

Comparing the phrasicon of teenagers in immersive and non-immersive settings: does input quantity impact range and accuracy?

Amélie Bulon & Fanny Meunier
CECL, UCL



Outline

1. Introduction
2. Methodology
3. Results
and
4. Discussion

1. INTRODUCTION



Assessing CLIL

Linguistic, cognitive and educational perspectives

- Main objective: investigate the influence of Content and Language Integrated Learning - and other educational and motivational factors - on the acquisition of an L2 (Hilgsmann et al. 2017)

- Subjects (**928** pupils)

(= data collection at T1)

(nb: longit.data recently collected)

	CLIL Dutch	Non-CLIL Dutch	CLIL English	Non-CLIL English
Primary school pupils (455)	175	71	103	106
Secondary school pupils (473)	141	114	104	114

- Data collected : linguistic tasks (written and oral; free productions and experimental tests), cognitive tests, socio-affective questionnaires, focus-groups interviews, etc.
- Our focus: phraseology -> comparison between CLIL and non-CLIL learners :
 - differences = ?
 - + interplay between cognitive, socio-emotional and pedagogical factors = ?

Features of CLIL

Brown, H., & Bradford, A. (2017)

Overview of CBI, CLIL, & EMI: Differing Approaches and Goals

Criteria	EMI	CLIL	CBI
Aims	Content mastery → L2 learning is often incidental. May have implicit language learning aims	Dual focus content and language learning → language learning aims are explicit	Language learning
Target groups	Nonnative English speakers, native English speakers	Nonnative English learners	Nonnative English learners
Teaching staff	Subject specialist; sometimes subject specialist and language specialist in collaboration	Language specialist (often in Britain, Japan, South America); subject specialist (often in Europe); team-teaching with both content and language specialists	Language specialist
Pedagogical approaches and teaching formats	Focus on development of subject knowledge, methods depend on those favored by the discipline and instructor; expert-oriented lecture is likely	Multi-modal interactive and learner centered approaches which support subject content and language learning; team teaching possible	Language learning tasks; methods depend on those favored by the instructor
The role of language	Language as a tool (vehicular function)	Language as a tool, subject and mediator (for constructing knowledge & sharing expertise)	Language as a subject

Criteria	EMI	CLIL	CBI
Expected learning outcomes	Subject knowledge acquisition	Integrated content and language competence development; focus on L2 productive and interactive skills	Language learning outcomes. Content learning is incidental, at a general knowledge level
Assessment	Students assessed on subject content knowledge	Students assessed on language proficiency / performance and subject content knowledge	Students assessed on language proficiency / performance

Note: Adapted from Unterberger & Wilhelmer (2011)



2. METHODOLOGY

- **Initial focus:** all types of phraseological units
- **Overall research hypothesis:**
 - since **CLIL** programs in FWB a) provide **greater exposure to target input** ('intensive exposure programmes', Muñoz 2012) than traditional L2 learning programs contexts ('limited-exposure contexts', Muñoz 2012), and b) aim at **both content and language competence development**
 - we expect the phraseological lexicon ('phrasicon') of **CLIL** students to display a **greater range of phraseological units** and to be **more accurate**

- **Subjects and data:**

- 90 CLIL and 90 non-CLIL French-speaking EFL learners
- 5th graders (upper sec. school) : age approx. 16
- written productions / non-content specific but everyday topics / emails to a friend (holidays or party)
- CLIL: 29,387 words / NON-CLIL: 23,747 words

Phrasicon: extraction and analysis

A three-step process

Step 1. Word clusters extracted using Wordsmith Tools (Scott 2012)

Step 2. List of clusters manually checked, pruned and organized per category of phraseme (referential phrasemes, textual phrasemes and communicative phrasemes; see Granger & Paquot 2008)

