

Dutch compound constructions in learner language:

Cross-linguistic influence and exposure effects

Kristel Van Goethem

In close collaboration with Isa Hendrikx



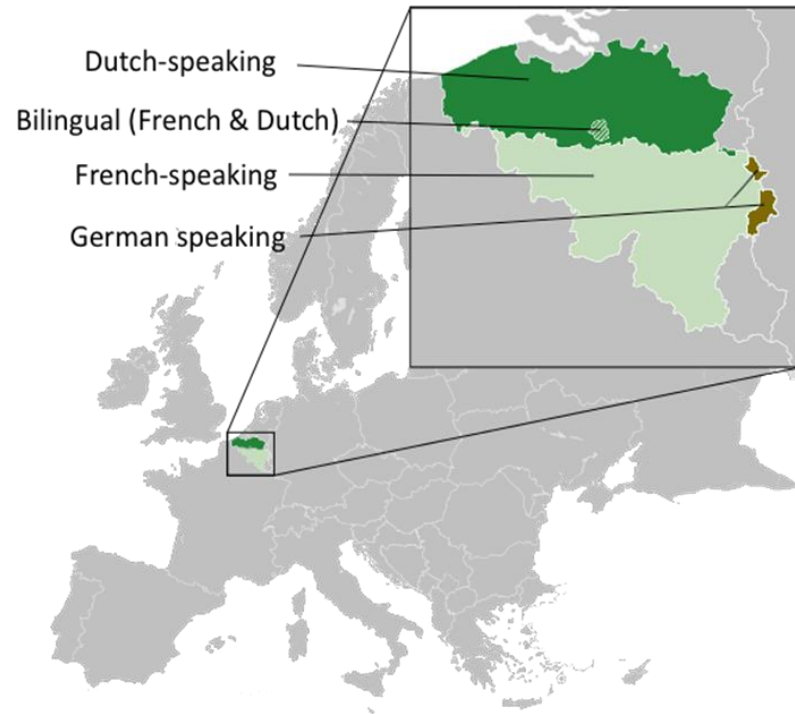
Hendrikx, I. & Van Goethem, K. 2024. Dutch compound constructions in additional language acquisition: a diasystematic-constructionist approach. *Constructions and Frames*.

« Foreign language learning is construction learning – what else? » (Herbst 2016)



- **BUT:** foreign (additional) language learning is a **complex** process!
 - What is the impact of **cross-linguistic differences** between the L1 and the additional language (especially differences in degree of analyticity)?
 - What is the impact of **exposure** on additional language learning?
- **Case study:** Acquisition of Dutch compound constructions by French-speaking learners in Belgium

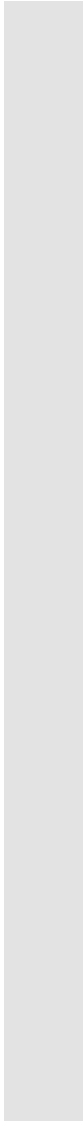
How will this talk challenge CxG?



“additional” language learning in Belgium

Outline

1. State of the art
 - Cross-linguistic differences: Dutch and French compound constructions
 - Exposure effects: CLIL
2. Diasystematic Construction Grammar
3. Aims of the study and research hypotheses
4. Corpus data and methods
5. A multilevel constructional analysis
6. Discussion and conclusions



1. State of the art

Cross-linguistic differences

Compounds and phrases in Dutch and French

- From a **CxG perspective**, “syntactic constructions may form an alternative to, or compete with the morphological expression of semantic and grammatical content” (Booij 2002: 301)
- In both Dutch and French, morphological and syntactic patterns are available to build **multi-word expressions**
 - Du. *wetenschapsbeleid* ‘science policy’ vs *wetenschappelijk beleid* ‘scientific policy’ (Booij 2019: 105)
 - Du. *frisdrank* ‘soft drink’ vs *sterke drank* ‘hard liquor’
 - Fr. *village(-)vacances* vs *village de vacances* ‘holiday resort’
- However, languages may significantly vary with respect to their **degree of analyticity** (cf. ‘Germanic / Romance Sandwich’, e.g. van Haeringen 1956, Hüning et al. 2006, Lamiroy 2011)
 - Dutch is a more synthetic language than French
 - **Dutch has a stronger tendency towards compounding than French** (Van Goethem 2009; Van Goethem & Amiot 2019)



'the free-wifi-on-the-bus-feeling'

'wet-hair-DIM-television'

Vanavond start "De droomfabriek" opnieuw



Media & Televisie

"Het wordt echte nattehcartjestelevisie". Bart Peeters en Gloria Monserec heropenen "De droomfabriek"

Examples of Dutch compounds

Cross-linguistic differences

Three important **differences** between compounding in Dutch and French:

1. **Productivity:** Compounding is more productive in Dutch than in French (Van Goethem 2009; Van Goethem & Amiot 2019)
 - Du. *badkamer* vs Fr. *salle de bains* 'bathroom'
 - Du. *zoetwaterzwembad* vs Fr. *piscine d'eau douce* 'freshwater pool'
2. **Word order:** Dutch has right-headed compounding, whereas French has left-headed compounding
 - Du. *bloemkool* vs Fr. *chou-fleur* 'cauliflower'
3. **Demarcation compound/phrase:** The distinction between compounds and (lexicalized) phrases is more clearly established in Dutch than in French (spelling, prosody, linking elements, loss of adjectival inflection)
 - Du. *frisdrank* vs Fr. *boisson rafraîchissante* 'soft drink'



What is the **impact of these cross-linguistic differences** on the acquisition of Dutch compounds by French-speaking learners?

Exposure effects: CLIL

- **Content and Language Integrated Learning (CLIL)** is a pedagogical approach promoted by the European Union to enhance widespread multilingualism (Coyle, Hood, & Marsh 2010)
- Although the CLIL approach has been extensively documented internationally (e.g., Ruiz de Zarobe et al. 2011, Rumlich 2016), its **impact on AL acquisition** remains unclear (Dalton-Puffer 2007)
- **Multidisciplinary project on CLIL** (2014-2019, UCLouvain & Université de Namur) exploring the linguistic, cognitive, educational and socio-affective effects of CLIL in French-speaking Belgium (cf. Hiligsmann et al. 2017; Van Mensel & Hiligsmann 2023)
 - General positive effect of CLIL, especially on AL Dutch
 - But limited number of case studies: phraseology (Bulon 2019), intensification (Hendrikx 2019)



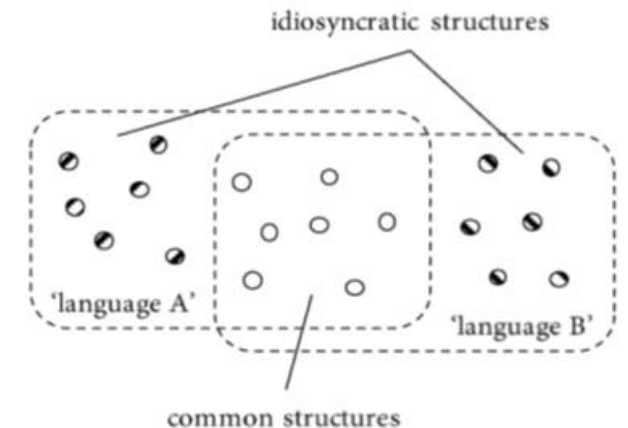
What is the **impact of CLIL** on the acquisition of Dutch compounds by French-speaking learners?



