

**Dutch compound constructions in additional language acquisition:
A diasystematic-constructionist approach**

Isa Hendrikx¹ & Kristel Van Goethem²

^{1.} *Université de Liège, Liège, Belgium*

^{2.} *F.R.S.-FNRS & Université catholique de Louvain, Louvain-la-Neuve, Belgium*

Suggestion for shortened title (for the running head): Dutch compound constructions in ALA

Abstract

Studies have demonstrated that Dutch has a much stronger tendency towards compounding than French (e.g., Du. *badkamer* vs Fr. *salle de bains* ‘bathroom’) when adopting a restrictive approach of compounding in which the presence of prepositions and/or internal inflection in multi-word expressions is considered evidence for their syntactic formation. The example above illustrates that Dutch compounding differs from French in another important aspect: while Germanic compounding is by definition right-headed, French has a general tendency towards left-hand headed compounds and phrases. In this study, we investigate the impact of these typological differences on the acquisition of Dutch nominal compounds by French-speaking learners in the context of multilingual Belgium. We provide an in-depth corpus analysis of the acquisition of Dutch compounds at different levels of abstraction (schematic and substantive compound constructions). Moreover, we investigate the impact of additional target-language input through CLIL programs (*Content and Language Integrated Learning*) on the acquisition of Dutch compounds by French-speaking learners of Dutch. The results are described and interpreted from the perspective of *Diasystematic Construction Grammar* (DCxG), which conceptualizes the linguistic competence of multilingual speakers as one integrated network of constructions, containing language-specific *idioconstructions* and shared *diaconstructions*.

Keywords: Compounding, French, Dutch, Diasystematic Construction Grammar (DCxG), Additional Language Acquisition (ALA), Content and Language Integrated Learning (CLIL), Learner Corpus Research (LCR)

1. Introduction

Schlücker (2019) points out the commonalities and differences between compounds as word-formation units and syntactically formed multi-word expressions in a wide set of European languages. Despite the formal differences, it is argued that both patterns may serve the same purpose and even enter into competition to do so (e.g., Dutch *wetenschapsbeleid* ‘science policy’ vs *wetenschappelijk beleid* ‘scientific policy’ (Booij 2019: 105)). Also in French, many examples of competition can be found between morphological and syntactical formations: for example, *village(-)vacances* coexists with *village de vacances* ‘holiday village, holiday resort’ (lit. village (of) holidays). *Village(-)vacances* is a compound, whereas *village de vacances* should be considered a phrasal structure when adopting a restrictive approach of compounding in which the presence of prepositions and/or internal inflection in multi-word expressions is considered evidence for their syntactic formation (Fradin 2009).

However, even within a same genealogical family, languages grammaticalize at a different pace and, consequently, they may significantly vary with respect to their degree of analyticity. English and French have been shown to be closest to the analytic pole of the respective “Germanic Sandwich” and “Romance Sandwich” clines, whereas Dutch and Spanish, for example, are relatively more synthetic languages (cf. van Haeringen 1956, Hüning *et al.* 2006, Lamiroy 2011, among others).

The cross-linguistic differences in degree of analyticity between French (analytic) and Dutch (more synthetic) are perceivable in the field of compounding. Van Goethem (2009) and Van Goethem & Amiot (2019) have demonstrated that Dutch has a stronger tendency towards compounding than French (e.g., Du. *Zoutwaterzwembad* vs Fr. *Piscine d’eau salée* ‘saltwater pool’; Du. *Verjaardagsfeest* vs Fr. *Fête d’anniversaire* ‘birthday party’). The examples show that Dutch compounding differs from French in another important aspect. While Germanic compounding is by definition right-headed, French, like other Romance languages, has a general tendency towards left-hand headed compounds and phrases (e.g., Du. *bloemkool* vs Fr. *chou-fleur* ‘cauliflower’). Thirdly, whereas most Germanic languages, including Dutch, establish a clear demarcation between compounds and (lexicalized) phrases on the basis of formal criteria (spelling, prosody, linking elements, loss of adjectival inflection), French compounds are not always easily distinguishable from syntactic expressions. Dutch *hoogspanning* ‘high voltage’, for instance, is clearly a compound (spelling as one word, compound stress on the first component, loss of adjectival inflection of *hoog*), whereas its French equivalent *haute tension* should be analyzed as a lexicalized phrase (see Schlücker,

2019, for a closer examination of the distinction between compounds and phrases in a series of European languages).

As a result, the term ‘compound’ is not always used in a consistent way in the literature on French morphology, and ‘compounding’ often refers to various types of complex lexical units regardless of the formation process, either morphological or syntactic (cf. Van Goethem 2009; Villoing, 2012; Van Goethem & Amiot 2019). From the restrictive point of view that we adopt here, only a limited set of patterns should be considered true compounds in French. Fradin (2009), for instance, concludes that only four productive compounding patterns should be retained in French: [Verb + Noun] (e.g., *brise-glace* ‘icebreaker’), [Adj + Adj] (e.g., *sino-coréen* ‘Sino-Korean’), [Noun + Noun] coordinates (e.g., *auteur-compositeur* ‘author-composer’) and [Noun + Noun] subordinates (e.g., *poisson-chat* ‘catfish’).¹

Dutch compounding, by contrast, is a very productive word-formation pattern and mainly serves to build complex nouns (e.g., *eetkamer* ‘dining room’), verbs (e.g., *overkomen* ‘lit. over-come, happen to’) and adjectives (e.g., *donkergeel* ‘dark yellow’). Dutch native speakers easily produce new compounds on the spot, and compounds may also easily be embedded into other compounds, which may lead to rather complex structures (e.g., *postzegelverzamelaarsbijeenkomst* ‘stamp collectors meeting’) (Don 2009; Booij 2019).

In this study, we investigate the impact of these typological differences on the acquisition of Dutch nominal compounds by French-speaking learners in the context of multilingual Belgium. Hiligsmann (1997) already noticed that the acquisition of Dutch compounds is a stumbling block for French-speaking learners. The present study will address the question more profoundly by means of a corpus-based constructionist analysis in which we examine the acquisition of the compounds at two levels of constructional analysis, namely at the schematic level (abstract patterns) and at the substantial level (concrete instantiations). Moreover, we investigate the impact of additional target-language input through CLIL programs (Content and Language Integrated Learning) on the acquisition of Dutch compounds by French-speaking learners of Dutch.

The corpus study is based on written corpus data collected within the context of a research project on CLIL in French-speaking Belgium (cf. Hiligsmann et al. 2017; Meunier et al. 2020). The data are annotated and analyzed for their formal make-up (compound length and structure), semantic function, accuracy, frequency and productivity (type token ratio, hapaxes).

¹ Villoing (2012: 36) adds to this a particular subclass of [Adj + Noun] compounds with a color adjective as head (e.g. *bleu-ciel* ‘sky blue’ (lit. blue-sky)).

From a theoretical point of view, the results are described and interpreted within the framework of Diasystematic Construction Grammar (DCxG) (among others Höder et al. 2021), which conceptualizes the linguistic competence of multilingual speakers as one integrated network of constructions, containing language-specific idioconstructions and shared diaconstructions.

The outline of the paper is as follows. In Section 2, we present a brief state of the art on Additional Language Acquisition (henceforth ALA)², with the focus on the acquisition of word formation and the potential impact of CLIL programs. Section 3 will present the theoretical framework of Diasystematic Construction Grammar and its implications for ALA and, more particularly, for the acquisition of compound constructions in an additional language (AL). Section 4 outlines the aims of the study and the research hypotheses. In Section 5, we present the corpus data and methods of our analysis in more detail. Section 6 presents the results of the corpus-based constructionist analysis and in Section 7, the main results are discussed and interpreted from the DCxG perspective. Finally, Section 8 contains the concluding remarks and the outlook for follow-up research.

2. Additional Language Acquisition and CLIL: a brief state of the art

2.1. Recent insights into Additional Language Acquisition

As already mentioned, this study takes a constructionist perspective on ALA that comes under the umbrella of usage-based approaches. According to Ortega (2013: 113-114), usage-based or emergentist theories share the following viewpoints. First, grammar learning is not rule-based (deductive), but driven by experience and exposure (inductive) (cf. Ellis & Cadierno 2009). For instance, constructions that occur frequently and cohesively (without structural changes) in a learners' input are expected to be acquired earlier and become more entrenched (Bybee 2010; Diessel 2016). Second, both input variables, such as frequency and salience, and learner variables, such as attention and ability of categorization, are key factors in language learning. Third, learner language and language development are characterized by variability and by a dynamic interaction of multiple variables.

² Following Ortega (2013) and Höder et al. (2021), we prefer to use the term *additional language* (AL) instead of second language (L2), third language (L3) and so on, because it is unclear to which extent L2 acquisition differs from L2+n acquisition. Moreover, in our corpus study based on French-speaking learners in Belgium, Dutch is not necessarily the L2 of the learners.

With respect to multilingualism, recent psycho- and neurolinguistic research has corroborated Cummins' (1979) hypothesis that multilingual speakers do not store their language knowledge in two separate compartments in the brain and that proficiency in one language is related to proficiency in another language (cf. the so-called "Common Underlying Proficiency" hypothesis). Recent studies show even stronger interrelatedness between bilinguals' language skills: when one language is being used, the other language is simultaneously activated (Hartsuiker, Pickering, & Veltkamp 2004; De Groot 2011; Hartsuiker & Bernolet 2017).³

The development of multilingualism during ALA has recently been conceptualized as a complex process of cognitive restructuring (e.g., Treffers-Daller & Tidball 2015). In that respect, the term *transfer* has been criticized as being too narrow to capture all cognitive processes that may take place in a learner's mind. Cook (2016: 24), for instance, argues that transfer restricts "L2 users to the position of cumulative monolinguals rather than seeing the richness of the L2 mind". He furthermore advances that learners "do not have a defective copy of the L2 in their minds so much as an interlanguage of their own making" (Cook 2016: 27).

In the same vein, Ellis & Cadierno (2009: 112) argue that AL acquisition is more complex than L1 acquisition, not only due to issues of quantity and quality of exposure, but also because "during development, L2 constructions are in direct competition with those of the learners' L1, and these may represent alternative ways of construing the same reality". Even though target-language input and teaching are assumed to foster AL acquisition, the L1-specific attentional biases may still result in *blind spots* for particular structures in the target language (Ellis & Sagarra 2011; Jach 2017).

2.2. Acquisition of word formation

Despite the large number of studies on ALA and Learner Language, little attention has been paid so far to the acquisition of morphological phenomena. In her (2011) state of the art, Lefer rightly states that contrastive morphology is largely under-researched and that little research in the domain of ALA acquisition has focused on word formation so far.

Nevertheless, several studies have emphasized the need for more explicit attention to word formation in foreign language pedagogy. For instance, Balteiro (2011) stresses that word-formation awareness and knowledge are crucial for learners' AL proficiency and creativity, and

³ For a detailed overview of previous (constructionist and non-constructionist) research into multilingualism, see Höder et al. (2021).

Jeon (2011) and Leontjev et al. (2016) prove that derivational knowledge is a significant predictor of AL reading comprehension and writing proficiency. Morphological awareness is well documented in English (e.g., Pasquarella et al. 2011), but – except for some fragmented studies (e.g., Deacon et al. 2007 for French) – it is still an uncharted territory for languages such as Dutch and French.