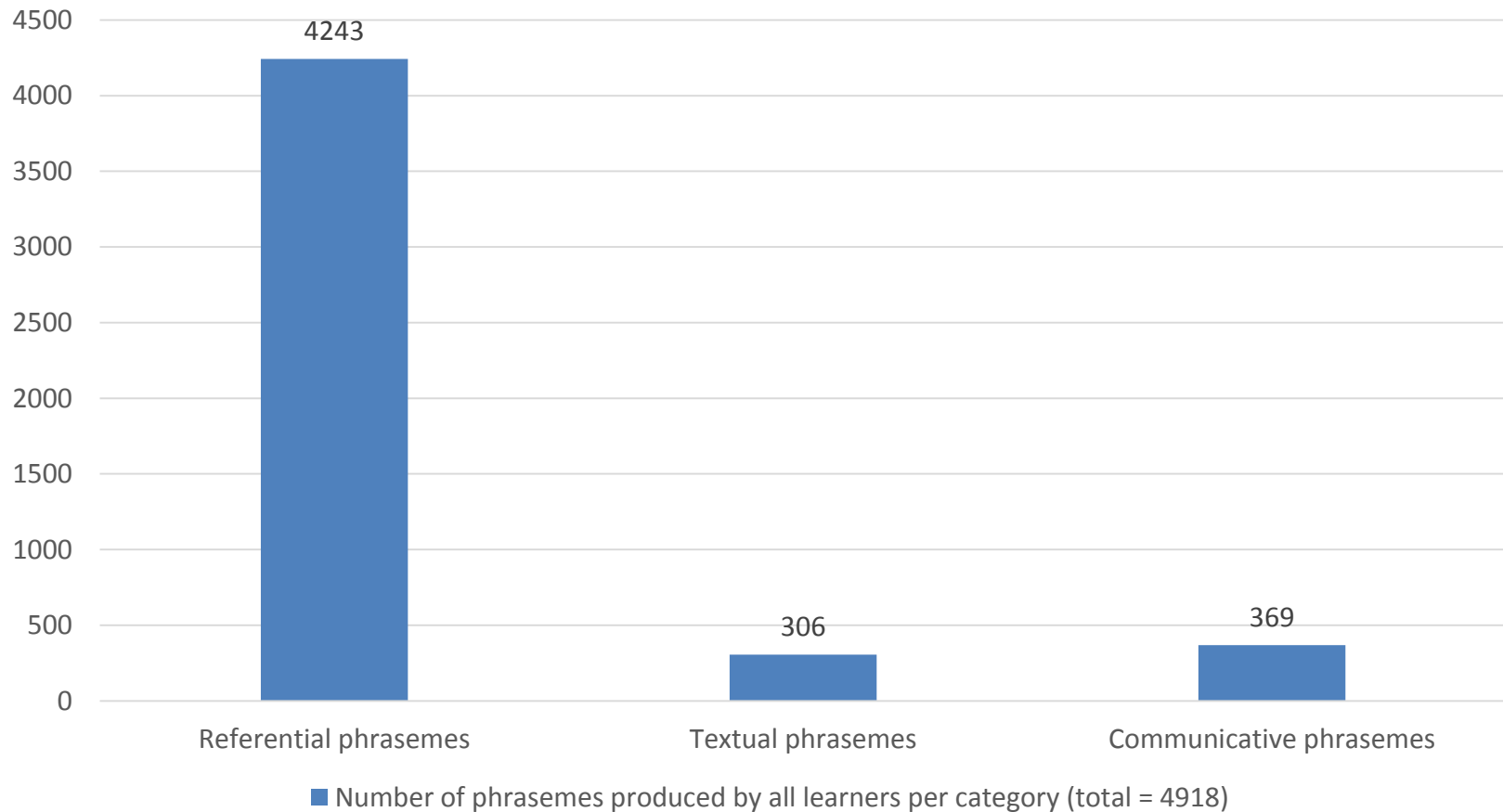
Step 3. Phraseological errors classified on the basis of Thewissen's (2008) and Hong et al.'s (2001) taxonomies

See handout for a detailed description and illustration of the taxonomies used for the classification of phrasemes and phraseological errors