2. Diasystematic Construction Grammar

Diasystematic Construction Grammar (DCxG)

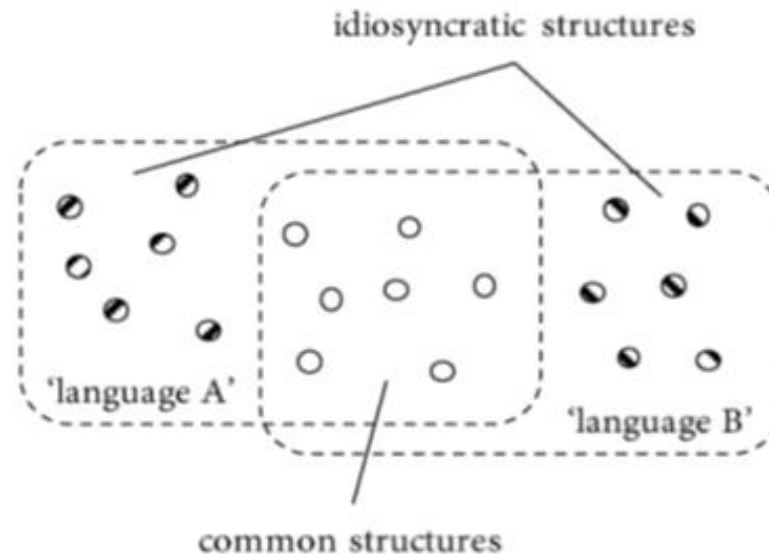
- DCxG framework developed by Höder (2012, 2014a, 2014b, 2018)
- **Integration of CxG into the study of multilingualism**
 - “multilingual speakers store and process all of their languages in terms of constructions that are organized into one common constructicon” (Höder et al. 2021: 311)
↔ earlier static conceptions of separate language systems in the bilingual’s mind
- The **multilingual constructicon**
 - contains “some constructions [that] are unspecified for language (such as abstract syntactic constructions and lexical concepts), while others (above all lexically and phonologically filled constructions) are language-specific” (Höder 2012: 247)
 - The common structures are called “**diaconstructions**”, the idiosyncratic structures are “**idioconstructions**”.



DCxG applied to AL acquisition

- DCxG first applied to **language contact** (Boas & Höder, 2018, 2021), but more recently also to **(additional) language acquisition** (Höder et al. 2021; Jach 2021; Van Goethem and Hendrikx 2021)
- **Dynamic and emergent view on AL learning:**
 - Dynamic process of constructional reorganization involving construction addition (entrenchment of newly acquired constructions) and deletion (so-called “disentrenchment” of redundant constructions) (Höder et al. 2021)
 - AL acquisition involves many cognitive reorganizational processes (interlingual identification, abstraction, generalization, categorization)
↔ ‘transfer’ too narrow to capture all these cognitive processes
- **Effects of cross-linguistic similarities and differences:**
 - Typologically similar languages have more shared ‘diaconstructions’ and learners can build on their L1 constructions to form AL productions
 - The greater the typological difference between the L1 and AL, the more reorganizational processes need to take place in the learners’ construction

Acquisition of Dutch compounds from a DCxG perspective



Language A = L1 French
Mostly phrasal constructions
[N1 Prep N2]_{NP}
(e.g., *salle de bains*)

Language B = AL Dutch
Mostly compound constructions
[N2 N1]_N
(e.g., *badkamer*)

• Reorganizational processes:

1. Dutch compounds ([N2 N1]_N, e.g., *badkamer*) need to be added to the learner's constructional network as **Dutch idioconstructions**
2. Some other constructions will have to be "disentrenched", evolving into **French idioconstructions** (e.g. [N1 Prep N2]_{NP}, e.g., *salle de bains* vs **kamer van bad*)
3. Some schematic phrasal constructions evolve into shared **diaconstructions**:
 - [N1 Prep Det N2]_{NP} (e.g., *fête pour mon anniversaire* / *feest voor mijn verjaardag*)
 - [A N]_{NP} (e.g., *meilleur ami* / *beste vriend*)

3. Aims of the study and research hypotheses

Aims of the study

- The **aims of this case study** are threefold:
 - i. to provide a **corpus analysis of the acquisition of Dutch compounds at multiple levels of abstraction** (schematic and substantive compound constructions), while exploring the impact of the **cross-linguistic differences** with French
 - ii. to investigate the impact of **additional target-language input through CLIL programs** on the acquisition of compound constructions, within the context of learning Dutch in French-speaking Belgium
 - iii. to describe and interpret the results from the perspective of **DCxG**
- Focus on **Dutch nominal compounds**

Research hypotheses

- **Research hypotheses:**
 - i. **Overgeneralization of French (phrasal) idioconstructions** (overgeneralized as shared FR-DU abstract diaconstructions):
 - Phrasal structures instead of compounds (**kamer van bad* instead of *badkamer*)
 - Left-headed compounds instead of right-headed compounds (**kamerbad*)
 - ii. **Positive impact of CLIL** on the acquisition of compound formation in AL Dutch



