets

Yet there is a high reward if NLP were successfully integrated

(1) Automatization of a number of tedious tasks such as resource creation

Exploit and enrich Native language

(2) Automatic Error Correction and feedback generation of open items

Analyse learner language

(3) Personalised teaching

Linguistic Complexity and Intelligent tutors

2. Technological readiness

Technological readiness (written language) depends on the field of application

(1) Native language

NLP is (partially) usable, requires integration

(2) Learner language

NLP requires development
should rely on insights of L2-acquisition and FLTL

(3) Model complexity (-> Intelligent tutors)

Requires an improvement in (2)
Work on integration of (1) and (2)

2.1 Native Language

NLP has evolved and improved dramatically the last 20 years
Some technologies are ready to use

Better results due to successful paradigmatic shift

From Rule-based NLP

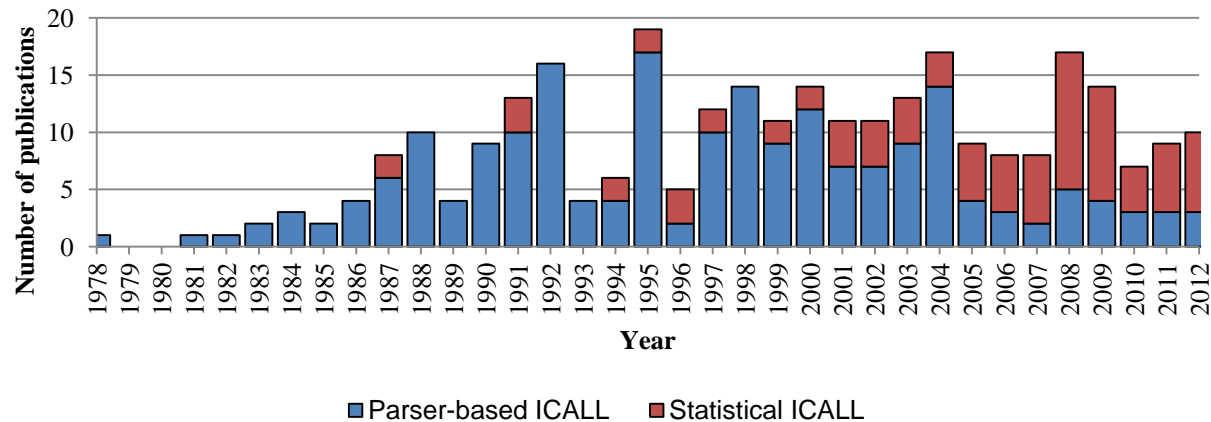
Manual engineering of grammar rules

To Statistical NLP

Learn from examples

use supervised machine learning techniques
to solve certain problems

Looking at the number of publications shown in this figure, parser-based CALL has dominated research up until 2005, but there seems to have been a shift in research interest and statistical error detection systems have since then taken the upper hand.



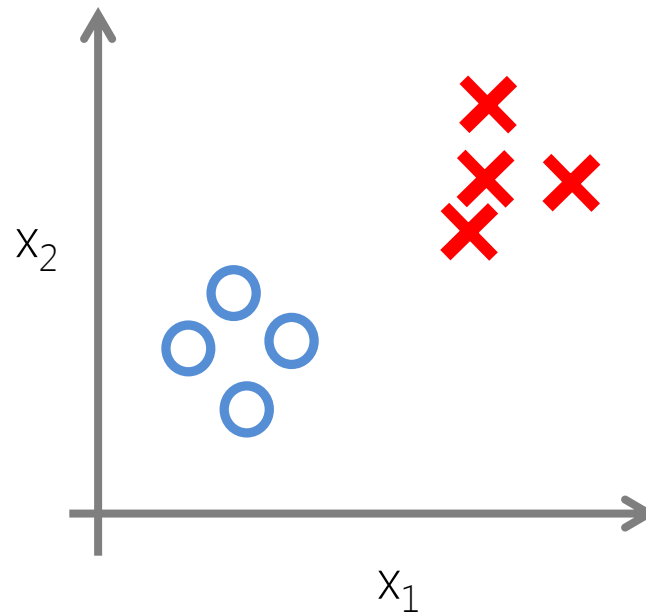
© PhD Ruben Lagatie

2.1.1 Statistical NLP

Supervised Learning

Learning from examples how to label them

Predict on the basis of this expert **model** how to label new examples



Examples of supervised learning

Sharapova beats Ivanovic to win Australian Open **Australian Open**
A year after being on the wrong end of one of Russian didn't **Open**
seven matches at **Great victory!** including wins the most-lopsided losses in a
Grand Slam final, Sharapova wrapped up her third major title with a 7-5, 6-3
victory over fourth-seeded Ana Ivanovic on Saturday. The 20-year-old over
~~three of the top four ranked players, erasing 12 months worth of painful~~
memories in the wake of her 6-1, 6-2 loss to Serena Williams last year. After
Ivanovic sprayed a forehand wide on match point, Sharapova dropped to her
knees and appeared to be **Sports** as she waved and blew kisses to
the crowd. Then she dropped her racket in her chair before heading to shake
hands and exchange high-fives with her father and supporters.

Date: **Declining job offer**
From: **Declining job offer**
To: **Declining job offer**
Subject: Re: Op zoek naar een medewerker
Hoi Sien,

4: **AMID** NIH cDNA clone, Links
Official Symbol: AMID and Name: apoptosis-inducing
factor (AIF)-like mitochondrion-associated inducer of death
[Homo sapiens]
Other Aliases: PRG3, RP11-367H5.2
Other Designations: 5430437E11Rik; apoptosis-inducing

Ik vrees dat **Non-spam** t moeten zoeken... ik kan het niet doen
en ik heb wa **Non-spam** enden en ex-collega's enzo maar helaas..

veel succes, *L'é tro bel cet voitur Voici tt ce ki me passione*
maarten *ds ma petite vi!!!é **Positive** I dotre*
*truk!!!Avou de Dc **sentiment***

ITEC 8-1-2009

6

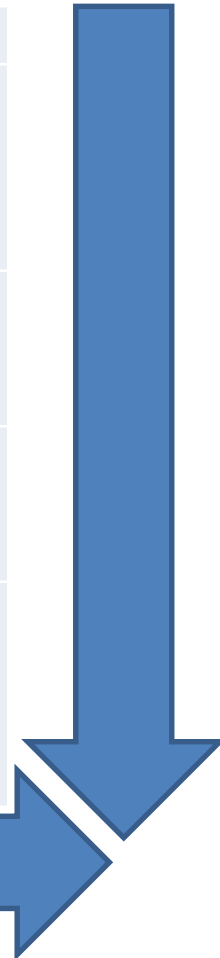
© Sien Moens (2009)

2.1.2 Native Language

	Form	Meaning
Word	morphological analysis Lemmatisation POS-tagging	
Phrase	Chunking	Named Entity Recognition
Sentence	Parsing	Sentiment Analysis Semantic Frame Labeling
Text	Language Identification Authorship attribution	Topic Detection Co-reference resolution Discourse analysis

Growing Complexity

	Form	Meaning
Word	morphological analysis Lemmatisation POS-tagging	
Phrase	Chunking	Named Entity Recognition
Sentence	Parsing	Sentiment Analysis Semantic Frame Labeling
Text	Language Identification Authorship attribution	Topic Detection Co-reference resolution Discourse analysis



Ongoing developments and improvements

- Combination of statistical and rule-based approaches
- Technically: the use of semantic representations is becoming mainstay and used in more complex NLP-tasks => a broader generalisation possible
- Discursive component (semantic text) is gaining importance
- *Unsupervised methods stand at the sideline in applications but are ready to make their entry (cf. semantic vector spaces – infra)*

2.2. Learner Language

Using predictive models trained on native language
will only work on advanced learners' output

Solution:

NLP: For beginning learners rely on a hypothesis
of what it is they meant to communicate
= Disambiguation of what the learner is trying to say

Goal is to detect errors: the difference between
hypothesis and produced utterance
= What is wrong?

Ambiguity of target hypothesis

e.g. “He hear you tell the history.”

-> He heard you speaking about the past.

- Many different hypotheses could be generated based on the provided output

I hear you tell your history

He heard you **told** a story

He hears you **telling** a story

...

Where is the mistake?

Multiple and simultaneous mistakes take place at all linguistic levels

Constraining open-ended output

Fully open-ended learner output hard to manage
-> constrain learner output (reduce possible answers)
by reducing the search space of the processing tools

Important role for design and use of CALL tools
The tools help to disambiguate the hypothesis

Task design should be structured towards specific type of output

Constraining learner language

Constraints...	on form	on meaning	
Explicit	<ul style="list-style-type: none"> multiple choice gap-filling list of words to reuse/reorder 	<ul style="list-style-type: none"> set content words translation reading comprehension questions 	<p>} E-Tutor (Heift, 2010)</p> <p>→ Tagarela (Amaral <i>et al</i>, 2011) CoSeC (Hahn & Meurers, 2012)</p>

INSTRUCTIONS
Build a sentence with the following words:

TASK
ich / gestern / kein- / Zeit / haben (simple past).

A O U B a o u

ih hatte gestern keine Zeit.

CHECK SOLVE NEXT

FEEDBACK
ih hatte gestern keine Zeit
There is a spelling mistake with ih. Suggestions: ich, ihm, ihn, or ihr

DICTIONARY

Módulos: 1 2 3 4 5 Atividades: 1 2

Leitura

Instrução

Leia o texto e responda às questões usando frases completas e o vocabulário apresentado no texto.

Quem é você?

Eu me chamo Patrícia Mattos, tenho quinze anos e moro em São Paulo. Eu estudo em uma escola pública e tenho muitos amigos.
Eu moro com minha mãe. Seu nome é Marta. Ela tem quarenta anos e é cozinheira em um restaurante de luxo.
Eu tenho um irmão. O nome dele é Claudio. Ele mora nos Estados Unidos e é músico. Ele toca Jazz e Blues. Ele é um excelente guitarrista.

Questões: 1 2 3 4 5 6 7 8
Próxima Questão (3)

Questão 2
Quantos anos ela tem?

Ela é quinze ano.

A A A A A E E I I O O O U U C C

Enviar

Análise:
Input: Ela é quinze ano.
I am not expecting the verb ser for this answer. Try using ter instead.
To see a possible answer, click [here](#).

Report Errors & Suggestions

Tagarela

Constraints...	on form	on meaning	
Explicit	<ul style="list-style-type: none"> • multiple choice • gap-filling • list of words to reuse/reorder 	<ul style="list-style-type: none"> • set content words • translation • reading comprehension questions 	E-Tutor (Heift, 2010) Tagarela (Amaral et al, 2011) CoSeC (Hahn & Meurers, 2012)
Implicit → <i>still too much variation to handle</i>		<ul style="list-style-type: none"> • task & context • dialogic interaction • open-ended conversations 	SPELL (Morton et al, 2012) (prototype) Reactive chatbots (e.g. Tutor Mike)



NLP4CALL



2016



YOU: Hi Mike!
 Mike: Hi! What can I do for you?
 YOU: Can you tell me something cool?
 Mike: What do you mean by "cool"?
 YOU: Something hype
 Mike: Where will we get our energy when we run out of oil?

Don't try to change the subject, Mike!

Enter

2.3. Model Linguistic Complexity and Learner Proficiency

Tutor combines a **language model** to measure linguistic complexity
&
a learner model to measure proficiency

Determine **complexity** at which linguistic level?

Lexical level: frequency in general corpora

Syntactic level: sentence length, parse structure

Measure complexity in what type of language?

- 1) in native language
- 2) in learner language : requires working NLP!

