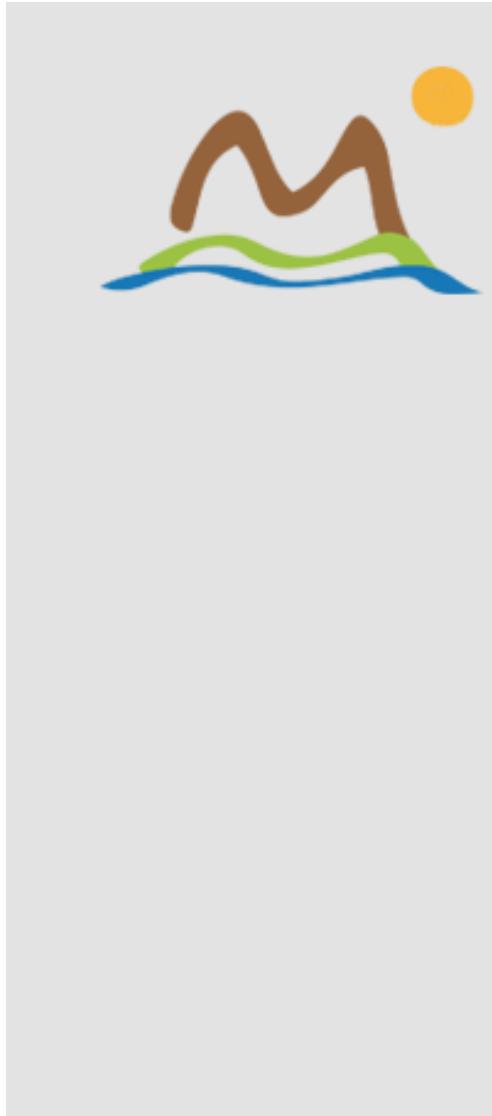


# Nominal compounds in Dutch and French

*Typological differences and additional language acquisition from a Diasystematic Construction Grammar perspective*

Kristel Van Goethem & Isa Hendrikx



# Introduction

# Introduction

- **Compound constructions in Dutch and French:**
  - In both Dutch and French, morphological and syntactic patterns are available to build lexical multi-word units
    - Du. *wetenschapsbeleid* 'science policy' vs *wetenschappelijk beleid* 'scientific policy' (Booij 2019: 105)
    - Fr. *village(-)vacances* vs *village de vacances* 'holiday resort'
  - 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)

# Introduction

- 3 important **differences** between Dutch and French compounding:
  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. *zoutwaterzwembad* vs Fr. *piscine d'eau salée* 'saltwater pool'
  2. **Word order:** Dutch has right-headed compounding, whereas French has left-headed compounding
    - Du. *postzegel* vs Fr. *timbre-poste* '(postage) stamp'
  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. *hoogspanning* vs Fr. *haute tension* 'high voltage'



**Research question:** What is the impact of these differences on the acquisition of Dutch compounds by French-speaking learners?

# Outline

- 1. State of the art
- 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

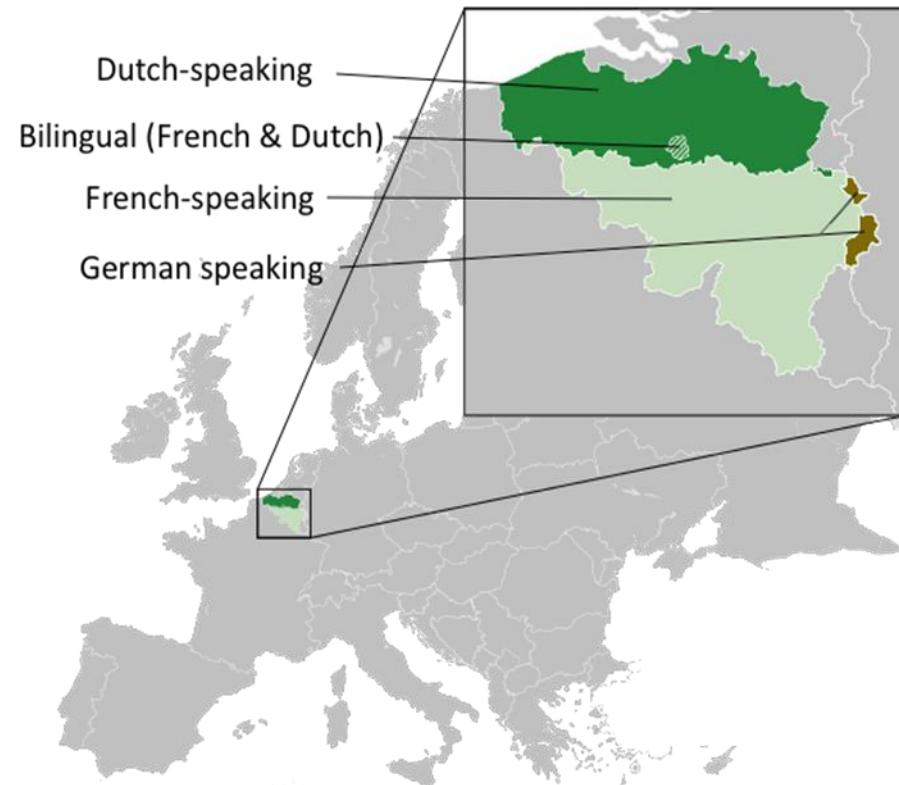
# State of the art: acquisition of word-formation

