

The relationship between L1 and L2 stance use in novice academic writing

La relación entre el uso de marcadores de posicionamiento en la
L1 y la L2 en la escritura académica de escritores noveles

Pauline Jadoulle – Magali Paquot
UCLouvain / Belgium

Abstract – Due to its combination of lexical, phraseological, and discursive features, stance is susceptible to cross-linguistic influence, especially among novice writers. However, research exploring the relationship between L1 and L2 stance use remains limited, and faces challenges related to attribution, sample size, and comparability. Additionally, much of this research fails to account for individual variation. This paper addresses these gaps by examining the relationship between L1 and L2 stance use in a corpus of paired L1-L2 texts by the same French-speaking students, supplemented by spot checks in a corpus of L1 English novice writing. The results reveal strong similarities in stance marker frequencies and lexical/phraseological preferences across L1 and L2 texts, suggesting that cross-linguistic influence shapes novice writers' stance use. The findings also highlight the importance of individual variation in stance use, which aligns with recent studies emphasizing the necessity of accounting for such variation when investigating L2 data.

Keywords – L2 writing; novice writing; L1-L2 relationship; cross-linguistic influence; stance

Resumen – Debido a la combinación de rasgos léxicos, fraseológicos y discursivos que lo caracterizan, el posicionamiento es susceptible a la influencia interlingüística, especialmente entre los escritores noveles. Sin embargo, las investigaciones que exploran la relación entre el uso del posicionamiento en la L1 y en la L2 siguen siendo limitadas y presentan dificultades relacionadas con la atribución, el tamaño de la muestra y la comparabilidad. Además, gran parte de estos estudios no tiene en cuenta la variación individual. Este artículo aborda estas lagunas mediante el análisis de la relación entre el uso del posicionamiento en la L1 y en la L2 en un corpus de textos emparejados en ambas lenguas escritos por los mismos estudiantes francófonos, complementado con comprobaciones puntuales en un corpus de escritura académica en inglés como lengua materna elaborado por escritores noveles. Los resultados revelan una marcada similitud en las frecuencias de los marcadores de posicionamiento y en las preferencias léxicas y fraseológicas entre los textos en L1 y L2, lo que sugiere que la influencia interlingüística condiciona el uso del posicionamiento por parte de los escritores noveles. Asimismo, los hallazgos ponen de relieve la importancia de la variación individual en el uso del posicionamiento, en consonancia con estudios recientes que subrayan la necesidad de tener en cuenta dicha variación en el análisis de datos de L2.

Palabras clave – escritura en L2; escritura de escritores noveles; relación entre la L1 y la L2; influencia interlingüística; posicionamiento



1. INTRODUCTION

Stance, understood as the linguistic ways writers express their assessment and commitment, and attitudes and value judgments towards the information presented, has been a focal point of interest in academic writing research for a number of decades (e.g. Hyland 2005a; Biber 2006). Studies have shown that stance is deeply embedded in the linguistic and cultural norms of academic writing, shaping standardized ways of expressing meanings such as certainty (e.g. *of course*), doubt (e.g. *possible*), and agreement (e.g. *acknowledgment*). These conventions have become so entrenched in academic writing practices that the lexical choices available to express such meanings are no longer arbitrary. Instead, these lexical choices often appear in formulaic or idiomatic patterns, with writers frequently relying on (more or less) fixed expressions to convey stance effectively and appropriately. For example, the expression *be considered as* conveys epistemic stance by presenting the following information as tentative rather than factual, and *it is interesting to* signals attitudinal stance by highlighting the perceived interest or relevance of a point. Moreover, stance-related lexical choices are part of a larger set of patterns that have evolved to fulfill specific discursive functions. Expressions like *this suggests that* and *it is possible that*, for instance, have become standard tools to convey hedging.

Stance draws upon established patterns in the L1 of academic writers, whether at the lexical level (word selection), the phraseological level (lexico-grammatical expressions), or the discourse level (functions and discourse conventions). These are key areas where cross-linguistic influence may be observed, especially among novice writers, for whom L1 patterns may interact with those of the target language. Lexical transfer is the most extensively documented form, with research highlighting the prevalence of phenomena such as borrowings or false cognates in novice writers' L2 texts (e.g. Odlin 1989; Swan 1997; Ringbom 2007; Jarvis 2009; Williams 2015). Research has nonetheless shown that cross-linguistic lexical similarity alone is not always sufficient to trigger cross-linguistic influence in L2 production, as word knowledge extends beyond the lexical level (e.g. Ringbom 2007; Paquot 2010; Shatz *et al.* 2023). For example, L2 writers often tap into the phraseological resources of their L1 when producing multi-word units such as collocations or lexical bundles (e.g. Paquot 2008; Yamashita and Jiang 2010; Rica Peromingo 2012; Cao and Badger 2021), underscoring the significant role of phraseological transfer (Paquot 2010, 2013, 2017). Transfer at the discourse level,

although less systematically or frequently investigated, is also an important phenomenon (Odlin 1989; Ringbom 2007; Paquot 2010). As Ringbom (2007: 64) explains, “L1 patterns in [discourse] are deeply entrenched in the learner’s mind and thus are more resistant to modification and development than grammatical or, especially, lexical patterns”. For instance, studies in contrastive rhetorics have consistently demonstrated that L2 writers’ rhetorical strategies, organizational patterns, and writing processes often mirror those of their L1 writing (e.g. Uzawa 1996; Kubota 1998; Wang and Wen 2002; Hirose 2003; Uysal 2008; van Weijen *et al.* 2009; Huang *et al.* 2011). More than either phraseological or discursive anchor, the interaction of both types appears particularly susceptible to transfer. Paquot’s (2013, 2014, 2017) analyses of lexical bundles in EFL learner writing revealed that cross-linguistic influence has a stronger impact on phraseological expressions fulfilling particular discursive functions. In particular, discourse organizers, which structure and connect parts of the text (e.g. *to sum up*), and stance markers, which express the writers’ commitment and attitude (e.g. *very difficult to*), are more strongly influenced than referential lexical bundles, which primarily denote entities (e.g. *the nature of the*).

Against this backdrop, cross-linguistic influence is expected to play a significant role in shaping L2 academic writers’ expression of stance. Yet, empirical evidence remains limited. This study aims to fill this gap by investigating the relationship between L1 and L2 stance use in novice academic writing. The following research question is addressed:

RQ. To what extent does novice writers’ use of stance in L2 academic writing resemble their use of stance in L1 academic writing?

Here, ‘novice writing’ is defined as “unpublished pieces of writing that have been written in educational or training settings, (often) for purposes of assessment” (Scott and Tribble 2006: 133). The remainder of this paper is organized as follows. Section 2 reviews previous research on stance and cross-linguistic influence in academic writing. Section 3 returns to the research question and presents the current study and its hypotheses. Sections 4 and 5 describe the data and methodology, respectively. Section 6 reports the results, which are discussed in Section 7. Finally, Section 8 summarizes the main findings, outlines the study’s limitations, and suggests directions for future research.

2. LITERATURE REVIEW

While academic writing has traditionally been characterized by an emphasis on objectivity and detachment, it is now widely recognized as inherently argumentative and deeply rhetorical. Stance markers are prominent features of this register, with scholars systematically employing them to successfully communicate with readers and convince them of their claims (Hyland 2005a; Biber 2006). For example, expert writers often rely on epistemic stance, using boosters (e.g. *definitely*) and hedges (e.g. *maybe*) to signal the degree of commitment to the propositions they advance. Hedging is especially frequent and enables academic writers to present claims as cautious and open to negotiation. In addition to hedges and boosters, expert writers also employ attitudinal markers (e.g. *important*, *interestingly*), which play a key role in conveying writers' value judgments. Despite the importance of stance, academic writing remains shaped by an apparent paradox, requiring writers to balance between objectivity and subjectivity.

For L2 English novices, stance in academic writing presents a significant challenge. For example, they often struggle to strike the right tone, sometimes adopting a style that is overly assertive or personal (e.g. Paquot 2010; Larsson *et al.* 2022). Various explanations have been proposed for such tendencies, including cultural influences and a lack of register awareness (e.g. Hyland and Tse 2005; Wang and Jiang 2018). For example, Chinese EFL learners are often described as avoiding overt self-representation in their writing due to the collectivist nature of Chinese culture (e.g. Wang and Jiang 2018). Less frequently, cross-linguistic influence has been suggested as a factor, stating that novice writers' stance use in L2 English reflects patterns from their L1 (e.g. Chen 2010; Jiang 2015). However, empirical evidence supporting this hypothesis is sparse, and Jarvis' (2000: 246) criticism that "L1 influence has been treated largely as a you-know-it-when-you-see-it phenomenon" is still valid – such influence is often discussed only as a post-hoc interpretation in discussion and conclusion sections. For example, after conducting her analysis, Chen (2010) speculated that differences in epistemic modality and rhetorical strategies between L1 and L2 writers might stem from differences between Chinese and English.

Two exceptions to this general lack of empirical focus on cross-linguistic influence are Liu and Thompson (2009) and Chung *et al.* (2021), both of which applied the Appraisal model to L1 and L2 novice texts produced by the same writers. This model, developed within Systemic Functional Linguistics, provides a framework for analyzing

evaluative meanings through three interacting systems, namely Attitude (comprising Affect, Judgment, and Appreciation), Graduation (comprising Force and Focus), and Engagement (comprising monogloss and heterogloss) (Martin 2000; Martin and White 2005). Liu and Thompson (2009) found similar frequencies in Appreciation items (i.e. evaluative items such as *beautiful* and *striking*) in L1 Chinese and L2 English argumentative writing, but observed fewer instances of Affect (e.g. *fear*) and Judgment (e.g. *skillfully*) items in L1 Chinese argumentative writing. They attributed this weaker presence of Affect and Judgment in Chinese to the rhetorical aim of maintaining social harmony. Chung *et al.* (2021) reported mixed findings, identifying both similarities and differences in the use of Appraisal resources in L1 Vietnamese and L2 English argumentative writing. One similarity was that Attitude items (e.g. *special*) appeared more frequently than Engagement (e.g. *according to X*) and Graduation (e.g. *very*) items. One difference reported was the stronger presence of adjectives modifying nouns in L1 Vietnamese (e.g. *reluctant responsibilities*). While these two studies yield interesting insights into L1-L2 relationships, their findings must be interpreted with caution due to the limited datasets they analyze. Liu and Thompson (2009) examined the argumentative texts of just one student in L1 Chinese and L2 English (one text per language), while Chung *et al.* (2021) analyzed 20 L1 Vietnamese and 20 L2 English paragraphs from the same students. Moreover, Chung *et al.* (2021) explicitly describe their approach as corpus-informed rather than corpus-based, meaning that their corpus serves “simply as a bank of examples to illustrate a theory” (McEnery and Hardie 2012: 17, as cited in Chung *et al.* 2021: 2). The results presented in Liu and Thompson (2009) and Chung *et al.* (2021) are therefore not generalizable, and must be considered in light of the methodologies and data used. Furthermore, as both studies focus on argumentative texts, little is known about how cross-linguistic influence shapes stance use in other registers required in university settings, such as empirical research writing.