3a. Results

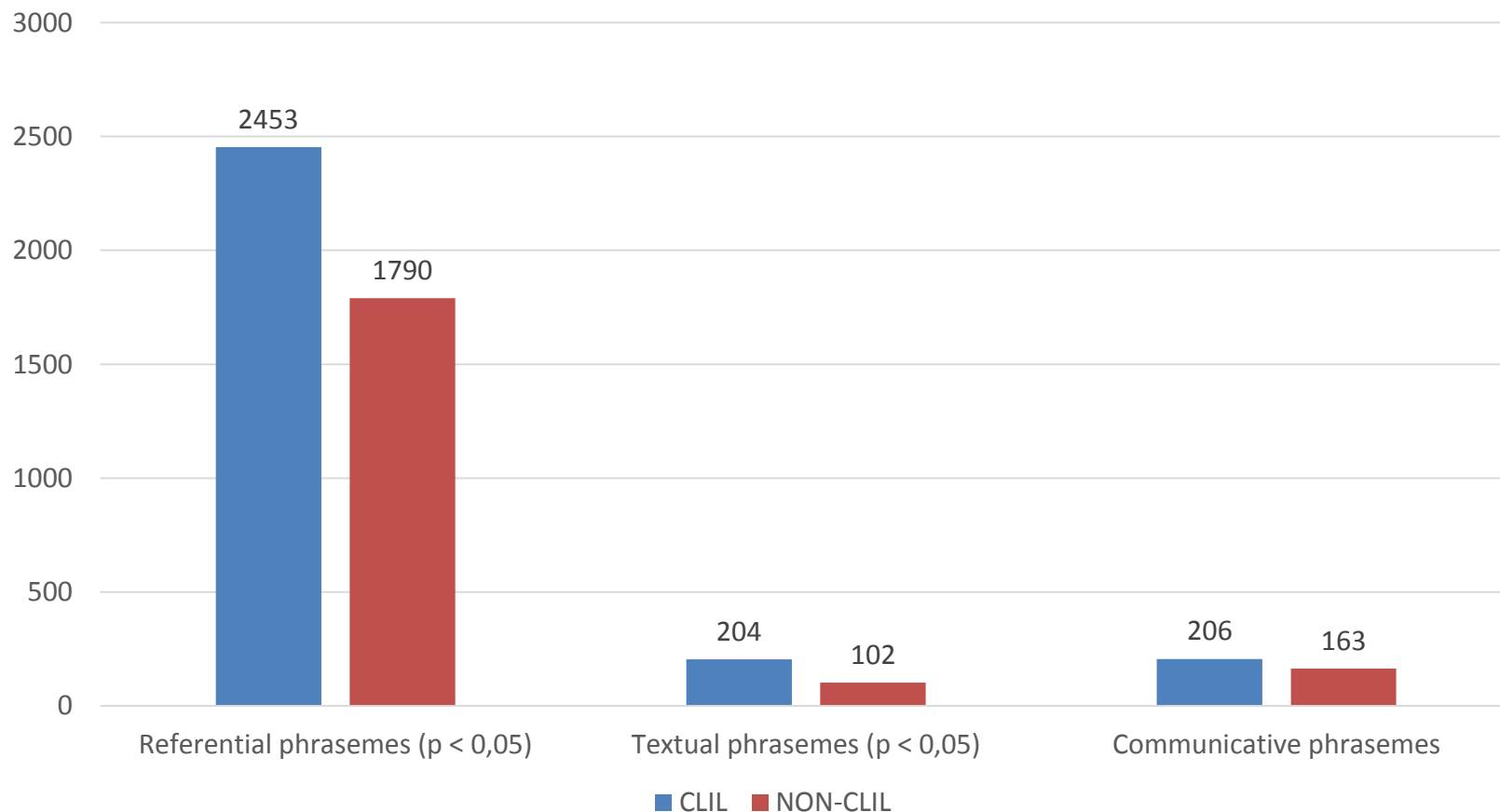
Range of phrasemes

Frequencies per category - global trends all learners



Frequencies per category - comparison CLIL/NON-CLIL

Comparison total number of phrasemes: CLIL > NON-CLIL ($p < 0,05$)