2.3. Content and Language Integrated Learning

As an additional criterion, we investigate in this study how the acquisitional process is impacted by additional target language exposure in the form of Content and Language Integrated Learning (CLIL), within the context of multilingual Belgium.

Belgium is internationally known as a multilingual country, comprising three official languages (Dutch, French, and German). Good language skills are highly valued in Belgium and are needed in the job market (Mettewie & Van Mensel 2009). Nevertheless, recent research conducted by the European Union indicates that foreign language teaching and learning in Belgium is on a downward spiral (e.g., Eurostat 2016).

In this regard, it is worth mentioning that the Dutch- and French-speaking communities in Belgium have a relatively high degree of autonomy in political, educational, and other matters, and it might come as a surprise to learn that teenagers living in the French Community are only rarely exposed to Dutch outside the school context. English, by contrast, is pervasive in everyday Belgian life through social media, music, television, etc. (Mettewie & Van Mensel 2020).

The pedagogical approach of CLIL has been promoted by the European Union to enhance widespread multilingualism (Coyle, Hood & Marsh 2010). In CLIL, several school subjects (e.g., history, geography, science, and mathematics) are taught in a language other than the students' L1 (Blondin 2003; Chohey-Paquet 2008; Hiligsmann et al. 2017). Although the CLIL approach has been extensively documented internationally (e.g., Ruiz de Zarobe et al. 2011; Rumlich 2016), its impact on SLA remains a subject of scholarly debate. Dalton-Puffer (2007) reports that, at least for English as a foreign language, morphology, receptive skills, vocabulary, and spoken fluency are skills that are positively affected by CLIL, while skills such as syntax, writing, and pragmatics are less clearly affected.

More than two decades after it was first implemented in French-speaking Belgium, the number of schools and students involved in CLIL programs continue to increase steadily.

Nonetheless, research on CLIL in French-speaking Belgium is still relatively scarce (e.g., Chopey-Paquet 2008; Beheydt 2014).

In order to gain a more complete view of the impact of CLIL on learning foreign languages (Dutch and English) in French-speaking Belgium, a multidisciplinary team of researchers from the Université catholique de Louvain and the Université de Namur have conducted a joint research project (2014-2019) exploring linguistic, cognitive, educational, and socio-affective avenues of research (cf. Hiligsmann et al. 2017; Van Mensel & Hiligsmann 2021). Students enrolled in CLIL secondary schools in the French-speaking Community of Belgium receive a considerable amount of additional target-language exposure, at least 4 hours per week more than their non-CLIL peers, which was expected to be beneficial for their foreign language acquisition. The project confirmed the beneficial impact of CLIL on the acquisition of Dutch and English in well-defined linguistic fields, such as phraseology (Bulon 2019) and intensification (Hendrikx 2019), with a more significant impact observed for Dutch than for English (see also Van Mensel & Hiligsmann 2021). However, its impact needs to be investigated for many more linguistic fields in order to be confirmed. This study contributes to filling this gap by assessing the impact of CLIL on the acquisition of Dutch compound constructions by learners in French-speaking Belgium.

3. Diasystematic Construction Grammar

3.1. General principles

In constructionist approaches, the mental grammar of language users is seen as a dynamic, hierarchical network of constructions, i.e., systematic and conventionalized form-meaning pairs, (Hoffmann and Trousdale 2013: 3) that range from fully schematic, i.e., abstract schemas with open slots, to fully substantive patterns, i.e., lexically filled constructions in which all slots are prespecified (cf. Goldberg 2006). Constructions are thought of as the fundamental units of speakers' cognitive linguistic competence and of L1 and AL acquisition. Learners are expected to first acquire substantive constructions as a whole, especially the ones that occur frequently in their input, before deriving more schematic constructions (Tomasello 2003). Moreover, the understanding of language as one integrated network of constructions, i.e., the so-called *constructicon*, without compartmentalized grammar and lexicon, implies that “[S]yntactic constructions may form an alternative to, or compete with the morphological expression of

semantic and grammatical content” (Booij 2002: 301; also Booij 2010, 2018). As such, the constructionist approach goes beyond the traditional lexicon-grammar divide and is particularly suited to study the competition between morphological and syntactic constructions.

More precisely, this study will be framed in the theory of *Diasystematic Construction Grammar* (DCxG), recently developed by Höder (2012, 2014a, 2014b, 2018), which integrates constructionist insights into the study of multilingualism. Following up on recent neuro- and psycholinguistic insights into multilingualism (see Section 2.1), Hilpert and Östman (2014: 139) argue that “[b]ilingual or multilingual speakers have the ability to form meta-generalizations that connect corresponding constructions from different languages in their repertoire”. In the same vein, DCxG assumes that “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). Grammatical knowledge is processed on the basis of the available input via processes of interlingual identification, abstraction, generalization, and categorization, regardless of language boundaries (Höder 2018: 37). The so-called *multilingual constructicon* (see Figure 1) 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 latter *idioconstructions*.

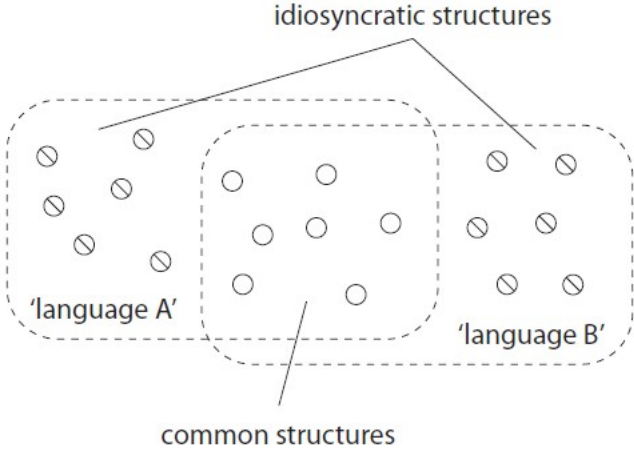


Figure 1: The multilingual constructicon (from Höder 2018: 44)

In such an integrated model, AL learning “amounts to an extension and reorganization of the constructicon, resulting not only in the gradual entrenchment of new constructions that represent (a learner variety of) the AL, but also in the modifications of the previously acquired

constructions and the links between them” (Höder et al. 2021: 311). The model requires a fundamental reinterpretation of established concepts such as *interference* (Lado 1957) and *interlanguage* (Selinker 1972), which imply a unidirectional view on language acquisition and a static conception of separate language systems in the bilingual’s mind, in terms of a dynamic and emergent constructional network representing the learner’s AL as an integrated part of his/her mental constructicon (Höder et al. 2021: 319).

3.2. Additional Language Acquisition from a DCxG perspective

Although Höder’s framework has so far mainly been applied to issues of language contact (cf. Boas & Höder, 2018, 2021), we consider it a promising analytical tool for studies on ALA, because it engenders an organic and emergent view of multilingual cognition, in line with recent cognitive insights. DCxG has only recently been applied to language acquisition (Höder et al. 2021; Jach 2021; Van Goethem & Hendrikx 2021). Van Goethem & Hendrikx (2021), for instance, showed how the DCxG approach entails a refreshing look at the acquisition of intensifying constructions by French-speaking learners of Dutch (see also Hendrikx 2019). The results of this study unveiled specific reorganizational processes that occur in the *diasystem* of French-speaking learners of Dutch, such as the overgeneralization of the particular schematic pattern with a boosting degree adverb (e.g., *heel moe* ‘very tired’) and the inaccurate use of specific intensifiers, such as the Dutch quantifier *veel* (**veel moe* ‘lit. much tired’).

To sum up, DCxG provides a dynamic view on ALA, and conceptualizes the constructions specific to each language as *tagged* for that particular language, in a similar manner to how constructions are tagged for particular social or contextual variables. Moreover, DCxG can elucidate the effect of cross-linguistic similarities and differences. Whereas typologically similar languages have a higher degree of diasystematicity (i.e., more abstract constructions are shared between the two languages, and learners can build on their L1 experiences and routines for their AL productions), typological differences between languages imply a lower degree of diasystematicity. The greater the typological difference between the L1 and AL, the more reorganizational processes need to take place in the learners’ diasystem.

Finally, DCxG accommodates the insights from usage-based linguistics, in particular regarding the effects of frequency of use and the saliency of the construction on the acquisition process (see Section 2.1), given that aspects such as frequency and saliency are also properties of constructions. Frequency of input is a crucial factor in ALA (Ellis & Wulff 2008: 1), as

frequent input fosters the “entrenchment” of AL constructions.⁴ Learners develop schemas by generalizing over more specific (substantial) items in their input. A construction that reoccurs frequently, that has been encountered recently, and which is perceptually salient, will be acquired earlier on, and become entrenched more deeply (Jarvis & Pavlenko 2008: 186; see also Long & Sato 1983 and Doughty 1991), although the association between these factors and their manifestation in the learners’ productions is mostly indirect, as shown in Jarvis (2002).

3.3. Acquisition of Dutch nominal compound constructions from a DCxG perspective

As argued in the introduction of this paper, compounds and phrases both exist in Dutch and French and may serve the same function, although they are built differently and show considerable differences in productivity. One of the basic assumptions of Construction Grammar lies precisely in the non-modular view of language, with lexicon and grammar as endpoints of a continuum. It is not surprising that many recent studies in the field of multi-word expressions are in the constructionist vein. In this approach, it can be assumed that both compounds and phrasal structures with a naming function can act as conventionalized form-meaning pairings or ‘constructions’.⁵

In the context of DCxG, we can summarize that ALA should be conceptualized as a dynamic process of constructional reorganization, involving construction addition (entrenchment of newly acquired constructions) and deletion (disentrenchment or inhibition of redundant constructions) (Höder et al. 2021). During the process of AL acquisition, new AL constructions become entrenched and labelled for structural, semantic, socio-linguistic and/or contextual properties (such as frequency of input), while at the same time previously acquired L1 constructions are either generalized or tagged for the L1. The main challenge for French learners of Dutch then consists in correctly integrating the Dutch idioconstructions (compounds) into their multilingual constructicon. In the remainder of this study, we will focus on the acquisition of Dutch nominal compounds.

When applying our case study to the DCxG framework, language A in Figure 2 corresponds to L1 French and language B to AL Dutch. Language A (French) contains mostly

⁴ Schmid (2015: 10) defines entrenchment “as the continuous routinization and re-organization of association, depending on exposure to and frequency of identical or similar processing events, subject to the exigencies of the social environment”.

⁵ Therefore, we should accept the existence of what Booij (2010: 190) calls ‘lexical phrasal constructions’, i.e. syntactic formations that should be stored as lexical units in the mental lexicon.

phrasal structures, such as $[N_1 \text{ Prep } N_2]_{NP}$ (e.g. *salle de bains* ‘bathroom, *fête d’anniversaire* ‘birthday party’), whereas Language B (Dutch) has a stronger tendency towards $[N_2 N_1]_N$ compounding (e.g. *badkamer* ‘bathroom’, *verjaardagsfeest* ‘birthday party’). When acquiring Dutch, the French-speaking learner will have to reorganize his/her construction in different ways. Compound nouns such as $[N_2 N_1]_N$ (e.g. *badkamer*, *verjaardagsfeest*), on the one hand, will have to be added to the learner’s multilingual constructional network as Dutch idioconstructions, not only because of their higher productivity in Dutch compared to French, but also because of their different word order. Some other constructions will have to evolve into French idioconstructions (e.g. $[N_1 \text{ Prep } N_2]_{NP}$, e.g., *salle de bains* vs **kamer van bad*). Finally, some constructions may undergo more complex processes of modification. For instance, while Dutch generally prefers nominal compounds to their syntactic equivalents, in certain cases the phrasal equivalent is also available in Dutch and should evolve into a shared schematic diaconstruction, e.g. $[N_1 \text{ Prep } \text{Det } N_2]$ (*feestje van/voor (mijn) verjaardag / fête pour (mon) anniversaire* ‘party of/for (my) birthday’). Hence, learners will need to develop awareness of the acceptability of different variants, which is closely related to frequency of use.