In sum, research that discusses cross-linguistic influence on the use of stance in L2 writing tends to suffer from one of two key limitations: (i) an ‘attribution problem’ (Singleton 1987), where claims rely on post-hoc interpretations without incorporating L1 data, or (ii) a reliance on qualitative analyses of limited datasets focused on argumentative writing. This is particularly surprising, considering that, since the early 2000s, researchers have repeatedly called for more empirical and large-scale analyses of cross-linguistic influence (e.g. Jarvis 2000; Gilquin 2008). To address these issues, Jadoulle (2023b)

supplemented a traditional comparison – between English expert writing, L1 English novice writing and French EFL learner writing – with a corpus of L1 French novice writing of approximately 370,000 words, specifically focusing on empirical research writing. Jadoulle (2023b) examined the use of hedges, revealing significant overlap in hedge use between L1 (French) novice writing and French EFL learner writing, thus providing empirical support for the hypothesis of cross-linguistic influence in stance use among novice writers. While this study represents a step forward, its methodology presents two key limitations. First, because the L1 French corpus consists of texts written by different individuals from those in the French EFL corpus, comparability issues arise. Many variables influence student writing, including age, gender, time spent abroad, other languages spoken, socio-economic background, reading/viewing habits, writing habits, personal beliefs, aptitude, and motivation (e.g. Paquot *et al.* 2023). Accounting for all these factors to construct perfectly matched corpora is virtually impossible. Differences in writer profiles between datasets could therefore impact findings on cross-linguistic influence. Second, studying L1 and L2 writing produced by different individuals does not account for individual variation. While comparing L2 writing and L1 writing from different writers offers insights into language transfer phenomena at a macro level, it does not capture how individual writers diverge in the extent to which they transfer stance from their L1 to their L2, or whether transfer takes on diverse forms across individuals, for example.

To address these methodological issues, the present study examines stance use among novice writers by using a corpus of L1 French and L2 English texts produced by the same individuals.

3. THE CURRENT STUDY

The study focuses on the relationship between L1 (French) and L2 (English) stance usage in academic writing by French-speaking novice writers. More specifically, as already indicated in Section 1, it addresses the following research question:

RQ. To what extent does novice writers' use of stance in L2 academic writing resemble their use of stance in L1 academic writing?

To this end, the study aims to (i) determine whether French-speaking students use stance markers with similar frequencies in their L1 (French) and L2 (English), and (ii) examine

whether they rely on similar lexical and phraseological preferences across the two languages.

As explained in Section 2, research has suggested a connection between L1 and L2 writing among students, particularly in terms of lexical, phraseological, or discursive patterns (e.g. Odlin 1989; Ringbom 2007; Jarvis 2009; Paquot 2013). Based on the state of the art, it can be hypothesized that novices are likely to replicate the stance mechanisms they use in their L1 when writing in their L2. The following hypotheses are therefore proposed:

H1: French-speaking novice writers show similar frequencies of stance markers in their L1 (French) and in their L2 (English) academic texts.

H2: French-speaking novice writers show similar lexical/phraseological preferences of stance markers in their L1 (French) and in their L2 (English) academic texts.

4. DATA

To answer the research question, the paper examines L1 and L2 data from the same writers – a methodology still rare in Learner Corpus Research due to the unavailability of suitable datasets. Section 4.1 describes the dataset and methodological decisions made to control for various learner-related variables that can affect cross-linguistic influence. To provide additional context for interpreting the findings on L1-L2 writing and to discuss cross-linguistic influence in relation to English academic writing conventions, the study also draws on a comparable corpus of L1 English writing (Section 4.2).

4.1. *The CARE*

The *Corpus of Academic French and English* (CARE) contains L1 French and L2 English academic writing (i.e. term papers and BA dissertations) produced by the same individuals accompanied with learner-related metadata. A variety of factors may interact with cross-linguistic influence, including language proficiency, prior writing experience and knowledge, and learners' communicative needs (Jarvis and Pavlenko 2008). To

account for these factors, methodological steps were taken to collect comparable data and relevant information.

Specifically, we collected textual data from UCLouvain (Belgium) French-speaking students who write in both (L1) French and (L2) English, mainly for linguistics and/or literature courses, at BA and/or MA level. Such L1-L2 data is scarce and challenging to collect due to several factors: (1) few students at UCLouvain write in both French and English, (2) those who do, do not typically produce the same types of texts in both languages (e.g. argumentative essays, term papers, cover letters; text lengths also vary widely, ranging from very short texts of 250 words to longer texts of up to 10,000 words), and/or (3) they often do not produce such texts at the same stages in their academic journey.

A strict selection of texts was necessary to ensure that the French and English texts were as comparable as possible in terms of text types. Specifically, two texts per student were selected – one in French and one in English. The aim was to match the texts as closely as possible in terms of level of study (e.g. third-year BA, MA), discipline (e.g. linguistics), text length, and purpose (report on empirical research). The final corpus contains data from 25 students. The English part of the corpus (CARE-EN) comprises 25 texts and 68,891 words, and the French part (CARE-FR) is made up of 25 texts and 80,830 words. An overview of the CARE corpus is provided in the supplementary materials¹ (Table S1).

Participating students completed a questionnaire providing information on their age, level of study, and reading and writing habits. The average age of CARE writers is 23.2 years old (SD = 2.9). On average, they have been studying at university for 4.3 years (SD = 1.7), and have studied English for 4.8 years (SD = 2.1) in high school and 3.5 years (SD = 1.2) at university. Almost all writers regularly engage with academic texts in English, with reported frequencies ranging from ‘often’ to ‘every day’ (for both reading and writing). Only a small minority (1 and 2 students) report reading and writing such texts ‘rarely’ or ‘never’. A similar pattern is observed in their experience with reading and writing academic texts in French, indicating that CARE writers are a relatively homogeneous group in terms of their academic literacy in both languages.

¹ The supplementary materials can be found here: <https://osf.io/zubfd/>

The proficiency level of CARE writers was determined based on the vocabulary section of the DIALANG test² and their responses to a self-assessment question (see supplementary materials, Section B). While self-assessment measures are often criticized for lacking precision in estimating L2 proficiency, research shows that they generally correlate with proficiency levels determined by independent tests (e.g. Flege *et al.* 2002; Tremblay 2011), and may actually have more validity than commonly assumed (e.g. Bowden 2016). In this study, no significant disparities were found between the two measures (an overview of the proficiency levels based on both tests is provided in the supplementary materials, Table S2). Overall, the majority of CARE writers demonstrate a C1 English proficiency level.

4.2. The BAWE & MICUSP

We also consulted a comparable corpus of L1 academic novice writing to observe whether patterns of stance use in L2 novice academic writing resemble those found in the L1 of the students, or whether they are already approximating what is produced by comparable L1 English novice writers. This approach is useful, though not without complexity, as it makes it possible to examine the interplay between cross-linguistic influence from the broader development of L2 proficiency and academic literacy. The L1 English corpus used in the study (namely the BAWE&MICUSP) comprises linguistics papers from the *British Academic Written English*³ (Heuboeck *et al.* 2008, BAWE) and from the *Michigan Corpus of Upper-level Student Papers*⁴ (Ädel and Römer 2012, MICUSP). Following Larsson (2016), the corpus includes texts classified as ‘research reports’ and ‘essays’ (BAWE), and ‘research papers’, ‘argumentative essays’,⁵ and ‘proposals’ (MICUSP). The resulting corpus consists of 72 academic papers written by native speakers of English, totaling 174,345 words.

² <https://wp.lancs.ac.uk/ltrg/projects/dialang-2-0/> (25 June, 2025). Note that the DIALANG is no longer available; the data collection took place between 2018 and 2022, which is when students took the test.

³ [BAWE \(British Academic Written English\) and BAWE Plus Collections](#) (6 January, 2026).

⁴ <https://elicorpora.info/main> (6 January, 2026).

⁵ MICUSP argumentative essays are pieces of research-based writing (examples of such essays are “National Identity and Language Education Policy” or “A Discussion of the VP-External and VP-Internal Subjects Hypotheses”).

5. METHOD

5.1. Operationalization of stance

This study is grounded in the stance frameworks proposed by Biber *et al.* (1999) and Hyland (2000, 2005a, 2005b), which conceptualize stance both semantically and syntactically.

Semantically speaking, stance includes two main categories, namely epistemic stance and attitudinal stance.⁶ Epistemic stance refers to writers' comments on the "status of information in a proposition" (Biber *et al.* 1999: 972). Particularly, it reflects writers' assessment and commitment regarding the certainty and reliability of that information (Biber *et al.* 1999; Hyland 2005a). Following Hyland (2000, 2005a, 2005b), epistemic stance is divided into boosters and hedges. Boosters are used to express the writers' level of certainty, involvement, and confidence towards the information presented. By contrast, hedges are used to express the writers' caution, uncertainty, and distance towards the information they put forward. Attitudinal stance, on the other hand, refers to writers' personal attitudes and value judgments. Some of the most prominent attitudinal meanings are importance (e.g. *crucially*, *important* (fr.)), relevance (e.g. *pertinent* (fr.), *appropriate*), necessity (e.g. *nécessaire* (fr.), *require*), and agreement/acknowledgment (e.g. *agree*, *admettre* (fr.)).

Syntactically speaking, this study focuses on metadiscursive stance markers, specifically those instances where a stance marker explicitly qualifies a proposition within the same sentence (Gray and Biber 2014).⁷ Three key grammatical categories are considered to express stance metadiscursively. It is the case of adverbials (e.g. *maybe*), complement clause constructions controlled by adjectives, nouns, and verbs (e.g. *likely/avantage* (fr.)/*assert* + complement clause), and modal verbs (e.g. *might*) (Biber *et al.* 2018). In this study, the focus is on adverbials and complement clauses controlled by adjectives, nouns, and verbs.

⁶ For time and space reasons, we excluded the category of 'presence' as described in Hyland (2005b), and focused on epistemic and attitudinal stance, as they more directly relate to commitment and attitude.

⁷ This is to be contrasted from studies adopting the Appraisal framework (e.g. Liu and Thompson 2009; Chung *et al.* 2021), which also consider non-metadiscursive devices in their analyses. For a discussion on this distinction between metadiscursive and non-metadiscursive stance, please refer to Jadoulle (2024) and Jadoulle (2023a).