3. Applications

Focus on written language input & output

Native Language : Enrich and Exploit

- 1. Resource generator*
- 2. Reading companion*
- 3. Exercise and test generator*

Learner Language : Finding Errors

- 4. Error detector, feedback generator and automatic scoring tool*
- 5. Writing aid*

Model Learner Behaviour : measure Complexity

- 6. Input provider*
- 7. Adaptive item sequencer*

-> Conceptual outline + applications from academic R&D

3.1 Resource Generator

What? Providing learners with easy access to authentic language use

How? Exploiting authentic (annotated) corpora

=> (Intermediate or) advanced learners become linguistic investigators
Requires well developed linguistic awareness

Used methods :

- Concordancing on (bilingual) corpora
- Search on terms and filter on specific categories (cfr. annotation)

3.1 Resource Generator

Projects:

- Concordancing on bilingual corpora
& more advanced search engines

REBECA-project (ITEC)

= Ressources électroniques bilingues extraites de corpus alignés

DPC-project (ITEC) <http://www.kuleuven-kulak.be/DPC/>

= Dutch Parallel Corpus

- Corpus-enriched learner dictionaries & grammars

BLF-project (Serge Verlinde) <http://ilt.kuleuven.be/blf/>

= Base Lexicale du Français

REBECA-project

5. Gezondheid en leefmilieu in België

Index

Nauwelijks een eeuw geleden leden duizenden mensen in ons land nog aan ziekten veroorzaakt door de slechte kwaliteit van het leef- en werkmilieu.

Die tijd is intussen voorbij. Tal van ziekten zijn onder controle. Maar vandaag heeft de **overheid** totaal andere gezondheidsproblemen, die zijn veroorzaakt door vervuiling door de industrie, het verkeer en door de menselijke activiteit in het algemeen.

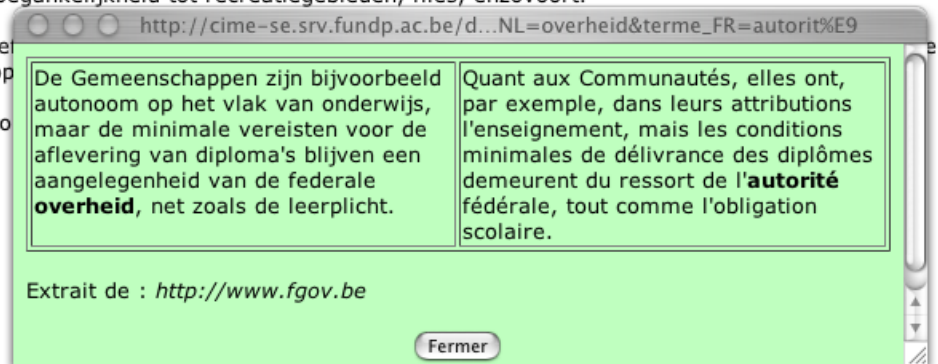
De opkomst van nieuwe chemische producten, nieuwe productieprocessen en technologieën en de vermenging van allerlei pollutiebronnen hebben hun weerslag op het klimaat, de kwaliteit van de lucht en de bodem, de biodiversiteit en de voedselketen. Vaak is het effect ervan pas na enkele jaren of zelfs decennia later zichtbaar.

Bovendien is de verstedelijking sterk toegenomen. In 2000 leefde ongeveer 80% van de bevolking in stedelijke gebieden. Dat heeft gevolgen. In heel wat steden duiken hoe langer hoe meer stressverschijnselen op die te maken hebben met het leefmilieu: ozonpieken, zware luchtvervuiling, toenemend lawaai, stijgende afvalproductie, moeilijkere toegankelijkheid tot recreatiegebieden, files, enzovoort.

En dan is er nog de maatschappelijke ongelijkheid. Die heeft ook factoren. Die factoren hebben een rechtstreekse invloed op de gezondheid.

De strijd tegen ziekte en vervuiling kan dus maar succesvol zijn als we de welzijn van de hele bevolking verbeteren.

bron: <http://www.belgium.be> - 19.09.2003



overheid (nom, de, overheden)	
... nder toezicht van alle hogere overheden , in het kader van de fe nales en étant subordonnées à toutes les autorités supérieures.
... angelegenheid van de federale overheid , net zoals de leerplicht mes demeurent du ressort de l' autorité fédérale, tout comme l'o ...
De overheid heeft een nieuw reglement uitgevaardigd.	Les autorités ont promulgué un nouveau règlement.
... erd in welke administratie en/of overheid daarbij betrokken is.	... r quelle administration et/ou pouvoir public est impliqué dans ...
... erd in welke administratie en/of overheid daarbij betrokken is.	... s intéressés de savoir quelle administration et/ou pouvoir publ ...
... hillende administraties en/of overheden die hierbij betrokken z r quelle administration et/ou pouvoir public est impliqué dans ...
... hillende administraties en/of overheden die hierbij betrokken z s intéressés de savoir quelle administration et/ou pouvoir publ ...
Administraties en overheden zullen elkaars gegevens zoveel mog ...	Les administrations et les autorités doivent partager et utili ...


DPC-project

tweetalig zoeken Frans - Nederlands

<p>woord1</p> <p>en <input type="text"/> op <input checked="" type="radio"/> woordvorm <input type="radio"/> lemma (* is jokerteken, ? is één karakter jokerteken)</p> <p>woordsoort</p> <p>Maak uw keuze ▼</p> <p>positie</p> <p>de positie is niet belangrijk ▼</p>	<p>alles wissen</p>	<p>woord1</p> <p><input type="text"/> op <input checked="" type="radio"/> woordvorm <input type="radio"/> lemma (* is jokerteken, ? is één karakter jokerteken)</p> <p>woordsoort</p> <p>Maak uw keuze ▼</p> <p>positie</p> <p>de positie is niet belangrijk ▼</p>
<p><input checked="" type="radio"/> EN <input type="radio"/> NIET</p>		<p><input checked="" type="radio"/> EN <input type="radio"/> NIET</p>
<p>woord2</p> <p>réalité <input type="text"/> op <input checked="" type="radio"/> woordvorm <input type="radio"/> lemma (* is jokerteken, ? is één karakter jokerteken)</p> <p>woordsoort</p> <p>Maak uw keuze ▼</p> <p>positie</p> <p>de positie is niet belangrijk ▼</p>		<p>woord2</p> <p><input type="text"/> by <input checked="" type="radio"/> woordvorm <input type="radio"/> lemma (* is jokerteken, ? is één karakter jokerteken)</p> <p>woordsoort</p> <p>Maak uw keuze ▼</p> <p>positie</p> <p>de positie is niet belangrijk ▼</p>
<p>maximum aantal ▼ woorden tussen de zoekwoorden: 0</p>		<p>maximum aantal ▼ woorden tussen de zoekwoorden: <input type="text"/></p>
	<p><input type="checkbox"/> random hits</p>	
	<p>toon 100 ▼ resultaten</p>	

Example: "en réalité" French - Dutch

Aantal resultaten: 114 waarvan er 114 getoond worden.

 Exporteer deze resultaten naar Excel.

<p>i C En réalité, il existe deux types de bokashi.</p>	<p>i C Er zijn eigenlijk twee soorten bokashi.</p>
<p>i C Il s'agit en réalité d'une entourloupe : le magazine contient des articles emphatiques mais creux, dans lesquels Boorman change tout simplement le nom de l'entreprise et du produit dans chaque numéro.</p>	<p>i C In werkelijkheid is het een zeepbel: het tijdschrift bevat hoogdravende, maar nietszeggende artikels waarin Boorman voor elk nummer gewoon de namen van bedrijf en product vervangt.</p>
<p>i C Donc, pour nous, l'essentiel était qu'il parle, tant il racontait des choses énormes, du plus amusant au point de vue politique au préjudiciable, par exemple quand, évoquant "les poitrines de poulet et les dos bossus", il parlait en réalité des anciens combattants...</p>	<p>i C Voor ons kwam het er dus op aan dat hij veel aan het woord kwam, want hij vertelde veel enormiteiten, soms amusant maar ook vaak schadelijk voor de politiek. Essentieel is zijn uitspraak over de kiekenborsten en de bochels als hij het in feite had over oud-strijders.</p>
<p>i C En réalité, la défense de la ville est constituée dans un premier temps de remparts en terre. Les murs en pierre seront dressés dans ces remparts, sur des arcades.</p>	<p>i C In werkelijkheid werden eerst aarden wallen opgetrokken en werd de stenen muur daarin gefundeerd, op bogen.</p>
<p>i C En réalité, au cours de la nuit, T'Serclaes escalade sournoisement le mur du jardin de son propre palais.</p>	<p>i C In werkelijkheid is T'Serclaes in het geniep en bij nacht en ontij over de tuinmuur van zijn eigen paleis geklommen.</p>
<p>i C « Cette inscription est gratuite : on donne ainsi l'impression que le bar n'est pas accessible à tout le monde, alors qu'en réalité ce n'est pas le cas. »</p>	<p>i C "Die inschrijving is gratis: je geeft zo het gevoel dat dit niet voor iedereen bereikbaar is, terwijl dit wel zo is."</p>
<p>i C En réalité, le bruissement perceptible est dû à la combinaison de diverses fréquences renforcées par la forme du coquillage.</p>	<p>i C In werkelijkheid horen we in zo'n schelp niets anders dan een ruistoon, die een mengeling is van allerlei frequenties die door de vorm van de schelp versterkt worden.</p>



www.franel.eu

Extension: video search

Exploitation of XML-based captions (transcriptions and translations) to render video selection more versatile and attractive:

lemmas are time tags indexes which form reference points to video selections



Lingu@tic Subtitles: FR005T200

FR005T200

V01	La cuisine du Nord ne se prend pas le chou sauf pour les recettes, comme cette entrée festive, l'aumônière de Saint-Jacques. Le réveillon aux saveurs d'ici, c'est une protestation : la grande cuisine du Nord veut exister et se faire reconnaître.
S01	« On n'est pas assez reconnu, je trouve, dans le nord, on va toujours dans le sud, comment dire, chercher les talents. On a des talents dans le sud alors que c'est pas vrai, dans le Nord on a des talents. On ne va pas assez fouiller dans le nord et on s'arrête toujours dans le bas de la France, quoi ».
V01	Régionalisme au fourneau, peut-être un peu, curiosité aussi, celle de composer et de proposer de nouvelles saveurs par exemple.
S01	« Je pratique les Saints-Jacques dans mon restaurant, d'une façon différente mais je prends les Saints-Jacques qui viennent du Nord, ça fait beaucoup de maraîchers et on en parle pratiquement jamais ».
S02	« On est quand même riche en gibier, en produits de la mer, en viandes, beaucoup de légumes finalement, le Nord propose beaucoup de maraîchers et on en parle pratiquement jamais ».
V01	Et justement ce sont les légumes qui égayent cette raie aux choux verts, elle ne dépareillera pas la suite du réveillon. Ce réveillon son livre de recettes, les chefs de la métropole réunis dans le club des tables gourmandes ont pris la plume et dans le titre : Recettes en Nord.

FR005T200_trad

V01	De kookkunst van Noord-Frankrijk maakt zich geen zorgen, behalve wat haar recepten betreft, zoals dit feestelijk voo gordelbeursje'. Meer dan een kennismaking met een oudejaarsfeestmaal met smaken van hier, is het een protestactie: wil bestaan en erkend worden.
-----	---

à 00:00:09.6800

a 00:00:24.8900|00:00:28.8400|00:00:59.9000|00:01:12.7000

alors 00:00:24.8900

assez 00:00:18.0300|00:00:31.9000

au 00:00:38.3000

aussi 00:00:38.3000

aux 00:00:09.6800|00:01:05.8100

bas 00:00:35.6500

beaucoup 00:00:24.8900|00:00:55.0000|00:00:59.9000

bonnes 00:00:31.9000

calais 00:00:51.1200

ce 00:01:05.8100

celle 00:00:42.0800

c'est 00:00:09.6800|00:00:24.8900

cette 00:00:06.6400|00:01:05.8100|00:01:12.7000

chefs 00:01:16.0000

chercher 00:00:18.0300|00:00:31.9000

chou 00:00:04.0000

choux 00:01:05.8100

club 00:01:16.0000

comme 00:00:06.6400

comment 00:00:18.0300

composer 00:00:42.0800

cuisine 00:00:04.0000|00:00:14.1800|00:01:12.7000

curiosité 00:00:38.3000

dans 00:00:18.0300|00:00:18.0300|00:00:24.8900|00:00:28.8400|

00:00:31.9000|00:00:35.6500|00:00:47.2800|00:01:16.0000|00:01:19.9500

Viewing video extracts based on lexical search in transcription and/or translation (challenge: search on semantically related words)



3.2 Reading Companion

What?