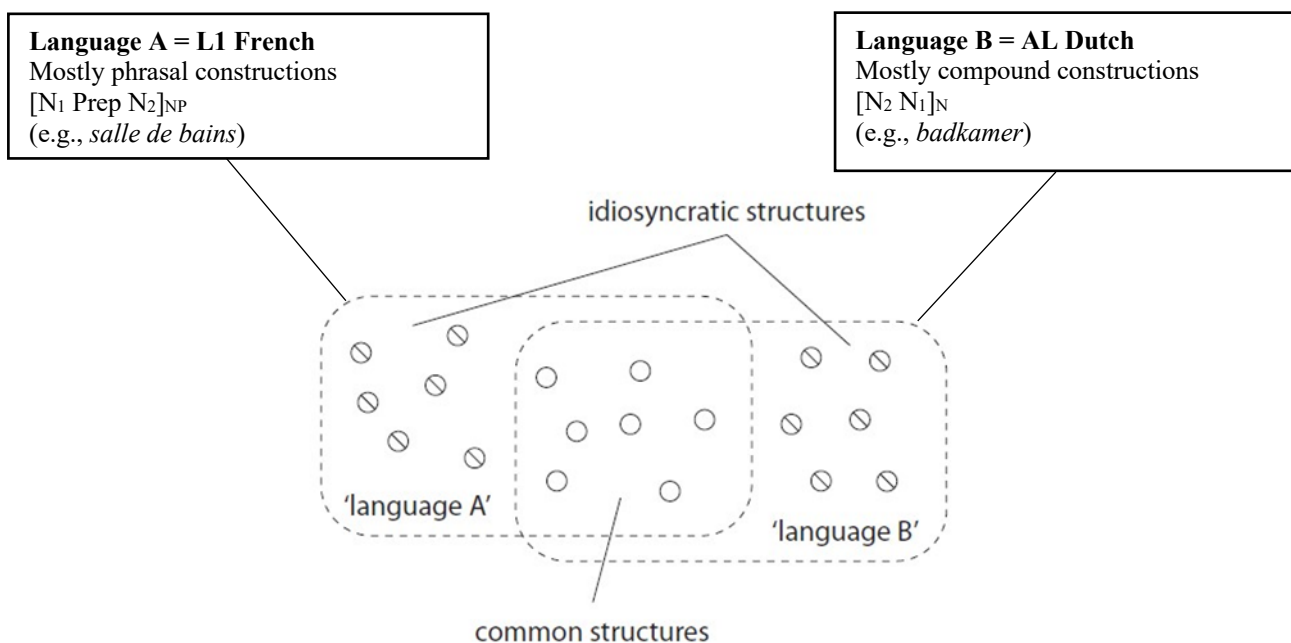


Figure 2: Acquisition of Dutch nominal compounds from a DCxG perspective

4. Aims of the study and research hypotheses

The aims of this study are threefold:

- (i) First, we aim to provide an in-depth analysis of the acquisition of compound nouns at multiple levels of abstraction (schematic and substantive compound constructions).⁶
- (ii) Second, we intend 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) Third, we wish to further explore the diasystematic-constructionist perspective, and in particular its application to AL acquisition. We will describe the results of our corpus analysis and interpret them from the perspective of DCxG, since this theory combines the assumptions of Construction Grammar with the most up-to-date insights into multilingualism and ALA.

We formulate two main research hypotheses:

- (i) First, based on the typological differences between the languages under study (Section 1), and following the insights from the DCxG approach (Section 3), we expect, among others, the learners to overgeneralize the French “phrasal” idioconstructions (e.g., **kamer van bad* instead of *badkamer*) and the word order typical of French compounding (e.g., **kamerbad*). We hypothesize that both phenomena result from the fact that learners will misuse their French idioconstructions as diaconstructions, shared by French and Dutch.
- (ii) Second, based on earlier research on the impact of CLIL on the acquisition of particular linguistic phenomena (cf. Section 2.3), we expect a positive impact of CLIL on the accuracy and productivity of the formed compounds in AL Dutch.

5. Corpus data and methods

5.1. Data collection and corpus sample

⁶ We also conducted a corpus study of the data at the semi-schematic level. However, the subcorpora turned out too small to derive statistically relevant results at this level. For completeness, we can add that the most frequent heads in the three datasets were *-feest* ‘party’ and *-vakantie* ‘holiday’, which is no surprise since these terms correspond to the specified topics of the writing task. The most frequent modifiers were *verjaardags-* ‘birthday’ and *lievelings-* ‘favorite’.

The written corpus used in this study is part of the *Multilingual Traditional Immersion and Native Corpus* (MULTINCo) (Meunier et al. 2020), which was collected between 2015 and 2017 within the research project on CLIL jointly conducted by the Université catholique de Louvain and the Université de Namur (cf. Section 2.3). MULTINCo contains spoken and written data from French-speaking learners in their L1 French and in their AL English or AL Dutch and written data from control groups of native-speakers of Dutch and native-speakers of English, as well as various background variables of the learners (such as a composite measure of their extracurricular AL input) (Meunier et al. 2020). MULTINCo also contains comparable samples of students enrolled in CLIL and traditional foreign language learning programs. The sampled schools are socio-economically and geographically diverse (Meunier et al. 2020).

The data collections took place at the Université catholique de Louvain (in Louvain-la-Neuve, Belgium). The students performed a variety of computer-administered tasks, including two writing exercises, in which they were asked to write two (fictional) e-mails to a friend about a party or their holidays (the topics were randomly assigned), one in their AL Dutch (or English), one in their L1 French. The participants had a maximum of 25 minutes to write each e-mail (min. 15 lines, in Lime Survey 2006⁷) and had no access to online dictionaries or other reference tools. The spell check facility on the computers used for the data collection was turned off. The native speakers followed the same procedure as the learners but wrote only one e-mail in their first language.

In this study we use the written data from 213 French-speaking learners of Dutch in the 6th year of secondary school and 59 L1 Dutch speakers (on average 17 years old). Data from the control group of Dutch native speakers were collected in the Netherlands and in Flanders (the Dutch-speaking community of Belgium).

The nouns in the resulting subcorpora were extracted automatically. Thereafter we manually extracted all compound nouns. Blends (e.g., *beer-pong*), derivational compounds (e.g., *minderjarige* ‘underaged child’), lexicalized compounds (e.g., *leeftijd* ‘lit. live-time, age’, *maandag* ‘Monday’) and proper names (e.g., *Nederland* ‘The Netherlands’) were not included in our selection of compound nouns.

Table 1 provides an overview of the number of participants, the number of nouns and the number of compound nouns per condition. It shows that although the samples considerably

⁷Lime Survey: <https://www.limesurvey.org/about-limesurvey/team-contributors>.

diverge in size, the proportion of compounds relative to the total number of nouns is quite stable across the three conditions (non-CLIL, CLIL, natives). The proportion of compound nouns relative to the total number of nouns reaches the highest percentage in the CLIL condition, although the differences between the three groups are not significant ($\chi^2 = 0.90, p = .34.$). This fact is also visualized in Figure 3, which shows the boxplots of the proportion of compounds per nouns per condition.

	AL Dutch non-CLIL	AL Dutch CLIL	Control corpus native Dutch
Number of participants	84	129	59
Number of nouns	3205	5964	2146
Number of compound nouns	240	476	155
Proportion compounds per nouns	7,49%	7,98%	7,22%

Table 1: Number of participants, nouns and compound nouns per condition

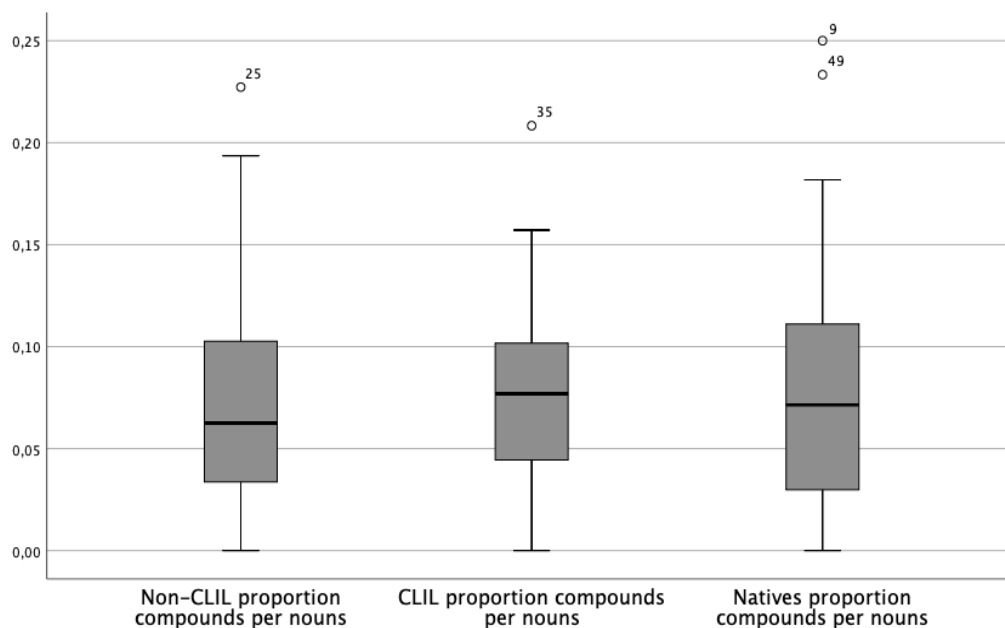


Figure 3: Proportion compounds per nouns per condition⁸

5.2. Data annotation and analysis

⁸ The outliers are e-mails containing a much greater proportion of compounds per nouns than the other e-mails in the subcorpus.

Once all the compound nouns in the three sub-corpora were identified, they were manually annotated for their formal (5.2.1) and semantic (5.2.2) properties, as well as for their accuracy (5.2.3).

5.2.1. Formal properties

In terms of formal characteristics of the compounds, we annotated the length of the compound and the compound structure. The categories of each variable and some examples are provided in Table 2.