5.2. Identification and disambiguation of stance markers

The identification of stance markers is based on Jadoulle (2023b), who developed a methodology to select stance markers based on five corpora of English and French academic writing (expert and novice). While Jadoulle (2023b) focuses on hedging, the same method was applied to all three semantic categories (namely boosters, hedges, and attitudinal markers), using the same corpora.⁸ Adverbials were identified by extracting adverbs (tagged as [tag='RB.?'] for English and [tag='*.R'] for French) and 2- and 3-grams (minimum frequency = 3) using Sketch Engine (Kilgarriff *et al.* 2014), and then manually selecting those items that met the semantic and syntactic definition of stance outlined in Section 5.1. Complement clause constructions were identified by extracting common adjectival, nominal, and verbal stance patterns (e.g. noun + *that*, verb + *que* (fr.)) using Sketch Engine (e.g. [tag='N.*'] [word='that|to'], [tag='V.*'] [word='que|qu|de|d']), and retaining those adjectives, nouns, and verbs that met the same semantic and syntactic criteria (e.g. *hypothesis*, *penser* (fr.)). The final selection of stance markers is presented in the Appendix. It contains 79 English adverbials, 58 French adverbials, 144 English adjectives, 88 French adjectives, 83 English nouns, 60 French nouns, 86 English verbs, and 76 French verbs.

These items were extracted from the CARE and the BAWE&MICUSP via Sketch Engine, then disambiguated in context following the disambiguation guidelines presented in Jadoulle (2023b), whereby adverbials are included when they modify a clause or sentence and are not part of verbal dependency, while complement clauses are included if they function metadiscursively to qualify propositional content. Note that some markers may be classified into various semantic categories depending on the context in which they are found. For example, the adjective *évident* (fr.) typically serves as a booster (see (1)), but when preceded by negation, it shifts to convey a meaning similar to *easy* and thus functions as an attitudinal marker (see (2)). Consequently, the tables presented in the Appendix do not present markers categorized as boosters, hedges, and attitudinal markers; such categorization is offered in the Results section (Section 6) and in the supplementary materials (Tables S3-S5).

⁸ Even though the selection process was not based on the CARE corpus, the chances of having missed important items are very slim, as the CARE-FR and the CARE-EN are much smaller than the corpora used, and contain very similar data to the corpora of L1 French and L2 English novice writing respectively (i.e. academic texts written by French-speaking students enrolled in a BA or MA program in languages and literatures at UCLouvain).

- (1) *Il est évident que les influences anglaises ne sont pas ressenties de la même manière dans toutes les régions du monde.* (CARE0006-FR-04)

It is **obvious** that English influences are not felt in the same way in all parts of the world.

- (2) *L'acte de transposer l'intrigue d'Hugo à une époque contemporaine de 2014 permet à Tommy de mettre en lumière l'intemporalité du propos de l'œuvre adaptée, tout en éclaircissant les subtilités du récit que le temps et les changements de mœurs ont rendu peu **évident** à saisir pour un public non averti.* (CARE0030-FR-01)

The act of transposing Hugo's plot to a time contemporary with 2014 allows Tommy to highlight the timelessness of the adapted work's subject matter, while at the same time clarifying the subtleties of the story that time and changing mores have made **difficult** for an uninformed audience to grasp.

Table 1 provides an overview of the absolute frequencies of adverbials and controlling adjectives, nouns, and verbs before and after disambiguation in the CARE-FR, CARE-EN, and BAWE&MICUSP. A total of 12,766 occurrences were disambiguated in context. Among the different grammatical categories, adjectives, and nouns include the highest number of excluded occurrences; only about 16 per cent and 10 per cent of the occurrences of adjectives and nouns respectively were kept after the disambiguation process. Most excluded occurrences of adjectives are instances in which the items are used predicatively or attributively (e.g. *appropriate translation*). Nouns are rarely complemented by a clause, which explains the small percentage of occurrences kept after disambiguation (e.g. excluded occurrence of *difficulty*: (...) *annotating learner language presents a few difficulties*).

CATEGORY	CARE-FR	CARE-EN	BAWE & MICUSP
Adverbials	Before: 133 After: 102	Before: 165 After: 92	Before: 503 After: 319
Adjectives	Before: 528 After: 99	Before: 593 After: 118	Before: 2,153 After: 422
Nouns	Before: 466 After: 66	Before: 603 After: 66	Before: 2,228 After: 234
Verbs	Before: 994 After: 425	Before: 1,075 After: 448	Before: 3,325 After: 1,910
Total	Before: 2,121 After: 692	Before: 2,436 After: 724	Before: 8,209 After: 2,885

Table 1: Absolute frequencies of stance markers before and after disambiguation in the CARE-FR, CARE-EN, and BAWE&MICUSP

5.3. Statistical methods

The CARE was mainly analyzed by means of the paired-samples t-test. One of the assumptions of the test is that the differences of the paired values are normally distributed (Gries 2013). This was checked by means of the Shapiro-Wilk test. The function ‘t.test()’ was used to compute the test in *R* (R Core Team 2023). The significance threshold was set at 0.05, and the effect size reported is Cohen’s *d*. Effect sizes of 0.2, 0.5, and 0.8 correspond to small, moderate, and large effects respectively (Cohen 1988).

Comparisons of specific marker occurrences (e.g. *interesting*) were conducted using the chi-square test of independence, as the frequency of these markers per text was too limited for the paired-samples t-test. We ran the test using the ‘chisq.test()’ function in *R* and, following Gries (2013), we used ϕ as the effect size. The outputs from the chi-square tests are provided in the supplementary materials (Tables S6 to S22).

Comparisons between the CARE-FR, CARE-EN, and BAWE&MICUSP were conducted by means of the Kruskal-Wallis H test (function: ‘kruskal.test()’ in *R*) (threshold of 0.05). This non-parametric test was chosen because the Shapiro-Wilk test indicated that the frequency distributions were not normally distributed. A post hoc Dunn’s test with Bonferroni-corrected *p*-values (‘dunn_test()’ in *R*) was then applied to identify which differences were significant.

6. RESULTS

This section presents the results of the analyses, focusing on stance marker frequencies (Section 6.1) and lexical/phraseological preferences (Section 6.2).

6.1. Frequencies of stance markers

Figures 1-6 present the results for the three semantic categories, with the French and English texts written by the same student considered as paired data points. The results show that CARE writers do not systematically use fewer or more epistemic markers in French vs. in English ($p = 0.992$ for boosters, $p = 0.1611$ for hedges). The boxplots presented in Figures 1 and 2 reveal that, while some writers use more boosters in English

than in French for example, others do the opposite. This variation supports the conclusion that no clear trend emerges at the group level.

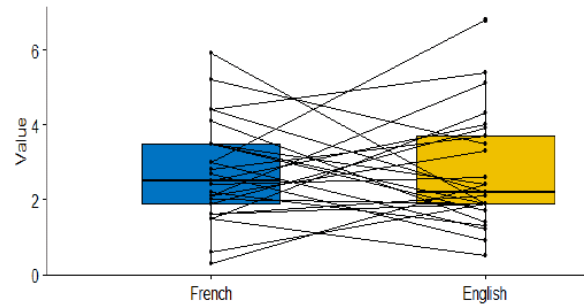


Figure 1: Booster frequencies in the CARE-FR and CARE-EN (occurrences per 1,000 words) ($t = 0.010128$, $df = 24$, $p\text{-value} = 0.992$; mean difference: 0.004; effect size: -0.01 (small))

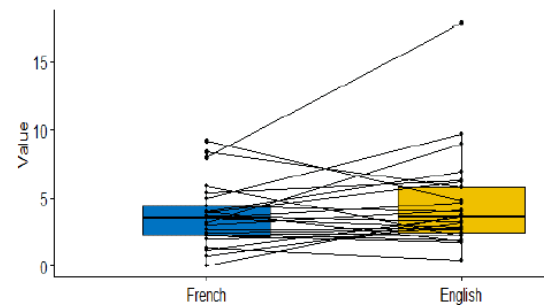


Figure 2: Hedge frequencies in the CARE-FR and CARE-EN (occurrences per 1,000 words) ($t = -1.4459$, $df = 24$, $p\text{-value} = 0.1611$; mean difference: -0.880; effect size: 0.29 (small))

Comparisons of booster and hedge frequencies across BAWE&MICUSP, CARE-EN, and CARE-FR reveal significant differences between the L1 English corpus and both subsets of the CARE (see Figures 3-4). For example, the median booster frequency in the BAWE&MICUSP is 7.3 boosters per 1,000 words, compared to 2.5 and 2.2 in the CARE-FR and CARE-EN respectively.

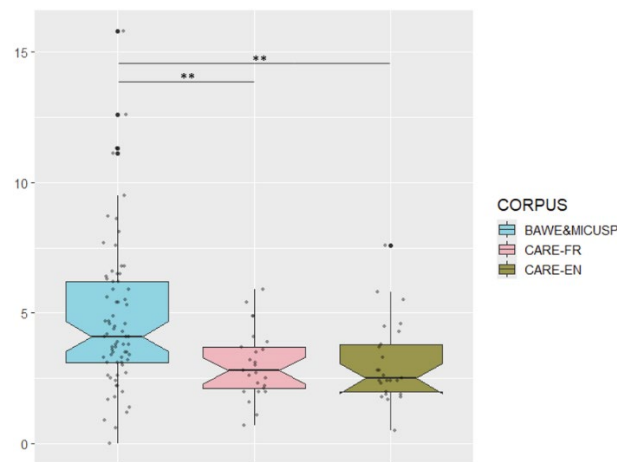


Figure 3: Booster frequencies in the BAWE&MICUSP, CARE-EN, and CARE-FR (occurrences per 1,000 words) (Kruskal-Wallis chi-squared = 16.462, $df = 2$, $p\text{-value} = 0.0002662$; Dunn post hoc test: $p\text{-adj} =$

0.00319 [BAWE&MICUSP vs. CARE-EN], $p_{\text{adj}} = 0.00472$ [BAWE&MICUSP vs. CARE-FR], $p_{\text{adj}} = 1$ [CARE-EN vs. CARE-FR])

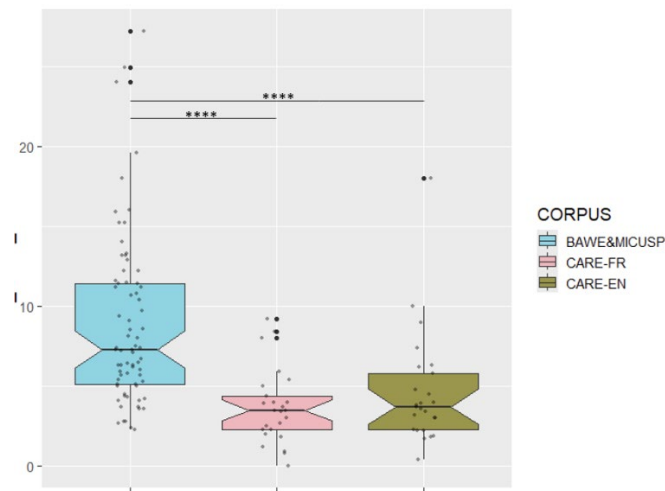


Figure 4: Hedge frequencies in the BAWE&MICUSP, CARE-EN, and CARE-FR (occurrences per 1,000 words) (Kruskal-Wallis chi-squared = 37.735, $df = 2$, $p\text{-value} = 6.398\text{e-}09$; Dunn post hoc test: $p_{\text{adj}} = 0.0000337$ [BAWE&MICUSP vs. CARE-EN], $p_{\text{adj}} = 0.000000384$ [BAWE&MICUSP vs. CARE-FR], $p_{\text{adj}} = 1$ [CARE-EN vs. CARE-FR])

The situation is slightly different for attitudinal markers, as it appears that French-speaking students use more attitudinal markers in their English texts compared to their French texts (see Figure 5; $p = 0.0342$). The difference is not marked however: the effect size is relatively small (0.4), and the mean difference is only -0.952, indicating that, on average, the English texts feature approximately 1 occurrence per 1,000 words more attitudinal markers than the French texts.