4. Corpus data and methods

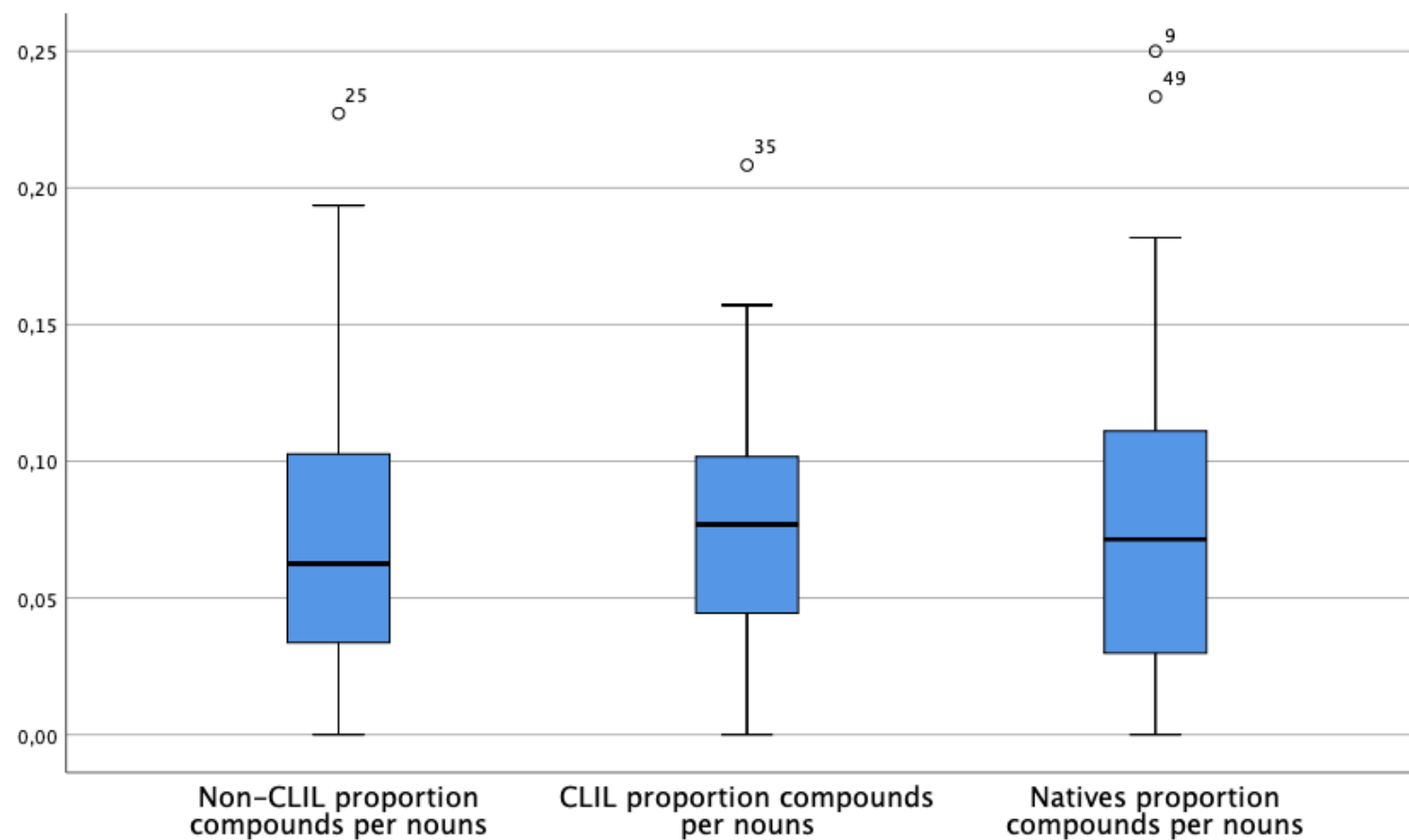
Data collection and corpus sample

- *Multilingual Traditional, Immersion and Native Corpus (MuTINCo)* (Meunier et al. 2023)
 - Written corpus collected in 2015-2017
 - 213 AL Dutch learners in 5 schools in French-speaking Belgium (CLIL + non-CLIL)
 - 59 L1 Dutch speakers (The Netherlands + Flanders)
 - 6th year of secondary school; on average 17 years old
 - Task: write an e-mail to a friend about a party or their holidays

	AL Dutch non-CLIL	AL Dutch CLIL	Native Dutch
Number of participants	84	129	59
Number of nouns	3205	5964	2146
Number of nominal compounds	240	476	155
Proportion nominal compounds per nouns	7,5%	8,0%	7,2%

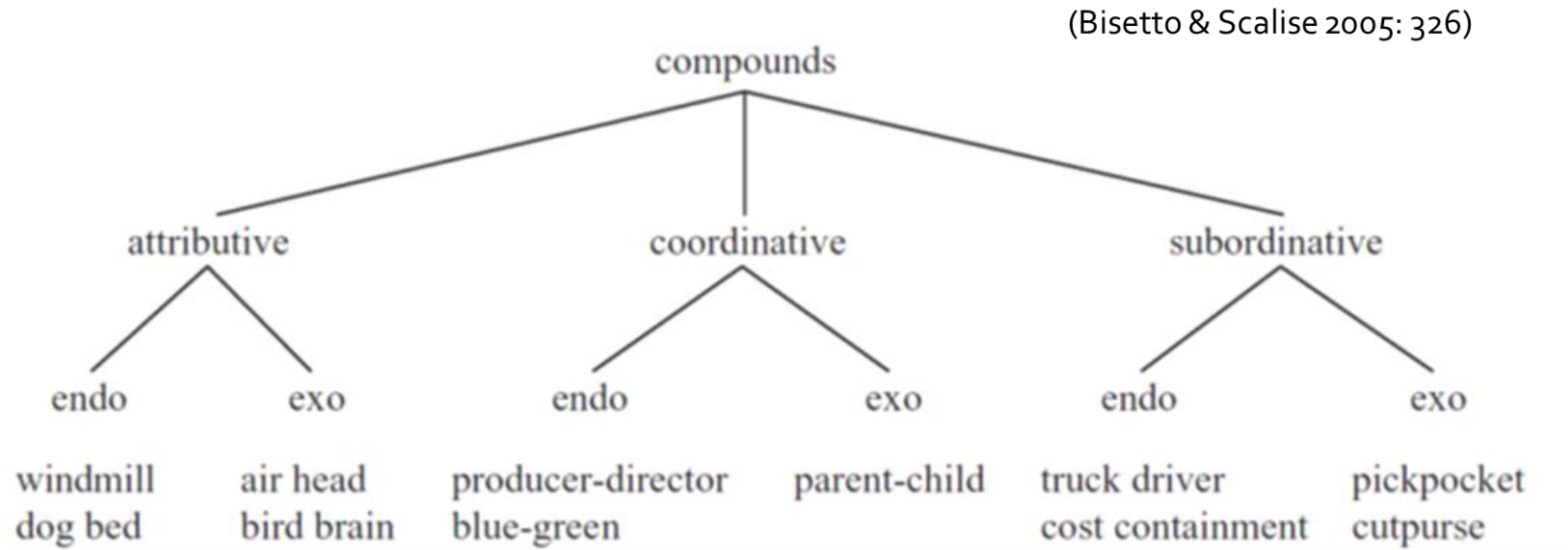
Data collection and corpus sample

Proportion nominal compounds/nouns



Data annotation

- **Formal properties:**
 - **Compound length** (number of lexical components)
 - **Compound structure** ([NN]_N, [VN]_N, [AN]_N, ...)
- **Semantic properties :**
 - **Compound function**



Data annotation

Accuracy:

- **Accuracy of the compound:** spelling mistakes, grammatical mistakes, lexical mistakes, linking morpheme, codeswitching (Fr/Eng), word order
- **Word-formation accuracy:** phrase instead of compound*, simplex instead of compound, etc.

***Checked in *Corpus Hedendaags Nederlands*: the compound had to be at least 10 times more frequent than the phrase to consider it “non-idiomatic”:**

- *mensen van mijn klas* (1 hit) vs *mijn klasgenoten* (123 hits) → non-idiomatic
- *in het centrum van de stad* (1076 hits) vs *in het stadscentrum* (956 hits) → accurate