Helping learners understand foreign-language input
By providing a different and enhanced perspective

How?

Enrich language input through annotation layers

- add information
- provide alternative ways of exploration
- serve as a bridge to connect more information

Formal linguistic information (Lemma) > Connect to dictionary information
Semantic information (NER) > Connect to encyclopedic information

Projects:

GLOSSER-project (John Nerbonne)

<http://www.let.rug.nl/glosser/Glosser>

lemmatization of the inflected forms
dictionary entry (cf. Van Dale)
examples of the word from corpora

iRead+ project (ITEC)

Lemmatization & POS-tagging
Named entity recognition (persons, organisations, locations)

GLOSSER-project

Home	Help	Morphology	Dictionary	Examples
Text index	On	Off	On	Off

DANS DEUX DOCUMENTS RESSUSCITÉS, L'ANTISÉMITISME DU III^e REICH.
L'autre face du génocide.

Mille cent trente-six pages pour raconter le sort des centaines de milliers de martyrs du génocide commis par l'occupant nazi en Union soviétique: c'est le fameux Livre noir, dont le texte intégral vient, pour la première fois, de paraître en français (1). Commandé en 1942 par le Comité antifasciste juif d'URSS aux journalistes; écrivains Ilya Ehrenbourg et Vassili Grossman, il rassemble les témoignages de correspondants de guerre soviétiques sur les massacres antisémites commis par les troupes allemandes et leurs collaborateurs. Mais l'oeuvre ne paraît pas. Car le Comité juif antifasciste, créé pendant la seconde guerre mondiale pour mobiliser les juifs du monde en faveur de l'Union soviétique, est victime, dès 1946, de la répression stalinienne contre le « cosmopolitisme », en premier lieu le « nationalisme juif » (lire, ci-dessous, l'article de Laurent Rucker). Ses dirigeants sont « jugés » en 1952, quand ils n'ont pas, comme le grand acteur Solomon Mikhoels, été assassinés avant. Saisi avant impression, Le Livre noir sera partiellement édité en Israël dans les années 70. Sa parution dans la traduction du texte russe complet dont Ilya Ehrenbourg avait corrigé les épreuves donne au lecteur français la mesure du rôle décisif joué par l'extermination des juifs soviétiques dans la mise en oeuvre de la « solution finale » planifiée lors de la conférence de Wannsee, en janvier 1942. C'est bien sur le front de l'Est que la SS a expérimenté l'[anéantissement](#) des juifs d'Europe. On l'oublie trop souvent.

par PHILIPPE BURREIN.

(1) Ilya Ehrenbourg et Vassili Grossman, *Le Livre noir*, col 1. « Hebraïca », Actes Sud-Solin, Arles-Paris, 1995, 1 136 pages, 285 F.

Le Monde diplomatique, décembre 1995, Page 7

Analysis:
anéantissement+Masc+SG+Noun

Lemmatization of inflected forms

anéantissement [aneãt ismã] 0.1 vernietiging => verdelging, verval 0.2 neerslachtigheid => verslagenheid, volkomen uitputting

Dictionary entry (Van Dale)

... autre que le capitaine Nicholl! Ils se battent ce matin au bois de Skersnav! J'ai tout appris de la bouche de Barbicane! S'il est tué, c'est l'**anéantissement** de nos projets! Il faut donc empêcher ce duel! Or, un seul homme au monde peut avoir assez d'empire sur Barbicane pour l'arrêter, et cet homme ...

'De la terre à la lune' - Jules Verne, Hetzel, Paris, 1865.

Examples from corpora



HOME
ABOUT
API
CONTACT US



Standard de Liège

Une odeur de viande grillée flotte dans l'air. On croit entendre le refrain de [Johnny Hallyday](#). Des grappes de jeunes se taquinent gentiment, écharpe rouge et blanche autour du cou, vareuse floquée [Van Damme](#). Dix mètres plus loin, trois quadras urinent de concert, le gobelet rempli de bière entre les dents. Un couple s'enlace, à l'ombre du stade rougeoyant de [Sclessin](#). De l'autre côté de la Meuse, le haut-fourneau d'[Ougrée](#) se dresse dans un ciel entre chien

[Gilles De Bilde](#), [Rodrigo Beenkens](#), [Pierre-Yves Jeholet](#), les [Wathelet père et fils](#)... Tous sont passés par la loge du groupe [Circus](#). Pour faire monter la température, [Philippe Favray](#) a sa spécialité : les « orgasmes », un cocktail de [Baileys](#), [Jet 27](#) et glace frappée.

Retour sur le terrain, pour un autre [Standard-Genk](#), le 24 février. Un dimanche glacial. De la neige fondante tombe sur [Liège](#), drue. Les supporters attendent le coup d'envoi engoncés dans leur manteau, bonnet sur la tête. Au même moment, une brochette d'invités prend l'apéritif, blottis

anesthésiera (anesthésier)
Verbe

Rendre insensible à la douleur par l'anesthésie.

Example sentence: "On employait autrefois, pour anesthésier les malades, un appareil relativement compliqué."

[Read more on Wiktionary](#)

entreprise un médecin du en téléphonant à match depuis une op froid ? Toi et tu trouves pas ?

Le froid anesthésiera aussi la partie. Après nonante minutes de jeu au ralenti, les joueurs se sépareront sur un score bloqué à zéro de chaque côté. Qu'importe, côté [VIP](#), la bonne humeur reste de mise. On croise un homme débinaire, le patron des briqueteries [Melissen](#). « Un des



the intelligent reading companion

- HOME
- ABOUT
- API
- CONTACT



Johnny Hallyday (Confidence: 1)



Jean-Philippe Smet, dit Johnny Hallyday, est un chanteur, compositeur et acteur français né le 15 juin 1943 à Paris. Avec plus de cinquante ans de carrière, il est l'un des plus célèbres chanteurs francophones et l'une des personnalités les plus présentes dans le paysage médiatique français, où plus de couvertures de magazines lui ont été consacrées.

<http://www.johnnyhallyday.com/>
[Read more on Wikipedia](#)

On croise un sosie de Johnny Hallyday. Des grappes de jeunes se taquinent gentiment, écharpe rouge et blanche autour du cou, vareuse floquée Van Damme. Dix mètres plus loin, trois quadras urinent de concert, le gobelet rempli de bière entre les dents. Un couple s'enlace, à l'ombre du stade rougeoyant de Sclessin. De l'autre côté de la Meuse, le haut-fourneau d'Ougrée se dresse dans un ciel entre chien

Gilles De Bilde, Rodrigo Beenkens, Pierre-Yves Jeholet, les Wathelet père et fils... Tous sont passés par la loge du groupe Circus. Pour faire monter la température, Philippe Favray a sa spécialité : les « orgasmes », un cocktail de Baileys, Jet 27 et glace frappée.

Retour sur le terrain, pour un autre Standard-Genk, le 24 février. Un dimanche glacial. De la neige fondante tombe sur Liège, drue. Les supporters attendent le coup d'envoi engoncés dans leur manteau, bonnet sur la tête. Au même moment, une brochette d'invités prennent l'apéritif, blottis bien au chaud dans la loge d'une grande entreprise liégeoise. Coupe de champagne en main, un médecin du CHU scrute le stade à travers la vitre, tout en téléphonant à l'un de ses amis, qui s'apprête à suivre le match depuis une tribune 2 ouverte à tous les vents. « Pas trop froid ? Toi et moi, là, cela fait un peu lutte des classes, tu trouves pas ? Le froid anesthésiera aussi la partie. Après nonante minutes de jeu au ralenti, les joueurs se sépareront sur un score bloqué à zéro de chaque côté. Qu'importe, côté VIP, la bonne humeur reste de mise. On croise un homme

Persons (64):

- Mario
- Louis Smal
- Josly Piette
- Frisaye
- Mercedes CLS
- Bob Claes
- Marc Wilmots
- Justine Henin
- Bernard Tapie
- Mpoku
- Roland Duchâtelet
- Jean-Michel Javaux
- M. Frisaye
- Benoît Poelvoorde
- Jean-Michel Zecca
- Wathelet
- Meusinvest
- Van Damme
- Philippe Favray
- Johnny Hallyday

Waarom **wachtte** Anderlecht zo lang? Net zoals in Genk **wachtte** ook Anderlecht lang om van trainer te **wisselen**. In wezen **valt** het toe te **juichen** dat clubs **kiezen** voor sportieve continuïteit. Maar soms **voel** je dat **doorgaan** met een trainer geen enkele zin **heeft**. Het **was** een opmerkelijke passage in de geschiedenis van Anderlecht, de intrede van John van den Brom. Een Nederlander met lef en flair, open en eerlijk, communicatief en toegankelijk, een adept van het offensief voetbal. De mensen **stonden** op de banken bij het bij momenten **sprankelende** spel dat Anderlecht toen **bracht**, de term revolutionair **lag** in de mond, zoiets **hadden** ze in het Astridpark al lang niet meer **gezien**. **Zou** John van den Brom er maandag aan **hebben** **gedacht**, op het moment dat hij

gebrek aan duidelijkheid dat uiteindelijk ook het huwelijk tussen Mario Been en RC Genk **deed** **stranden**? Net zoals in Genk **wachtte** ook Anderlecht lang om van trainer te **wisselen**. In wezen **valt** het toe te **juichen** dat clubs **kiezen** voor sportieve continuïteit. Maar soms **voel** je dat **doorgaan** met een trainer geen enkele zin **heeft**. Dat **was** bij Genk zo en het **was** nu bij Anderlecht niet anders. De vraag **blijft** waarom de bestuurlijke top van de kampioen zolang **wachtte**, terwijl er wedstrijden **waren** dat bestuurslui voortijdig naar huis **gingen**. Omdat ze **hoopten** op een ommekeer? Omdat ze op het einde van het seizoen in schoonheid afscheid **wilden** **nemen**? Maar welke basis **was** er nog om op die kentering te **rekenen**? De wanvertoning van **afgelopen** zaterdag op OH Leuven **was** nog maar zelden **vertoond**, het

Selectie woordsoort (85)

Woordsoort

Werkwoord

Features

WVORM

PVTIJD

PVAGR

POSITIE

BUIGING

3.3 Exercise and test generator

What? Assist the teacher with the creation and selection of exercises
Based on authentic material

How? Create (mostly) closed-form exercises on:

- morpho-syntactic correctness
- lexical restrictions
- semantic appropriateness

based on the analysis of L1 text materials
and/or on the analysis of learner errors

For? From beginning to advanced learner

Projects :

Morpho-syntactic

iRead+ project (ITEC)

VIEW-project (Detmar Meurers) <http://sifnos.sfs.uni-tuebingen.de/VIEW>

= Visual Input Enhancement of the Web

Lexical restrictions:

Alfalex-project (Serge Verlinde) <http://ilt.kuleuven.be/publicaties/tools.php>

iRead+ Exercise generator: Three step model

- Explore
 1. Application highlights target items
 2. Learner recognizes target items and clicks on target items
- Practice
 - Fill gaps or multiple choice activities
 - Feedback
- Play
 - Target items, drill & practice
 - Time constraint





Visual Input Enhancement of the Web (VIEW)

Home

About VIEW

Getting Started

Topics and Activities

Articles

Determiners

Gerunds vs. Infinitives

Noun Countability

Phrasal Verbs

Prepositions

Ser vs. Estar

Wh-Questions

Firefox Extension

Changelog

Feedback

Articles

The **articles** topic just highlights the use of the words *a*, *an*, *the*. If you'd like to practice with more types of articles and determiners, try the **determiners** topic.

- **Color**

In the text, VIEW shows you the articles *a*, *an*, *the* in **orange**.

- **Click**

Click on each article you find in the text and VIEW will show you whether you guessed **right** or **wrong**.

- **Multiple Choice**

Choose an article from the drop-down list and VIEW will show you whether you guessed **right** or **wrong**. You can get help by clicking on [?](#).

- **Practice**

Fill in the blanks with the **correct** article. If you guessed **wrong** or you find it too difficult you can get help by clicking on [?](#).

Try it out:

This form lets you get an idea of what VIEW can do, but it may have trouble accessing or displaying some pages. For the best results, use the **VIEW firefox extension**.