- Little attention has been paid to the **acquisition of word-formation** (Lefer 2011)
- **Morphological awareness** is well documented in English (e.g. Pasquarella et al. 2011), but it is still an **uncharted territory for languages such as Dutch and French** (except for some fragmented studies, e.g. Deacon et al. 2007 for French)
- Nevertheless, several studies emphasize the **need for more explicit attention to word-formation in foreign language pedagogy** (a.o. Balteiro 2011, Jeon 2011, Leontjev et al. 2016)

# State of the art: 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), AL English:
    - Positive effect on **morphology**, receptive skills, vocabulary and spoken fluency
    - Less clear effect on syntax, writing and pragmatics
- **Multidisciplinary ARC-CLIL project** (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)
  - General positive effect of CLIL, especially on AL Dutch
  - But limited number of case studies: phraseology (Bulon 2019), intensification (Hendrikx 2019)

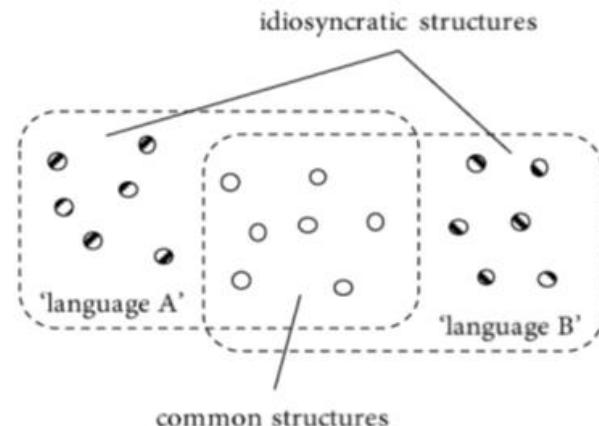
# State of the art: CLIL in French-speaking Belgium



## 2. Diasystematic Construction Grammar

# Diasystematic Construction Grammar (DCxG)

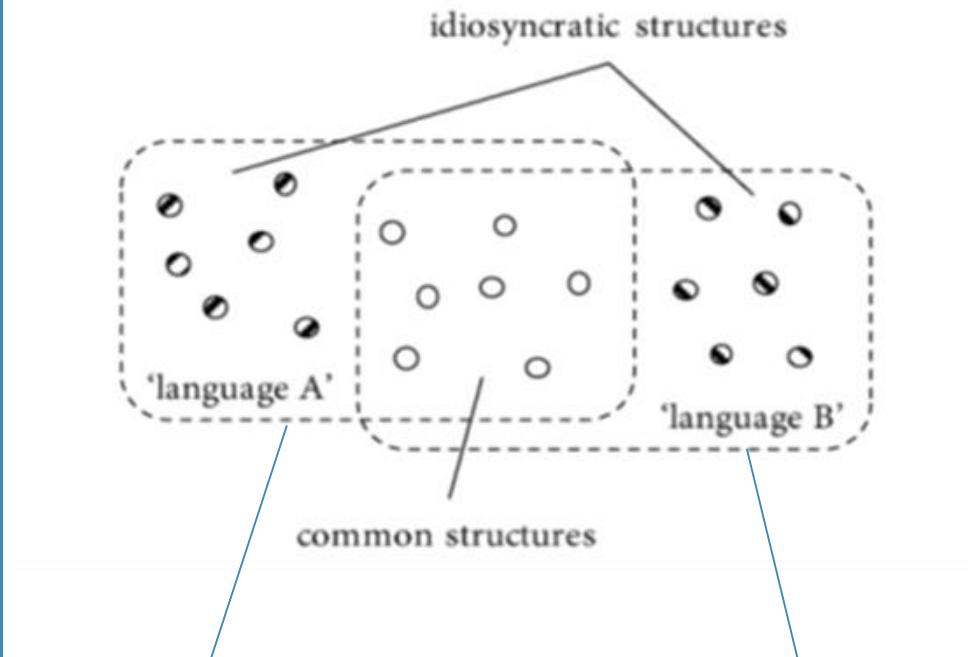
- DCxG recently developed by Höder (2012, 2014a, 2014b, 2018)
- **Integration of CxG into the study of multilingualism**, in line with most recent cognitive insights on multilingualism
  - “multilingual speakers store and process all of their languages in terms of constructions that are organized into one common construction” (Höder et al. 2021: 311) ↔ earlier static conceptions of separate language systems in the bilingual’s mind
- **The multilingual construction**
  - 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)
  - **Effects of cross-linguistic similarities and differences:**
    - Typologically similar languages have more shared ‘diaconstructions’ and learners can build on their L1 constructions for their AL productions
    - The greater the typological difference between the L1 and AL, the more reorganizational processes need to take place in the learners’ constructicon.