Comparison CLIL/NON-CLIL per sub-category

Number of phrasemes produced by CLIL vs NON-CLIL per sub-category

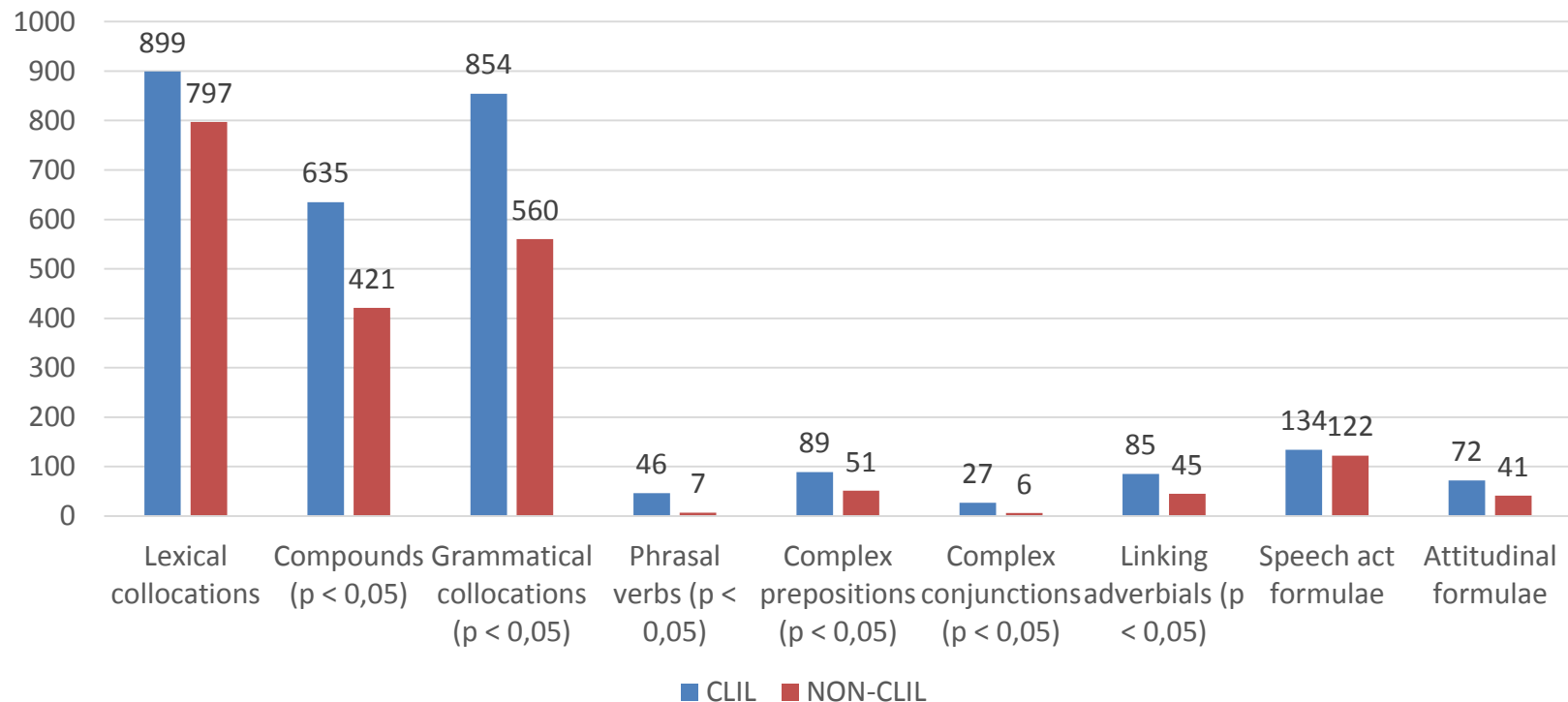


Illustration: Phrasal verbs

CLIL (13)	wake up; end up; break up; find out; sum up; clean up; pick up; turn out; throw up; go on; come up; go out (with someone); run away
NON-CLIL (2)	wake up; go out (with someone)

- Emergence criterion (Pallotti 2007: 366): **a point in time corresponding to the first systematic and productive use of a structure**
- ‘First systematic use’ = a moment in interlanguage development in which there are signs of regular, constant use of the structure
- Strictly speaking, ‘first’ should only apply to longitudinal data (see future work section).
- However, Pallotti (2007: 367) extends the concept of emergence to cross-sectional designs where emergence is seen as:
 - at least ***minimal evidence for systematic and productive use*** of a structure
 - “whether this level has, chronologically, just been attained or has characterized the interlanguage for a long time cannot of course be established with such data. In this sense, ***‘emergence’ means ‘presence’, that is the structure has emerged (at some time) and is now being used systematically and productively.***