URL:

Activity type: Colorize Click Multiple Choice Practice

[Go!](#) (opens in a new window)

Some example sites with colorizing:

- **Wikipedia** (computer-assisted language learning)
- **The Guardian** (car-free cities)
- **TechCrunch** (Woot vs. AP)
- **The Onion** (on This American Life)
- **Public Literature** (Emma by Jane Austen)

Alfalex (KU Leuven – Serge Verlinde)

Exercice sur les collocations

••• Complétez les cadres à l'aide des informations données entre parenthèses à la fin des phrases. •••

Si vous trouvez la phrase trop difficile, vous pouvez obtenir une nouvelle phrase en cliquant sur le lien [modifier] en fin de phrase. Vous pouvez obtenir la première lettre de la forme du verbe à compléter en cliquant sur le lien [première lettre] en fin de phrase. Cette aide vous coûte 2 points sur 4.

1 S'il voulait signifier ainsi que la nation tout entière serait à l'image d'une embarcation dont les voiles flasques attendent misérablement qu'Eole veuille bien les gonfler, il _____ erreur.

{verbe support de *erreur*, présent}
[modifier] [première lettre]

2 C'est là qu'ils _____ l'élevage de chiens boxers, qu'ils exploitent leur garage et tentent de maintenir un style de vie simple, reposant sur le travail.

{verbe support de *élevage*, présent}
[modifier] [première lettre]

3 L'Institut national de la santé et de la recherche médicale et l'Agence nationale d'accréditation et d'évaluation de la santé _____ leur concours au dispositif.

{verbe support de *concours*, futur}
[modifier] [première lettre]

4 L'entendre déclarer qu'elle n' _____ pas de talent pour la vie privée se révèle également pertinent à la compréhension de sa démarche.

{verbe support de *talent*, présent}
[modifier] [première lettre]

5 L'information a _____ jusqu'au palais de Justice de Neufchâteau.

{verbe support de *information*, participe passé}

accueil
mes scores
dico personnel

exercices

morphologie
collocations

conjugaison
synonymes

dérivation
schémas
actanciels

genre
traduction
F > NL

prépositions
traduction
NL > F

Exercice sur les collocations

••• Complétez les cadres à l'aide des informations données entre parenthèses à la fin des phrases. •••

Si vous trouvez la phrase trop difficile, vous pouvez obtenir une nouvelle phrase en cliquant sur le lien [modifier] en fin de phrase.
Vous pouvez obtenir la première lettre de la forme du verbe à compléter en cliquant sur le lien [première lettre] en fin de phrase. Cette aide vous coûte 2 points sur 4.

1 S'il voulait signifier ainsi que la nation tout entière serait à l'image d'une embarcation dont les voiles flasques attendent misérablement qu'Eole veuille bien les gonfler, il _____ erreur.

{verbe support de *erreur*, présent}
[modifier] [première lettre]

2 Mais ce dernier en a _____ l'élevage pour mieux l'inviter à sa table

{verbe support de *élevage*, participe passé}
[modifier] [première lettre]

3 L'Institut national de la santé et de la recherche médicale et l'Agence nationale d'accréditation et d'évaluation de la santé _____ leur concours au dispositif.

{verbe support de *concours*, futur}
[modifier] [première lettre]

4 L'entendre déclarer qu'elle n' _____ pas de talent pour la vie privée se révèle également pertinent à la compréhension de sa démarche.

{verbe support de *talent*, présent}
[modifier] [première lettre]

5 L'information a _____ jusqu'au palais de Justice de Neufchâteau.

{verbe support de *information*, participe passé}

[accueil](#)

[mes scores](#)

[des nouvelles](#)

[exercices](#)

[morphologie](#)

[collocations](#)

[conjugaison](#)

[synonymes](#)

[dérivation](#)

[schémas
actanciels](#)

[genre](#)

[traduction
F > NL](#)

[prépositions](#)

[traduction
NL > F](#)

https://www.kuleuven.be/alfalex/collocations2.php?session=401a49b21da9eba29a94c178d274d7b78&cleinfo=3257&numphras...

3. Applications

Focus on written language input & output

Native Language : Enrich and Exploit

1. *Reading companion*
2. *Resource generator*
3. *Exercise and test generator*

Learner Language : Finding Errors

4. *Error detector, feedback generator and automatic scoring tool*
5. *Writing aid*

Model Learner Behaviour : measure Complexity

6. *Input provider*
7. *Adaptive item sequencer*

-> Conceptual outline + applications from academic R&D

3.4 Error detector, feedback generator and automatic scoring tool

What?

Automated analysis of learner output and generation of feedback

Not limited to closed exercises (MC, fill-in-the-blank, etc.) but also (semi-)open practice tasks (translation, correction, rephrasing, etc.)

How?

1) Approximate (or fuzzy) string matching (ASM)

-> anticipate all potential well-formed and ill-formed learner responses (with inclusion of regular expressions)

Right answer: Il faut absolument qu'ils participent au rallye.
[Il faut [absolument/coûte que coûte/à tout prix] qu'ils participent au rallye / ils sont obligés de participer au rallye / leur participation au rallye est [absolument/tout à fait] [indispensable/nécessaire/obligatoire]] .

2) Parser-based (malrules, constraint relaxation, etc.) and/or statistical and machine learning methods

Towards (half-)open activities with automated analysis and feedback



Fleat VI Harvard University
August 11-15, 2015



Bridging the gap between closed and open items
OR
how to make CALL more intelligent

Piet Desmet & Bert Wylín



NLP4CALL



2016



Projects:

Task design: require the learner to use certain language material

e-tutor-project (Trude Heift) <http://www.e-tutor.sfu.ca>
translation

Tagarela-project (Detmar Meurers) <http://sifnos.sfs.uni-tuebingen.de/tagarela>
answer should include certain words

Task design: offers clues towards a limited set of possible answers

Dialog Dungeon (ITEC)


gamified written dialog tasks

focus on specific set of problems

e.g. past tenses, subject-verb agreement, etc.

E-tutor project

the etutor
GERMAN LANGUAGE TUTORING



INTRODUCTION CONTENTS STRUCTURE: Ex. 6 Kapitel 12

No.: 1 2 3 4 5

INSTRUCTIONS

Translate the following:

Our dog isn't afraid of your (familiar, singular) cat, Frank.

Ä Ö Ü ß ä ö ü

Unsere Hund hat keinen Angst von deiner Katze. Frank

Check Solve Skip

FEEDBACK HISTORY GRAMMAR HELP DICTIONARY ERROR PROFILE

Unsere Hund hat keinen Angst von deiner Katze Frank

Carefull A mistake with the possessive pronoun **unsere**. You need the masculine inflection here.

» See help on: [possessive pronoun](#)

The etutor | Help | Logout

Copyright 2003-11 T. Heift. All rights reserved.

Tagarela-project

THE TAGARELA SYSTEM  THE OHIO STATE UNIVERSITY
ICALL RESEARCH GROUP
ГРУППА ИССЛЕДОВАНИЙ
THE OHIO STATE UNIVERSITY

Listening Reading Description Fill-in-Blanks Rephrasing Vocabulary Home Logout

Módulos: [1](#) [2](#) [3](#) [4](#) [5](#) Atividades: [1](#) [2](#)

Leitura

Instrução 

Leia o texto e responda às questões usando frases completas e o vocabulário apresentado no texto.

Quem é você?



Meu nome é João de Castro. Eu sou morador do Rio de Janeiro. Tenho trinta anos e estudo medicina na UFRJ.

Meu pai se chama Antônio e minha mãe Júlia. Meu pai é jornalista e trabalha em um grande jornal local chamado "O Globo".

Minha mãe é professora de francês. Além do francês, ela também fala italiano e inglês. Ela trabalha em uma escola particular.

Questões: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#)

Questão 1 **Análise:**

Onde mora João?

à	á	â	ã	ä	é	ê	í	ó	ô	õ	ú	ü	ç
A	Á	Â	Ã	Ä	E	Ê	I	O	Ó	Ô	Ú	Ü	Ç

Dialog Dungeon gamified written dialogue tasks

@Frederik Cornillie

The screenshot displays the 'Dialog Dungeon' interface. At the top, it says 'Detective stories' with a progress bar at 0 and a user profile for 'dialogdungeon.tester5 police constable'. The main content area is titled 'THE MURDERS IN THE RUE MORGUE'. Below the title, there is a list of tasks: '1. EXTRAORDINARY MURDERS' (with a blue arrow and three locked padlocks), '2. ???', '3. ???', and '4. ???'. A 'Review the evidence' button is at the bottom of this list. To the left, there is a 'feedback' button. The chat area shows a sequence of messages: a system message 'You are in your apartment, reading a book on 4th-century Arabic logic, when suddenly the police prefect knocks and enters.', a message from a character (Dupin) saying 'Dupin, I bring terrifying news.', a user response in a green bubble saying 'Type something to keep the dialogue going.' with a 'send' button and a 'cancel' button, and a message from another character saying 'A horrific event! Here, the Gazette de l'Inde et de l'Asie...'. Red callout boxes highlight the 'Type something to keep the dialogue going.' message with the text 'semi-open written exercise' and the 'send' button with the text 'support: link to responses of peers'.

Detective stories



dialogdungeon.testers5
police constable

THE MURDERS IN THE RUE MORGUE

1. EXTRAORDINARY MURDERS →



- 2. ???
- 3. ???
- 4. ???

Review the evidence

feedback

You are in your apartment, reading a book on 4th-century Arabic logic, when suddenly the police prefect knocks and enters.



Dupin, I bring terrifying news.



Other detectives said:

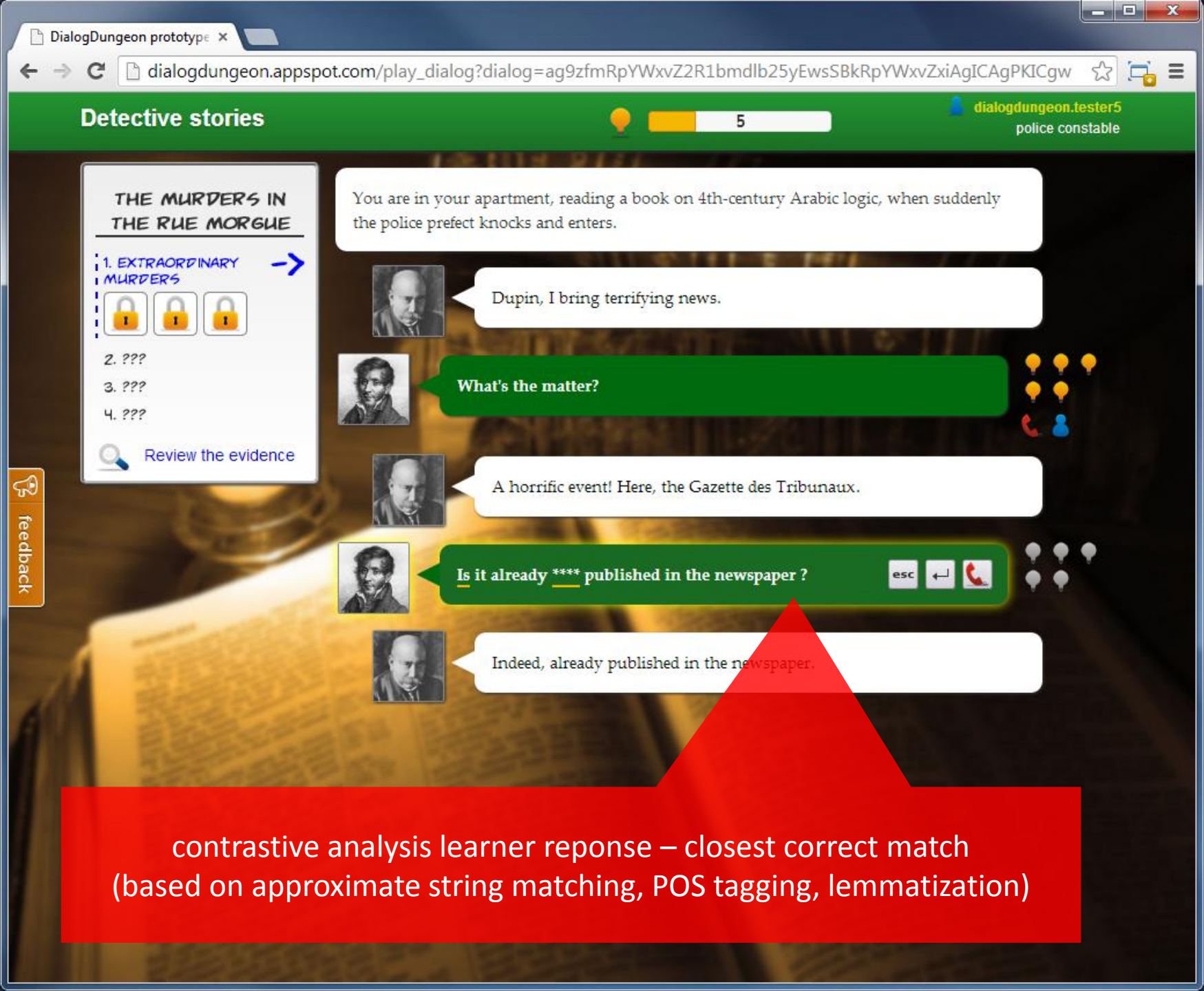
- What happened?
- Tell me more.
- What's the matter?