# DCxG applied to the acquisition of Dutch compounds



**Language A = L1 French**  
Mostly phrasal constructions  
[N<sub>1</sub> Prep N<sub>2</sub>]<sub>NP</sub>  
(e.g. *salle de bains*)

**Language B = AL Dutch**  
Mostly compound  
constructions  
[N<sub>2</sub> N<sub>1</sub>]<sub>N</sub> (e.g. *badkamer*)

- During AL acquisition:
  1. Dutch compounds ([N<sub>2</sub> N<sub>1</sub>]<sub>N</sub>, e.g. *badkamer*) will have to be added to the learner's multilingual constructional network and have to be tagged as **Dutch idioconstructions**
  2. Some other constructions will have to be "disentrenched", evolving into **French idioconstructions** (e.g. [N<sub>1</sub> Prep N<sub>2</sub>]<sub>NP</sub>, e.g. *salle de bains* vs *\*kamer van bad*)
  3. Some schematic phrasal constructions evolve into shared **diaconstructions**: e.g. [N<sub>1</sub> Prep Det N<sub>2</sub>]<sub>NP</sub> (e.g. *fête pour mon anniversaire / feest voor mijn verjaardag*)

### 3. Aims of the study and research hypotheses

# Aims of the study

- The aims of this 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)
  - 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 idioconstructions (mistagged as FR-DU 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 accuracy and productivity of the compounds and word-formation in AL Dutch

# 4. Corpus data and methods

# Data collection and corpus sample

- *Multilingual Traditional Immersion and Native Corpus* (MuTINCo) (Meunier et al. 2020)
  - Written corpus collected in 2015-2017 from 213 AL Dutch learners in 5 schools in French-speaking Belgium and 59 L1 Dutch speakers (17 Y.O.)
  - The participants were asked to write an e-mail to a friend about a party or their holidays

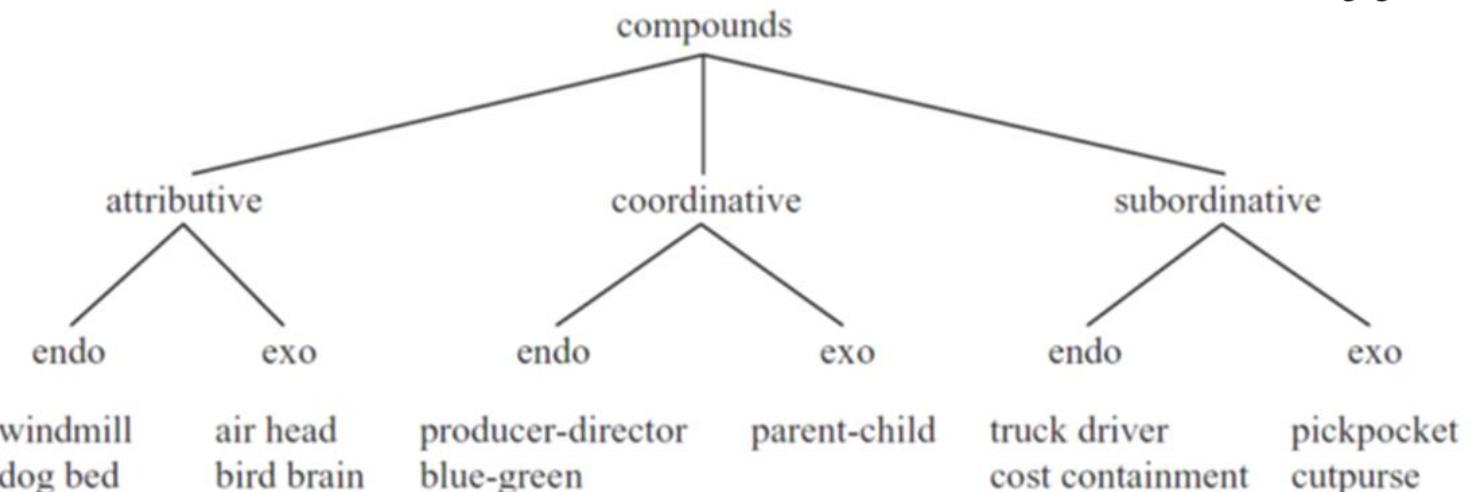
	AL Dutch non-CLIL	AL Dutch CLIL	Native Dutch
<b>Number of participants</b>	84	129	59
<b>Number of nouns</b>	3205	5964	2146
<b>Number of compound nouns*</b>	240	476	155
<b>Proportion compounds per nouns</b>	0,075	0,080	0,072