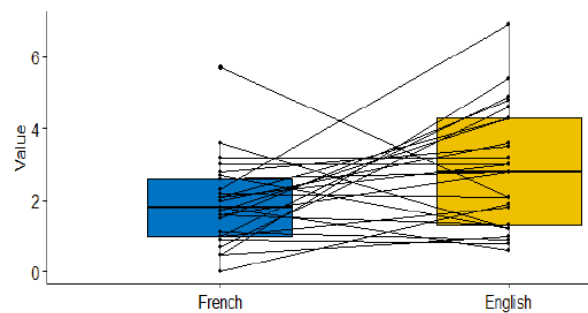


Figure 5: Attitudinal marker frequencies in the CARE-FR and CARE-EN (occurrences per 1,000 words) ($t = -2.2457$, $df = 24$, $p\text{-value} = 0.0342$; mean difference: -0.952; effect size: 0.4 (small-moderate))

The frequencies of attitudinal markers appear similar between the CARE-EN and the BAWE&MICUSP (see Figure 6). The only significant difference identified by the Kruskal-Wallis test is observed between the BAWE&MICUSP and the CARE-FR, with medians of 2.5 vs. 1.8 occurrences per 1,000 words, respectively.

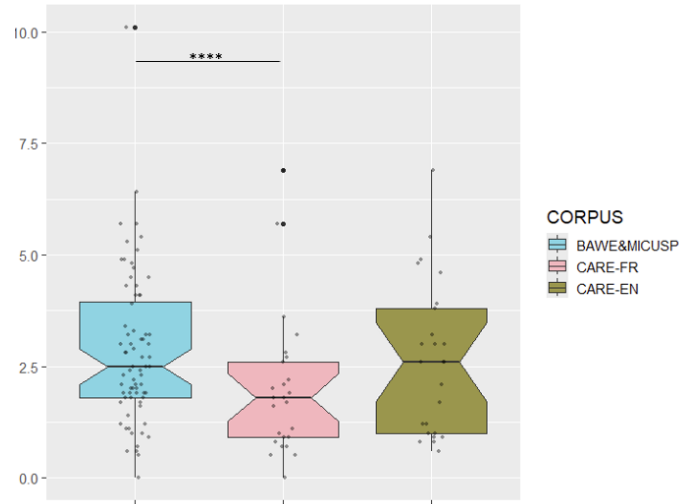


Figure 6: Attitudinal marker frequencies in the BAWE&MICUSP, CARE-EN, and CARE-FR (occurrences per 1,000 words) (Kruskal-Wallis chi-squared = 7.0632, $df = 2$, p -value = 0.02926 ; Dunn post hoc test: $p_{adj} = 1$ [BAWE&MICUSP vs. CARE-EN], $p_{adj} = 0.0237$ [BAWE&MICUSP vs. CARE-FR], $p_{adj} = 0.268$ [CARE-EN vs. CARE-FR])

6.2. Lexical and phraseological preferences

This section presents the results regarding the lexical and phraseological preferences for boosters, hedges, and attitudinal markers. Due to the low frequency of markers per text, detailed pairwise quantitative analyses of lexical preferences were not feasible. Instead, group-level comparisons focus on the top 10 boosters, hedges, and attitudinal markers. Spot checks of overall trends in the BAWE&MICUSP are also conducted to compare the findings on L1-L2 writing with those from L1 English student writing.

6.2.1. Boosters

The top 10 boosters in the CARE-FR vs. the CARE-EN are presented in Table 2, together with their absolute and relative frequencies (see supplementary materials, Table S3 for the whole table). It appears that the two corpora have a shared set of boosters. Out of the top 10 markers, we note the seven pairs of equivalents *voir*|*see*, *montrer*|*show*, *savoir*|*know*, *démontrer*|*demonstrate*, *bien sûr*|*of course*, *conclure*|*conclude*, and *prouver*|*prove* (the cross-linguistic equivalents were operationalized by considering those items that were offered as the first translation of each other in Linguee (see Jadoulle 2024 for more information)). Out of these pairs, and as revealed by the chi-square tests (see supplementary materials, Tables S6-S12), four have similar frequencies in the CARE-FR and in the CARE-EN (namely *savoir*|*know*, *démontrer*|*demonstrate*, *bien*

sûr|of course, and *conclure*|conclude). The favored semantic area in both corpora is ‘demonstration’, with the French markers *voir*, *montrer*, *démontrer*, *constater*, and *prouver* (see (3)), and the English markers *show*, *demonstrate*, *prove*, and *see* (see (4)).

- (3) *Comme leur absence des dictionnaires, cela **prouve** que ces huit termes ne sont pas utilisés par les Français.* (CARE0007-FR-04)

Like their absence from dictionaries, this **proves** that these eight terms are not used by the French.

- (4) ... Ahmad persist in thinking that the society never truly accepts their differences, as **proven** by the difficulty for Jack to answer when Ahmad asks him to name one Arab-American. (CARE0010-EN-02)

	L1 FRENCH WRITING (CARE-FR)			L2 ENGLISH WRITING (CARE-EN)		
	Booster	Abs. freq.	Rel. freq.	Booster	Abs. freq.	Rel. freq.
1	<i>voir</i>	43	5.3	<i>show</i>	37	5.4
2	<i>montrer</i>	25	3.1	<i>demonstrate</i>	22	3.2
3	<i>savoir</i>	21	2.6	<i>prove</i>	20	2.9
4	<i>démontrer</i>	19	2.4	<i>see</i>	18	2.6
5	<i>constater</i>	14	1.7	<i>know</i>	12	1.7
6	<i>bien sûr</i>	12	1.5	<i>of course</i>	8	1.2
7	<i>affirmer</i>	10	1.2	<i>clearly</i>	6	0.9
8	<i>établir</i>	7	0.9	<i>assert</i>	4	0.6
9	<i>conclure</i>	6	0.7	<i>clear</i>	4	0.6
10	<i>prouver</i>	6	0.7	<i>conclude</i>	4	0.6
Total		163	20.2		135	19.6

Table 2: Top 10 boosters and their absolute and relative frequencies (per 10,000 words) in the CARE-FR and CARE-EN

Nine of the top 10 boosters are shared between the CARE-EN and the BAWE&MICUSP (see Table S3, supplementary materials). These items also closely resemble the boosters (in the form of cross-linguistic equivalents) found in both the CARE-EN and the CARE-FR. In fact, seven of these items are identical: *show*, *demonstrate*, *prove*, *see*, *know*, *of course*, and *conclude*.

The CARE-FR and CARE-EN are small in size and therefore present a relatively low number of occurrences of specific boosters, which makes it difficult to draw solid conclusions on the phraseology of those markers (e.g. *conclure* (fr.) occurs with an absolute frequency of 6 instances in the CARE-FR). The analyses conducted nonetheless suggest that, in addition to sharing a similar set of boosters, CARE writers also tend to use those markers as part of the same constructions cross-linguistically. For example, *montrer* (fr.) and *show* are mainly followed by a *que*- or *that*-clause (e.g. *The study shows that [...]*) rather than included in a *comme*- or *as*-clause (e.g. *As shown in [...]*). As

illustrated in Table 3, in addition to sharing similar frequencies in the CARE-FR and in the CARE-EN, *démontrer* (fr.) and *demonstrate* are used as part of the same grammatical structures; the most frequent category being ‘*démontrer|demonstrate* introducing a *que*- or *that*-clause’ for instance (see (5) and (6)). Investigating the agents associated with *démontrer|demonstrate* also revealed similar tendencies in both corpora, with for example the combination of ‘inanimate entity + *démontrer|demonstrate*’ (see (5)) occurring with a frequency of 0.7 and 0.5 occurrences per 10,000 words in French and English respectively (see Table 4).

(5) Table 1 and Table 2 **demonstrate** that both markers are more frequent in French ... (CARE0026-EN-01)

(6) *Nos observations corroborent la théorie établie par Ariel en 1990 qui a **démontré** que la présence de pronoms clitiques ... coïncide avec des référents très accessibles ...* (CARE0019-FR-06)

Our observations corroborate the theory established by Ariel in 1990 who **demonstrated** that the presence of clitic pronouns ... coincides with highly accessible referents ...

STRUCTURE	CARE-FR	CARE-EN
<i>démontrer demonstrate</i> included in a <i>comme as</i> -clause	0.6	0.9
<i>démontrer demonstrate</i> introducing a <i>que that</i> -clause	1.1	1.2
<i>démontrer demonstrate</i> introducing a subordinate interrogative clause	0.4	0.9
other	0.3	0.2

Table 3: Relative frequencies of structures associated with *démontrer|demonstrate* in the CARE-FR and CARE-EN (per 1,000 words)

AGENT	CARE-FR	CARE-EN
author(s)	0.6	1.5
inanimate entity (e.g. ‘travail’ (fr.), ‘results’)	0.7	0.5
first-person pronoun	0.2	0.1
second-person pronoun	0.5	0.1
no agent (passive voice)	0.1	0.1

Table 4: Relative frequencies of agents associated with *démontrer|demonstrate* in the CARE-FR and CARE-EN (per 1,000 words)

6.2.2. Hedges

Table 5 gives an overview of the top 10 hedges in the CARE-FR and the CARE-EN (see supplementary materials, Table S4 for the whole table). *Sembler|seem* and *considérer|consider* hold a central place as hedging devices in the CARE-FR and CARE-

EN. They also exhibit similar frequencies in French and in English (see supplementary materials, Tables S13 and S14). We also note the presence of *penser*|*think* and *probablement*|*probably* in the two top 10 hedges (and the two pairs present similar frequencies across the CARE-FR and the CARE-EN; see supplementary materials, Tables S15 and S16). A key difference between the two corpora nonetheless appears to be the prevalence of verbs in the top 10 markers of the CARE-EN (eight verbs, namely *seem*, *argue*, *consider*, *claim* (v.), *appear*, *believe*, *think*, *see*), in comparison with a more varied set of hedges in the CARE-FR (five verbs, namely *sembler*, *considérer*, *penser*, *estimer*, and *proposer*; three adverbials, namely *peut-être*, *sans doute*, and *probablement*; and two nouns, namely *hypothèse* and *impression*).