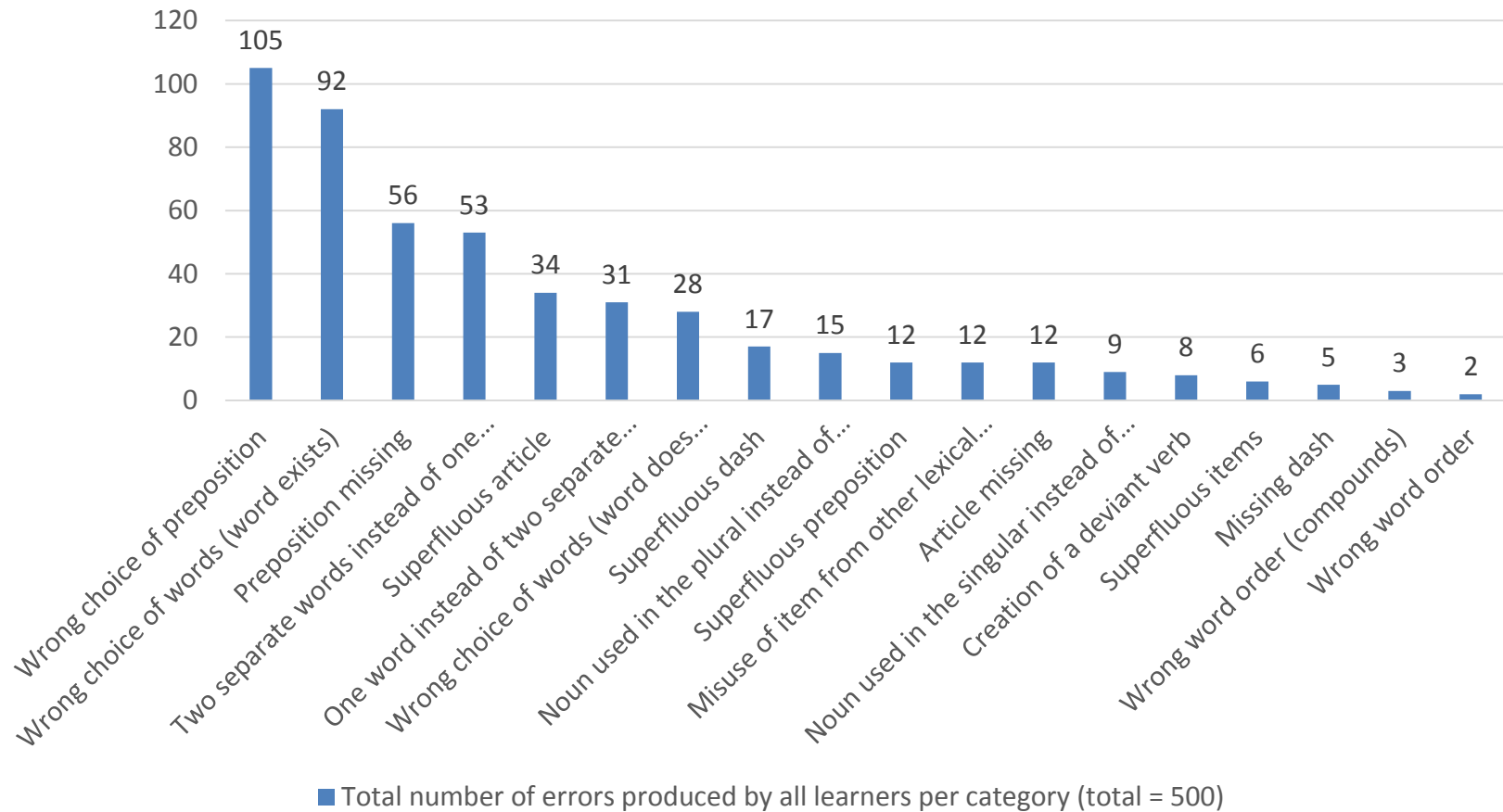
The ‘apperceived input-comprehended input-intake-integration-output’ (Gass, 1997) or ‘input-interaction-output’ process (Gass & Mackey, 2006)

- Phrasal verbs in **output** (learner's overt manifestation of L2 knowledge) may be influenced, i.a., by frequency in input
- In line with recent findings: Nicaise’s 2015 study of teacher talk (native teachers in CLIL vs non-native teachers in non-CLIL settings) :
 - 4 to 1 ratio in the relative frequencies of phrasal verbs produced by native teachers during classroom teacher talking time (*native teachers display a tendency to resort more naturally to phrasal verbs* (Nicaise 2015: 185)).