5. A multilevel constructional analysis

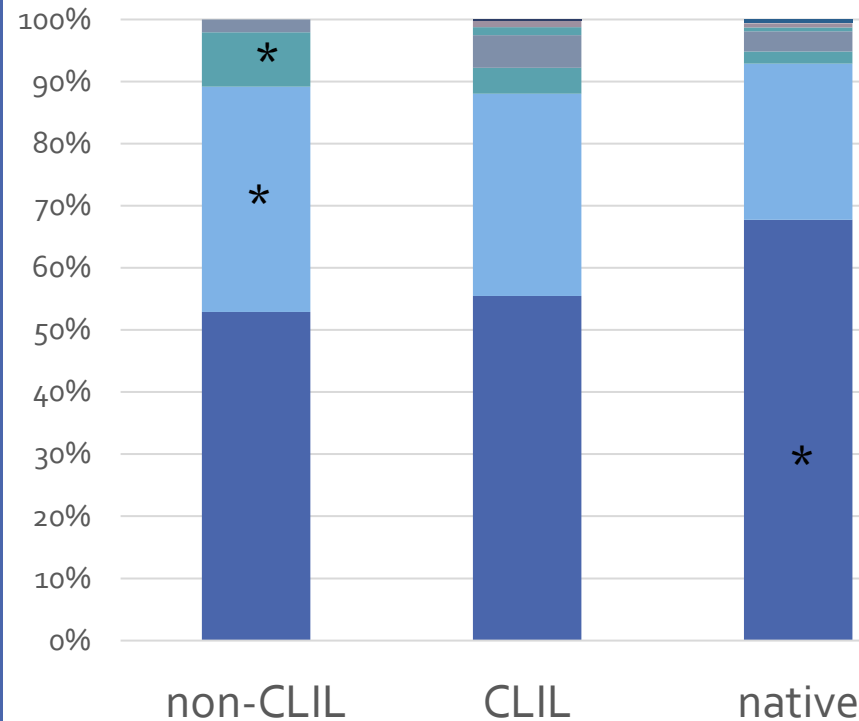
Schematic level: formal properties

Compound length	AL Dutch non-CLIL	AL Dutch CLIL	Native Dutch
2 components	227 (94,58%)	443 (93,07%)	139 (89,68%)
3 components	10 (4,17%)*	33 (6,93%)	16 (10,32%)
4 components	3 (1,25%)	0	0
Totals	240	476	155

- **2 lexical components:**
 - *Er waren veel **nachtclub** en het was niet duur.* (non-CLIL) ['night club']
- **3 lexical components:**
 - *Ik dacht vandaag weer terug aan de leuke avond die we gisteren hadden op de **kerstboomverbranding** en in de winterbar* (Native) ['lit. Christmas tree-burning']

Schematic level: formal properties

Compound structure



[N+N]_N

- Daarna trekken we meer het **regenwoud** in en gaan we daar op safari. (native) ['rainforest']

[V+N]_N

- Waren we in een mooie hotel met een **zwembad** (...) (non-CLIL) ['swimming pool']

[A+N]_N

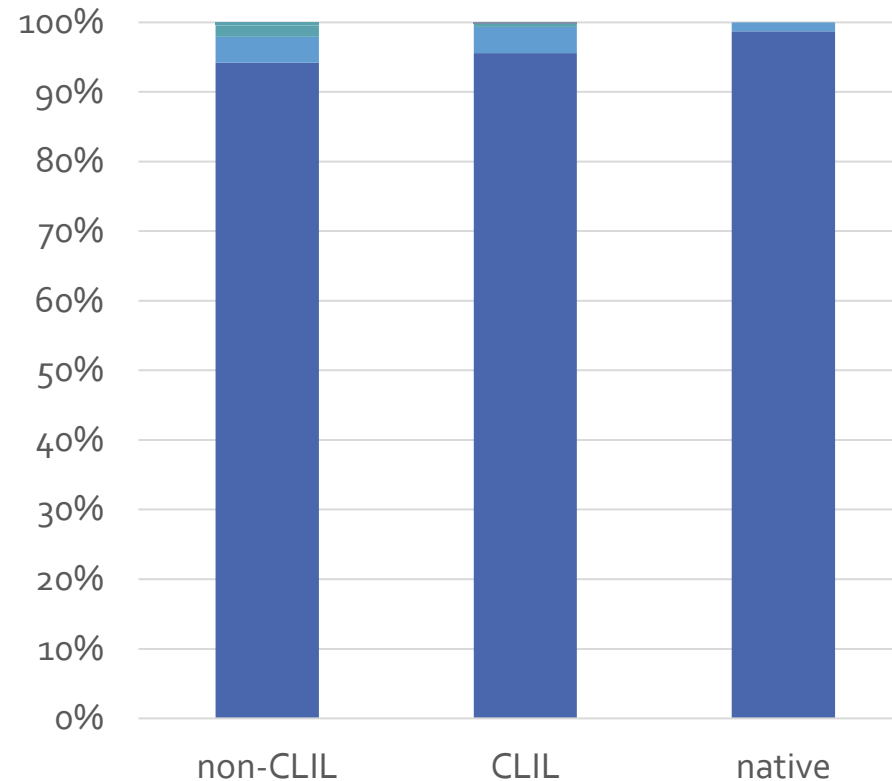
- Ze heeft een auto gehad van haar ouders, een nieuwe gsm van zijn **grootouders** en ze heeft veel geld gehad (non-CLIL) ['grandparents']

[Prep/Adv+N]_N

- Het **tegendeel** is onwaarschijnlijk, niet waar? (CLIL) ['lit. against part; opposite']

Schematic level: semantic properties

Compound function



Attributive endocentric

- *het was de verjaardag van mijn **buurjongen**.* (non-CLIL) ['neighbour boy']

Attributive exocentric

- *We konden ook alcohol drinken zelfs als we nog **snotneuzen** zijn, ik moet toch aanvaarden dat ik vandaag een kater heb* (CLIL) ['lit. snot noses; brats']

Subordinative endocentric

- *Daarna hebben we de GSM gedragen met een **haardroger** en we hebben zijn geld op de chauffage gezet.* (CLIL) ['hairdryer']

Subordinative exocentric

- *Er waren veel mensen in de straten en wij waren bang voor **pickpocketen**.* (CLIL)