\*Excluded: blends, derivational compounds, lexicalized compounds, proper names

# Data annotation

- **Formal properties:**
  - Compound length
  - Compound structure
- **Semantic properties :**
  - Compound function

(Bisetto & Scalise 2005: 326)



- **Accuracy:**
  - Accuracy of the compound
  - Word-formation accuracy

# 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	9 (3,75%)	33 (6,93%)	16 (10,32%)
4 components	3 (1,25%)	0	0
coordination	1 (0,42%)	0	0
<b>Totals</b>	<b>240</b>	<b>476</b>	<b>155</b>

## 2 components:

- *Er waren veel **nachtclub** en het was niet duur.* (non-CLIL) 'night club'

## 3 components:

- *Ik dacht vandaag weer terug aan de leuke avond die we gisteren hadden op de **kerstboomverbranding** en in de winterbar (Controls) 'lit. Christmas tree-burning'*

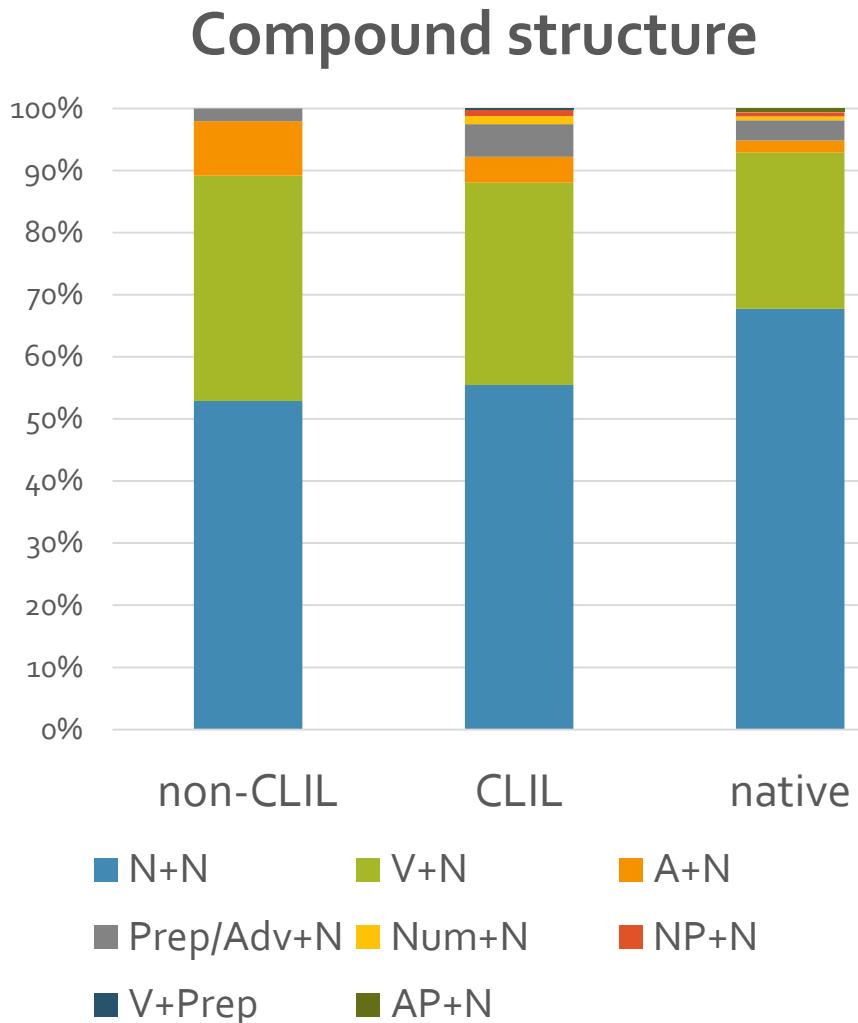
## 4 components:

- *We hebben een groot **voetbal wedstrijd** met anderen jongeren georganiseerd.* (non-CLIL) 'lit. football game'

## Coordination:

- *het best was dat zijn heel goed speelde ( ze deden meestal **rock- en popcovers**, ook met een paar van hun eigen liedjes).* (non-CLIL) 'rock and pop cover songs'

# Schematic level: formal properties



#### [N+N]<sub>N</sub>

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

#### [V+N]<sub>N</sub>

- *Maar als je een goed schuilplaats vind, dan ben jij in zekerheid, lol! (CLIL) 'lit. hiding place'*