See more ...



A horrific event! Here, the *Gazette des Tribunaux*.

learning support:
responses of
peers



Detective stories

5

dialogdungeon.testers5
police constable

THE MURDERS IN THE RUE MORGUE

1. EXTRAORDINARY MURDERS



- 2. ???
- 3. ???
- 4. ???

Review the evidence

You are in your apartment, reading a book on 4th-century Arabic logic, when suddenly the police prefect knocks and enters.



Dupin, I bring terrifying news.



What's the matter?



A horrific event! Here, the Gazette des Tribunaux.



Is it already **** published in the newspaper?



Indeed, already published in the newspaper.

feedback

contrastive analysis learner reponse – closest correct match
(based on approximate string matching, POS tagging, lemmatization)

Detective stories



dialogdungeon.tester5
police constable

THE MURDERS IN THE RUE MORGUE

1. EXTRAORDINARY MURDERS



- 2. ???
- 3. ???
- 4. ???

Review the evidence

feedback

You are in your apartment, reading a book on 4th-century Arabic logic, when suddenly the police prefect knocks and enters.



Dupin, I bring terrifying news.



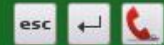
What's the matter?



A horrific event! Here, the Gazette des Tribunaux.



Is it already **** published in the newspaper?



Well done ! Top detectives used **a verb** here, like you did, but they used another one. Look carefully for clues. Or possibly, the word you used could not be recognized -- check your spelling.



Indeed, already published in the newspaper.

supportive prompt with metalinguistic hints
(based on POS tagging)

3.5 Writing aid

What? Support the learner in writing a functional, well-formed text

How? No automated correction, but suggestions to help the learner

Real-time, online, assistance on

- lower-order skills (spelling) , lexical (including MWE)
- Lexico-grammatical, Grammatical skills
- Semantic, Pragmatic analysis

Through

- On-the-fly prompts
- Q&A interactive system
- Post-writing checks

Projects:

Interactive Language Toolbox-project (Serge Verlinde)

<https://ilt.kuleuven.be/inlato>

Bon patron, SpellCheckPlus & SpanishChecker (Nadaclair Language Technologies)

<http://bonpatron.com>; <http://spellcheckplus.com>; <http://spanishchecker.com>

Glosser-project (Univ. of Sydney) <http://www.glosserproject.org/en/>

Interactive Language Toolbox

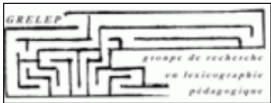
KU LEUVEN

zoek

Leuven Language
Institute

Interactive Language Toolbox

(Almost) everything you always wanted to know about... texts, words and word combinations in Dutch, English and French



- Dutch
- English
- French

Marie est une des 250 jeunes qui vont partir à l'étranger avec AFS. Mes parents et moi ont parlé longtemps de la possibilité d'étudier à l'étranger. Quand j'étais petite, nous voyagions beaucoup, à Portugal et aux Etats-Unis par exemple. Je le trouvais toujours intéressant de pouvoir

clear search form

reload page

Click on a button to access information.

Help me

- understand my text
- translate my text
-

KU Leuven Association only

correct my text

correct my text

in 6 steps:

- Words: spelling, morphology, ...
- Sentences: subject-verb, ...
- Common errors: word choice, ...
- Noun: agreement, position of adjectives, ...
- Verb: auxiliaires, imparfait/passé composé, ...
- Prepositions and word combinations

- analyse word frequencies

Take a tour

- search for words and word combinations: [click here](#)
- text revision: [click here](#)

Interface

- in English
- en français
- in het Nederlands

FAQ

[click here](#)

More... (French only)

- Discover vocabulary
- Exercises

Tip

- Words combine together. Discover how.
- Learn how to find good translations.

Personal database

KU LEUVEN kulak

NLP4CALL

itec

2016

mec



BonPatron.com

correcteur de grammaire et d'orthographe



Tapez ou collez votre texte en français ici et cliquez le bouton "Vérifier le texte".

« je » est féminin ?

je suis francophone ?

Conjugaison des verbes



Nouveau !
[Guide de grammaire](#)

Vérifier le texte

[un exemple](#) | [effacer](#)

Placez la souris au-dessus du texte formaté pour voir paraître des explications.

English

grammaire:

à modifier à vérifier

orthographe:

à modifier à vérifier

Français



Le nom de **mon mère** est Linda et elle **est 40 ans**. Elle **était** née **au Chine** et a **imigrée** ici. J'ai **deux soeur** et **une frère**. Pendant **mon enfance**, j'ai **habité maison** au **centre ville**. Il y a cinq ans, mes parents ont **décidé** de **changer** de maison. J'**étais peur** parce que **je voulais pas** perdre mes amis. C'était **difficile** au début. Un jour, mon père m'a dit : "il **faut** **que tu fais** plus d'effort pour rencontrer des amis, essaye de **faire parti** d'un club à l'école. J'**ai décidé** d'essayer le club de danse. **Toute de suite**, j'ai rencontré **ma ami Paul et maintenant** nous passons **beaucoup des jours** ensemble. **Chaque samedi on va ensembles** au centre d'achat et on a un **bon temps** ! Je suis content d'avoir **déménagé** !

Ce nom est féminin. Le mot qui précède doit aussi être au féminin, par ex.: "Paul loves his mother" = *Paul adore sa mère*.

« je » est féminin ?

je suis francophone ?

Conjugaison des verbes



Nouveau !
Guide de grammaire

Vérifier le texte

[un exemple](#) | [effacer](#)

Placez la souris au-dessus du texte formaté pour voir paraître des explications.

English

grammaire:

à modifier à vérifier

orthographe:

à modifier à vérifier

Français

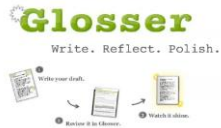
Le nom de **mon mère** est Linda et elle **est 40 ans**. Elle **était** née **au Chine** et a **imigrée** ici **au Edmonton** en 1959, où elle a rencontré mon père. J'ai **deux soeur** et **une frère**. Pendant **5 ans**, **on a resté** dans une **petit maison** au **centre ville**. Il y a cinq ans, mes parents ont décidé **de déménagé** et ils **ont nous** dit qu'on va changer de maison. J' **étais peur** parce que **je voulais pas** perdre mes amis. C'était **difficile** au début. Un jour, mon père m'a dit : "il **faut que tu fais** plus d'effort pour rencontrer des amis, essaye de **faire parti** d'un club à l'école. J' **ai décider** d'essayer le club de danse. **Toute de suite**, j'ai rencontré **ma ami Paul et maintenant** nous passons **beaucoup des jours** ensemble. **Chaque samedi on va ensembles** au centre d'achat et on a un **bon temps** ! Je suis content d'avoir **déménagé** !

Les expressions de quantité sont suivies de la préposition *de*, par ex.: *il y a beaucoup de personnes* (sauf si on précise un nom spécifique, par ex.: *beaucoup des gens que je connais*).

Conjugaison des verbes

Nouveau !
Guide de grammaire

Vérifier le texte
un exemple | effacer



Glosser project (Sydney)

About

Glosser is a web-based framework for providing automated feedback. The framework allows writers to use different writing tools such as Wiki's, cloud-based Google Docs or MS Word. The documents are then processed to produce a wide range of types of feedback spanning collaborative or individual writing activities, visual and text modalities, feedback can be on surface or content features, on the writing product (the final document) or the process. This range of possibilities requires an architecture that is highly flexible and extensible.

Glosser provides different forms of feedback, each appropriate for specific writing genres. Current feedback 'tools' include:

Spelling. We have integrated iSpell so after a document is harvested the tool displays it with the misspelled words underlined. When clicking on the word the student can see a list of suggestions to fix the problem. This tool is a good example of the data flow from harvesting to user interaction so it is described in more detail later.

Flow. Two paragraphs are said to flow when they talk about similar topics. They do not flow when the semantic distance between them is too large. Glosser's tool displays warnings when the distance between two paragraphs is greater than a threshold (that can be changed by the user). This tool is a good example of how single document LSA can be used and is discussed in more detail later on.

Flow Map. This map provides a visualization of how the parts of a document relate and follow each other, producing a visual representation of a quality measures based on the average semantic distance between parts of a document (O'Rourke, et al., 2011). The visualization is not only useful to students but has also been shown to help tutors mark essays more efficiently and reliably (O'Rourke, et al., 2011).

Topics. This tool (shown in Figure 4) shows a table with the key topics developed in the document, the three most important sentences referring to that topic and the identification of the person who wrote most about it. Default trigger questions include: "Are the ideas used in the essay relevant to the question?", "Are the ideas developed correctly?".

Topic Map. The topic map tool is used to identify the structure of topics developed in the document. The tool identifies keywords using a network-based analysis then maps them to a visual representation. If two keywords are used in the same sentence, they are linked in the visualization. The structure of topics can then be subjectively inferred from the semantics of the visualization. The sentences that cover more than one of these topics can be seen when the link between them is shown on a rollover. The interactive visualization tool is implemented as a Java applet...

Concept Map. The topic map tool uses a naïve approach, where ---- . The concept map tool incorporates an automated concept map generator discussed in more detail elsewhere (Jorge Villalon & Calvo, 2009; J. Villalon & Calvo, 2011).

Participation. This tool is aimed at supporting collaborative writing activities and/or the teaching the awareness of the writing process.

Automatic Questions. Questions of different types can be generated from a single document. We have developed an approach to automatic question generation (Liu, et al., 2010) integrated into this tool. Since the questions currently available are produced from citations found in the document, this tool is mostly useful for activities involving literature reviews.

Glosserproject is kindly supported by YourKit. YourKit, LLC is the creator of innovative and intelligent tools for profiling Java and .NET applications, and they support open source projects with licenses for their full-featured **profilers**. Take a look at YourKit's leading software products: [YourKit Java Profiler](#) and [YourKit .NET Profiler](#).

3. Applications

Focus on written language input & output

Native Language : Enrich and Exploit

1. *Reading companion*
2. *Resource generator*
3. *Exercise and test generator*

Learner Language : Finding Errors

4. *Error detector, feedback generator and automatic scoring tool*
5. *Writing aid*

Model Learner Behavior : measure Complexity

6. *Input provider*
7. *Adaptive item sequencer*

-> Conceptual outline + applications from academic R&D

3.6 Input provider

- What? Support learners in searching for texts adapted to
Their current proficiency level and interest
Optimising for reaching the next level
- How? (Semi-)automated selection of comprehensible & authentic text material
For vocabulary and reading practice
- Linguistic analysis of readability and complexity
Analysis of formal, lexical and semantic complexity
 - Operationalize proficiency and link to complexity
 - Perform syntactic and lexical text simplification/elaboration
 - Determine topical interest of texts

Projects:

REAP-project (Carnegie Mellon) <http://reap.cs.cmu.edu>

“Reader-Specific Lexical Practice for Improved Reading Comprehension”

SATO-project (François Daoust) <http://www.ling.uqam.ca/sato>

“Système d’analyse de texte par ordinateur”

+ SATO-Calibrage

Automated formal analysis of a corpus (existing or personal)

Aim: automatically identify the lexical items in the text that are complex for specific learners

Measures for the prediction of word complexity

1. Linguistically motivated measures
2. Resource-based measures
3. Learner-based measures



1. Linguistically motivated measures

= identify word characteristics that are good predictors of word difficulty, especially *psycholinguistic variables* (Shardlow 2013; Gala et al. 2014)

- *lexical sophistication vs. basic vocabulary list*
- *word frequency (SUBTLEX, Lexique3, ...)*
- *word sense ambiguity (WordNet synset)*
- *complex orthographical patterns*
- *frequency of morphological affixes*
- ...