Substantive level: frequency

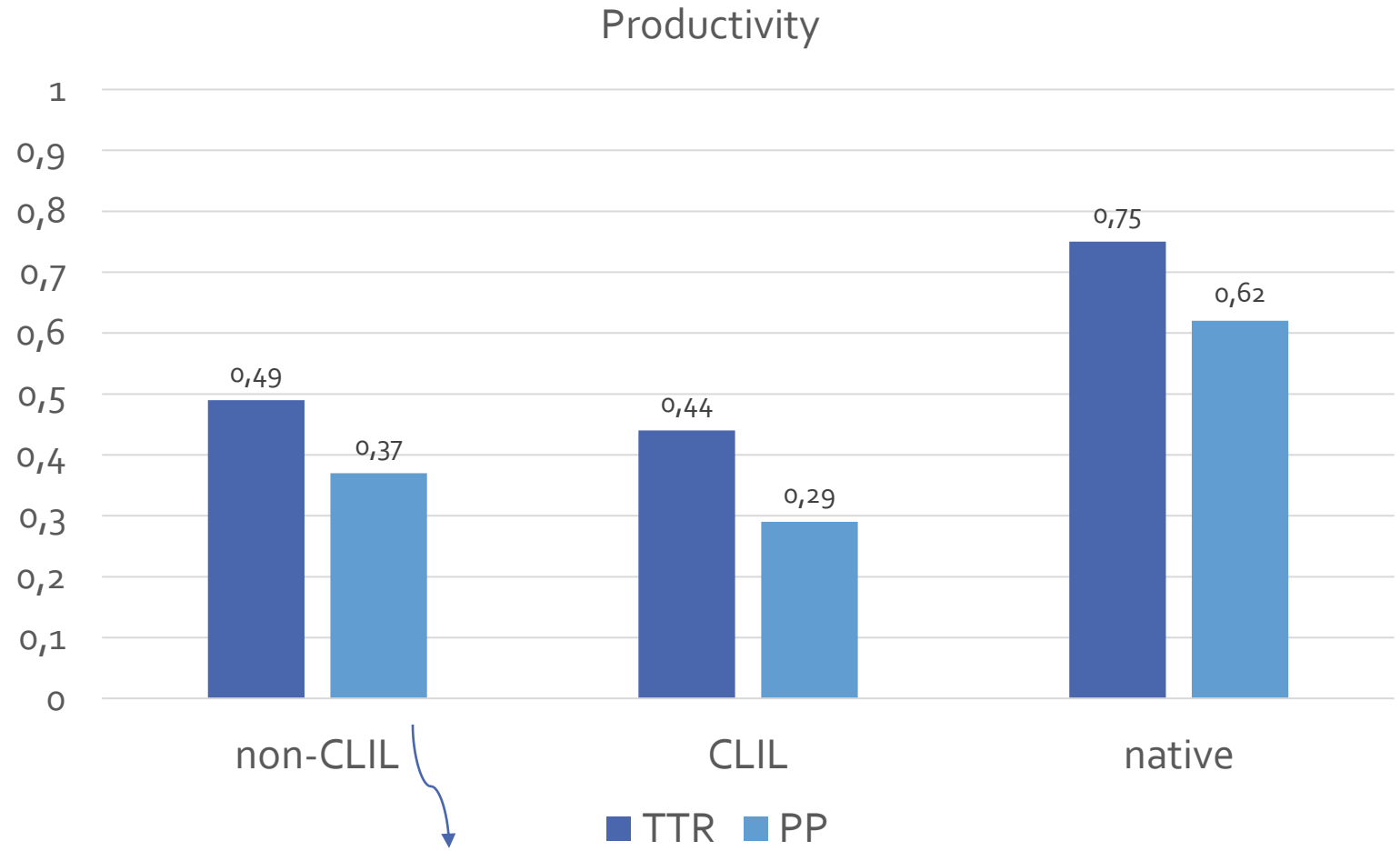
Most frequent compounds per condition

Top 5	AL Dutch non-CLIL	AL Dutch CLIL	Native Dutch
1	<u>zwembad</u> (27) 'swimming pool'	<u>verjaardag</u> (54) 'birthday'	<u>zwembad</u> (9) 'swimming pool'
2	<u>verjaardag</u> (22) 'birthday'	<u>zwembad</u> (52) 'swimming pool'	<u>verjaardag</u> (6) 'birthday'
3	<u>vliegtuig</u> (19) 'aeroplane'	<u>vliegtuig</u> (17) 'aeroplane'	<u>verjaardagsfeest</u> (5) 'birthday party'
4	**aeroport (6)	*Paasvakantie 'Easter holidays' (10), <u>verjaardagsfeest</u> (10)	°dansvloer (4) 'dance floor'
5	*voetbal (5) *volleybal (5) <u>verjaardagsfeest</u> (5)	*ziekenhuis (9) 'hospital'	°halfuur (3) 'half hour' <u>vliegtuig</u> (3) 'aeroplane' °zonsondergang 'sunset' (3)
Totals	240	476	155

- ° = occurs in native corpus, but not in the learner corpora
- * = occurs in learner corpora, but not in native corpus
- ** = misuse (codeswitching)

Substantive
level:
productivity

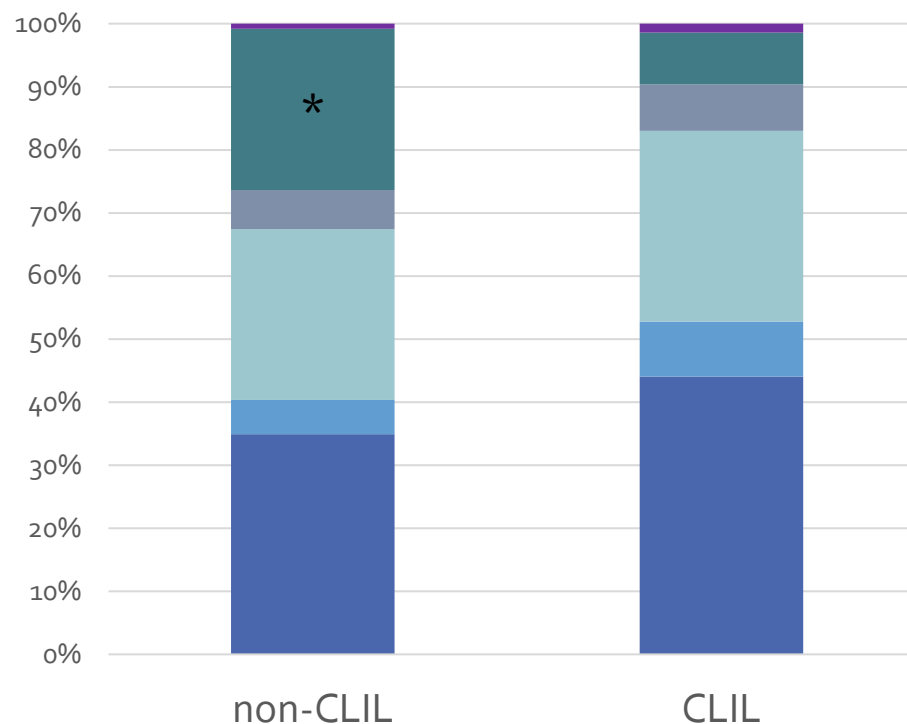
Productivity of compounds per condition



60% inaccuracies
(especially codeswitches)

Substantive level: accuracy

Accuracy of the compound



	Non-CLIL	CLIL
Total number of mistakes in compounds	129 (53,75%)	218 (45,80%)
Total number of compound nouns with mistakes	111 (46,25%)	188 (39,50%)
Total number of compound nouns	240	476

Spelling mistake

- *De vlucht was so lang en de **vlichttuig** was niet leuk.* (non-CLIL) [vliegtuig 'plane']

Grammatical mistake (gender/number)

- *Iedereen houdt van **vuurwerken** (...)* (CLIL) [vuurwerk 'fireworks']

Codeswitching (FR/ENG)

- *Het koek kwam van mijn vriend die een **aspirant-pâtissier** is.* (non-CLIL)
- *Er was een babyfoot en een **ball van football*** (non-CLIL)
- *Haar mama had een lekker **verjaardagsgateau** gemaakt (...)* (CLIL) [verjaardagstaart 'birthday cake']

Lexical mistake

- *Bovendien waren er geen proper **voetweg** zoals het uitgelegt werd in onze guide.* (CLIL) [wandelpad 'hiking trail']