Formal properties	Categories of the variable	Examples
Compound length	Number of lexical components	2 lexical components: <i>voetbal</i> ‘football’
		3 lexical components: <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>grootouders</i> ‘grandparents’
	[Prep/Adv+N] _N	<i>buitenland</i> ‘lit. outside-land; abroad’
	[Num+N] _N	<i>vijftal</i> ‘about five’
	[NP+N] _N	<i>onderwaterwereld</i> ‘lit. under water world, the universe under the water’

Table 2: Formal annotation of the compounds

5.2.2. Semantic properties

For the semantic annotation, we refer to the classification of compounds proposed by Bisetto and Scalise (2005), who distinguish three different compound functions. Whenever two nouns form a compound in which the non-head expresses an attribute, literal or metaphorical, of the head, it is classified as an *attributive compound*, for instance *windmill* and *air head* in English (Bisetto & Scalise 2005: 10). The second compound function that Bisetto and Scalise (2005)

identify is coordination. *Coordinate compounds* consist of two elements that are implicitly tied by the conjunction *and*; these compounds have more than one head: a *singer songwriter* is both a *singer* and a *songwriter* (Bisetto & Scalise 2005: 10). Thirdly, a compound can be considered as *subordinate* “whenever there is a ‘complement’ relation between the two constituents” (Bisetto & Scalise 2005: 9). In *ballet dancer* for instance, *ballet* is the complement of the deverbal head *dancer*. Compounds of all these three classes can be either endocentric or exocentric (depending on the presence vs. absence of a head constituent) (Bisetto & Scalise 2005: 10).

While Bisetto and Scalise (2005) focus on English compounds, their classification is also applicable to Dutch compounds, as shown by the examples in Table 3.

Compound function	Example
Attributive endocentric	<i>windmolen</i> ‘windmill’
Attributive exocentric	<i>leeghoofd</i> ‘lit. empty-head; air head’
Coordinative endocentric	<i>hotel-restaurant</i>
Coordinative exocentric	<i>ouder-kind(relatie)</i> ‘parent-child (relationship)’
Subordinative endocentric	<i>kurkentrekker</i> ‘lit. corks-puller; corkscrew’
Subordinative exocentric	<i>zakkenroller</i> ‘lit. pockets-picker; pickpocket’

Table 3: Semantic annotation of the compounds

5.2.3. Accuracy

The data were also annotated for their accuracy. We distinguished between the accuracy of the formed compound nouns (Table 4) and the accuracy of the word formation (Table 5). The latter type involves for instance the use of a phrase or a simplex word in cases where a compound is (more) appropriate.

As illustrated in Table 4, we distinguish different types of “inaccuracies”⁹ in the compound, related to spelling, lexicon, grammar (e.g., incorrect gender or number), the use of the linking morpheme, the word order or code switching. In the case of codeswitching, either the entire compound is used in another language than Dutch (mostly French or English) (e.g., *rhetoreis [laatstejaarsreis]* < Fr. *voyage rhéto* ‘senior trip’) or only part of the compound (e.g.,

⁹ By the term “inaccuracies” we refer to a range of non-idiomatic expressions that should be situated on a scale of acceptability: as rightly pointed out by one of the reviewers, some expressions are acceptable, but not idiomatic (among others because of a lower frequency of use in standard Dutch), while others are clear deviations from standard Dutch (e.g. spelling and word order mistakes).

verjaardagsgateau [*verjaardagstaart*] ‘birthday cake’). All examples in Tables 4 and 5 are taken from our corpus (see Section 5.1). The correct forms are indicated between square brackets.

Accuracy of the compound	Example
Spelling	* <i>vliechtuig</i> [<i>vliegtuig</i>] ‘airplane’
Lexicon	* <i>lievelingseetje</i> [<i>lievelingseten</i>] ‘favorite food’
Grammar (gender, number)	* <i>dit e-mail</i> [<i>deze e-mail</i>] ‘this email’ * <i>mijn nieuwe zonnebrilen</i> [<i>mijn nieuwe zonnebril</i>] ‘my new sunglasses’
Linking morpheme	* <i>verjaarsdag</i> [<i>verjaardag</i>] ‘birthday’
Word order	* <i>schoolmuziek</i> [<i>muziekschool</i>] ‘music school’
Codeswitching (to French or English)	<i>pasta met *basilic sauce</i> [<i>basilicumsaus</i>] * <i>football</i> [<i>voetbal</i>]

Table 4: Accuracy of the compound

Word-formation accuracy	Example
Phrase instead of compound	* <i>vakantie van Pasen</i> [<i>Paasvakantie</i>] ‘Easter Holidays’
Simplex instead of compound	<i>een *stap in Padoue</i> [<i>uitstap</i>] ‘excursion’
Compound instead of phrase	* <i>vrijtijd</i> [<i>vrije tijd</i>] ‘spare time’

Table 5: Word-formation accuracy

Table 5 shows the different types of unidiomatic word-formation: the use of a phrase instead of a compound, the use of a simplex instead of a compound, and the use of a compound instead of a phrase. In order to qualify a phrase as ‘inaccurate’, we used the *Corpus Hedendaags Nederlands*¹⁰ as a reference corpus. We set as a threshold that the corresponding compound had to be at least 10 times as frequent as the phrase in this reference corpus to consider the latter an inaccurate use. For instance, the occurrence *mensen van mijn klas* ‘people from my class’ (1 hit) was considered inaccurate because the corresponding compound construction *mijn klasgenoten* ‘my classmates’ (123 hits) was more than ten times as frequent. Conversely, the phrase *in het centrum van de stad* ‘in the center of the city’ (1076 hits) was not considered

¹⁰ <https://ivdnt.org/corpora-lexica/corpus-hedendaags-nederlands/>

inaccurate since it occurs even with a higher frequency than *in het stadscentrum* ‘in the city center’ (956 hits). As explained before (footnote 9), these non-idiomatic uses should be seen as instances on a scale of acceptability in which some cases are just much less frequent than the native use, but should not be considered true mistakes.

6. A multilevel constructional corpus analysis

The annotated compounds were analyzed at the schematic (6.1) and the substantive (6.2) constructional level.

6.1. Schematic level

At the schematic level, we first present the formal properties of the compounds in the three subcorpora (structure and length), before discussing the semantic characteristics of the compounds.

6.1.1. Formal properties

a) Compound structure

Table 6 and Figure 4 show the proportions of different compound structures in the three subcorpora. We observed that the $[N+N]_N$ compounds (example (1)) are “the norm” in the native corpus (67,74%), while the learners use them significantly¹¹ less frequently (non-CLIL 52,92%; CLIL 55,46%).

The second most frequent compounds are $[V+N]_N$ compounds (example (2)) and these are used more frequently by the learners (non-CLIL 36,25%; CLIL 32,56%) than by the natives (25,16%). While the overrepresentation of $[V+N]_N$ is significant in the e-mails of the non-CLIL learners in comparison to the native data ($\chi^2 = 5.33, p = 0.02$), we observed no significant overrepresentation of $[V+N]_N$ in the e-mails of the CLIL learners ($\chi^2 = 3.01, p = 0.08$). As will be shown in Section 6.2.1, the high frequency of $[V+N]_N$ compounds in the dataset can be

¹¹ We observed significant fewer $[N+N]_N$ constructions in the non-CLIL e-mails compared to the natives ($\chi^2 = 8.54, p = 0.00$), as well as significant fewer $[N+N]_N$ constructions in the CLIL e-mails compared to the natives ($\chi^2 = 7.26, p = 0.01$). No significant differences in the use of $[N+N]_N$ between CLIL/ non-CLIL ($\chi^2 = 0.42, p = 0.52$) were found.

explained by the high frequency of some particular compounds, typically used in the contexts of holidays or parties, namely *zwembad* ‘swimming pool’, *verjaardag* ‘lit. celebrate-day; birthday’ and *vliegtuig* ‘lit. flying-object; plane’.

The [A+N]_N constructions (example (3)) are also used more frequently by the learners, especially by the non-CLIL learners (non-CLIL 8,75%; CLIL 4,20%), in comparison to the native speakers (1,94%). Again the overrepresentation of [A+N]_N is significant in the non-CLIL data ($\chi^2 = 7.66, p = 0.01$), while there is no significant difference between the use of [A+N]_N in the e-mails of the CLIL learners and the native speakers ($\chi^2 = 1.71, p = 0.19$).

Except [Prep/Adv+N]_N compounds (example (4)), less than 2% of the compounds in the three subcorpora has another structure, such as [Num+N]_N, [NP+N]_N, [V+Prep]_N or [AP+N]_N. Note that none of these inter-group differences is significant.¹²

- (1) *Daarna trekken we meer het **regenwoud** in en gaan we daar op safari.*¹³ (Native)
‘Then we move more into the **rainforest** and go on safari there.’
- (2) *Maar als je een goed **schuilplaats** vind , dan ben jij in zekerheid , lol !* (CLIL) ‘But if you find a good **hiding place**, then you are in security, lol!’
- (3) *Ze heeft een auto gehad van haar ouders, een nieuwe gsm van zijn **grootouders** en ze heeft veel geld gehad* (non-CLIL) ‘She has had a car from her parents, a new cell phone from his **grandparents** and she has had a lot of money’
- (4) *De camel van Alexis heeft alle **nagerechten** gegeten.* (CLIL) ‘Alexis’ camel ate all the **desserts** (lit. after-dish).’

Compound structure	AL Dutch non-CLIL	AL Dutch CLIL	Control corpus native Dutch
[N+N] _N	127 (52,92%)	264 (55,46%)	105 (67,74%)*
[V+N] _N	87 (36,25%)*	155 (32,56%)	39 (25,16%)
[A+N] _N	21 (8,75%)*	20 (4,20%)	3 (1,94%)
[Prep/Adv+N] _N	5 (2,08%)	25 (5,25%)	5 (3,23%)
[Num+N] _N	0	6 (1,26%)	1 (0,65%)

¹² We observed no significant differences in the use of [Prep/Adv+N]_N between the non-CLIL learners and the natives ($\chi^2 = 0.50, p = 0.48$). We observed no significant differences in the use of [Prep/Adv+N]_N between the CLIL learners and the natives ($\chi^2 = 1.06, p = 0.30$).

¹³ Examples are taken verbatim from the corpus including inaccuracies in language use, punctuation, and spelling.

[NP+N] _N	0	5 (1,05%)	1 (0,65%)
[V+Prep] _N	0	1 (0,21%)	0
[AP+N] _N	0	0	1 (0,65%)
Totals	240	476	155

Table 6: Compound structure per condition (statistically significant differences are indicated by *)

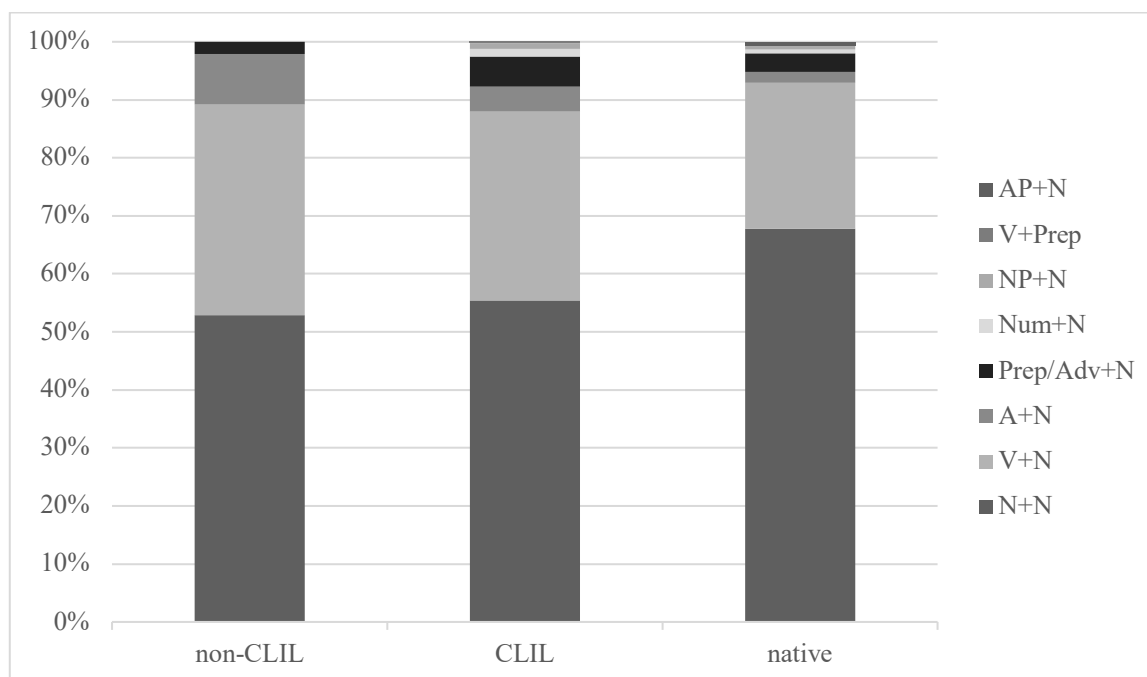


Figure 4: Compound structure per condition

In sum, with respect to compound structure, the learners use significantly fewer [N+N]_N compounds than the natives. Nevertheless, the CLIL learners' use of compounds is more similar to the natives' use of compounds than the non-CLIL learners' use. Concretely, the CLIL learners use more [N+N]_N constructions and fewer [V+N]_N constructions than the non-CLIL learners.