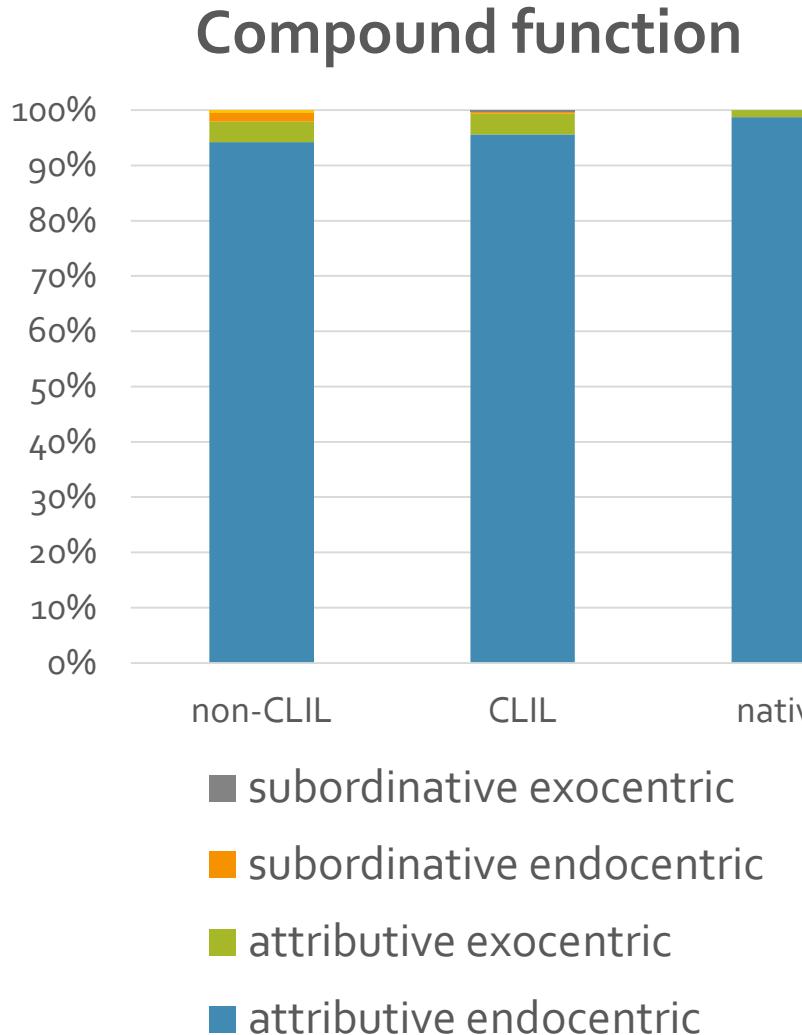
#### [A+N]<sub>N</sub>

- *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]<sub>N</sub>

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

# Schematic level: semantic properties



#### Attributive endocentric

- *het was de verjaardag van mijn buurjongen.* (Non-CLIL) 'neighbor 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 gedrogen 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 and productivity