Mistake in the linking morpheme

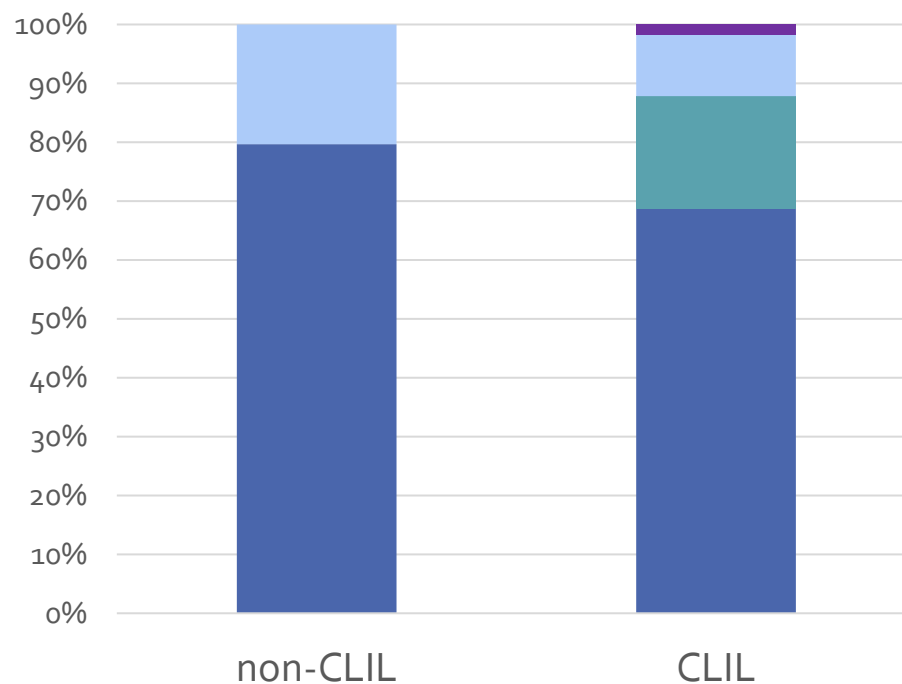
- *Daarna hebben we een **aardbeienstaart** gegeten dat de koker had gemaakt.* (CLIL) [aardbeientaart 'strawberry cake']

Mistake in word order

- *hij was in dezelfde **schoolmuziek** als jij en jullie kennen elkaar sinds ongeveer zes jaren .* (CLIL) [muziekschool 'music school']

Substantive level: accuracy

Word-formation accuracy



Phrase instead of compound

- *ik kan niet goed zwemmen maar ik had 2 **lessen van zwembad** (...).* (non-CLIL) [zwemlessen 'swimming classes']

Compound instead of phrase

- *dat was een idee op mijn zusje en mijn **bestvriend** Floriane.* (non-CLIL) [beste vriend 'best friend']

Simplex instead of compound

- *Daarna zijn we naar Venesie gegaan met een **stap** in Padoue waar we de stad, de katedraal en de klooster hebben bezocht.* (CLIL) [uitstap 'trip']

Compound instead of simplex

- *Ik was de **kookman**, ik weet het als het best idee was want ik heb nooit koken van mijn hele leven* (CLIL) [kok 'cook']

	Non-CLIL	CLIL
Total number of word-formation mistakes	19	68
Total number of nouns	3205	5964
Proportion word-formation mistakes/nouns	0.59%	1.14%*



6. Conclusions and discussion

Summary of the results

Multilevel constructional analysis:

➤ Schematic level:

- ***Compound length:***
 - Learners use 2-component compounds more frequently than natives; natives use more often longer (3-component) compounds
- ***Compound structure:***
 - $[N+N]_N$ compounds are predominant in the native and learner datasets
 - Compared to the natives, the learners use more often $[V+N]_N$ compounds (cf. high frequency of *verjaardag* 'birthday', *zwembad* 'swimming pool')
- ***Compound function:***
 - Attributive endocentric compounds are predominant in the native and learner datasets

Summary of the results

Multilevel constructional analysis:

➤ Substantive level:

- **Productivity:**
 - Lower productivity (both TTR and PP) of compounds in learner group
 - Natives use a greater variety of compounds (e.g., *dansvloer* 'dance floor', *zonsondergang* 'sunset') (despite a smaller dataset!)
- **Accuracy:**
 - Mostly spelling and grammatical mistakes in compound use (+ codeswitching in non-CLIL group)
 - 'phrase instead of compound' is the most frequent word-formation mistake in learner group

Discussion

Cf. Research hypotheses

(1) Overgeneralization of French (phrasal) idioconstructions?

- ✓ **Phrasal structures instead of compounds**
 - 'phrase instead of compound' (e.g., **lessen van zwembad*) is indeed the most frequent word-formation mistake in the learner group
 - (= generalized as FR-DU diaconstruction instead of FR idioconstruction)
 - However, also cases of overgeneralization of Dutch compound cxn (compound instead of phrase/simplex, e.g., **bestvriend*, **kookman*)
- ✗ **Left-headed compounds instead of right-headed compounds**
 - Difference in word order between FR/DU does not seem to be an obstacle (few mistakes)
 - (= DU right-headed word order pattern correctly added as DU idioconstruction by the learners)
 - ~ saliency of distinctive word order + learnt implicitly through the acquisition of frequent compounds

Discussion

(2) Positive impact of CLIL on compound formation in AL Dutch?

- Only partial positive impact of CLIL:
 - Many mistakes or non-idiomatic uses in both groups!
 - At the **schematic** level:
 - Results of the CLIL group are closer to the results of the natives with respect to ***compound length and structure***
 - At the **substantive** level:
 - Higher ***productivity*** scores in non-CLIL group, but affected by more misused types and hapaxes
 - More ***misuse in compound formation*** in non-CLIL group
 - Significantly more ***word-formation inaccuracies*** in CLIL group ('simplex instead of compound' and 'compound instead of simplex')

Discussion

Interpretation of these results?

- The **non-CLIL** learners still heavily **rely on their L1 constructions** to create compounds (significantly more codeswitching, e.g., **aeroport*)
- Especially the **CLIL** learners seem to go beyond their L1 constructions and gradually **entrench the AL Dutch idioconstructions**
 - not only at the **substantive level** (frequent compounds, e.g., *ziekenhuis* 'hospital', *vliegtuig* 'plane')
 - but also at the level of the **abstract compound schema** $[N_2N_1]_N$ to creatively form new compounds, with varying success however (e.g., **voetweg*, **kookman*, **verjaardagsgateau*)



Conclusions

- Our case study has shown:
 - **The complexity of additional language learning**
 - Some cross-linguistic differences pose more problems than others (especially the syntax-morphology issues)
 - Besides CLIL exposure, many other factors play a role (topic of the task, input from other languages, etc.)
 - Difficulty to draw a line between mistakes and creativity
 - **The benefits of the (D)CxG approach to study AL acquisition:**
 - The CxG approach allows us to treat morphological and syntactic constructions on a par and to explore them at different levels of abstraction
 - The DCxG approach allows us to do justice to the complexity of AL acquisition
 - Learner corpus research framed in DCxG provides indirect access to the cognitive restructuring processes taking place in the minds of multilingual speakers