b) Compound length

As shown in Table 7, both the native speakers and the learners use mostly compounds consisting of two components (example (5)), although the learners use these shorter compounds relatively

more frequently (non-CLIL 94,58%; CLIL 93,07%) than the natives (89,68%).¹⁴ While the native speakers use significantly more 3-component compounds (10,32%) (example (6)) in comparison to the non-CLIL learners (3,75%) ($\chi^2 = 6.86, p = 0.01$), there is no significant difference in the use of 3-component compounds between the CLIL learners (6,93%) and the natives ($\chi^2 = 1.86, p = 0.17$). Note that we also observed three occurrences of 4-component compounds in the non-CLIL data (example (7)).

(5) *Er waren veel **nachtclub** en het was niet duur.* (non-CLIL) ‘There were many **nightclub** and it was not expensive’

(6) *Ik dacht vandaag weer terug aan de leuke avond die we gisteren hadden op de **kerstboomverbranding** en in de winterbar* (Native) ‘I thought back today to the fun night we had yesterday at the **Christmas tree burning** and at the winter bar’

(7) *We hebben een groot **voetbal wedstrijd** met anderen jongeren georganiseerd.* (non-CLIL). ‘We organized a big **soccer game** with other teenagers.’

Compound length	AL Dutch non-CLIL	AL Dutch CLIL	Control corpus 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

Table 7: Compound length (statistically significant differences are indicated by *)

In sum, the native speakers of Dutch use relatively more long compounds (3 components) than the learners, especially more than the non-CLIL learners. Nevertheless, the longest compounds were produced by the non-CLIL learners.

6.1.2. Semantic properties

As shown in Table 8 and Figure 5, the analysis of the compounds in terms of their semantic function yields rather uniform results: attributive endocentric compounds (example (8)) are “the

¹⁴ No significant difference in the use of 2-components compounds in the data of the non-CLIL learners and the natives ($\chi^2 = 3.33, p = 0.07$). No significant difference in the use of 2-components compounds in the data of the CLIL learners and the natives ($\chi^2 = 1.88, p = 0.17$).

native norm” (98,71%) and also predominant in the learner data (non-CLIL 94,58%; CLIL 95,59%). Yet, the CLIL learners use these compounds in more native-like proportions: whereas there is a significant difference between the non-CLIL learners and the native speakers in their use of attributive endocentric compounds ($\chi^2 = 4.39, p = 0.04$), there is no significant difference between the CLIL learners and the natives ($\chi^2 = 3.24, p = 0.07$).

The second most frequent compound function consists of attributive exocentric compounds (example (9)), which are used slightly more frequently by the learners (non-CLIL 3,75%; CLIL 3,78%) than by the native-speakers (1,29%).¹⁵ The learners also use a small number of subordinative endocentric compounds (example (10)) (non-CLIL 1,67%; CLIL 0,42%). Moreover, the CLIL subcorpus also contains one subordinative exocentric compound (example (11)).

- (8) *het was de verjaardag van mijn buurjongen.* (Non-CLIL) ‘it was my **neighbor** boy's birthday.’
- (9) *En dan rook ge echt zo de geur van vuurwerk.* (Native) ‘And then you really smelled the smell of fireworks.’
- (10) *Daarna hebben we de GSM gedrogen met een haardroger en we hebben zijn geld op de chauffage gezet.* (CLIL) ‘Then we dried the cell phone with a **hair dryer** and we put his money on the heater’
- (11) *Er waren veel mensen in de straten en wij waren bang voor pickpocketen.*
[zakkenrollers] (CLIL) ‘There were a lot of people in the streets and we were afraid of pickpockets.’

Compound function	AL Dutch non-CLIL	AL Dutch CLIL	Control corpus native Dutch
Attributive endocentric	227 (94,58%)*	455 (95,59%)	153 (98,71%)
Attributive exocentric	9 (3,75%)	18 (3,78%)	2 (1,29%)
Subordinative endocentric	4 (1,67%)	2 (0,42%)	0
Subordinative exocentric	0	1 (0,21%)	0
Coordinative exocentric	0	0	0
Totals	240	476	155

¹⁵ No significant difference was observed in the use of attributive exocentric compounds between non-CLIL learners and natives ($\chi^2 = 2.10, p = 0.15$). Additionally, no significant difference was found in the use of attributive exocentric compounds between CLIL learners and natives ($\chi^2 = 2.10, p = 0.15$).

Table 8: Compound function per condition (statistically significant differences are indicated by *)

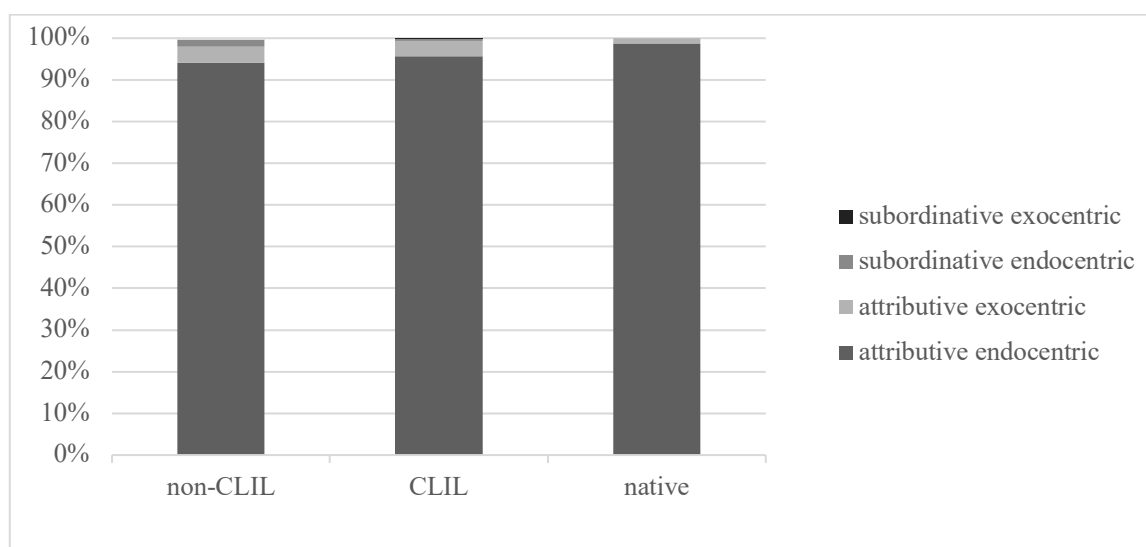


Figure 5: Compound function per condition

6.2. Substantive level

At the substantive constructional level, we examined the frequency and productivity of the compound constructions as well as the accuracy of use of these constructions in the different subcorpora.

6.2.1. Frequency and productivity

Table 11 provides an overview of the five most frequent compounds used in each dataset. Not surprisingly, most of these frequent compounds are closely related to the central topics of the writing tasks (an email about a party or about holidays). It turns out that the same two compounds are used most frequently in the three datasets: *zwembad* ‘swimming pool’ and *verjaardag* ‘birthday’ (in this order in the non-CLIL and control corpus; in the inverse order in the CLIL corpus). Nonetheless, three compounds occur in the top five of the control data, but not in the top five of the learner corpora: *dansvloer* ‘dance floor’, *halfuur* ‘half hour’ and *zonsondergang* ‘sunset’. In turn, the learner corpora contain a few frequent compounds that do not figure in the top five of the most frequent compounds of the control corpus: **aeroport* ‘airport’, *voetbal* ‘football’ and *volleyball* ‘volleyball’ in the non-CLIL data and *Paasvakantie*

‘Easter holidays’ and *ziekenhuis* ‘hospital’ in the CLIL data. Note that **aeroport* is a French compound noun, the Dutch equivalent compound being *vliegveld* ‘airport’.

Top 5	AL Dutch non-CLIL	AL Dutch CLIL	Control corpus native Dutch
1	<i>zwembad</i> (27) ‘swimming pool’	<i>verjaardag</i> (54) ‘birthday’	<i>zwembad</i> (9) ‘swimming pool’
2	<i>verjaardag</i> (22) ‘birthday’	<i>zwembad</i> (52) ‘swimming pool’	<i>verjaardag</i> (6) ‘birthday’
3	<i>vliegtuig</i> (19) ‘aeroplane’	<i>vliegtuig</i> (17) ‘aeroplane’	<i>verjaardagsfeest</i> (5) ‘birthday party’
4	<i>*aeroport</i> (6) ‘airport’	<i>Paasvakantie</i> (10) ‘Easter holidays’, <i>verjaardagsfeest</i> (10) ‘birthday party’	^o <i>dansvloer</i> (4) ‘dance floor’
5	<i>voetbal</i> (5) ‘football’, <i>volleybal</i> (5) ‘volleyball’, <i>verjaardagsfeest</i> (5) ‘birthday party’	<i>ziekenhuis</i> (9) ‘hospital’	^o <i>halfuur</i> (3) ‘half hour’, <i>vliegtuig</i> (3) ‘aeroplane’, ^o <i>zonsondergang</i> (3) ‘sunset’
Total number of compounds	240	476	155

Table 11: Most frequent compounds per condition (number of tokens indicated between brackets)

^o= occurs in native corpus, but not in the learner corpora

* = inaccurate use

Table 12 and Figure 6 present the type-token ratio (TTR) and the potential productivity (PP) or hapax legomena ratio (Baayen 2009) of the compounds per condition. The compounds in the control corpus show the highest productivity both in terms of the type-token ratio and the hapax legomena ratio. Surprisingly, the compounds in the non-CLIL dataset display a slightly higher

productivity than the compounds in the CLIL dataset. However, the productivity measures appear to be affected by the considerable number of lexical inaccuracies or cases of codeswitching in the non-CLIL hapaxes (see also Section 6.2.2 on Accuracy). Almost 60% of the total of hapax legomena in the non-CLIL dataset contains lexical inaccuracies or codeswitching (53/89); in the CLIL group this proportion is lower (74/139 or 53,24%). The proportion of codeswitching in the hapaxes in particular is significantly higher in the non-CLIL subcorpus than in the CLIL subcorpus ($\chi^2 = 7.97$ $p = 0.00$) (see example (12)).