	L1 FRENCH WRITING (CARE-FR)			L2 ENGLISH WRITING (CARE-EN)		
	Hedge	Abs. freq.	Rel. freq.	Hedge	Abs. freq.	Rel. freq.
1	<i>sembler</i>	70	8.7	<i>seem</i>	71	10.3
2	<i>considérer</i>	62	7.7	<i>argue</i>	40	5.8
3	<i>peut-être</i>	20	2.5	<i>consider</i>	35	5.1
4	<i>penser</i>	16	2.0	<i>claim</i> (v.)	18	2.6
5	<i>estimer</i>	13	1.6	<i>appear</i>	15	2.2
6	<i>proposer</i>	13	1.6	<i>believe</i>	15	2.2
7	<i>sans doute</i>	13	1.6	<i>probably</i>	15	2.2
8	<i>hypothèse</i>	12	1.5	<i>think</i>	14	2.0
9	<i>probablement</i>	11	1.4	<i>see</i>	12	1.7
10	<i>impression</i>	8	1.0	<i>likely</i> (adj.)	9	1.3
Tota		238	29.4		244	35.4
l						

Table 5: Top 10 hedges and their absolute and relative frequencies (per 10,000 words) in the CARE-FR and CARE-EN

The commonalities in the use of the top 10 hedges between the CARE-FR and the CARE-EN are therefore less pronounced than those observed for boosters. By contrast, the most frequent hedges are notably similar to those found in L1 English novice writing (see Table S4, supplementary materials). Specifically, the CARE-EN and BAWE&MICUSP share eight out of the top 10 hedges (i.e. *seem*, *consider*, *claim* (v.), *appear*, *believe*, *think*, *see*, *likely* (adj.)).

6.2.3. Attitudinal markers

As Table 6 shows, in a similar way as for boosters (and to some extent hedges), the two corpora (i.e. CARE-FR and CARE-EN) present a shared set of attitudinal markers, with the following groups of items: *intéressant*|*interesting*(ly), *difficile*|*difficult*|*hard*, *nécessaire*|*necessary*, *difficulté*|*difficulty*, *important*|*important*, and *réaliser*|*realize*. In

addition, these pairs are usually used with similar frequencies (it is the case of *intéressant*|*interesting*, *difficile*|*difficult*, *nécessaire*|*necessary*, *difficulté*|*difficulty*, and *réaliser*|*realize*; see supplementary materials, Tables S17-S22). The shared attitudinal adjectives (i.e. *intéressant* (fr.), *interesting*, *difficile* (fr.), *difficult*, *nécessaire* (fr.), *necessary*, *important* (fr.), and *important* (en.)) are mainly used as part of an anticipatory-*it* structure (see (7)). *Intéressant* (fr.) and *interesting* (see (8)) both occupy a central place as attitudinal markers in the CARE-FR and in the CARE-EN. In addition, *difficile*|*difficult* and *difficulté*|*difficulty* rank high in the top 10 attitudinal markers. One main difference between the CARE-FR and the CARE-EN concerns the frequency of *important*|*important* in L1 French and in L2 English by novice writers; the marker is far more common in the CARE-EN than in the CARE-FR.

- (7) It is **necessary** to start with the beginning and examine how the main themes are introduced ... (CARE0012-EN-04)

- (8) *Il serait donc **intéressant** de poursuivre la question dans une recherche future.* (CARE0006-FR-04)

It would therefore be **interesting** to pursue this question in future research.

L1 FRENCH WRITING (CARE-FR)				L2 ENGLISH WRITING (CARE-EN)		
	Attitudinal marker	Abs. freq.	Rel. freq.		Abs. freq.	Rel. freq.
1	<i>intéressant</i>	21	2.6	<i>important</i>	15	2.2
2	<i>difficile</i>	15	1.9	<i>interesting</i>	13	1.9
3	<i>juger</i>	12	1.5	<i>difficulty</i>	10	1.5
4	<i>nécessaire</i>	8	1.0	<i>realize</i>	10	1.5
5	<i>difficulté</i>	7	0.9	<i>difficult</i>	8	1.2
6	<i>important</i>	5	0.6	<i>interestingly</i>	8	1.2
7	<i>réaliser</i>	5	0.6	<i>hard</i>	7	1.0
8	<i>avantage</i>	4	0.5	<i>necessary</i>	7	1.0
9	<i>se contenter</i>	4	0.5	<i>worth</i>	7	1.0
10	<i>convenir</i>	4	0.5	<i>surprising</i>	6	0.9
Total		85	10.5		91	13.2

Table 6: Top 10 attitudinal markers and their absolute and relative frequencies (per 10,000 words) in the CARE-FR and CARE-EN

The top 10 attitudinal markers of the CARE-EN share five items with the top 10 attitudinal markers of the BAWE&MICUSP (i.e. *important*, *interesting*, *difficulty*, *difficult*, and *necessary*) (see Table S5, supplementary materials). The attitudinal lexical preferences of the CARE-EN are arguably more similar to those of the CARE-FR (see above).

7. DISCUSSION

The analyses presented in Section 6 revealed interesting findings with respect to the relationship between L1 and L2 stance use in novice academic writing. Section 7.1 addresses the extent to which French-speaking students present similar stance use in their L1 and L2 texts, and Section 7.2 expands on the concept of individual variation, which also emerged as a key factor.

7.1. L1-L2 similarities

The comparisons between the CARE-FR and the CARE-EN, coupled with spot checks in the BAWE&MICUSP, suggest that there is a link between L1 and L2 stance use by the same individuals.

First, the CARE-FR and the CARE-EN show similar frequencies of epistemic markers, partly supporting Hypothesis 1 and suggesting that French-speaking students tend to use stance markers with similar frequencies across languages. These results point to the presence of transfer of function in L2 English novices' stance use (Paquot 2010). For instance, they indicate that the amount of hedging CARE writers resort to in English actually mirrors the amount of hedging they use in French. By contrast, the L1-L2 similarities in epistemic marker frequencies were not observed for attitudinal markers, as the English texts from the CARE contain more occurrences of attitudinal markers than the French texts. This, however, might be explained by the CARE-EN having a higher proportion of literature texts compared to the CARE-FR despite our efforts to prepare comparable datasets (see Section 4). Stance usage displays cross-disciplinary differences in academic writing, even across so-called 'soft' knowledge disciplines (e.g. Gray 2015; Yoon and Römer 2020). In comparison with linguistics texts, literature texts primarily consist of interpretative qualitative analyses with "less control of variables" (Hyland 2005b: 187) and a plurality of possible interpretations, possibly leading writers to take a more personal stance (see (9)). Furthermore, the themes addressed in literature texts often carry a more pronounced attitudinal dimension compared to those in linguistics texts, delving into sometimes societal and historical issues like racial discrimination (CARE0006-EN-02) or World War I (CARE0012-EN-04).

- (9) David is an exceptional "mecha" who is able to feel emotions and more **importantly** to feel love, just as a human child would do. (CARE0024-EN-01)

The comparisons with the BAWE&MICUSP indicate higher frequencies of epistemic markers in the latter compared to the two subsets of the CARE, especially in the case of hedges. By contrast, attitudinal marker frequencies are similar between the CARE-EN and the BAWE&MICUSP. These findings suggest that, while French-speaking students' use of epistemic markers in English does not align with the patterns typically observed in English writing, CARE writers might show some level of adaptation to English academic conventions in their use of attitudinal stance.

Section 6 also highlighted shared sets of boosters and attitudinal markers (and to some extent hedges) across the CARE-FR and the CARE-EN. This supports Hypothesis2 and indicates that French-speaking students employ cross-linguistic correspondents in their French and English texts. Examples include pairs of markers like *prouver*|*prove*, *bien sûr*|*of course*, *considérer*|*consider*, *penser*|*think*, *intéressant*|*interesting*, and *difficulté*|*difficulty* (see (10) and (11)). Among these markers, many appear with similar frequencies in the two subsets, such as *démontrer*|*demonstrate*, *bien sûr*|*of course*, *penser*|*think*, and *difficulté*|*difficulty*. This suggests that the frequency of an L1 marker is transferred to its L2 equivalent, providing evidence of transfer of frequency (Paquot 2010). Moreover, the writers present similar phraseological uses of markers like *démontrer* (fr.) and *demonstrate* (which are used in combination with similar agents for instance) and share a favored semantic area for boosters, namely 'demonstration'. In addition to transfer of frequency, traces of transfer of phraseological patterns are also visible (Paquot 2010). In other words, L2 English novices appear to transfer the patterns they commonly use in their L1 to their L2. The relatively small size of the CARE, however, limits in-depth analyses of transfer of lexical preferences at the individual level. A larger corpus would allow for a deeper understanding of whether different individuals present similar forms of transfer (e.g. transfer of frequency, phraseological patterns), for example.

- (10)a. *Comme leur absence des dictionnaires, cela **prouve** que ces huit termes ne sont pas utilisés par les Français.* (CARE0007-FR-04)

Like their absence from dictionaries, this **proves** that these eight terms are not used by the French.

- b. Jack Levy might argue that people's origins do not really matter, for they can all be considered as Americans: "Irish-Americans, African-Americans, Jewish-Americans; there are even Arab Americans" (301), but Ahmad persist in thinking that the society never truly accepts their differences, as **proven** by

the difficulty for Jack to answer when Ahmad asks him to name one Arab-American (302). (CARE0010-EN-02)

- (11)a. Une **difficulté** importante dans le cadre de la linguistique, note Do Amaral, est que la recherche dans le domaine n'a pas encore été en mesure de fournir une théorie qui puisse rendre compte de tous les aspects du langage humain ... (CARE0011-FR-03)

A major **difficulty** in linguistics, notes Do Amaral, is that research in the field has not yet been able to provide a theory that can account for all aspects of human language ...

- b. During the annotation process, we encountered the following **difficulties**: ... (CARE0011-EN-02)

Beyond reflecting their own L1 stance lexical/phraseological preferences, L2 writers also show tendencies that align with L1 English novice writers. For instance, boosting lexical preferences overlap not only between the CARE-EN and the CARE-FR, but also between the CARE-EN and the BAWE&MICUSP corpora. This suggests that the students' L1 may have a facilitating effect on their use of English stance lexical/phraseological preferences.