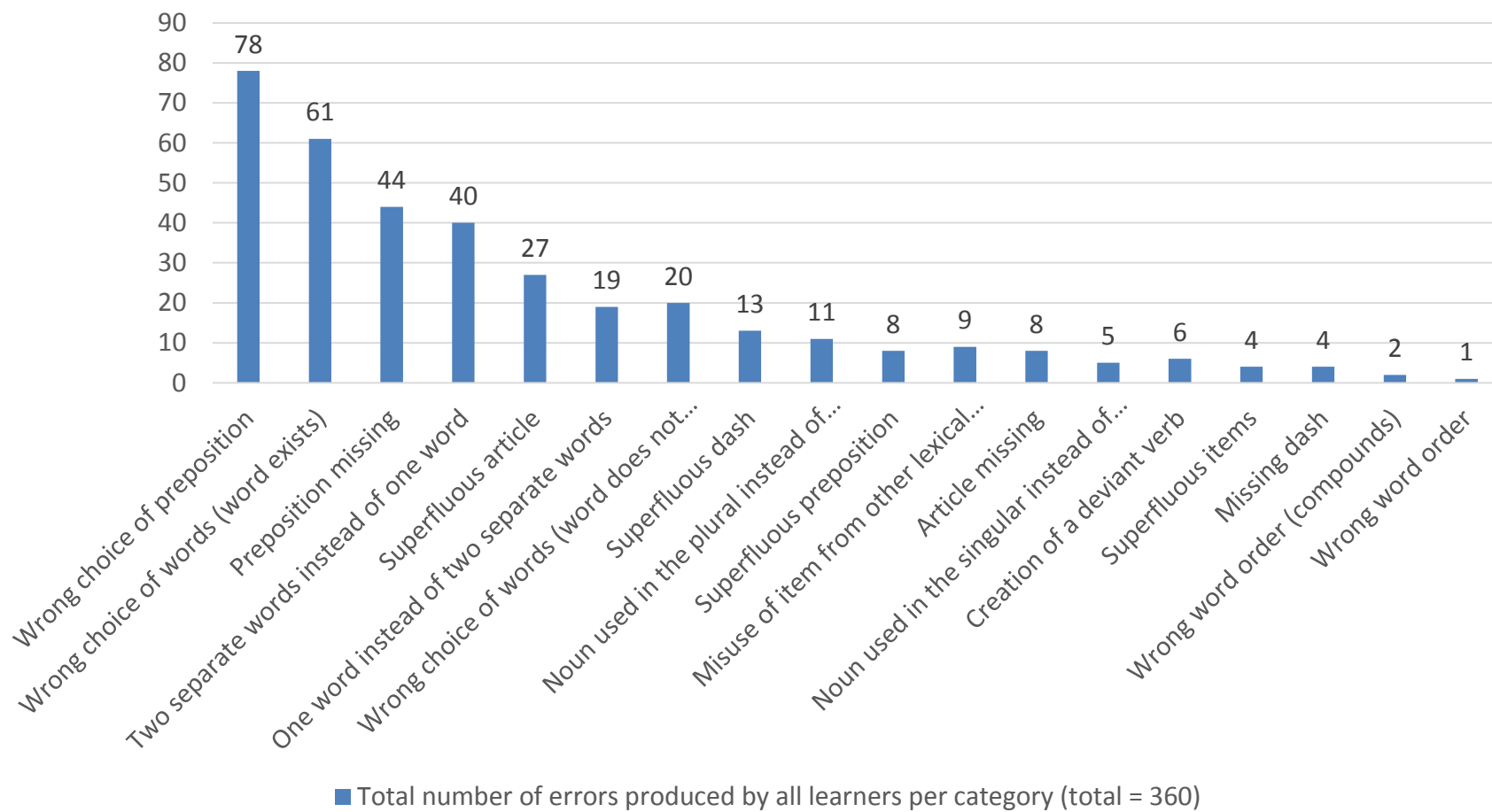
3b. Results

Phraseological errors

Frequencies per category - global trends all learners



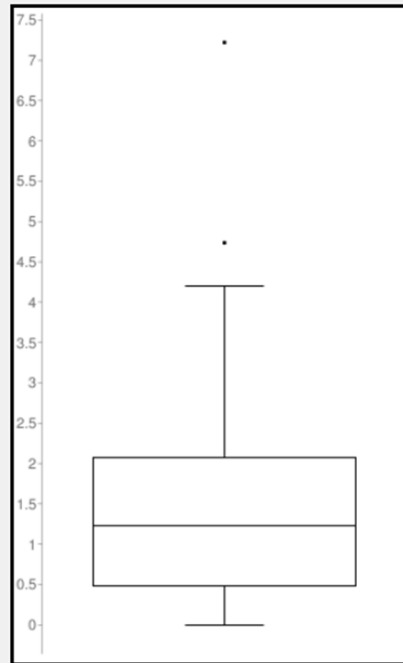
Frequencies per 200 words - global trends all learners



Errors per 200 words CLIL vs NON-CLIL

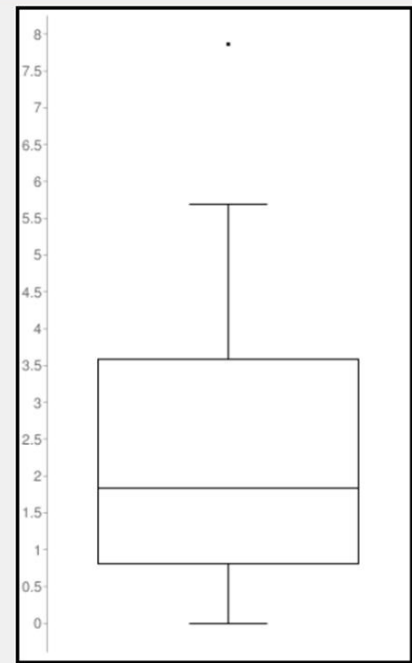
Population size: 90
Median: 1.22704005
Minimum: 0
Maximum: 7.21649485
First quartile: 0.49041175
Third quartile: 2.0729115975
Interquartile Range: 1.5824998475
Outliers: 7.21649485 4.73684211

[Mean](#)
[Geometric mean](#)
[Harmonic mean](#)
[Median](#)
[Mode](#)
[Standard Deviation](#)
[Variance](#)
[Mean absolute deviation](#)
[Range](#)
[Interquartile range](#)
[Quartiles](#)
[All dispersion data](#)



Population size: 90
Median: 1.832167832
Minimum: 0
Maximum: 7.860262009
First quartile: 0.814134344
Third quartile: 3.5918134375
Interquartile Range: 2.7776790935
Outlier: 7.860262009

[Mean](#)
[Geometric mean](#)
[Harmonic mean](#)
[Median](#)
[Mode](#)
[Standard Deviation](#)
[Variance](#)
[Mean absolute deviation](#)
[Range](#)
[Interquartile range](#)
[Quartiles](#)
[All dispersion data](#)