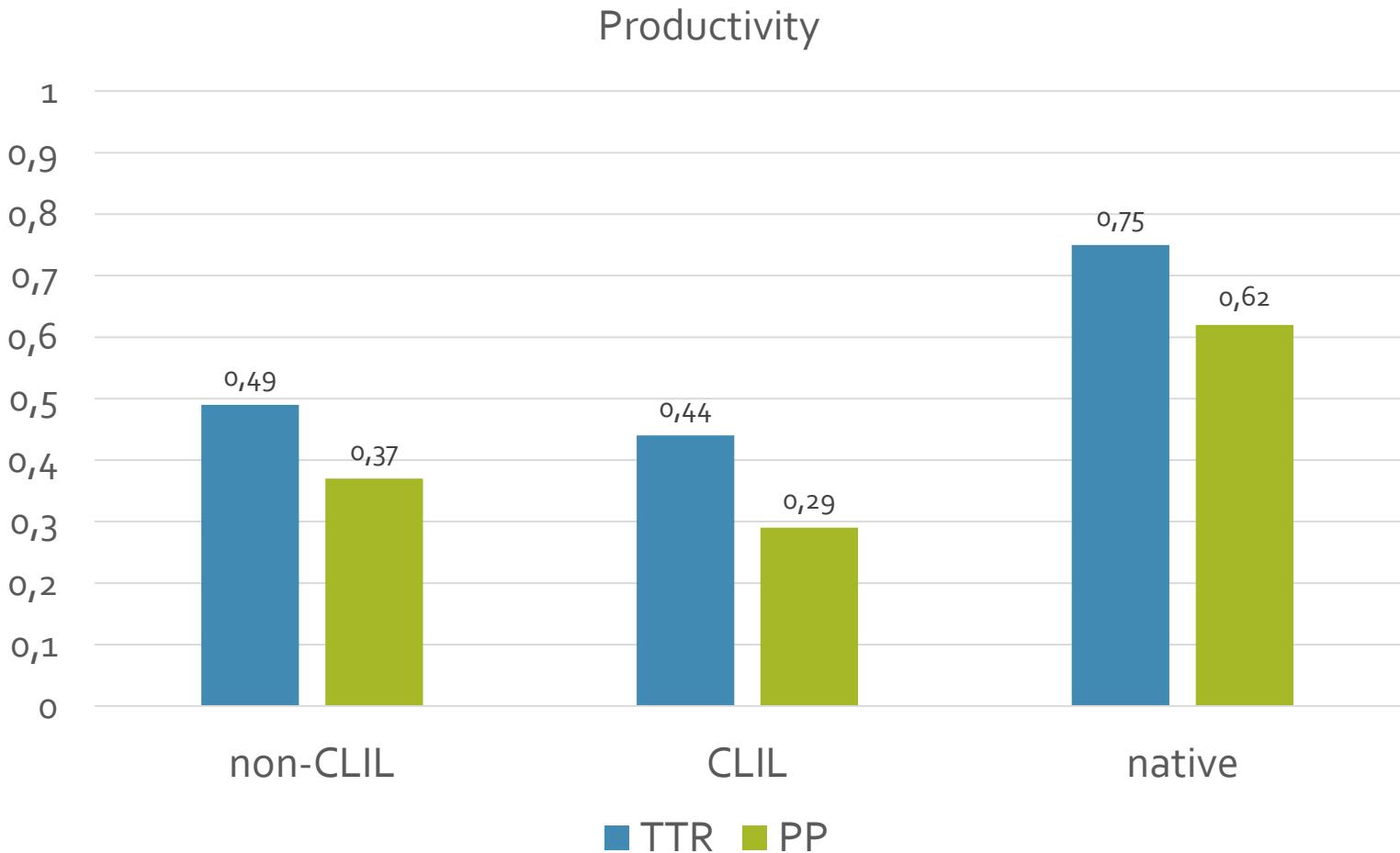
## Most frequent compounds per condition

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

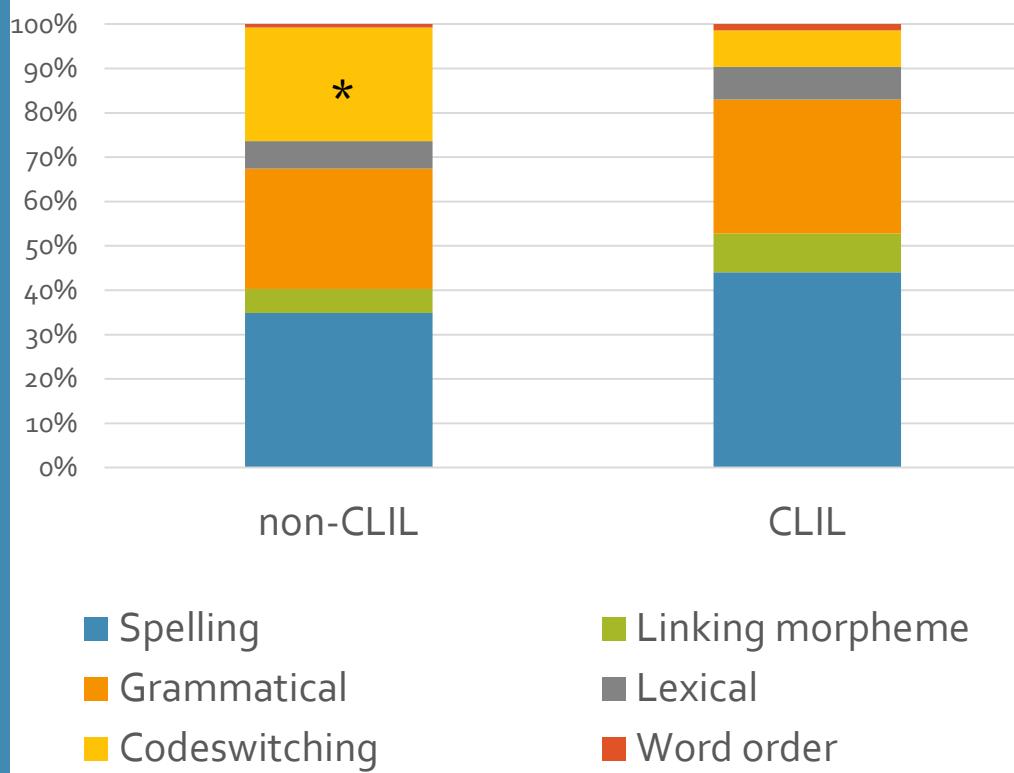
# Substantive level: frequency and productivity