References

- Balteiro, I. (2011). Awareness of L1 and L2 word-formation mechanisms for the development of a more autonomous L2 learner. *Porta Linguarum*, 15. 25-34.
- Bisetto, A. & Scalise, S. (2005). The classification of compounds. *Lingue e Linguaggio* 4:2. 319-332.
- Blondin, C. (2003). L'immersion linguistique dans l'enseignement fondamental en communauté française de Belgique: l'état de la question. *Journal de l'immersion*, 25(2), 19-31.
- Boas, H. C. & Höder, S. (Eds). (2018). *Constructions in Contact*. Constructional perspectives on contact phenomena in Germanic languages. Amsterdam/Philadelphia: John Benjamins. vi+316pp.
- Boas, H. C. & Höder, S. (Eds). (2021). *Constructions in Contact 2. Language change, multilingual practices, and additional acquisition* (CAL 30). Amsterdam/Philadelphia: John Benjamins.
- Boas, Hans C. & Höder, S. (2021). Widening the scope: Recent trends in constructional contact linguistics. In Hans C. Boas & Steffen Höder (eds.), *Constructions in Contact 2. Language change, multilingual practices, and additional language acquisition* (Constructional Approaches to Language 30), 2–13. Amsterdam/Philadelphia: Benjamins.
- Booij, G. (2002). Constructional idioms, morphology, and the Dutch lexicon. *Journal of Germanic Linguistics*, 14(4), 301-329.
- Booij, G. (2010). *Construction Morphology*. Oxford: Oxford University Press.
- Booij, G. (ed.) (2018). *The Construction of Words. Advances in Construction Morphology. (Studies in Morphology 4)*. Dordrecht: Springer.
- Booij, G. (2019). *The Morphology of Dutch*, Second, revised edition. Oxford UK: Oxford University Press.

References

- Chohey-Paquet, M. (2008). CLIL in French-Speaking Belgium: Transforming Paradox into Potential. In C. M. Coonan (Ed.), *CLIL e l'Apprendimento delle Lingue. Le Sfide del Nuovo Ambientato di Apprendimento* (pp. 239-250). Venezia: Libreria Editrice Cafoscarina.
- Coyle, D., Hood, P., & Marsh, D. (2010). CLIL: Content and language integrated learning. Cambridge: Cambridge University Press.
- Deacon, S. H., L. Wade-Woolley & J. Kirby. (2007). Crossover: The role of morphological awareness in French immersion children's reading. *Developmental Psychology* 43:3. 732-746.
- Dalton-Puffer, C. (2007). *Discourse in Content and Language Integrated (CLIL) Classrooms*. Amsterdam / Philadelphia: John Benjamins.
- Hendrikx, I. & Van Goethem, K. 2024. Dutch compound constructions in additional language acquisition: a diasystematic-constructionist approach. *Constructions and Frames*.
- Herbst, Th. 2016. Foreign language learning is construction learning – what else? Moving towards Pedagogical Construction Grammar. In S. De Knop & G. Gilquin (eds.), *Applied Construction Grammar*, 21–52. Berlin/Boston: De Gruyter.
- Hiligsmann, Ph. (1997). Linguïstische aspecten en pedagogische implicaties van de tussentaal van Franstalige M.O.-leerders van het Nederlands. Unpublished PhD thesis. Université de Liège.
- Hiligsmann, Ph., 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 in the French-speaking Community of Belgium: linguistic, cognitive and educational perspectives. *Cahiers du GIRSEF*, 109, 1-24. Retrieved from https://cdn.uclouvain.be/groups/cms-editors-girsef/demey/Cahier_109_final.pdf.
- Höder, S. (2012). Multilingual constructions: a diasystematic approach to common structures. In K. Braunmüller & C. Gabriel (Eds.), *Multilingual individuals and multilingual societies*. Amsterdam/ Philadelphia: Benjamins. 241-257.

References

- Höder, S. (2014a). Constructing diasystems: Grammatical organisation in bilingual groups. In T. A. Åfarli & B. Maehlum (Eds.), *Studies in Language Companion Series 154* (pp. 137-152). Amsterdam: John Benjamins Publishing Company.
- Höder, S. (2014b). Convergence vs. divergence from a diasystematic perspective. In: K. Braunmüller,
- Höder, S. (2018). Grammar is community-specific. Background and basic concepts of Diasystematic Construction Grammar. In: H. Boas & S. Höder (Eds.), *Constructions in Contact. Constructional perspectives on contact phenomena in Germanic languages*. Amsterdam/Philadelphia: Benjamins. 37-70.
- Hüning, M., Vogl, U., Van der Wouden, T., & Verhagen, A. (2006). *Nederlands tussen Duits en Engels. Handelingen van de workshop aan de Freie Universität Berlin*. Leiden: Stichting Neerlandistiek Leiden.
- Jach, D. (2021). Something I was dealing with. Preposition Placement in Multilingual Constructicons. In Boas, Hans C. & Steffen Höder (Eds.), *Constructions in Contact 2. Language change, multilingual practices, and additional language acquisition*. Amsterdam, The Netherlands: Benjamins, 340–374.
- Jeon, E. H. (2011). Contribution of morphological awareness to second-language reading comprehension. *The Modern Language Journal* 95:2. 217-235.
- Lamiroy, B. (2011). Degrés de grammaticalisation à travers les langues de même famille. *Mémoires de la Société de linguistique de Paris*, 19, 167-192.
- Lefter, M.-A. (2011). Contrastive word-formation today: Retrospect and prospect. *Poznan Studies in Contemporary Linguistics*, 47(4), p. 645-682.
- Leontjev, D., Huhta, A. & K. Mäntylä. (2016). Word derivational knowledge and writing proficiency: How do they link? *System* 59. 73-89.

References

- Meunier, F., Hendrikx, I., Bulon, A., Van Goethem, K. & Naets, H. (2023). MULTINCo: Multilingual Traditional Immersion and Native Corpus. Better-documented multi-literacy practices for more refined SLA studies. *International Journal of Bilingual Education and Bilingualism* 26:5. 572-589.
- Pasquarella, A., Chen, X., Lam, K., Luo, Y. C. & Ramirez, G. (2011). Cross-language Transfer of morphological awareness in Chinese-English bilinguals. *Journal of Research in reading* 34 (1), 23-42.
- Ruiz de Zarobe, Y., Sierra, J. M., & del Puerto, F.G. (Eds.). 2011. *Content and Foreign Language Integrated Learning. Contributions to Multilingualism in European Contexts*. Bern: Peter Lang.
- Rumlich, D. (2016). *Evaluating bilingual education in Germany CLIL students' general English proficiency, EFL self-concept and interest*. Frankfurt am Main: Peter Lang.
- Schlücker, B. (Ed.). (2019). *Complex lexical units: compounds and multi-word expressions*. Berlin: De Gruyter.
- Van Goethem, K. (2009). Choosing between A+N compounds and lexicalised A+N phrases: The position of French in comparison to Germanic languages. *Word Structure*, 2, 241-253.
- Van Goethem, K. & D. Amiot. (2019). Compounds and multi-word expressions in French. In: B. Schlücker (Ed.), *Complex lexical units: compounds and multi-word expressions*. Berlin: De Gruyter. 127-152.
- Van Goethem, K. & Hendrikx, I. (2021). Intensifying constructions in second language acquisition: A diasystematic-constructionist approach. In H. C. Boas & S. Höder (Eds.) *Constructions in Contact 2: Language change, multilingual practices, and additional acquisition (CAL 30)*. Amsterdam/Philadelphia: John Benjamins. 375-428.
- Van Haeringen, C. B. (1956). *Nederlands tussen Duits en Engels*. The Hague: Servire.
- Van Mensel, L. & Hilgsmann, Ph. (2023) Assessing CLIL: a multidisciplinary approach. *International Journal of Bilingual Education and Bilingualism* 26:5. 529-533.