Data is from: ☒ Population ☐ Sample

CLIL : significantly fewer errors

- One-tailed t-test

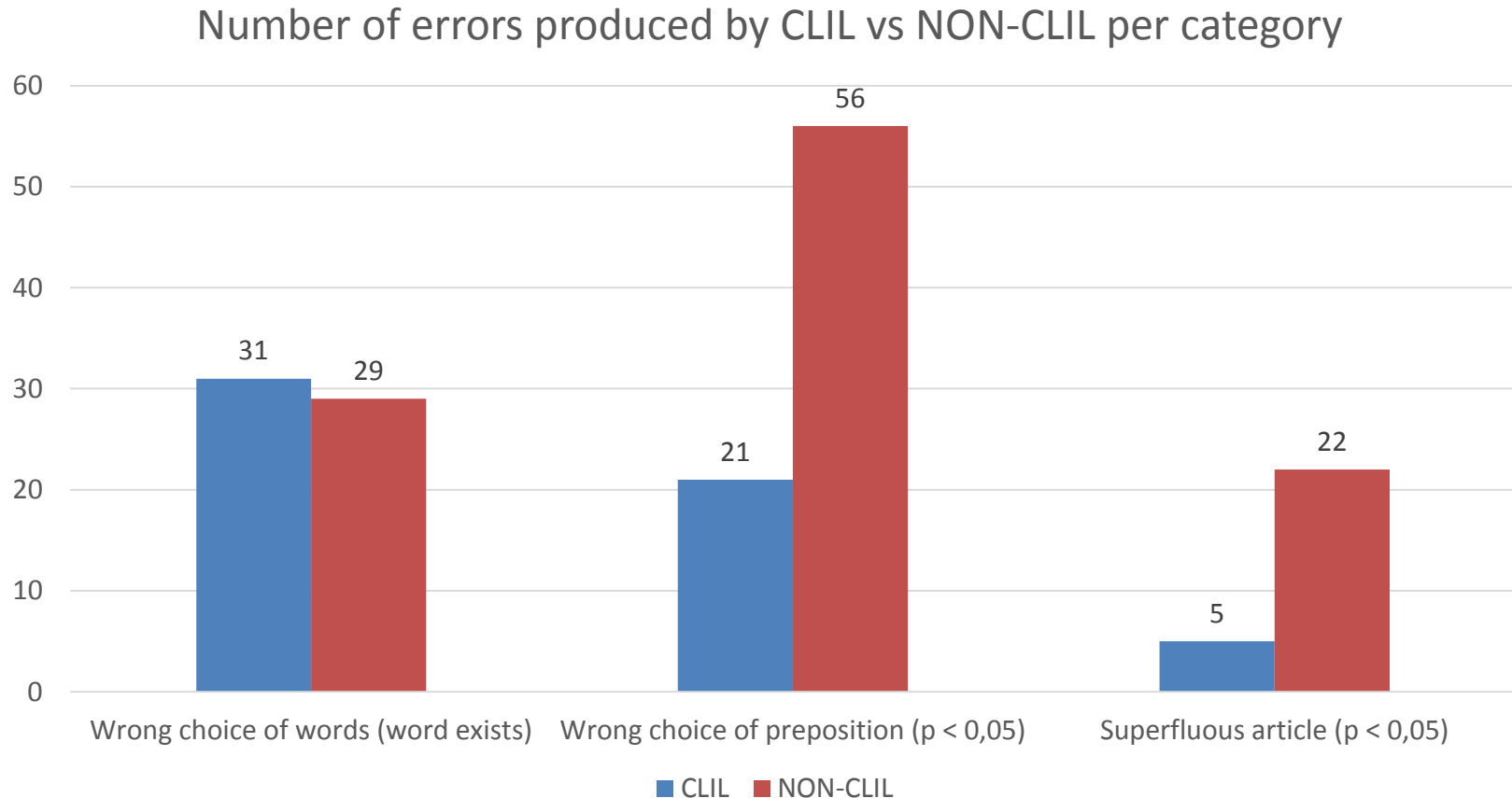
- t -value -3.56009
- p -value .000238.

