« UNE ZONE INFLAMMATOIRE SÉQUELLAIRE D'OESOPHAGITE **PEPTIQUE** »: FUNCTIONAL ANALYSIS OF THE EXPRESSION OF **CAUSE/EFFECT PATTERNS IN MEDICAL RECORDS**



Agathe PIERSON¹ Louise-Amélie COUGNON² (1) PhD Student (FSR), CENTAL, IL&C, UCLouvain (2) FNRS Post-doc, CENTAL, IL&C, UCLouvain

INTRODUCTION

The medical language is admitted as a *sublanguage* "Certain proper subsets of the sentences of a language may be closed under some or all of the operations defined in the language, and thus constitute a sublanguage of it" (Harris 1971).

- Sublanguage:
- Circumscribed subject area
- Common vocabulary
- Common habits of word usage (Hirschman & Sager 1982)



- Causality is a phenomenon that reaches all these linguistic levels (Nazarenko 2002).
- *Causality* is everywhere in language because it helps to (Gross 2009):
 - dispel uncertainties
 - understand links between various events or states
 - structure the universe around us



RESEARCH QUESTIONS

- 1. Are there **special features** relating to the use of causal expressions in medical language?
- 2. Are there **contextual factors** that have an influence on the use of a particular type of causal expression in written medical language?
- 3. Which **function** does each type of causal expression correspond to, according to these factors, in medical language?

59%

CORPUS: iMediate

- 226,106 anonymised medical texts divided into 1,000 patient's medical files between 1996 and 2014
 - → 88 million words
- Collected in six services of a Brussels hospital (gastroenterology, abdominal surgery, bariatric surgery, MRI, ultrasound and scanners)
- Gathering different sub-genres of medical texts (Patient Discharge Summary, Radiology Reports, etc.)



each

STEP 2

Creating a French reference corpus

on the basis of the composition of the CRFC (Siepmann, Bürgel & Diwersy 2016)

	Informal spoken	Spoken	
-		Pseudo-spoken	
	Stage plays and film scripts		
	Film and daily soap subtitles		1 million words for each componant
	Text messages		
	Discussion forums (Doctissimo)		
	Formal oral (speeches, news)	Pseudo-written	
	Academic texts	Written	Sum: 11 million words
	Non-academic texts		
	Prose fiction		
	Diaries and blogs		
	Newspapers		

STEP 3

Comparing between medical professional language and general language Manual annotation of one sample of both corpora (WebAnno 2014)

	Number of words	WebAnno	Edited by R		
iMediate corpus	2,773	0.96	0.69		
Reference corpus (RC)	62,522	0.97	0.7		
Intra-annotator agreement					

	General categories of causal expressions	Precise categories of causal expressions
iMediate corpus	χ ² = 67.880 (p = 0.001)	χ ² = 934.190 (p = 0.001)
Reference corpus	Cramer's V = 0.141 (p = 0.001)	Cramer's V = 0.523 (p = 0.001)

Chi-squared and Cramer's V between iMediate corpus and French reference corpus

STEP 5

Determining whether the 12 contextual factors have an impact on the types of causal expressions or not.

 20°

12%





the patient • Number of documents • Patient's age





disease

Temporality of the disease



□ Gastroenterology

Bariatric surgery

□ MRI

Ultrasound

Scanners

■ Abdominal surgery

Factors related to the document

- Sub-genre of the text
 - Date of writing
 - Place of writing
 - Service
- Length of the text

Using statistic tests between contextual variables and the 'causal expression' variable

STEP 6

Identifying relevant factors and hypotetical functions for each type of causal expression, depending on the context.

	Categories of Explicit causal expressions				
Services	χ² = 2473.022 (p = 0.001) Cramer's V = 0.169 (p = 0.001)				
Statistical tests for the variables Services and Explicit causal expressions <u>Conclusion:</u> The variable « Services » is relevant, but only 3% (V ²) of explicit causal expressions are influenced by the type of service from which the file comes.					
or instance: urgery (16 ar nction, such ntervention su	There is an anomalous use of goal conjunctions in the services of abdominal and bariatric nd 25%) compared to the other services (bewteen 4 and 11%) \rightarrow It might be a clue of a specific as a substantiating function of a past medical treatment. <i>upra-claviculaire D pour lever compression vasculaire / Supraclavicular D surgery intervention to lift</i>				

vascular compression

CONCLUSION & PROSPECTS

Research Questions:

1. Are there special features relating to the use of causal expressions in medical language?

Answer: There are special features relating to the use of causal expressions in medical language \rightarrow the variability of the variability of the use of causal expressions (at the micro level). This supports the necessity of a linguistic study of causal expressions in French medical language.

2. Are there contextual factors that have an influence on the use of a particular type of causal expression in written medical language? Answer in progress: Requiring the inclusion of methods of disambiguation and incorporation of implicit causal (October 2017)

3. Which function does each type of causal expression correspond to, according to these factors, in medical language? Answer in progress: Requiring graphs that extract all types of causal expressions (November 2017). This will allow us to identify global functions (and no partial ones) in order to have a global insight of the functioning of causal expressions in French medical language. To confirm our hypothesis: field survey with a questionnaire (LimeSurvey) that will be submitted to 150 physicians (January 2018)



15th International Pragmatics Conference, Belfast (Nothern Ireland), 16-21 July 2017