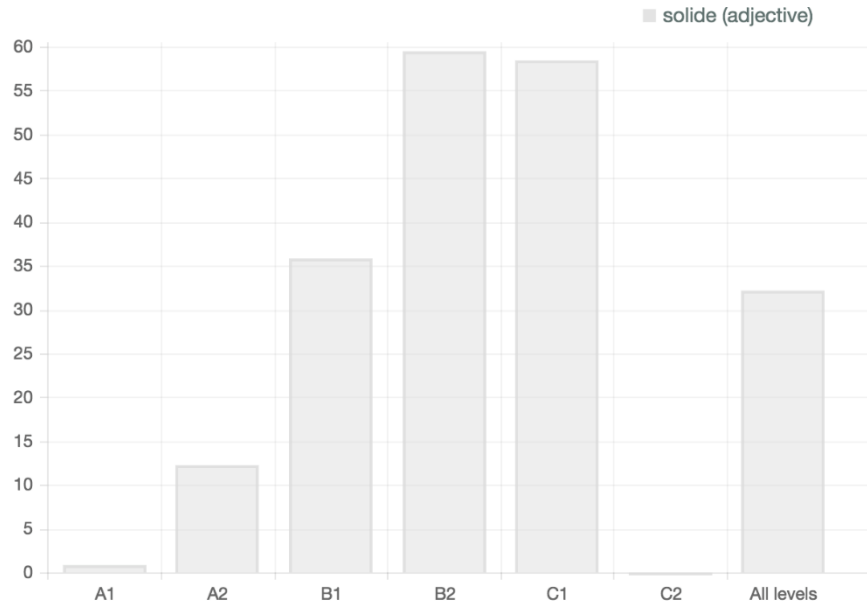


2. Resource-based measures

FLELEX: A graded lexical resource
<http://cental.uclouvain.be/flelex/>

- linking word frequency and word complexity
- dictionary of frequencies observed along a difficulty scale
 - L2 : CEFR-graded texts and textbooks (François et al., 2014)

Frequencies by CEFR levels for the word solide*.



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

Main issues:

- What is the word's CEFR level? In other words, what level targets the word as « core vocabulary » to be acquired at this level ?
- How can we identify this level?
 - first occurrence
 - threshold the Xth percentile
 - ...
- How can we distinguish core and peripheral vocabulary?
 - complex words observed in A1 text might be fortuitous (i.e. peripheral vocabulary)
- How to deal with phraseology?

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 How-to ▾

Enter a text:

Or select a text file:

Geen bestand gekozen

N.B.: Currently, only .txt files are supported (max. 2MB).

Which is your CEFR level? A1 ▾

Which tagger? TreeTagger ▾



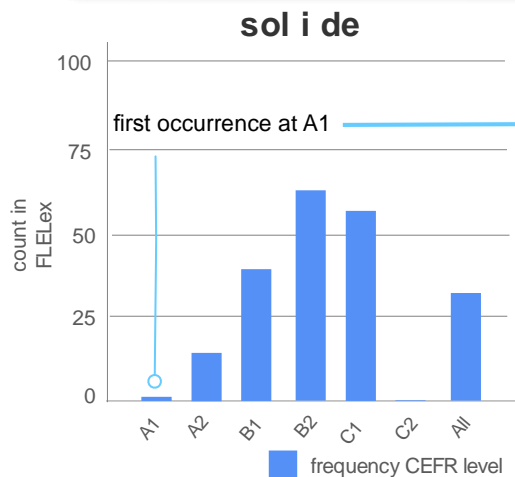
Webmaster: CENTAL (Centre de traitement automatique du langage)
Collège Erasme, 1 place Blaise Pascal, B-1348 Louvain-la-Neuve (Belgique)



Lexical competence prediction — Expert model

Predictive model based on a
graded lexical resource

Predict known and unknown words in a new text
for a given *level*



Archimède Level A2

Archimède est un grand scientifique grec de Syracuse. Il est un **physicien**, mathématicien et ingénieur. Né en 287 av. J. -C. et mort en 212 av. J. -C. à Syracuse, il a vécu en Sicile (île actuellement italienne).

Œuvre scientifique.

Archimède fut le premier à mettre en équations **mathématiques** les faits faciles à observer dans la nature. Ses livres démontrèrent à l'aide des figures **géométriques** et des proportions, comment sont fait les équilibres, ou les centres de **gravité** des corps **solides**, quelles longueurs doivent avoir des **leviers** pour soulever les poids.

Annotations: **physicien** (unknown B1), **mathématiques** (B1 unknown), **géométriques** (unknown OOV), **gravité** (unknown B1), **solides** (known A1), **leviers** (unknown OOV).

Tack, A., François, T., Ligozat, A.-L., & 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 Resources and Evaluation (LREC'16)* (pp. 230–236). Portorož, Slovenia.



3. Subjective measures

= identify words that are found difficult by a specific learner

Problem: expert model gives one-sided predictions

- identical predictions for all learners having the same level
- OOV words (out-of vocabulary) either because of (a) technicality or (b) editorial choice

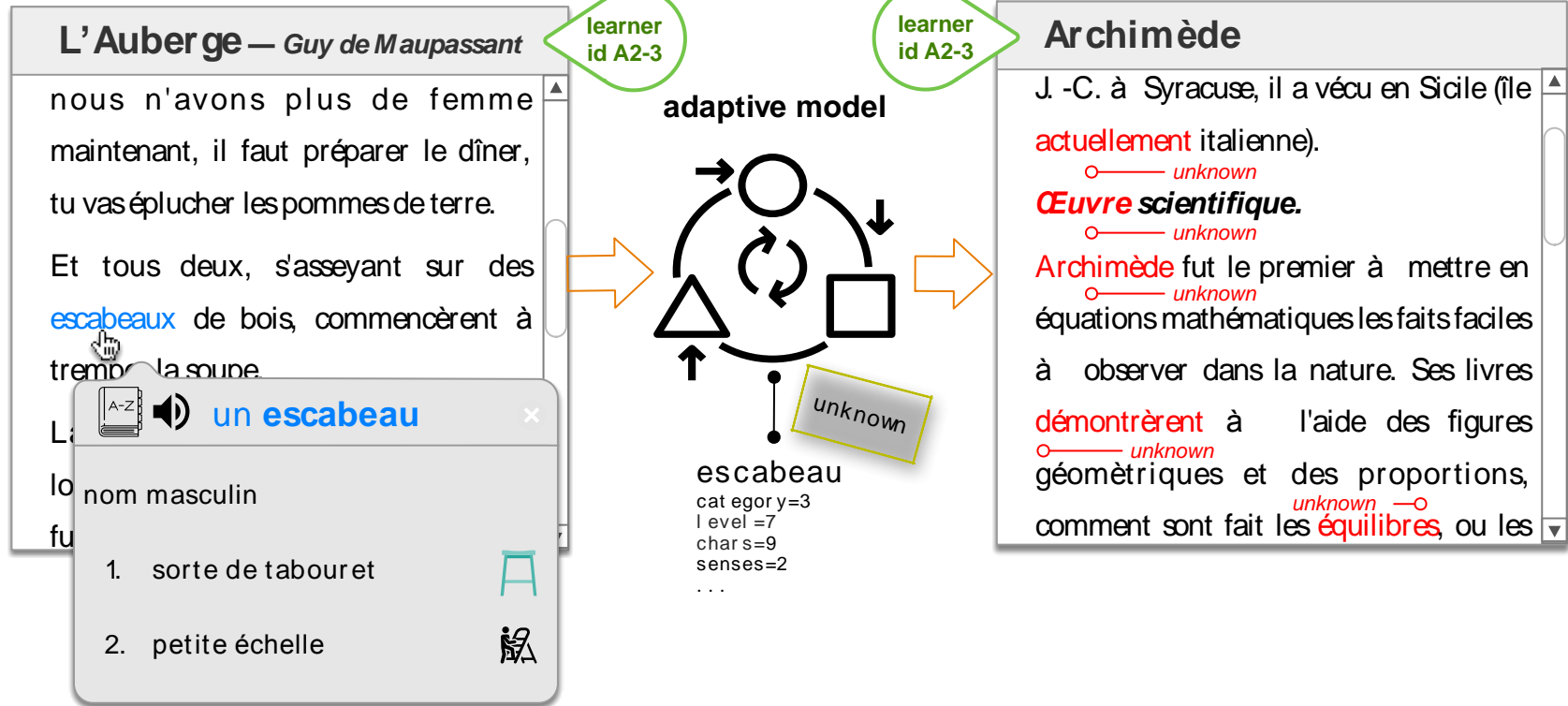
Towards a more fine-grained & personalized prediction

- integrated model of word complexity features
 - psycholinguistic + resource-based + subjective measures (e.g. cognates)
- adapt predictions according to learner feedback
 - triggered by a learner clicking on a difficult word
 - incremental learning of feature weights

Lexical competence prediction — Learner model

Predictive model based on learner input

Predict known and unknown words in a new text for a given learner



Tack, A., François, T., Ligozat, A.-L., & Fairon, C. (2016). Modèles adaptatifs pour prédire automatiquement la compétence lexicale d'un apprenant de français langue étrangère. In *Actes de la 23ème Conférence sur le Traitement Automatique des Langues Naturelles (TALN'16)* (pp. 221–234). Paris, France.

3.7 Adaptive Item Sequencer

a) **Adaptivity: to what?**

- difficulty level of the tasks
- learner profile (prior knowledge, motivation, cognitive load, interests & preferences)
- context (time and place, device, etc.)

(b) **What to adapt?**

- adaptive form representation (form in which content is presented to the learner, e.g. dynamically generated hypermedia pages)
- adaptive content representation (e.g. with or without learner support)
- adaptive curriculum sequencing (e.g. selection of items in function of difficulty level, learner profile, etc.)

(c) **How to implement adaptivity?**

- full program control (via reasoning component 'if.. then...' rules)
- full learner control
- shared control

Wauters, K., Desmet, P., Van Den Noortgate, W. (2010). Adaptive Item-Based Learning Environments Based on the Item Response Theory: Possibilities and Challenges. *Journal of Computer Assisted Learning*, 26 (6), 549-562.