(12) *Het koek kwam van mijn vriend die een **aspirant-pâtissier** is.* (non-CLIL) ‘The cake came from my friend who is an **aspiring pastry chef**.’

Productivity	AL Dutch non-CLIL	AL Dutch CLIL	Control corpus native Dutch
TTR	118/240 = 0,49	208/476 = 0,44	116/155=0,75
PP	89/240 = 0,37	139/476 = 0,29	96/155=0,62
Totals	240	476	155

Table 12: Productivity of compounds per condition

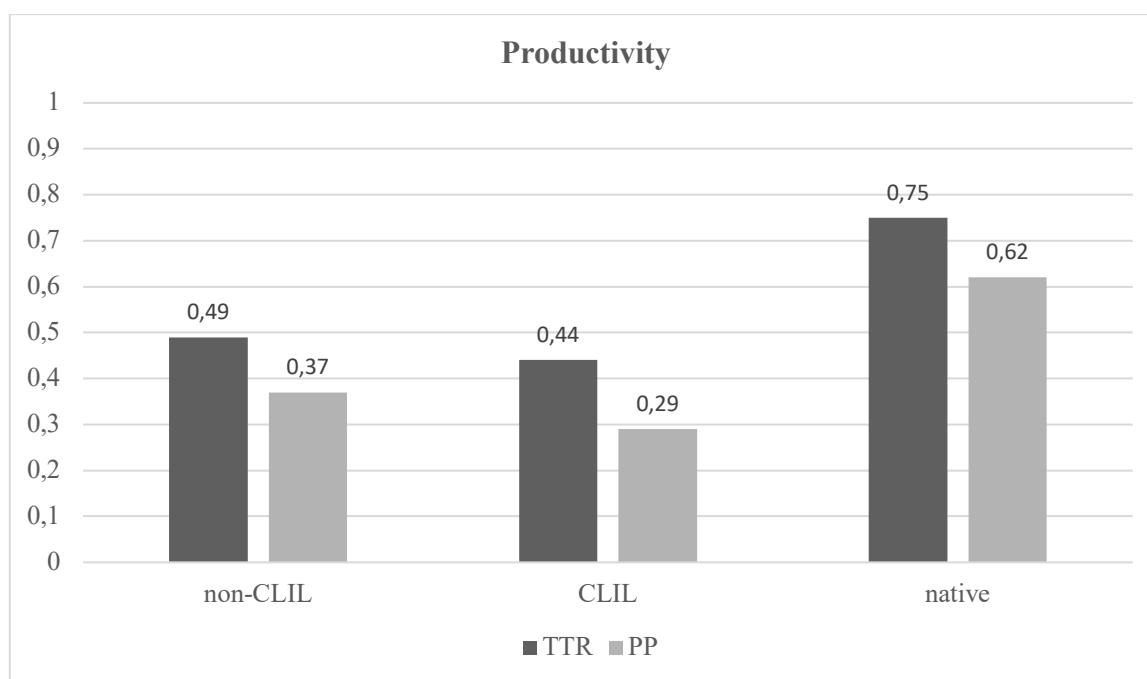


Figure 6: Productivity of compounds per condition (TTR and PP)

6.2.2. Accuracy

In addition to analyzing the frequency and the productivity of the compounds, we analyzed the accuracy of the compounds and of the word-formation strategies.

a) Inaccurate use in the compound constructions

As indicated in Section 4.3, we annotated the compounds for spelling mistakes, lexical inaccuracies, codeswitching, grammatical mistakes, mistakes in the use of the linking morpheme and mistakes in the word order. Table 13 and Figure 7 show the frequencies and proportions of these different categories in both learner groups.

Inaccurate use in the compound nouns	AL Dutch non-CLIL	AL Dutch CLIL
Spelling mistake	45 (18,75%)	96 (20,17%)
Lexical mistake	8 (3,33%)	16 (3,36%)
Codeswitching	33 (13,75%)*	18 (3,78%)*
Grammatical mistake	35 (14,58%)	66 (13,87%)
Linking morpheme	7 (2,92%)	19 (3,99%)
Word order	1 (0,42%)	3 (0,63%)
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

Table 13: Inaccurate use in the compound nouns per condition (statistically significant results are indicated by *)

Overall, the non-CLIL students produce slightly more inaccurate compounds than the CLIL students (respectively 46,25% and 39,50%), but both groups of learners seem to encounter the same difficulties. Spelling mistakes make up most of the observed inaccuracies (non-CLIL 18,75% vs. CLIL 20,17%) (example (13)). Grammatical mistakes make up 14,58% of the inaccuracies in the non-CLIL data and 13,87% of the inaccurate use in the CLIL compounds (example (14)). The third most frequent case of inaccurate use consists of codeswitching

(example (15)). Only in this category we observe a significant difference between both learner groups, the non-CLIL students producing significantly more codeswitching (non-CLIL 13,75% vs. CLIL 3,78%) ($\chi^2 = 24.00, p < 0.00001$). Lexical inaccuracies are equally frequent in the non-CLIL data (3,33%) and the CLIL data (3,36%) (example (16)). A few compounds contain misuse of the linking morpheme (non-CLIL 2,92% vs. CLIL 3,99%) (example (17)). Finally, contrary to our expectations, we find very few mistakes in the word order of the compounds (non-CLIL 0,42% vs. CLIL 0,63%) (example (18)).

- (13) *We zullen met de auto gaan omdat mijn moeder voor het **vliegtruig [vliegtuig]** bang is?* (Non-CLIL) ‘We will go by car because my mother is afraid of the airplane?’ (Non-CLIL)
- (14) *Iedereen houdt van **vuurwerken [vuurwerk]**, en een beetje in de frisse , zwarte nacht stappen en naar de sterren kijken* (CLIL) ‘Everyone loves fireworks, and going out into the fresh, black night and looking at the stars.’
- (15) *Er was een **babyfoot [tafelvoetbal]** en een ball van football.* (Non-CLIL) ‘There was a table soccer and a football’
- (16) *Bovendien waren er geen proper **voetweg [wandelpad]** zoals het uitgelegt werd in onze guide.* (CLIL) ‘Besides, there were no proper hiking path as it was explained in our guide.’
- (17) *Daarna hebben we een **aardbeienstaart [aardbeientaart]** gegeten dat de koker had gemaakt.* (CLIL) ‘Then we ate a strawberry cake that the cook had made.’
- (18) *een nieuwe jas , een **album foto [fotoalbum]** een fototoestel, nieuwe zonnebrillen, ...* (CLIL) ‘a new coat , a photo album, a camera, new sunglasses, ...’ (CLIL)

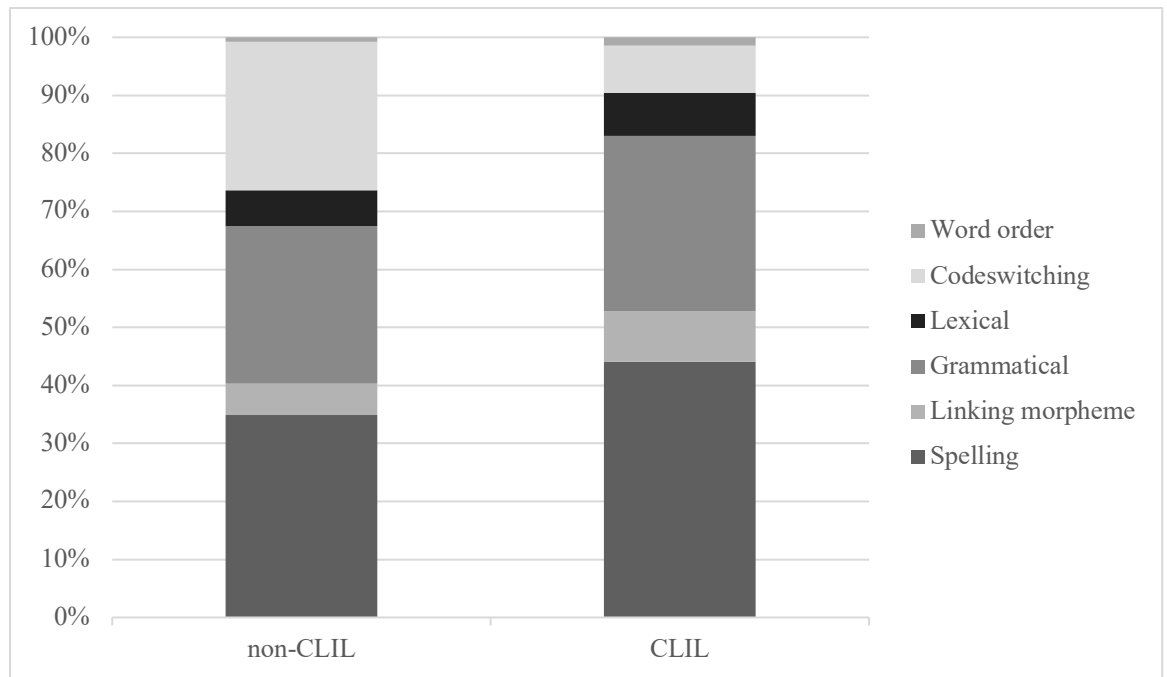


Figure 7: Inaccurate use in the compound nouns per condition

b) Word-formation inaccuracies

If we now turn to unidiomatic instances of word formation (see Table 14 and Figure 8), we observe that the most frequent phenomenon consists of the use of phrases instead of compounds (example (19)), as expected (slightly more but not significantly so in the non-CLIL data¹⁶). The second most frequent inaccurate use in word formation consists of the use of compounds instead of phrases (example (20)), again slightly more but not significantly so in the non-CLIL data¹⁷. This latter phenomenon was however not expected and, instead, points to overgeneralization of the Dutch idioconstruction in the form of a compound. A final recurring case of inaccurate word formation, but only in the CLIL dataset, is the use of a simplex word instead of a compound: example (21) implies simplification of the Dutch compound construction (*stap* instead of *uitstap*). The learner data contain only one occurrence of the use of a compound instead of a simplex (example (22)).

¹⁶ We observed no significant difference between the CLIL and non-CLIL groups in their use of phrases instead of compounds: $\chi^2 = 3.18$; $p = .07$.

¹⁷ We observed no significant difference between the CLIL and non-CLIL groups in their use of compounds instead of phrases: $\chi^2 = 0.01$; $p = .92$. Not significant at $p < .05$.

(19) *ik kan niet goed zwemmen maar ik had 2 lessen van zwembad [zwemlessen] (...)* (non-CLIL) ‘I can't swim well but I had 2 swimming lessons [lit. lessons from swimming pool] (...)’

(20) *dat was een idee op mijn zusje en mijn bestvriend [beste vriend] Floriane.* (non-CLIL) ‘that was an idea of my sister and my best friend Floriane’

(21) *Daarna zijn we naar Venesie gegaan met een stap [uitstap] in Padoue waar we de stad , de katedraal en de klooster hebben bezocht .* (CLIL) ‘Then we went to Venice with a trip to in Padua where we visited the city, the cathedral and the convent.’