## Productivity of compounds per condition



# Substantive level: accuracy

## Misuse in the compound nouns



	Non-CLIL	CLIL
Total number of mistakes	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

- We zullen met de auto gaan omdat mijn moeder voor het **vliegtrug** bang is? (Non-CLIL)

### Lexical mistake

- Bovendien waren er geen proper **voetweg** zoals het uitgelegd werd in onze guide. (CLIL)

### Code switching\*

- Ik heb ook een **lakeside** gedaan . (Non-CLIL)

### Grammatical mistake

- Iedereen houdt van **vuurwerken**, en een beetje in de frisse , zwarte nacht stappen en naar de sterren kijken (CLIL)

### Mistake in the linking morpheme

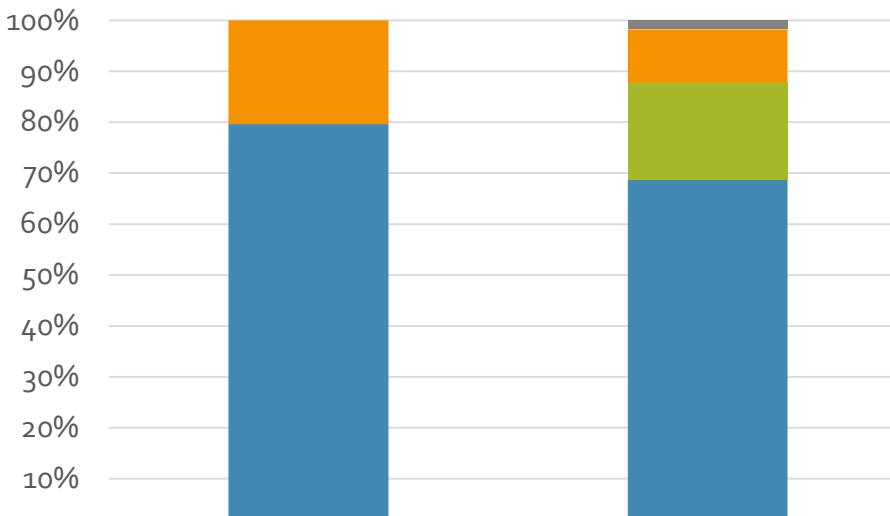
- Daarna hebben we een **aardbeienstaart** gegeten dat de koker had gemaakt . (CLIL)

### Mistake in the word order

- een nieuwe jas , een **album foto** een fototoestel, nieuwe zonnebrillen, ... (CLIL)

# Substantive level: accuracy

## Word-formation mistakes



- Compound instead of simplex
- Compound instead of phrase
- Simplex instead of compound
- Phrase instead of compound

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

### Phrase instead of compound

- we zijn in een mooie hotel met 2 sterren gebleven .* (non- CLIL)

### Compound instead of phrase

- dat was een idea op mijn zusje en mijn bestvriend Floriane.* (non- CLIL)

### Simplex instead of compound

- ... het was voor de 18 jaar van mijn vriend Felix ...* (CLIL) (verjaardag)

### Compound instead of simplex

- Ik was de kookman , ik weet het als het best idée was want ik heb nooit koken van mijn hele leven* (CLIL)

**\*Checked in *Corpus Hedendaags Nederlands*: the compound had to be at least 10 times as frequent as the phrase to consider it misuse:**

e.g. *mensen van mijn klas* (1 hit) vs *mijn klasgenoten* (123 hits) → misuse

e.g. *in het centrum van de stad* (1076 hits) vs *in het stadscentrum* (956 hits) → no misuse

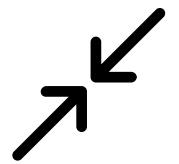
## 6. Discussion and conclusions

# Conclusions

## Cf. Research hypotheses

- (1) **Overgeneralization of French idioconstructions**  
(mistagged as FR-DU diaconstructions):
- ✓ Phrasal structures instead of compounds
    - 'phrase instead of compound' is indeed the most frequent word-formation mistake the in learner group  
(= mistagged as FR-DU diaconstruction instead of FR idioconstruction)
    - However, also cases of overgeneralization (compound instead of phrase/simplex)  
(= mistagged as DU idioconstructions)
  - ✗ 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 correctly tagged as DU idioconstruction by the learners)

# Conclusions



## (2) Positive impact of CLIL on accuracy and productivity of the compounds and word-formation in AL Dutch

- Better entrenchment of Dutch compounds in CLIL-group with respect to compound structure and accuracy, but not with respect to productivity and frequency
  - However, non-CLIL learners have a higher proportion of misused hapaxes
- Significantly more codeswitching in non-CLIL group (= overgeneralization of FR substantive idioconstructions)
- Significantly more word-formation mistakes in CLIL group
  - Typical word-formation mistakes of CLIL group: 'simplex instead of compound' and 'compound instead of simplex' (= linguistic creativity?)