3.7 Adaptive Item Sequencer

What about AI?

a) **Adaptivity: to what?**

- linguistic complexity -> language agent (expert system)
- item difficulty -> IRT
- learner profile (prior knowledge, motivation, cognitive load, interests & preferences) -> student agent

(b) **What to adapt?**

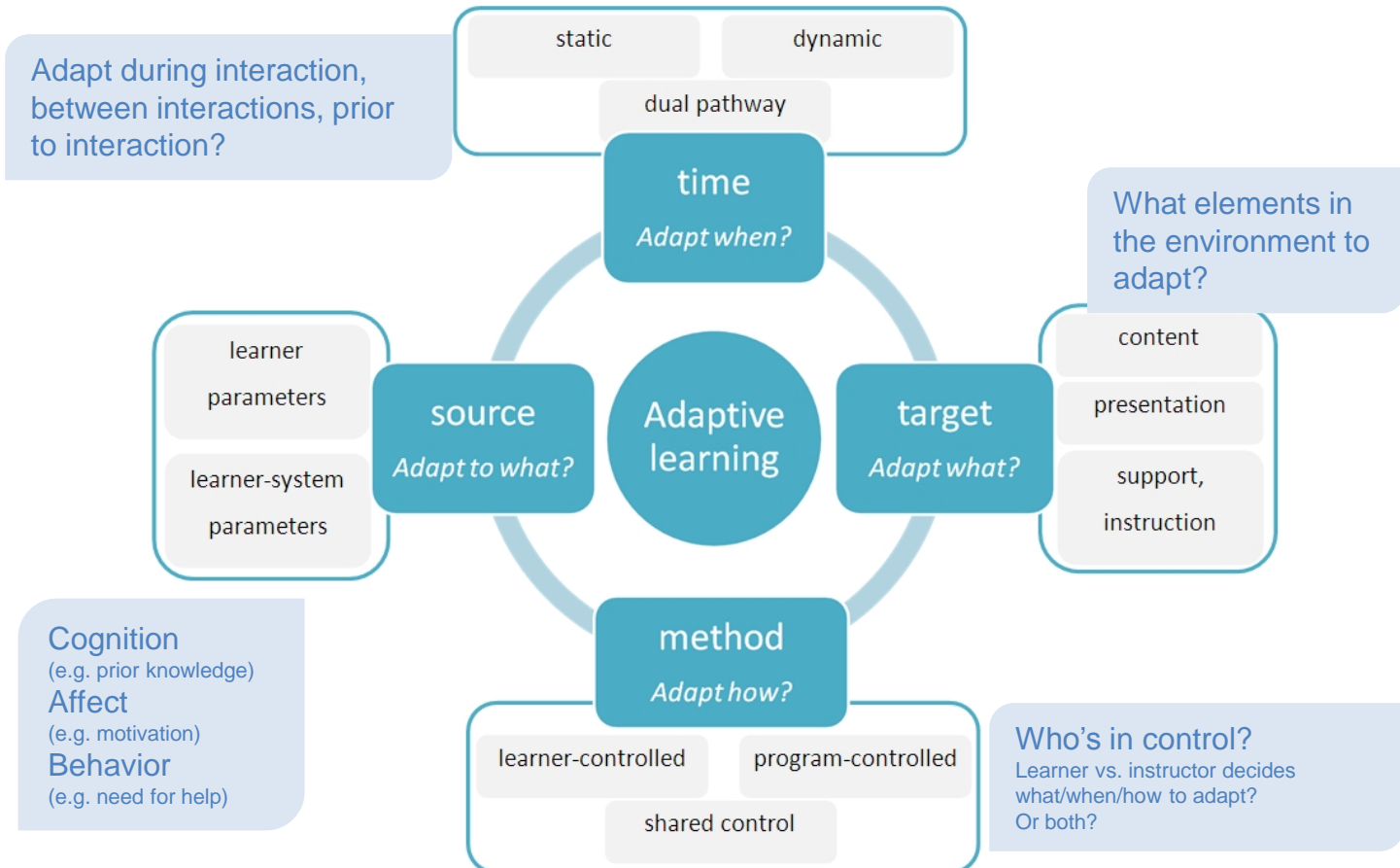
- adaptive curriculum sequencing (e.g. selection of items in function of difficulty level, learner profile, etc.) -> tutor agent

(c) **How to implement adaptivity?**

- full or shared control -> tutor agent

Beuls, Katrien (2013) *Processing, learning and tutoring of Spanish verb morphology*. Brussels: VUB. (PhD)

4D-model of adaptive instruction



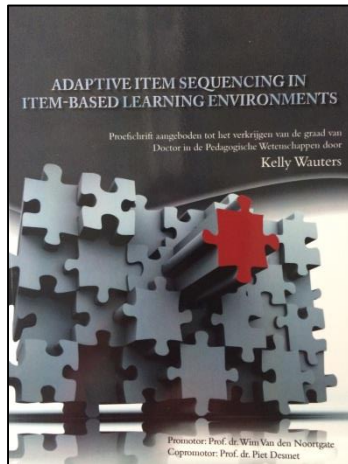
Vandewaetere, Desmet & Clarebout 2011 / Vandewaetere & Clarebout, 2012



http://www.slideshare.net/piet_desmet/2015-0522-presentatiecalicodesmet-vandewaetere-def

ADAPTIVE LEARNING

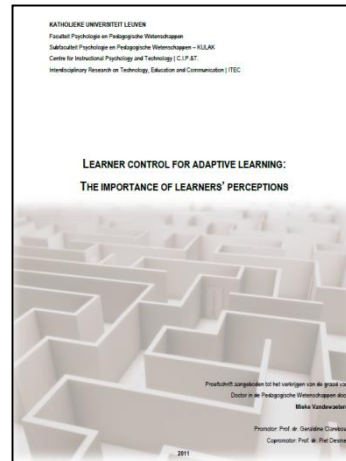
“HARNESSING COLLECTIVE INTELLIGENCE IN ORDER TO MAKE E-LEARNING ENVIRONMENTS ADAPTIVE”
(IOF KNOWLEDGE PLATFORM KU LEUVEN, 2007-2011)



ADAPTIVITY ON **ITEM** LEVEL

IRT FOR ADAPTIVE LEARNING
DYNAMIC LEARNING CURVES

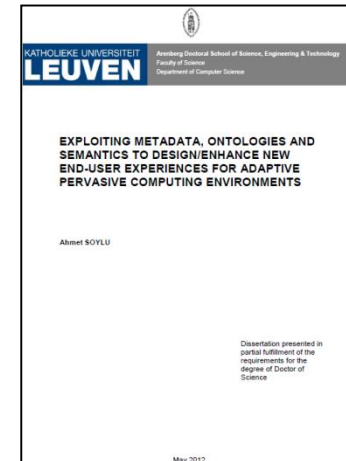
DR. KELLY WAUTERS
(2012)



ADAPTIVITY ON **LEARNER** LEVEL

LEARNER CONTROL AND OPEN LEARNER
MODELS
TOOL USE AND HELP-SEEKING

DR. MIEKE VANDEWAETERE
(2011)



ADAPTIVITY ON **CONTEXT** LEVEL

ONTOLOGIES FOR ADAPTIVE LEARNING
PERVASIVE, CONTEXT-AWARE SYSTEMS

DR. AHMET SOYLU
(2012)

4. Future Developments

4.1. Native Language

Unsupervised learning: Understand examples

Learn structure, explore

4.2. Learner Language

Towards a different model of interlanguage

Using CALL applications for gradual error detection

4.3. Measuring linguistic complexity

Individualized Lexical complexity

Link known with unknown vocabulary items

4.1 Native Language

Unsupervised Learning on Native Language

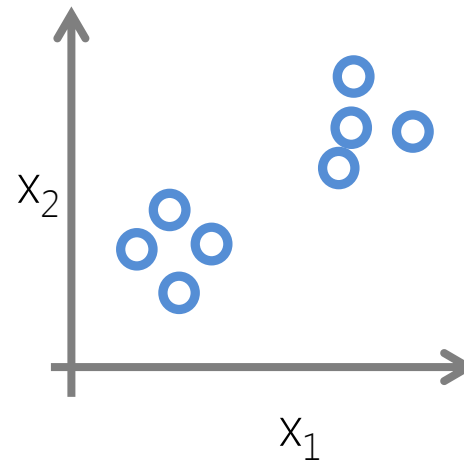
Description of examples through more finegrained clustering without labels

Interpretation by the learner himself

~ Concordancers

~ Similar words

~ Similar constructions



Shifts the responsibility of interpretation to the learner himself
But based on a much more informative clustering of examples

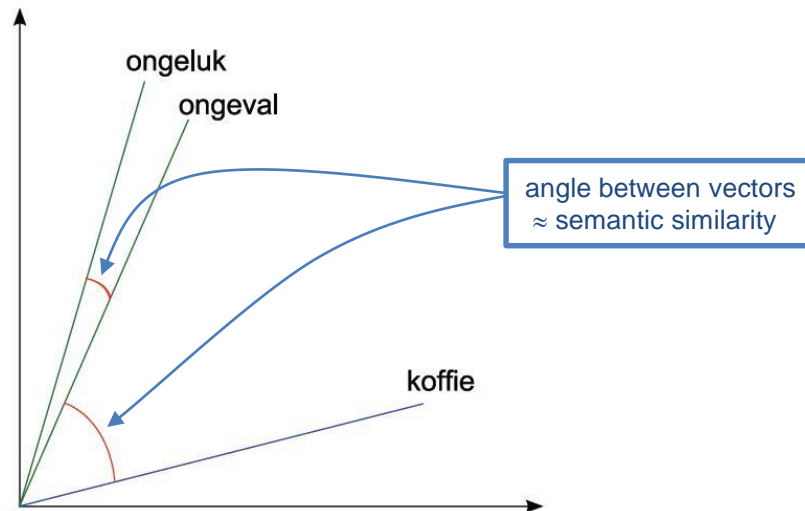
Semantic Vector Spaces

At the word level: Semantic Vector Spaces

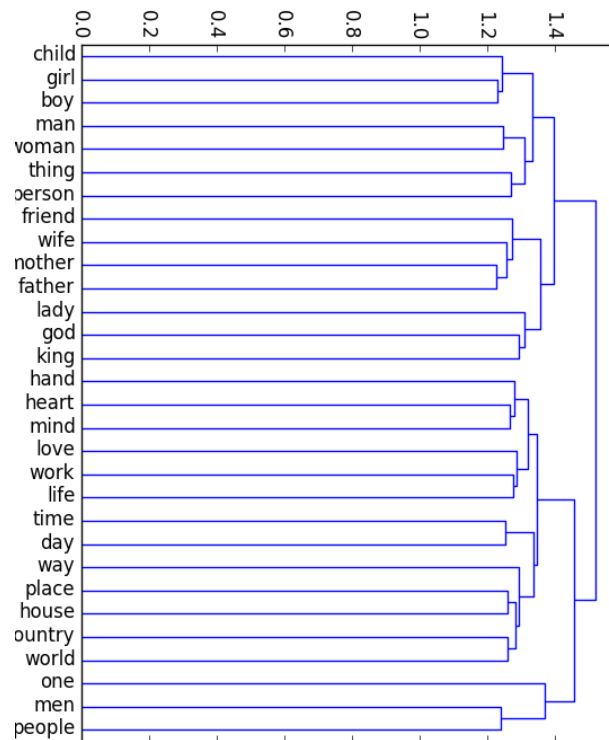
"You shall know a word by the company it keeps" (Firth, 1957)

- 1) Paradigmatic characteristics of words can be retrieved on the basis of their syntagmatic behaviour.
- 2) Map the usage in context of a word in a high dimensional space (e.g. with contexts as dimensions, and #occurrences as value)
- 3) Proximity in space signals **semantic relatedness**
Even relations can be represented as vectors in space

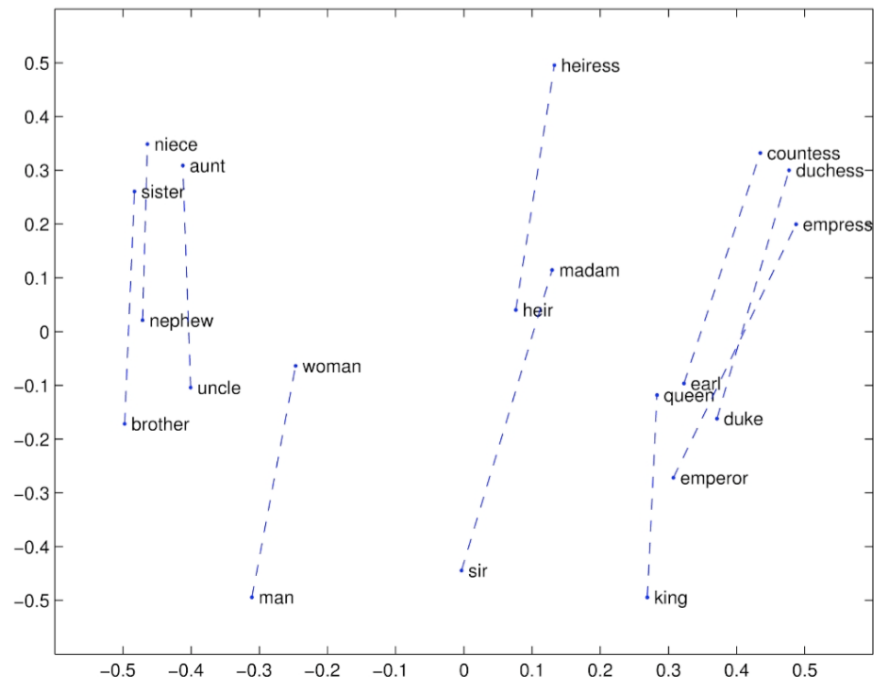
Unsupervised semantic comparisons



As the basis of Resource Creation



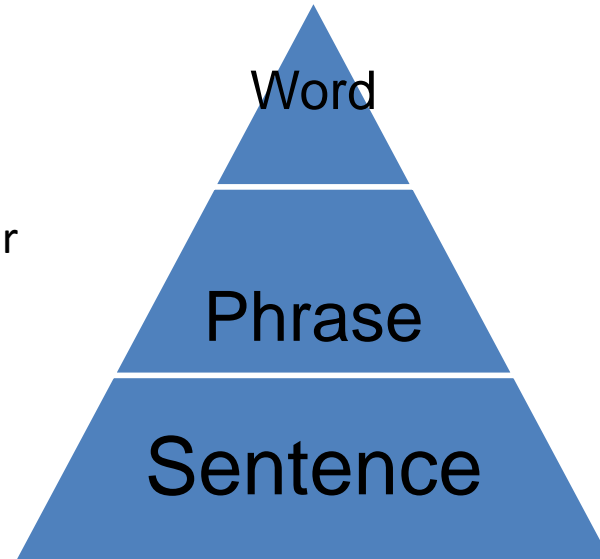
For visualising relationships



Semantic Vector Spaces

A promising future beyond the word level

The **same representation** for more complex linguistic structures using the principles of **compositionality**.