(22) *Ik was de kookman [kok], ik weet het als het best idee was want ik heb nooit koken van mijn hele leven* (CLIL) I was the cook (lit. cooking man), I don't know if it was the best idea because I have never cooked of my whole life’

Inaccurate word formation	AL Dutch non-CLIL	AL Dutch CLIL
Phrase instead of compound	15 (0,47%)	47 (0,79%)
Simplex instead of compound	0	13 (0,22%)
Compound instead of phrase	4 (0,12%)	7 (0,12%)
Compound instead of simplex	0	1 (0,02%)
Total number of word-formation mistakes	19	68
Total number of nouns	3205	5964
Proportion mistakes/nouns	0.59%	1.14%*

Table 14: Inaccurate word formation per condition (statistically significant differences are indicated by *)

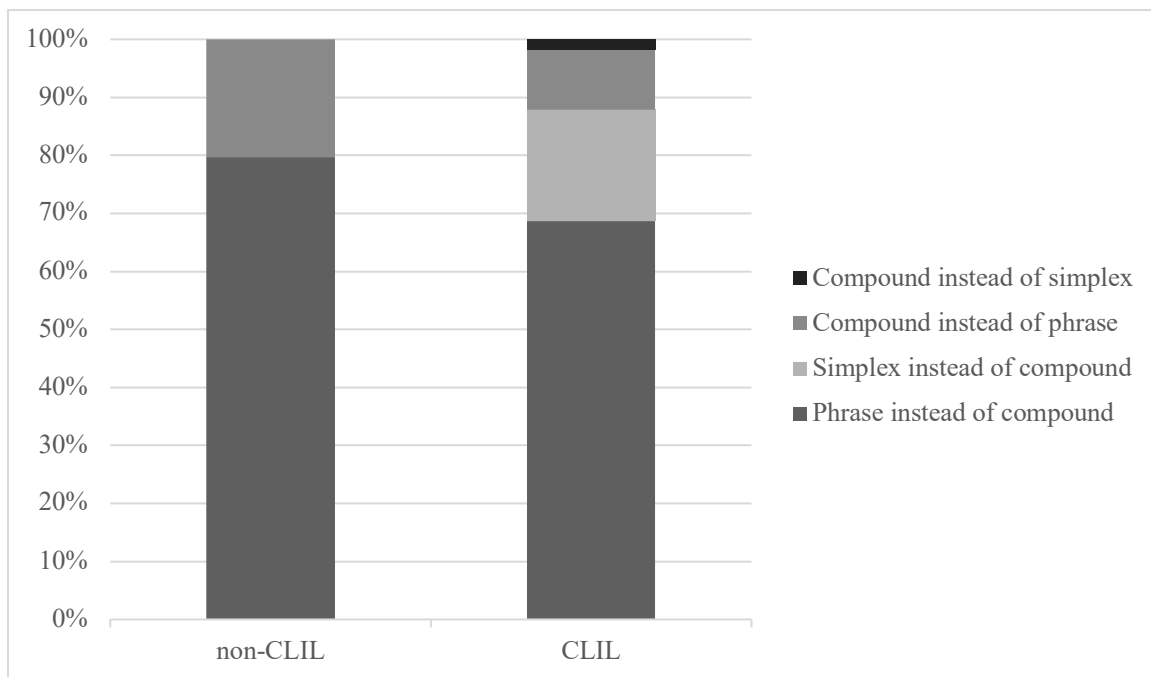


Figure 8: Inaccurate word formation per condition

In sum, considering the accuracy of word formation, we observe that the non-CLIL group produces proportionally more often a “phrase instead of compound”, as expected. However, in both groups we also observe the inverse phenomenon: “compound instead of phrase”. This may be related to phenomena such as hypercorrection or overgeneralization. Finally, the CLIL students produce overall more ($\chi^2 = 6.65; p = .01$) and more diverse word-formation inaccuracies than the non-CLIL students. The CLIL students also use simplexes instead of compounds and compounds instead of simplexes, for instance. This diversity in difficulties that the CLIL students encounter in forming words may indicate that the CLIL students are more at ease to employ their linguistic creativity (e.g., *kookman* instead of *kok*). In the next section we will interpret these results in more depth.

7. Discussion of the results

In his study we analyzed the acquisition of compound nouns at two different levels of abstraction (schematic and substantive compound constructions), exploring the application of the diasystematic-constructionist perspective to AL acquisition. Moreover, we investigated the impact of additional target-language input through CLIL programs on the acquisition of compound constructions.

Adopting a diasystematic-constructionist approach, we expected an overgeneralization of French idioconstructions, i.e., phrasal patterns. Concretely, we assumed to observe a significant overuse of phrasal structures in contexts where a compound is more appropriate in Dutch. The corpus analyses show that the use of a Dutch phrase instead of a compound is indeed the most frequent word-formation inaccuracy in the learner corpora. This implies that the learners misinterpreted some of the constructions as French-Dutch phrasal diaconstructions, which instead should be tagged as idioconstructions typical of French. Moreover, the compound construction, typical of Dutch, does not seem to be sufficiently entrenched yet. The difficulty for the learners lies, among other things, in the fact that in some cases both phrases and compounds are correct in Dutch (e.g., *verjaardagsfeest* ‘birthday party’ vs *feest voor mijn verjaardag* ‘party for my birthday’), but this is not a general rule (e.g., *vrije tijd* ‘spare time’ but not **vrijtijd*). In many cases, Dutch compounds should be used instead of phrases (e.g., *zwemlessen* ‘swimming classes’ vs **lessen van zwemmen/zwembad* ‘classes of swimming/swimpool’). This means that learners have to form French-Dutch diaconstructions in some cases relying on a shared schematic phrasal pattern, while in others they need to resort to a specific Dutch idioconstruction to create the proper compound.

While the use of a phrase instead of a compound was the most frequent word-formation inaccuracy that we found in the learner data, we encountered also cases of overgeneralization, where learners used a compound instead of a phrase or instead of a simplex word. From a diasystematic-constructionist perspective, we interpret these misused compounds as shared diaconstructions that are mistagged as Dutch idioconstructions.

Another expected overgeneralization of a French idioconstruction, misinterpreted as French-Dutch diaconstruction, consists of the use of left-headed compounds instead of right-headed compounds. Yet this hypothesis could not be corroborated. The difference in word order between French and Dutch does not seem to be a great obstacle for the learners in this study, since we observed very few word-order mistakes in the compound formation. This implies that the Dutch right-headed word order is correctly tagged by the learners. A possible explanation for the fact that word order seems to be acquired with more ease, whereas the choice for the phrasal construction or the compound construction was not entirely mastered yet, may be related to the frequency and saliency of these aspects, which both play a crucial role in the acquisition of a particular target-language feature, as discussed in Sections 2 and 3. Since constructions that occur frequently and cohesively in a learners’ input are expected to be acquired earlier, we hypothesize that the correct use of phrasal or compound constructions is harder to acquire for learners than the use of right-headed compounds, because the phrasal and

the compound constructions co-occur in Dutch, while left-headed compounds do not exist at all in Dutch. Hence, even with a limited amount of exposure to Dutch as an AL, learners will more quickly observe that in their target language compounds are right-headed while their first language exclusively forms left-headed compounds.

The second aim of our study was to investigate the impact of additional target-language input through CLIL programs. However, our corpus results only show a partial positive impact of CLIL on the accuracy and the productivity of the compounds and word-formation in AL Dutch. More specifically, we observed a better entrenchment of the newly acquired Dutch compound constructions in the CLIL-group at the schematic level, that is to say, the results of the CLIL group are closer to the ones of the natives with respect to the compound length, structure and function. However, this is not the case at the substantive level, that is to say, with respect to productivity and frequency. At the latter level, the non-CLIL students produce a greater variety of compounds. However, the measures of variety and productivity appeared to be affected by a great number of inaccuracies in the hapax legomena in the non-CLIL subcorpus.

Zooming in on the accuracy of the compounds in the different learner corpora, we registered significantly more codeswitching in the non-CLIL group, a specific example of overgeneralization of French (but also English) substantive idioconstructions. Somewhat surprisingly, we found however significantly more word-formation inaccuracies in the CLIL group than in the non-CLIL group. Typical non-idiomatic word-formation in the CLIL data concerns the use of simplexes instead of compounds and compounds instead of simplexes. These unexpected utterances could be illustrative of the learners developing linguistic creativity. As explained in Section 3.1, constructionist theories on language acquisition assume that learners first acquire substantive constructions as a whole (e.g., *vliegtuig* ‘airplane’, *ziekenhuis* ‘hospital’) before deriving more schematic patterns. The latter abstract patterns may subsequently be used to create new compounds (e.g., **kookman* ‘cook man’). Our results appear to indicate that the CLIL learners go beyond the use of substantive constructions which they acquired as a whole. More often than the non-CLIL learners, they have started to productively apply the abstract compound construction to form new concrete outcomes, with varying success however. According to this hypothesis, the CLIL learners would have added besides the Dutch substantive idioconstructions (concrete frequent compounds) also the Dutch schematic idioconstruction compounding pattern $[N_1 N_2]_N$ to their construction, and experiment with filling in this schematic construction at the substantive level.

8. Conclusion and outlook

In this study we investigated the AL acquisition of compound nouns at two different levels of abstraction and analyzed the impact of additional target-language input through CLIL programs on the acquisition of these constructions.

The study was framed in the theoretical framework of Diasystematic Construction Grammar, which conceptualizes the linguistic competence of multilingual speakers as one integrated network of constructions, containing language-specific *idioconstructions* and shared *diaconstructions*. As Dutch has a much stronger tendency towards compounding than French, we expected an overgeneralization of French idioconstructions in the learners' productions, which would be illustrated by a significant overuse of phrasal structures in contexts where a compound should be used. The results showed indeed that the the most frequent word-formation inaccuracy in the learner corpora involved the use of phrases instead of compounds. Similarly, we expected to observe the use of left-headed compounds (as in French) instead of right-headed compounds in the learners' productions, but this hypothesis was not corroborated. The fact that our first hypothesis was confirmed, but not the second, is probably related to differences in frequency and saliency of these target-language features. Both phrases and compounds are commonly used in Dutch – hence learners must become aware of the contexts in which either of these constructions is preferred, and in which both are possible – while left-headed compounds are completely absent in Dutch, enabling the learners to master this difference earlier on in the acquisition process.

The second goal of our study was to analyze the impact of additional target-language input through CLIL programs. Our results show that while CLIL appears to have no effect at the substantive level with respect to productivity and frequency, it positively affects the acquisition of compounds at the schematic level, that is to say, compound length, structure and function. Moreover, we suggest that the greater proportion of non-idiomatic word-formation in the CLIL productions indicates that the CLIL learners are one step ahead of the non-CLIL learners, having integrated the schematic nominal compound pattern and experimenting with the productivity of these schematic constructions at the substantive level.

Embedding the results of our study into the DCxG framework has several advantages. First, the constructionist perspective invites us to approach additional language acquisition at different levels of abstraction, and our study highlights that the results found for one level are not systematically transposable to another one. Therefore, it allows for a more fine-grained

analysis of the acquisition of specific linguistic structures. Second, combining Construction Grammar with the most recent insights into multilingualism makes DCxG an interesting model to explore additional language acquisition. As corpus linguists we do not have direct access to the cognitive processes taking place in the learner's mind, but the empirical corpus analysis enables us to describe the particular formal and semantic properties of the learner data and to compare the results obtained by two different learner groups, CLIL and non-CLIL learners. Hence, learner corpus research framed in DCxG provides indirect access to the cognitive restructuring processes taking place in the minds of language learners.