These results align with previous research highlighting the importance of cross-linguistic transfer among novice writers (e.g. Hirose 2003; Cao and Badger 2021). For example, as previously mentioned, extensive research on rhetorical and organizational strategies and writing processes has revealed numerous similarities across L1 and L2 texts by the same individuals (e.g. Hirose 2003; van Weijen *et al.* 2009). The observed L1-L2 relationship in terms of both stance use and rhetorical patterns/processes supports the claim that the writing features of novice writers in an L1 and in an L2 are interconnected on multiple levels, ranging from textual organization to more stylistic and pragmatic aspects. It should nonetheless be noted that the analysis only looked at L1 French writers. The conclusions made on cross-linguistic influence could be reinforced and/or nuanced by adding comparisons between French-speaking students and other novice populations across various L1 backgrounds (Jarvis 2000).

7.2. Individual variation

Results also highlight individual variation, emphasizing the unique value of the CARE corpus for examining L1-L2 writing relationships at the individual level. For instance,

while CARE writers generally exhibit similar frequencies of boosters in L1 French and L2 English, individual patterns vary: some writers use boosters more frequently in English than in French, while others do the opposite.⁹ Existing research on individual variation reports similar findings, highlighting the fact that despite observable overall tendencies in L2 data, variation exists, with certain individuals diverging from group tendencies (e.g. Gries and Wulff 2021; Jadoulle 2024). Jadoulle (2024) combined a corpus-based study of the CARE corpus and case studies of three individuals' stance use in L1 French and L2 English, and found that while writers tend to use a similar number of hedges in the two languages for example, one individual whose productions were closely examined actually adopts a more hedge-heavy tone in English than in French, showing "honesty, politeness, caution, and deference to the opinions of others" (Hinkel 2003: 327) (see (12)). Jadoulle (2024) concluded that this writer exhibits signs of language-specific register awareness, aligning with the tendency for English academic writing to display more hedges than French academic writing (Vold 2008).

- (12) Aslam thus **suggests** that the harm done to Afghanistan has been imported into it by foreign nations and for futile and selfish reasons, which **seems** to partially explain the population's hateful feelings towards the West. (CARE0019-EN-04)

Given the absence of significant differences in English proficiency levels among CARE writers (see Section 4), other factors likely contribute to the variation in L1-L2 stance use. One possible explanation is the varying levels of academic literacy among writers, as familiarity with the academic sphere and its cross-linguistic and cross-cultural nuances may differ. Less experienced students, for example, may rely more on their L1 literacy when writing in the L2. While we gathered some background information on students' academic reading and writing habits, it did not allow for a systematic exploration of the relationship between cross-linguistic influence and the development of L1 and L2 academic literacy. We believe that investigating this relationship in more detail is the necessary next step. To achieve this, however, there is a pressing need to collect L1-L2 corpora enriched with extensive metadata, including more detailed information on students' academic literacy in both languages.

⁹ We checked whether this was because, in some cases, the paired texts belong to different disciplines (linguistics vs. literature), but found that this variation occurs even within the same discipline.

8. CONCLUSION

Given the interplay of lexical, phraseological, and discursive properties, stance is arguably particularly susceptible to cross-linguistic influence, especially among novice writers (e.g. Ringbom 2007; Paquot 2013). However, research on the relationship between L1 and L2 stance use remains limited, often hindered by issues of attribution, sample size, comparability, and a lack of consideration for individual variation. This study addresses these gaps by examining the relationship between L1 and L2 stance use in paired L1-L2 texts produced by the same French-speaking students, supplemented by spot checks in a corpus of L1 English novice writing.

Comparisons between the CARE-FR and the CARE-EN revealed similar frequencies of stance markers and lexical/phraseological preferences, indicating a clear link between L1 and L2 stance use in novice writing. These findings suggest that the use of stance markers is affected by cross-linguistic influence. While some tendencies observed across L1-L2 texts in the CARE differ from typical patterns in L1 English writing (e.g. epistemic marker frequencies), others align more closely (e.g. boosting lexical preferences). In such cases, the learners' L1 may serve as a facilitating factor for their L2 English writing. The findings are also consistent with those of Jadoulle (2023b), who investigated the use of hedges in L1 French and L2 English texts produced by separate sets of novice writers, revealing many similarities across the two sets. However, the present study advances the methodology of Jadoulle (2023b) by examining L1 and L2 writing from the same individuals, rather than separate groups (see Section 2). This approach strengthens the validity of the conclusions about cross-linguistic effects in L2 writing.

The results also emphasize the importance of individual variation in the stance use of French-speaking student writers, with, for example, writers using higher frequencies of boosters in French than in English, and others doing the opposite. These findings are in line with recent research insisting on the importance of taking individual variation into account when investigating L2 data (e.g. Gries and Wulff 2021; Jadoulle 2024).

Other linguistic phenomena would also benefit from being explored from an L1-L2 perspective. A crucial first step is the collection of data that enables such analyses, specifically larger corpora containing paired L1 and L2 texts from the same individuals. Additionally, there is a need for L1-L2 corpora that include a broader range of language pairs, as well as more detailed metadata on L1-L2 academic literacy. However, the field

is moving in this direction, with the development of corpora such as the *Aachen Corpus of Academic Writing* (ACAW), which contains L1 German and L2 English writing by the same individuals (Ströbel *et al.* 2020).

As more L1-L2 corpora are developed, researchers will be better equipped to understand how cross-linguistic influence shapes stance in the development of L2 writing at the individual level. This deeper understanding will ultimately inform more effective, individualized teaching, enabling educators to better support novice writers as they navigate the complexities of academic discourse.

REFERENCES

- Ädel, Annelie and Ute Römer. 2012. Research on advanced student writing across disciplines and levels. Introducing the Michigan Corpus of Upper-level Student Papers. *International Journal of Corpus Linguistics* 17/1: 81–92. <https://doi.org/10.1075/ijcl.17.1.01ade>
- Biber, Douglas. 2006. *University Language. A Corpus-Based Study of Spoken and Written Registers*. Amsterdam: John Benjamins.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *The Longman Grammar of Spoken and Written English*. London: Longman.
- Biber, Douglas, Jesse Egbert and Meixiu Zhang. 2018. Lexis and grammar as complementary discourse systems for expressing stance and evaluation. In María de los Ángeles Gómez González and J. Lanchlan Mackenzie eds. *The Construction of Discourse as Verbal Interaction*. Amsterdam: John Benjamins, 201–226.
- Bowden, H. Wood. 2016. Assessing second-language oral proficiency for research. *Studies in Second Language Acquisition* 38: 647–675. <https://doi.org/10.1017/s0272263115000443>
- Cao, Dung and Richard Badger. 2021. Cross-linguistic influence on the use of L2 collocations: the case of Vietnamese learners. *Applied Linguistics Review* 14/3: 421–446. <https://doi.org/10.1515/applirev-2020-0035>
- Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences* (second edition). Hillsdale: Erlbaum.
- Chen, Hsin-I. 2010. Contrastive learner corpus analysis of epistemic modality and interlanguage pragmatic competence in L2 writing. *Arizona Working Papers in SLA & Teaching* 17: 27–51.
- Chung, Tieu-Thuy, Luyen-Thi Bui and Peter Crosthwaite. 2021. Evaluative stance in Vietnamese and English writing by the same authors: A corpus-informed appraisal study. *Research in Corpus Linguistics* 10/1: 1–30. <https://doi.org/10.32714/ricl.10.01.01>
- Flege, J. Emil, Ian A. MacKay and Thorsten Piske. 2002. Assessing bilingual dominance. *Applied Psycholinguistics* 23: 567–598. <https://doi.org/10.1017/s0142716402004046>
- Gilquin, Gaëtanelle. 2008. Combining contrastive and interlanguage analysis to apprehend transfer: Detection, explanation, evaluation. In Gaëtanelle Gilquin,

- Szilvia Papp and María Belén Díez-Bedmar eds. *Linking up Contrastive and Learner Corpus Research*. Amsterdam: Rodopi, 3–33.
- Gray, Bethany. 2015. *Linguistic Variation in Research Articles. When Discipline Tells Only Part of the Story*. Amsterdam: John Benjamins.
- Gray, Bethany and Douglas Biber. 2014. Stance markers. In Karin Aijmer and Christoph Rühlemann eds. *Corpus Pragmatics: A Handbook*. Cambridge: Cambridge University Press, 219–248.
- Gries, Stefan. 2013. *Statistics for Linguistics with R. A Practical Introduction* (second edition). Berlin and Boston: De Gruyter Mouton.
- Gries, Stefan and Stefanie Wulff. 2021. Examining individual variation in learner production data: A few programmatic pointers for corpus-based analyses using the example of adverbial clause ordering. *Applied Psycholinguistics* 42: 279–299. <https://doi.org/10.1017/s014271642000048x>
- Heuboeck, Alois, Jasper Holmes and Hilary Nesi. 2008. *The BAWE Corpus Manual. British Academic Written English Corpus (BAWE)* | Coventry University.
- Hinkel, Eli. 2003. *Teaching Academic ESL Writing: Practical Techniques in Vocabulary and Grammar*. Mahwah: Lawrence Erlbaum.
- Hirose, Keiko. 2003. Comparing L1 and L2 organizational patterns in the argumentative writing of Japanese EFL students. *Journal of Second Language Writing* 12: 181–209. [https://doi.org/10.1016/s1060-3743\(03\)00015-8](https://doi.org/10.1016/s1060-3743(03)00015-8)
- Huang, Xiaoyu, Xueying Liang and Effie Dracopoulos. 2011. A study on the relationship between university students' Chinese writing proficiency and their English writing proficiency. *English Language Teaching* 4/2: 55–65. <https://doi.org/10.5539/elt.v4n2p55>
- Hyland, Ken. 2000. *Disciplinary Discourses. Social Interactions in Academic Writing*. Harlow: Longman.
- Hyland, Ken. 2005a. *Metadiscourse: Exploring Interaction in Writing*. London: Continuum.
- Hyland, Ken. 2005b. Stance and engagement: A model of interaction in academic discourse. *Discourse Studies* 7/2: 173–192. <https://doi.org/10.1177/1461445605050365>
- Hyland, Ken and Polly Tse. 2005. Hooking the reader: A corpus study of evaluative *that* in abstracts. *English for Specific Purposes* 24: 123–139. <https://doi.org/10.1016/j.esp.2004.02.002>
- Jadoulle, Pauline. 2023a. An investigation of stance in French native novice academic writing. *Papers of the Linguistic Society of Belgium* 17: 76–106. <https://doi.org/10.61430/mqro4392>
- Jadoulle, Pauline. 2023b. L1 novice writing as a missing piece in the Learner Corpus Research puzzle: The case of hedging. *International Journal of Learner Corpus Research* 9/2: 181–215. <https://doi.org/10.1075/ijlcr.00033.jad>
- Jadoulle, Pauline. 2024. *Investigating Noviceness and Non-Nativeness in Academic Writing: A Cross-Linguistic Approach to Stance*. Louvain-la-Neuve: Université catholique de Louvain.
- Jarvis, Scott. 2000. Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. *Language Learning* 50: 245–309. <https://doi.org/10.1111/0023-8333.00118>
- Jarvis, Scott. 2009. Lexical transfer. In Aneta Pavlenko ed. *The Bilingual Mental Lexicon: Interdisciplinary Approaches*. Bristol: Multilingual Matters, 99–124.
- Jarvis, Scott and Aneta Pavlenko. 2008. *Crosslinguistic Influence in Language and Cognition*. London: Routledge.