# 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. (2019). The Morphology of Dutch, Second, revised edition. Oxford UK: Oxford University Press.
- Chopey-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 Ambiente 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.
- Hiligsmann, Ph. (1997). Linguistische aspecten en pedagogische implicaties van de tussentaal van Franstalige M.O.-leerders van het Nederlands. Unpublished PhD thesis. Université de Liège.

# References

- 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](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. Braumüller & C. Gabriel (Eds.), *Multilingual individuals and multilingual societies*. Amsterdam/ Philadelphia: Benjamins. 241-257.
- 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. Braumüller, 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.

# References

- Lefer, 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.
- Meunier, F., Hendrikx, I., Bulon, A., Van Goethem, K. & H. Naets. (2020). MultINCo: Multilingual Traditional Immersion and Native Corpus. Better-documented multi-literacy practices for more refined SLA studies. In L. Van Mensel & Ph. Hiligsmann (Eds.) *Assessing CLIL: A multidisciplinary approach [special issue]*. *Journal of Bilingual Education and Bilingualism*. 1-19.
- 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.



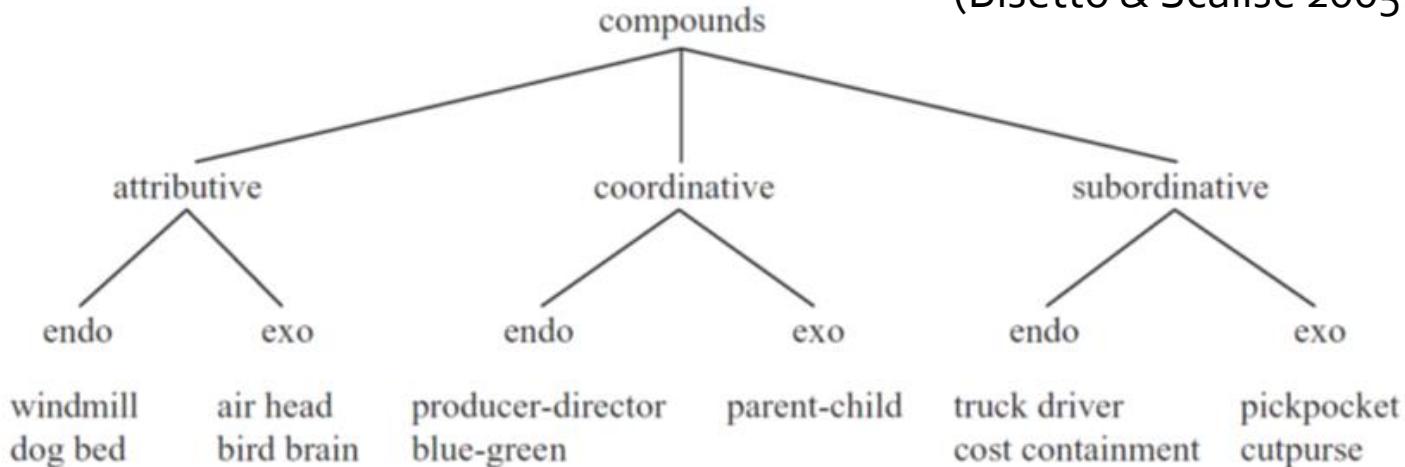
Cette photo par Auteur inconnu est soumis à la licence [CC BY](#)

# 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>lievelings<b>eetje</b></i> ( <i>lievelingseten</i> )
Grammar mistakes (gender/number)	* <i>dit e-mail</i> * <i>mijn nieuwe zonnebrilen</i>
Linking morpheme	* <i>verjaarsdag</i>
Word order	* <i>schoolmuziek</i> ( <i>muziekschool</i> )
Codeswitching (FR/ENG)	<i>pasta met *basilic sauce</i> <b>*football</b>

# Data annotation: word-formation accuracy

Word-formation accuracy	Example
Phrase instead of compound*	*vakantie van Pasen (Paasvakantie)
Simplex instead of compound	een *stap in Padoue (uitstap)
Compound instead of phrase	*vrijtijd (vrije tijd)

**\*Checked in *Corpus Hedendaags Nederlands*: the compound had to be at least 10 times as frequent as the phrase to consider it misuse:**  
e.g. *mensen van mijn klas* (1 hit) vs *mijn klasgenoten* (123 hits) → misuse  
e.g. *in het centrum van de stad* (1076 hits) vs *in het stadscentrum* (956 hits) → no misuse