4.2. Modeling Learner Language

Now starts from native language model
= basis of error analysis

Native language is a fully developed linguistic **system**

E.g. POS-tags: What defines a POS-category knowing the full system?

- Distributional restrictions : N usually follows Det/Adj
- Morphological restrictions: N + plural s
- Lexical restrictions:

⇒ Resolves to ONE POS-tag
Thus possible to train a pos-tagger

4.2.1. Towards a different model of interlanguage

However: POS might be a non-existing construct for a learner
(Diaz et al., 2010)

“He is joy” (He is happy)
⇒ ‘joy’ is not permitted in the slot.

Central : learner does **not understand/know** part(s) of the native system

Solution: **no full POS resolution**

But model each of the **system-components** separately

- Distributional restrictions
- Morphological restrictions
- Lexical restrictions

⇒ Does not necessarily converge to a single POS-tag

⇒ Partial information can then be used for corrections/feedback

4.2.2. Gradual error detection

CALL tools provide unseen (and underused) possibilities to

- Capture L2 production
- Advance towards fully open output

How? Ask for immediate feedback from the learner

Have the tool disambiguate the output at production times
A constant closed feedback loop

Progressive Feedback

In a progressive manner: the algorithm does not propagate wrong assumptions

- 1) orthography
- 2) POS-disambiguation
- 3) Morphology
- 4) Argument structure
- 5) ...

At each step construct and adjust the hypothesis

Have the system come as close as possible to verification

- Helping the user with analogous examples from corpora
- Through explanation of linguistic regularities
- ...

Check 1 of 6: words (spelling, morphology, ...)

Possible errors found for these items: click on the link(s)

Les mots

- [Orthographe](#) (spelling), [morphologie et influences de l'anglais et du néerlandais](#)

Check 2 of 6

Start check 2 of 6

Orthographe, morphologie et influences de l'anglais et du néerlandais

example = correct this word with the information on the right.

possibilité = check for possible errors for this word using the information on the right.

Marie est une des 250 jeunes qui vont partir à l'**étranger** avec AFS. Mes parents et moi ont parlé longtemps de la **possibilité** d'étudier à l'étranger. Quand j'étais **petite**, nous voyagions beaucoup, à Portugal et aux Etats-Unis par **exemple**. Je le trouvais toujours intéressant de pouvoir découvrir des nouvelles cultures et de faire des contacts avec la population.

A 2010 Hugo passait quelques mois à la Mexique. Avec vivre une année dans une autre culture, mon regard sur le monde est changé **fortement**. Les différences entre des cultures peuvent peut-être **causer** des conflits, mais il n'y a aucun problème ce qui tu ne peut pas **solver**.

CORRECT YOUR TEXT

- Influences de l'anglais ou du néerlandais

example



an example



= un exemple

solver



to solve



= résoudre

CHECK, IF NECESSARY, CORRECT YOUR TEXT

- Orthographe des mots

Check 2 of 6: sentences (subject-verb, ...)

Check these items for the whole text: click on the links

Les phrases

- [Accord sujet-verbe: contrôle général](#)
- [Place des compléments et emploi des prépositions](#)

Check 3 of 6: common errors (word choice, ...)

Possible errors found for these items: click on the link

Fautes fréquentes

- [Construction des phrases](#)
- [Confusions de mots](#)
- [Registre](#)
- [Problèmes de traduction](#)

Check 4 of 6: noun (agreement, position of adjectives, ...)

Possible errors found for these items: click on the link(s)

Les noms/substantifs

- Le genre: [masculin et féminin](#)
- La [place de l'adjectif](#)

Les noms de pays

- Le genre et [emploi des prépositions](#)

Check 5 of 6

Start check 5 of 6

Check 6 of 6: prepositions and word combinations

Possible errors found for these items: click on the link(s)

Les prépositions

- L'emploi de la préposition [après un verbe suivi d'un infinitif](#)
- L'emploi de la préposition [après un verbe, un substantif, un adjectif](#)

Les combinaisons de mots

- [Amélioration des combinaisons de mots](#)
- [Sugqestions de combinaisons de mots](#)

Extension du vocabulaire

- [Adjectifs à sens négatif](#)

Check 5 of 6: verb (auxiliaires, imparfait/passé composé, ...)

Possible errors found for these items: click on the link(s)

Les verbes

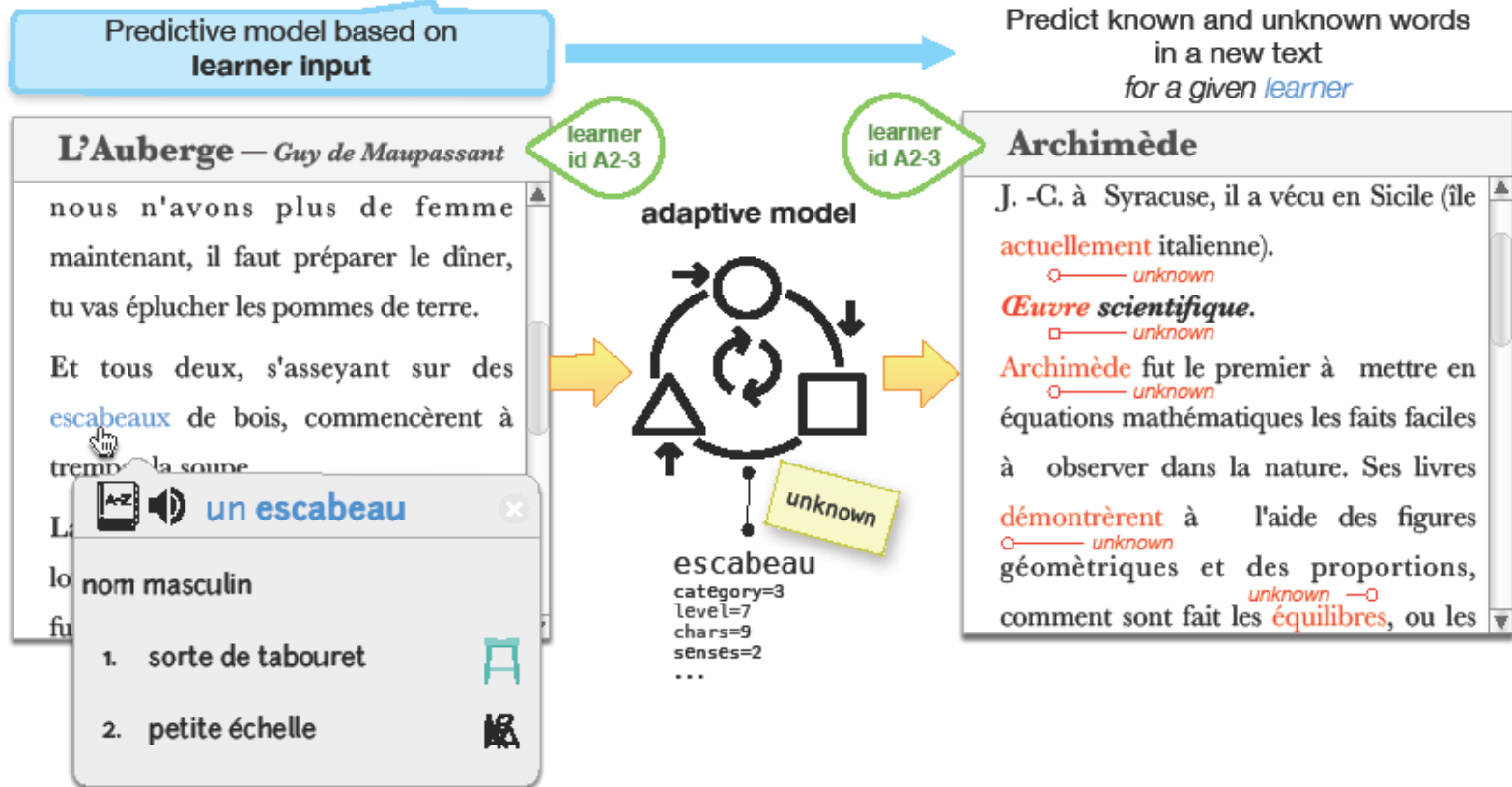
- L'accord sujet-verbe
- L'emploi du gérondif
- L'emploi du [subjonctif](#)
- L'emploi de [l'imparfait et du passé composé](#)
- L'emploi des [auxiliaires](#)

4.3 Intelligent tutors and linguistic complexity

Steps away from procedures that rely on “one-size fits all”
= A truly individualized learning program

- Capture and store L2 input and production per learner
- Determine known and unknown vocabulary items

Lexical competence prediction — Learner model



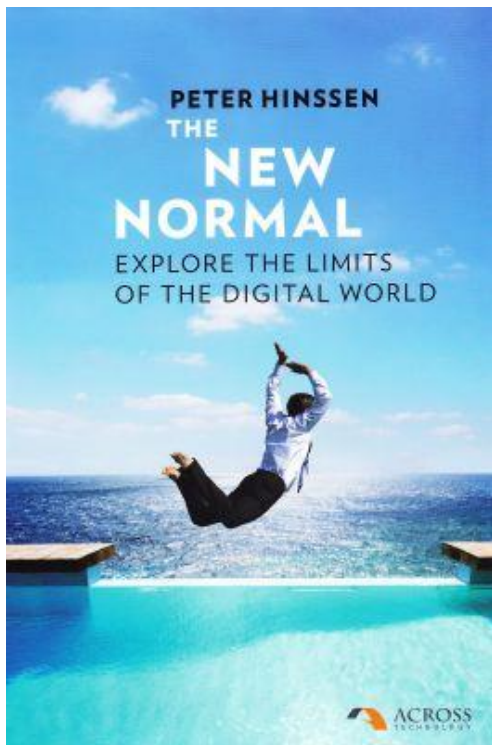
Steps away from procedures that rely on “one-size fits all”
= A truly individualized learning program

Additional possibilities:

Link unknown **vocabulary items** to
known vocabulary items:

Thesauri, dictionaries

Semantic Vector Spaces



Conclusion...

NLP and CALL

Provide a world of possibilities

- | | |
|--------------------------|------------------------|
| Analyse Native language | - Resource creation |
| Analyse Learner language | - Exercise Correction |
| Measure complexity | - Intelligent tutoring |

Provide the technological means to create ready to use applications

Future possibilities include visualization to enhance understanding, immediate feedback systems, individualized learning systems

=> Warranted qualified optimism about NLP-enhanced CALL

Contact

Piet Desmet

Piet.Desmet@kuleuven.be



be.linkedin.com/in/pietdesmet



[@PietDesmet](https://twitter.com/PietDesmet)

ITEC

www.kuleuven.be/itec