- Jiang, Feng. 2015. Nominal stance construction in L1 and L2 students' writing. *Journal of English for Academic Purposes* 20: 90–102. <https://doi.org/10.1016/j.jeap.2015.07.002>
- Kilgariff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubiček, Vojtěch Kovař, Jan Michelfeit, Pavel Rychly and Vít Suchomel. 2014. The Sketch Engine: Ten years on. *Lexicography* 1/1: 7–36. <https://doi.org/10.1007/s40607-014-0009-9>
- Kubota, Ryuko. 1998. An investigation of L1-L2 transfer in writing among Japanese university students: Implications for contrastive rhetoric. *Journal of Second Language Writing* 7/1: 69–100. [https://doi.org/10.1016/s1060-3743\(98\)90006-6](https://doi.org/10.1016/s1060-3743(98)90006-6)
- Larsson, Tove. 2016. *The Introductory it Pattern in Academic Writing by Non-Native-Speaker Students, Native-Speaker Students and Published Writers: A Corpus-Based Study*. Uppsala: Uppsala University.
- Larsson, Tove, Randi Reppen and Tülay Dixon. 2022. A phraseological study of highlighting strategies in novice and expert writing. *Journal of English for Academic Purposes* 60: 101179. <https://doi.org/10.1016/j.jeap.2022.101179>
- Liu, Xinghua and Paul Thompson. 2009. Attitude in students' argumentative writing: A contrastive perspective. *Language Studies Working Papers* 1: 3–15.
- Martin, James R. 2000. Beyond exchange: APPRAISAL systems in English. In Susan Hunston and Geoff Thompson eds. *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press, 142–175.
- Martin, James R. and Peter Robert R. White. 2005. *The Language of Evaluation. Appraisal in English*. Houndmills: Palgrave Macmillan.
- McEnery, Tony and Andrew Hardie. 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.
- Odlin, Terence. 1989. *Language Transfer: Cross-linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.
- Paquot, Magali. 2008. Exemplification in learner writing: A cross-linguistic perspective. In Sylviane Granger and Fanny Meunier eds. *Phraseology in Foreign Language Learning and Teaching*. Amsterdam: John Benjamins, 101–119.
- Paquot, Magali. 2010. *Academic Vocabulary in Learner Writing: From Extraction to Analysis*. London: Continuum.
- Paquot, Magali. 2013. Lexical bundles and L1 transfer effects. *International Journal of Corpus Linguistics* 18/3: 391–417. <https://doi.org/10.1075/ijcl.18.3.06paq>
- Paquot, Magali. 2014. Cross-linguistic influence and formulaic language: Recurrent word sequences in French learner writing. In Leah Roberts, Ineke Vedder and Jan Hulstijn eds. *EuroSLA Yearbook*. Amsterdam: John Benjamins, 216–237.
- Paquot, Magali. 2017. L1 frequency in foreign language acquisition: Recurrent word combinations in French and Spanish EFL learner writing. *Second Language Research* 33/1: 13–32. <https://doi.org/10.1177/0267658315620265>
- Paquot, Magali, Alexander König, Egon W. Stemle and Jennifer-Carmer Frey. 2023. Core metadata schema for learner corpora. *Open Data @ UCLouvain*, V1. <https://doi.org/10.14428/dvn/4cdx3p>
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
- Rica Peromingo, Juan Pedro. 2012. Corpus analysis and phraseology: Transfer of multi-word units. *Linguistics and the Human Sciences* 6: 321–343. <https://doi.org/10.1558/lhs.v6i1-3.321>
- Ringbom, Håkan. 2007. *Cross-linguistic Similarity in Foreign Language Learning*. Bristol: Multilingual Matters.

- Scott, Mike and Christopher Tribble. 2006. *Textual Patterns. Key Words and Corpus Analysis in Language Education*. Amsterdam: John Benjamins.
- Shatz, Itamar, Theodora Alexopoulou and Akira Murakami. 2023. Examining the potential influence of crosslinguistic lexical similarity on word-choice transfer in L2 English. *PLoS ONE* 18/2.
- Singleton, David. 1987. The fall and rise of language transfer. In Jim Coleman and Rixard Towell eds. *The Advanced Language Learner*. CILT, 27–53.
- Ströbel, Marcus, Elma Kerz and Daniel Wiechmann. 2020. The relationship between first and second language writing: Investigating the effects of first language complexity on second language complexity in advanced stages of learning. *Language Learning* 70/3: 732–767. <https://doi.org/10.1111/lang.12394>
- Swan, Michael. 1997. The influence of the mother tongue on second language vocabulary acquisition and use. In Norbert Schmitt and Michael McCarthy eds. *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press, 156–180.
- Tremblay, Annie. 2011. Proficiency assessment standards in Second Language Acquisition research: “Clozing” the gap. *Studies in Second Language Acquisition* 33/3: 339–72. <https://doi.org/10.1017/s0272263111000015>
- Uysal, H. Hande. 2008. Tracing the culture behind writing: Rhetorical patterns and bidirectional transfer in L1 and L2 essays of Turkish writers in relation to educational context. *Journal of Second Language Writing* 17: 183–207. <https://doi.org/10.1016/j.jslw.2007.11.003>
- Uzawa, Kozue. 1996. Second language learners’ processes of L1 writing, L2 writing, and translation from L1 into L2. *Journal of Second Language Writing* 5/3: 271–294. [https://doi.org/10.1016/s1060-3743\(96\)90005-3](https://doi.org/10.1016/s1060-3743(96)90005-3)
- van Weijen, Daphne, Huub van den Bergh, Gert Rijlaarsdam and Ted Sanders. 2009. L1 use during L2 writing: An empirical study of a complex phenomenon. *Journal of Second Language Writing* 18: 235–250. <https://doi.org/10.1016/j.jslw.2009.06.003>
- Vold, Eva Thue. 2008. *Modalité épistémique et discours scientifique. Une étude contrastive des modalisateurs épistémiques dans des articles de recherche français, norvégiens et anglais, en linguistique et médecine*. Unpublished doctoral dissertation, University of Bergen.
- Wang, Jingjing and Feng Jiang. 2018. Epistemic stance and authorial presence in scientific research writing. Hedges, boosters and self-mentions across disciplines and writer groups. In Pilar Mur-Dueñas and Jolanta Šinkūnienė eds. *Intercultural Perspectives on Research Writing*. Amsterdam: John Benjamins, 195–216.
- Wang, Wenyu and Qiufang Wen. 2002. L1 use in the L2 composing process: An exploratory study of 16 Chinese EFL writers. *Journal of Second Language Writing* 11: 225–246. [https://doi.org/10.1016/s1060-3743\(02\)00084-x](https://doi.org/10.1016/s1060-3743(02)00084-x)
- Williams, John. 2015. The bilingual lexicon. In John Taylor ed. *The Oxford Handbook of the Word*. Oxford: Oxford University Press, 493–507.
- Yamashita, Junko and Nan Jiang. 2010. L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly* 44/4: 647–668. <https://doi.org/10.5054/tq.2010.235998>
- Yoon, Hyung-Jo and Ute Römer. 2020. Quantifying disciplinary voices: An automated approach to interactional metadiscourse in successful student writing. *Written Communication* 37: 208–244. <https://doi.org/10.1177/0741088319898672>

Corresponding author

Pauline Jadoulle

Université catholique de Louvain

ILC

1348 Louvain-la-Neuve

Belgium

E-mail: pauline.jadoulle@uclouvain.be

received: June 2025

accepted: December 2025

APPENDIX¹⁰

<i>absolutely</i>	<i>imperatively</i>	<i>of course</i>	<i>strikingly</i>
<i>accurately</i>	<i>importantly</i>	<i>ostensibly</i>	<i>supposedly</i>
<i>admittedly</i>	<i>in theory</i>	<i>paradoxically</i>	<i>surely</i>
<i>apparently</i>	<i>incontestably</i>	<i>patently</i>	<i>surprisingly</i>
<i>arguably</i>	<i>indisputably</i>	<i>perhaps</i>	<i>theoretically</i>
<i>certainly</i>	<i>indubitably</i>	<i>plausibly</i>	<i>to be sure</i>
<i>clearly</i>	<i>inevitably</i>	<i>possibly</i>	<i>to my knowledge</i>
<i>compulsorily</i>	<i>interestingly</i>	<i>potentially</i>	<i>to our knowledge</i>
<i>conceivably</i>	<i>legitimately</i>	<i>presumably</i>	<i>trivially</i>
<i>crucially</i>	<i>likely</i>	<i>probably</i>	<i>unarguably</i>
<i>curiously</i>	<i>logically</i>	<i>promisingly</i>	<i>understandably</i>
<i>definitely</i>	<i>luckily</i>	<i>purportedly</i>	<i>undoubtedly</i>
<i>demonstrably</i>	<i>manifestly</i>	<i>reasonably</i>	<i>unfortunately</i>
<i>doubtless</i>	<i>maybe</i>	<i>remarkably</i>	<i>unmistakably</i>
<i>doubtlessly</i>	<i>naturally</i>	<i>resolutely</i>	<i>unquestionably</i>
<i>eventually</i>	<i>necessarily</i>	<i>rightly</i>	<i>unreasonably</i>
<i>evidently</i>	<i>no doubt</i>	<i>seemingly</i>	<i>unremarkably</i>
<i>fortunately</i>	<i>notably</i>	<i>self-evidently</i>	<i>unsurprisingly</i>
<i>hopefully</i>	<i>obligatorily</i>	<i>significantly</i>	<i>without any doubt</i>
<i>ideally</i>	<i>obviously</i>	<i>strangely</i>	

Table A1: English adverbials selected

¹⁰ The lists of markers presented in the Appendix include more markers than those found in the CARE, as they are derived from other – larger – corpora (see Jadouille 2023b).