The present study complements a previous one on the acquisition of intensifying constructions from a DCxG perspective (Van Goethem & Hendrikx 2021). We invite fellow researchers to adopt a similar approach to further explore and develop the cross-fertilization of Construction Grammar, Learner Corpus Research and Additional Language Acquisition. As for our part, we intend to explore new case studies in the field of word formation such as adjectival compounds and complex verbs in order to possibly validate or refine the results of the present study.

9. References

- Baayen, H. (2009). Corpus linguistics in morphology: Morphological productivity. In A. Lüdeling & M. Kytö (Eds), *Corpus linguistics. An international handbook* (pp. 900-919). Berlin & New York: Mouton de Gruyter.
- 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.
- Beheydt, L. (2014). Immersieonderwijs en contrastieve taalkunde. In L. Degand, Ph. Hiligsmann, L. Rasier, M. Sergier, S. Vanasten & K. Van Goethem (Eds.), *In het teken van identiteit. Taal en cultuur van de Nederlanden* (pp. 161-178). Louvain-la-Neuve: Presses Universitaires de Louvain.
- 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., & Höder, S. (Eds.) (2018). *Constructions in Contact. Constructional perspectives on contact phenomena in Germanic languages (Constructional Approaches to Language, 24)*. Amsterdam / Philadelphia: John Benjamins.
- Boas, H. C. & Höder, S. (eds). (2021). *Constructions in Contact 2. Language change, multilingual practices, and additional acquisition*. Amsterdam / Philadelphia: John Benjamins Publishing Company.
- 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). Compounds and multi-word expressions in Dutch. In B. Schlücker (Ed.), *Complex Lexical Units: Compounds and Multi-Word Expressions* (pp. 95-126). Berlin, Boston: De Gruyter.
- Bulon, A. (2019). *The acquisition of phraseological units by French-speaking learners of English and Dutch in CLIL and NON-CLIL settings: exposure effects on range and accuracy*. Unpublished PhD thesis. Université catholique de Louvain, Louvain-la-Neuve, Belgium.
- Bybee, J.L. (2010). *Language, Usage and Cognition*. New York: Cambridge University Press.
- 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.
- Cook, E.P.V. (2016). Transfer and the relationship between the languages of multi-competence. In R.A. Alonso (Eds.) *Crosslinguistic Influence in Second Language Acquisition* (pp. 24-37). Bristol: Multilingual Matters.
- Corpus Hedendaags Nederlands*. <http://chn.inl.nl/>
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge: Cambridge University Press.
- Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism, 19*, 121-129.
- Dalton-Puffer, C. (2007). *Discourse in Content and Language Integrated (CLIL) Classrooms*. Amsterdam / Philadelphia: John Benjamins.

- De Groot, A. (2011). *Language and Cognition in Bilinguals and Multilinguals. An Introduction*. New York and Hove: Psychology Press.
- Diessel, H. (2016). *Frequency and lexical specificity in grammar*. New York: Oxford University Press.
- Doughty, C. (1991). Second language instruction does make a difference: Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition*, 13, 431-469.
- Don, J. (2009). IE, Germanic: Dutch. In: Lieber, Rochelle and Štekauer, Pavol (Eds.), *The Oxford handbook of compounding*. Oxford: Oxford University Press. 370-385.
- Ellis, N. C., & Cadierno, T. (2009). Constructing a Second Language. Introduction to the Special Section. *Annual Review of Cognitive Linguistics*, 7, 111-139.
- Ellis, N. C., & Sagarra, N. (2011). Learned attention in adult language acquisition. *Studies in Second Language Acquisition*, 33(4), 589-624.
- Ellis, N. C., & Wulff, S. (2008). Usage-based approaches to second language acquisition. In B. VanPatten & J. Williams (Eds.), (2nd Edition). *Theories in Second Language Acquisition: An Introduction* (pp. 75-93). New York & London: Routledge.
- Eurostat. (2016). More than 80% of primary school pupils in the EU were studying a foreign language in 2014. Retrieved from <https://ec.europa.eu/eurostat/documents/2995521/7662394/3-23092016-AP-EN.pdf/57d3442c-7250-4aae-8844-c2130eba8e0e> [Last accessed 11-01-2019]
- Fradin, B. (2009). IE, Romance: French. In: Lieber, Rochelle and Štekauer, Pavol (Eds.), *The Oxford handbook of compounding*. Oxford: Oxford University Press. 417-435.
- Goldberg, A. (2006). *Constructions at Work: the nature of generalisation in language*. Oxford: Oxford University Press.
- Hartsuiker R. J. & Bernolet, S. (2017). The development of shared syntax in second language learning. *Bilingualism: Language and Cognition*, 20(2), 219-234.
- Hartsuiker, R.J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science*, 15(6), 409-414.
- Hendriks, I. (2019). *The acquisition of intensifying constructions in Dutch and English by French-speaking CLIL and non-CLIL students: Cross-linguistic influence and exposure effects*. Unpublished PhD thesis. Louvain-la-Neuve: Université catholique de Louvain.
- Hilgsmann, Ph. (1997). *Linguïstische aspecten en pedagogische implicaties van de tussentaal van Franstalige M.O.-leerders van het Nederlands*. Genève: Droz.

- Hiligsmann, Ph., Van Mensel, L., Galand, B., Mettwie, L., Meunier, F., Szmalec, A., Van Goethem, K., Bulon, A., De Smet, A., Hendriks, 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.
- Hilpert, M., & Östman, J.-O. (2014). Reflections on constructions across grammars. *Constructions and Frames*, 6(2), 137-142.
- Höder, S. (2012). Multilingual constructions: a diasystematic approach to common structures. In: K. Braunmüller & C. Gabriel (Eds.), *Multilingual individuals and multilingual societies* (pp. 241-257). Amsterdam / Philadelphia: John Benjamins.
- 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 / Philadelphia: John Benjamins.
- Höder, S. (2014b). Convergence vs. divergence from a diasystematic perspective. In K. Braunmüller, S. Höder & K. Köhl (Eds.), *Stability and divergence in language contact: Factor and mechanisms* (pp. 39-60). Amsterdam / Philadelphia: John Benjamins.
- 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* (pp. 37-70). Amsterdam / Philadelphia: John Benjamins.
- Höder, S., Prentice, J. & Tingsell, S. (2021). Additional language acquisition as emerging multilingualism: A Construction Grammar approach. In Hans C. Boas & Steffen Höder (Eds.) *Constructions in Contact 2: Language change, multilingual practices, and additional acquisition* (pp. 309-337). Amsterdam / Philadelphia: John Benjamins Publishing Company.
- Hoffmann, T., & Trousdale, G. (2013). *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press.
- 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. (2017). Usage-Based Approach to Preposition Placement in English as a Second Language: Preposition Placement in L2 English. *Language Learning*, 68(1), 271-304.
- Jach, D. (2021). *Something I was dealing with*. Preposition placement in multilingual constructions. In Hans C. Boas & Steffen Höder (Eds.) *Constructions in Contact 2:*

- Language change, multilingual practices, and additional acquisition* (pp. 339-374). Amsterdam / Philadelphia: John Benjamins Publishing Company.
- Jarvis, S. (2002). Topic continuity in L2 English article use. *Studies in Second Language Acquisition*, 24, 387-418.
- Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York: Routledge.
- Jeon, E. H. (2011). Contribution of morphological awareness to second-language reading comprehension. *The Modern Language Journal*, 95(2), 217-235.
- Lado, R. (1957). *Linguistics across cultures. Applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.
- Lamiroy, B. (2011). Degrés de grammaticalisation à travers les langues de même famille. *Mémoires de la Société de linguistique de Paris*, 19, 167-192.
- Lefer, M.-A. (2011). Contrastive word-formation today: Retrospect and prospect. *Poznan Studies in Contemporary Linguistics*, 47(4), 645-682.
- Leontjev, D., Huhta, A. & K. Mäntylä. (2016). Word derivational knowledge and writing proficiency: How do they link? *System*, 59, 73-89.
- Long, M., & Sato, C. (1983). Classroom foreigner talk discourse: Forms and functions of teachers' questions. In H. Seliger & M. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 268–285). Rowley, MA: Newbury House.
- Mettewie, L., & Van Mensel, L. (2009). Multilingualism at all costs. Language use and language needs in business in Brussels. *Sociolinguistica*, 23, 131-149.
- Mettewie, L., & Van Mensel, L. (2020). Understanding foreign language education and bilingual education in Belgium: a (surreal) piece of cake. *International Journal of Bilingualism and Bilingual Education*, 23. DOI: 10.1080/13670050.2020.1768211.
- Meunier, F., Hendriks, I., Bulon, A., Van Goethem, K. & Naets, H. (2020). MulTINCo: Multilingual Traditional Immersion and Native Corpus. Better-documented multi-literacy practices for more refined SLA studies. In L. Van Mensel & Ph. Hilgsmann (Eds.) *Assessing CLIL: A multidisciplinary approach* [special issue]. *Journal of Bilingual Education and Bilingualism*. 1-19.
- Ortega, L. (2013). *Understanding Second Language Acquisition*. London & New York: Routledge.
- Pasquarella, A., X. Chen, K. Lam, Y.C. Luo & G. Ramirez. (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, Barbara. (Ed.). (2019). *Complex lexical units: compounds and multi-word expressions*. Berlin: De Gruyter.
- Schmid, H. (2015). A blueprint of the Entrenchment-and-Conventionalization Model. In P. Uhrig & Th. Herbst (Eds.), *Yearbook of the German Cognitive Linguistics Association 3* (pp. 3-27). Berlin: De Gruyter Mouton.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10(3), 209-231.
- Tomasello, M. (2003). *Constructing a language*. Boston: Harvard University Press.
- Treffers-Daller, J. & Tidball, F. (2015). Can L2 learners learn new ways to conceptualise events? Evidence from motion event construal among English-speaking learners of French. In P. Guijarro-Fuentes, K. Schmit & N. Müller (Eds.), *The acquisition of French in multilingual contexts* (pp. 145-184). Bristol: Multilingual Matters.
- Van Goethem, K. & Amiot, D. (2019). *Compounds and multi-word expressions in French*. In Barbara Schlücker (Ed.), *Complex lexical units. Compounds and multi-word expressions* (pp. 127-152). Berlin: De Gruyter.
- 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)* (pp. 375-428). Amsterdam/Philadelphia: John Benjamins.
- 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 Haeringen, C. B. (1956). *Nederlands tussen Duits en Engels*. The Hague: Servire.
- Van Mensel, L. & Hiligsmann, Ph. (Eds.) (2021). Assessing CLIL: a multidisciplinary approach. *Thematic issue. International Journal of Bilingual Education and Bilingualism*, DOI: 10.1080/13670050.2021.1935441
- Villoing, F. (2012). French compounds. *Probus*, 24, 29-60.

Isa Hendrikx
Rue du Pitteurs 20
4020 Liège, Belgium

E-mail: isa.hendrikx@uliege.be
Orcid: <https://orcid.org/0000-0002-6441-9746>

Kristel Van Goethem
Place Blaise Pascal 1/L3.03.33
1348 Louvain-la-Neuve, Belgium

E-mail: kristel.vangoethem@uclouvain.be
Orcid: <https://orcid.org/0000-0003-2576-9011>