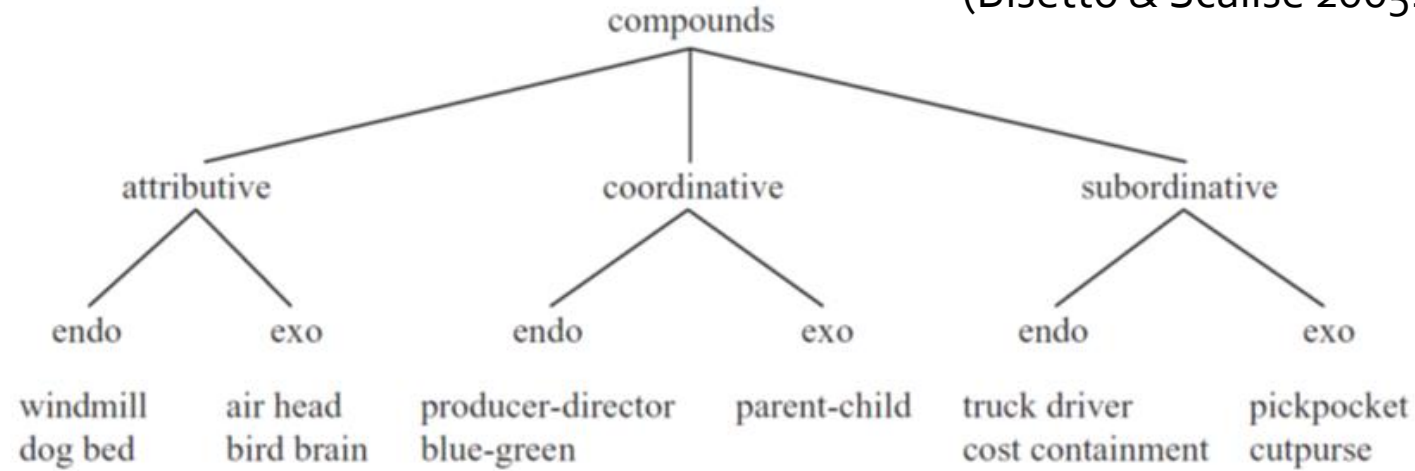
Cette photo par Auteur inconnu est soumis à la licence [CC BY](#)

Data annotation: formal properties

Formal properties	Categories of the variable	Examples
Compound length	Number of lexical components	2: <i>voetbal</i> 'football'
		3: <i>voetbalclub</i> 'football club'
Compound structure	$[N+N]_N$	<i>telefoonnummer</i> 'phone number'
	$[V+N]_N$	<i>surfplaatsje</i> 'surfing site'
	$[A+N]_N$	<i>rondtrip</i> 'round trip'
	$[Num+N]_N$	<i>vijftal</i> 'about five'
	$[NP+N]_N$	<i>onderwaterwereld</i> 'lit. under water world; the universe under the water'

Data annotation: semantic properties

(Bisetto & Scalise 2005: 326)



Compound function	Example
Attributive endocentric	<i>windmolen</i> 'windmill'
Attributive exocentric	<i>leeghoofd</i> 'air head'
Coordinative endocentric	<i>hotel-restaurant</i>
Coordinative exocentric	<i>ouder-kind(relatie)</i> 'parent-child (relationship)'
Subordinative endocentric	<i>kurkentrekker</i> 'corkscrew'
Subordinative exocentric	<i>zakkenroller</i> 'pickpocket'

Data annotation: accuracy of the compound

Accuracy of the compound	Example
Spelling mistakes	* <i>vliechtuig</i>
Lexical mistakes	* <i>lievelingseetje</i> (lievelingseten)
Grammar mistakes (gender/number)	* <i>dit e-mail</i> * <i>mijn nieuwe zonnebrilen</i>
Linking morpheme	* <i>verjaarssdag</i>
Word order	* <i>schoolmuziek</i> (muziekschool)
Codeswitching (FR/ENG)	<i>pasta met *basilic sauce</i> * <i>football</i>

Substantive level: accuracy

Word-formation accuracy	Example
Phrase instead of compound*	* <i>vakantie van Pasen</i> (Paasvakantie)

***Checked in *Corpus Hedendaags Nederlands*: the compound had to be at least 10 times more frequent than the phrase to consider it “non-idiomatic”:**
e.g. *mensen van mijn klas* (1 hit) vs *mijn klasgenoten* (123 hits) → non-idiomatic e.g.
in het centrum van de stad (1076 hits) vs *in het stadscentrum* (956 hits) → accurate

Semi-
schematic
level:
frequency and
productivity

Variation in the modifier position – TTR

	AL Dutch non- CLIL	AL Dutch CLIL	Control corpus native Dutch
[X-feest]	6 tokens, 2 types (TTR = 0,33)	14 tokens, 4 types (TTR = 0,28)	5 tokens, 1 type (TTR = 0,20)
[X-vakantie]	8 tokens, 4 types (TTR= 0,5)	14 tokens, 3 types (TTR = 0,21)	3 tokens, 2 types (TTR = 0,67)

Semi-
schematic:
frequency and
productivity

Variation in the head position - TTR

	AL Dutch non-CLIL	AL Dutch CLIL	Control corpus native Dutch
[verjaardags-X]	6 tokens, 2 types (TTR = 0,33)	14 tokens, 4 types (TTR = 0,28)	5 tokens, 1 type (TTR = 0,20)
[lievelings-X]	8 tokens, 8 types (TTR= 1)	6 tokens, 6 types (TTR = 1)	1 token, 1 type (TTR = 1)

Substantive level: accuracy hapaxes

Accuracy hapaxes

	non-CLIL		CLIL	
Spelling	20	22,47%	34	24,46%
Grammatical	12	13,48%	16	11,51%
Lexical	6	6,74%	15	10,79%
Linking morpheme	3	3,37%	11	7,91%
Codeswitching	12	13,48%*	4	2,88%
Word order	1	1,12%	2	1,44%
Compound instead of phrase	4	4,49%	0	0
Total mistakes in hapaxes	58		82	
Hapaxes with mistakes	53	59,55%	74	53,24%
Total hapaxes	89		139	

- **Spelling:** *We kwamen op **zaterdag avond** aan a een hele dag rijden . (CLIL)*
- **Codeswitching:** *Het koek kwam van mijn vriend die een **aspirant-pâtissier** is. (non-CLIL)*
- **Compound instead of phrase:** *Als cadeaus heb il twee mooie parfume van een goede **artisanaalwinkel** die ik goed kend. (non-CLIL)*

Outlook

- **Perspectives for future research:**
 - Impact of CLIL on accuracy while controlling for general vocabulary knowledge and extracurricular target-language input
 - Other morphological constructions: adjectival compounds, complex verbs, derivations
 - Follow-up project: comparison of AL learners with novice translators