- Mann-Whitney-Wilcoxon test

$U_A =$	5211.5	$P_{(1)}$	$P_{(2)}$
$z =$	-3.32	0.0005	0.0009

Comparison CLIL/NON-CLIL :

3 categories as illustrations



Examples

- Wrong choice of existing word => mostly verbs in lexical collocations and nouns in compounds
 - VERB + NOUN: *fished** a fish, *pass** holidays, *made** fun, had an *argue**
 - compounds: *alcohol** games, *attraction(s)** park, tennis *field**
- Wrong choice of preposition => mostly in grammatical collocations and complex prepositions
 - PREP + NOUN (more in CLIL): *in** the airport, *on** the afternoon/evening
 - VERB + PREP (more in NON-CLIL): think *at**, tell *for**, giving sth *at** sb
 - complex prepositions: *at** the top of
- Superfluous article
 - influence of L1 (French): all *the** night (CLIL), *the** next week (CLIL), all *of the** night (NON-CLIL), all *of the** day (NON-CLIL), *the** last night (NON-CLIL), *the** next year (NON-CLIL), eat *a** barbecue (NON-CLIL)

Grammaticality and acceptability errors (James 2013)

General trend at this stage of the analysis

- GRAMMATICALITY ERROR

= formally inexistent phraseme (very close approximation or no resemblance)

ex: « She is in love *for** him » [with]; « He *made a balad** » [went for a walk]

- ACCEPTABILITY ERROR

= formally existing phraseme but inappropriately used in context

ex: « She *participated in the party** » [attended the party] (in reference to the fact that she came to the party but was not actively involved in its organization, etc.)

- Both CLIL and NON-CLIL learners produced more grammaticality errors than acceptability errors

4. DISCUSSION

Re. RANGE

- CLIL learners significantly produced more phrasemes overall, and this in almost all (sub-)categories
- Greater variety in the CLIL group
- No statistical difference was found between the groups regarding:
 - lexical collocations
 - communicative phrasemes (both speech act and attitudinal formulae)

Re. ERRORS

- NON-CLIL learners significantly produced more errors overall
- NON-CLIL learners significantly made more mistakes in the choice of prepositions and the use of articles (overuse)
- Both groups of learners made mistakes in the selection of verbs in lexical collocations and nouns in compounds
- Both groups of learners produced more grammaticality errors than acceptability errors

Future work

- Look at the sub-categories within lexical and grammatical collocations
- Longitudinal data collected now = > further analysis of the data (from emergence to acquisition?)
- Oral data has been collected: comparison speech vs writing
- Ongoing multivariate and regression analyses: cognitive influence, input measures, etc.

References 1/2

- Brown, H., & Bradford, A. (2017). EMI, CLIL, & CBI: Differing approaches and goals. In P. Clements, A. Krause, & H. Brown (Eds.), *Transformation in language education*. Tokyo: JALT.
- Granger, S. & Paquot, M. (2008). Disentangling the phraseological web. In S. Granger & F. Meunier (Eds.), *Phraseology : an interdisciplinary perspective*. John Benjamins, Amsterdam/Philadelphia, 27-49.
- Hiligsmann, P., Van Mensel, L., Galand, B., Mettewie, L., Meunier, F., Szmalec, A., Van Goethem, K., Bulon, A., De Smet, A., Hendrikx, I. & Simonis, M. (2017) Assessing Content and Language Integrated Learning (CLIL) in French-speaking Belgium: linguistic, cognitive, and educational perspectives. In *Cahier de Recherche du Girsef*, n° 109, 1-24.
- Gass, S. (1997) *Input, Interaction, and the Second Language Learner*, Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Gass, S. & Mackey, A. (2006) Input, Interaction and Output. An overview. In K. Bardovi-Harlig & Z. Dörnyei (eds). *Themes in SLA. AILA Review* 19, 3–17.
- Hong, A. L., Rahim, H. A., Hua, T. K. & Salehuddin, K. (2001). Collocations in Malaysian English learners' writing: A corpus-based error analysis. 3L; Language, Linguistics and Literature, *The Southeast Asian Journal of English Language Studies*, 17 (special issue), 31-44.

References 2/2

- James C. (2013). *Errors in language learning and use: Exploring error analysis*. London and New York: Routledge.
- Muñoz, C. (Ed.) (2012). *Intensive exposure experiences in second language learning* (Vol. 65). Multilingual Matters.
- Nicaise, E. (2015). Highlighting linguistic features of native English teacher talk as a reference framework for French-speaking EFL teachers : a corpus study. UCL, Louvain-la-Neuve.
- Pallotti, G. (2007) An Operational Definition of the Emergence Criterion. *Applied Linguistics* 28/3: 361–382.
- Scott, M. (2012). *WordSmith Tools*. Liverpool: Lexical Analysis Software.
- Thewissen, J. (2008). The phraseological errors of French-, German-, and Spanish speaking EFL learners: Evidence from an error-tagged learner corpus. *Proceedings from the 8th Teaching and Language Corpora Conference* (TaLC 8). In Associação de Estudos e de Investigação Científica do ISLA-Lisboa, 300-306.

Q&A time

- Please contact us

amelie.bulon@uclouvain.be

fanny.meunier@uclouvain.be