<i>acceptable</i>	<i>effective</i>	<i>misguided</i>	<i>simple</i>
<i>accurate</i>	<i>effortful</i>	<i>misleading</i>	<i>simplistic</i>
<i>advisable</i>	<i>effortless</i>	<i>natural</i>	<i>straightforward</i>
<i>afraid</i>	<i>encouraging</i>	<i>necessary</i>	<i>strange</i>
<i>apparent</i>	<i>essential</i>	<i>needless</i>	<i>striking</i>
<i>appropriate</i>	<i>evident</i>	<i>nice</i>	<i>suggestive</i>
<i>arguable</i>	<i>fair</i>	<i>normal</i>	<i>sure</i>
<i>astonishing</i>	<i>false</i>	<i>notable</i>	<i>surprised</i>
<i>aware</i>	<i>far-fetched</i>	<i>noteworthy</i>	<i>surprising</i>
<i>beneficial</i>	<i>fascinating</i>	<i>noticeable</i>	<i>tempted</i>
<i>captivating</i>	<i>fitting</i>	<i>obvious</i>	<i>tempting</i>
<i>careful</i>	<i>fortunate</i>	<i>odd</i>	<i>tricky</i>
<i>certain</i>	<i>good</i>	<i>painless</i>	<i>true</i>
<i>challenging</i>	<i>happy</i>	<i>plausible</i>	<i>unaware</i>
<i>clear</i>	<i>hard</i>	<i>possible</i>	<i>unclear</i>
<i>complex</i>	<i>hasty</i>	<i>preferable</i>	<i>uncontentious</i>
<i>conceivable</i>	<i>helpful</i>	<i>probable</i>	<i>undeniable</i>
<i>confident</i>	<i>imperative</i>	<i>problematic</i>	<i>understandable</i>
<i>conscious</i>	<i>implausible</i>	<i>productive</i>	<i>unexplainable</i>
<i>content</i>	<i>important</i>	<i>profitable</i>	<i>unfortunate</i>
<i>contradictory</i>	<i>impractical</i>	<i>promising</i>	<i>unlikely</i>
<i>convenient</i>	<i>inadequate</i>	<i>prone</i>	<i>unnecessary</i>
<i>convinced</i>	<i>inappropriate</i>	<i>provocative</i>	<i>unreasonable</i>
<i>correct</i>	<i>inconceivable</i>	<i>prudent</i>	<i>unsure</i>
<i>critical</i>	<i>inevitable</i>	<i>quick</i>	<i>urgent</i>
<i>crucial</i>	<i>instructive</i>	<i>reasonable</i>	<i>useful</i>
<i>curious</i>	<i>interesting</i>	<i>relevant</i>	<i>useless</i>
<i>dangerous</i>	<i>judicious</i>	<i>reluctant</i>	<i>valid</i>
<i>debatable</i>	<i>justified</i>	<i>remarkable</i>	<i>valuable</i>
<i>desirable</i>	<i>keen</i>	<i>reputed</i>	<i>vital</i>
<i>difficult</i>	<i>key</i>	<i>revealing</i>	<i>well(-)established</i>
<i>disingenuous</i>	<i>legitimate</i>	<i>ridiculous</i>	<i>well(-)known</i>
<i>doubtful</i>	<i>likely</i>	<i>right</i>	<i>wise</i>
<i>dubious</i>	<i>logical</i>	<i>safe</i>	<i>worth</i>
<i>eager</i>	<i>meaningful</i>	<i>sensible</i>	<i>worthwhile</i>
<i>easy</i>	<i>mindful</i>	<i>significant</i>	<i>wrong</i>

Table A2: English adjectives selected¹¹

¹¹ Some of the items (adverbs, adjectives, nouns and verbs) were mistagged (e.g. the verb *demand* was tagged as a noun). These items were moved to the proper table.

<i>acknowledg(e)ment</i>	<i>criticism</i>	<i>knowledge</i>	<i>problem</i>
<i>affirmation</i>	<i>difficulty</i>	<i>likelihood</i>	<i>proof</i>
<i>agreement</i>	<i>doubt</i>	<i>misconception</i>	<i>reluctance</i>
<i>argument</i>	<i>drawback</i>	<i>mistake</i>	<i>requirement</i>
<i>assertion</i>	<i>duty</i>	<i>necessity</i>	<i>simplicity</i>
<i>assumption</i>	<i>evidence</i>	<i>need</i>	<i>stance</i>
<i>assurance</i>	<i>exaggeration</i>	<i>no question</i>	<i>standpoint</i>
<i>awareness</i>	<i>failure</i>	<i>obstinacy</i>	<i>suggestion</i>
<i>belief</i>	<i>feeling</i>	<i>occasion</i>	<i>supposition</i>
<i>bias</i>	<i>guarantee</i>	<i>opinion</i>	<i>surprise</i>
<i>caveat</i>	<i>hope</i>	<i>opportunity</i>	<i>suspicion</i>
<i>certainty</i>	<i>hypothesis</i>	<i>particularity</i>	<i>temptation</i>
<i>challenge</i>	<i>illusion</i>	<i>perception</i>	<i>theory</i>
<i>chance</i>	<i>imperative</i>	<i>perspective</i>	<i>thesis</i>
<i>claim</i>	<i>impression</i>	<i>point of view</i>	<i>thought</i>
<i>concern</i>	<i>inevitability</i>	<i>position</i>	<i>unlikelihood</i>
<i>conclusion</i>	<i>insight</i>	<i>possibility</i>	<i>value</i>
<i>confirmation</i>	<i>interest</i>	<i>potential</i>	<i>view</i>
<i>conjecture</i>	<i>interpretation</i>	<i>preference</i>	<i>viewpoint</i>
<i>contention</i>	<i>intuition</i>	<i>presumption</i>	<i>wonder</i>
<i>conviction</i>	<i>judg(e)ment</i>	<i>probability</i>	

Table A3: English nouns selected

<i>accept</i>	<i>consider</i>	<i>imply</i>	<i>realis/ze</i>
<i>acknowledge</i>	<i>contend</i>	<i>infer</i>	<i>recognis/ze</i>
<i>admit</i>	<i>convince</i>	<i>interpret</i>	<i>require</i>
<i>advocate</i>	<i>declare</i>	<i>judge</i>	<i>reveal</i>
<i>affirm</i>	<i>deem</i>	<i>know</i>	<i>see</i>
<i>agree</i>	<i>demand</i>	<i>lose sight</i>	<i>seem</i>
<i>allege</i>	<i>demonstrate</i>	<i>maintain</i>	<i>show</i>
<i>appear</i>	<i>deny</i>	<i>make sense</i>	<i>speculate</i>
<i>appreciate</i>	<i>deserve</i>	<i>manage</i>	<i>strike</i>
<i>argue</i>	<i>doubt</i>	<i>object</i>	<i>submit</i>
<i>assert</i>	<i>establish</i>	<i>perceive</i>	<i>suggest</i>
<i>assume</i>	<i>fail</i>	<i>posit</i>	<i>suppose</i>
<i>assure</i>	<i>feel</i>	<i>postulate</i>	<i>surmise</i>
<i>attest</i>	<i>find</i>	<i>prefer</i>	<i>suspect</i>
<i>believe</i>	<i>forget</i>	<i>presume</i>	<i>take for granted</i>
<i>claim</i>	<i>guarantee</i>	<i>pretend</i>	<i>testify</i>
<i>complain</i>	<i>guess</i>	<i>promise</i>	<i>think</i>
<i>concede</i>	<i>hesitate</i>	<i>propose</i>	<i>venture</i>
<i>conclude</i>	<i>hope</i>	<i>protest</i>	<i>view</i>
<i>confess</i>	<i>hurt¹²</i>	<i>prove</i>	<i>worry</i>
<i>confirm</i>	<i>hypothesis/ze</i>	<i>purport</i>	
<i>conjecture</i>	<i>imagine</i>	<i>question</i>	

Table A4: English verbs selected

¹² *It wouldn't hurt to.*

<i>à juste titre</i>	<i>éventuellement</i>	<i>irréductiblement</i>	<i>probablement</i>
<i>absolument</i>	<i>évidemment</i>	<i>légitimement</i>	<i>raisonnablement</i>
<i>apparemment</i>	<i>fatalement</i>	<i>logiquement</i>	<i>résolument</i>
<i>assurément</i>	<i>forcément</i>	<i>malheureusement</i>	<i>rien d'étonnant</i>
<i>bien entendu</i>	<i>heureusement</i>	<i>manifestement</i>	<i>sans doute</i>
<i>bien sûr</i>	<i>idéalement</i>	<i>naturellement</i>	<i>sans aucun doute</i>
<i>censément</i>	<i>immanquablement</i>	<i>nécessairement</i>	<i>significativement</i>
<i>certainement</i>	<i>impérativement</i>	<i>obligatoirement</i>	<i>supposément</i>
<i>certes</i>	<i>impérieusement</i>	<i>ostensiblement</i>	<i>sûrement</i>
<i>clairement</i>	<i>incontestablement</i>	<i>oui</i>	<i>théoriquement</i>
<i>crucialement</i>	<i>indéniablement</i>	<i>paradoxalement</i>	<i>valablement</i>
<i>curieusement</i>	<i>indubitablement</i>	<i>peut-être</i>	<i>visiblement</i>
<i>définitivement</i>	<i>inéluctablement</i>	<i>possiblement</i>	<i>vraisemblablement</i>
<i>étonnamment</i>	<i>inévitablement</i>	<i>potentiellement</i>	
<i>étrangement</i>	<i>ironiquement</i>	<i>prétendûment</i>	

Table A5: French adverbials selected

<i>absurde</i>	<i>curieux</i>	<i>indispensable</i>	<i>préférable</i>
<i>abusif</i>	<i>délicat</i>	<i>inintéressant</i>	<i>prétentieux</i>
<i>admissible</i>	<i>difficile</i>	<i>injuste</i>	<i>primordial</i>
<i>agaçant</i>	<i>éclairant</i>	<i>injustifié</i>	<i>probable</i>
<i>aisé</i>	<i>enclin</i>	<i>intéressant</i>	<i>productif</i>
<i>amusant</i>	<i>erroné</i>	<i>inutile</i>	<i>raisonnable</i>
<i>anodin</i>	<i>essentiel</i>	<i>judicieux</i>	<i>possible</i>
<i>ardu</i>	<i>étonnant</i>	<i>juste</i>	<i>regrettable</i>
<i>banal</i>	<i>étrange</i>	<i>justifié</i>	<i>remarquable</i>
<i>certain</i>	<i>évident</i>	<i>légitime</i>	<i>révéléateur</i>
<i>clair</i>	<i>facile</i>	<i>logique</i>	<i>risqué</i>
<i>complexe</i>	<i>faux</i>	<i>loisible</i>	<i>significatif</i>
<i>compliqué</i>	<i>fondamental</i>	<i>malaisé</i>	<i>simple</i>
<i>compréhensible</i>	<i>fortuit</i>	<i>manifeste</i>	<i>soucieux</i>
<i>condamné</i>	<i>frappant</i>	<i>mieux</i>	<i>souhaitable</i>
<i>connu</i>	<i>hasardeux</i>	<i>naturel</i>	<i>sûr</i>
<i>conscient</i>	<i>heureux</i>	<i>nécessaire</i>	<i>surprenant</i>
<i>contestable</i>	<i>impératif</i>	<i>normal</i>	<i>surpris</i>
<i>contraignant</i>	<i>important</i>	<i>obligatoire</i>	<i>susceptible</i>
<i>convaincant</i>	<i>improbable</i>	<i>opportun</i>	<i>utile</i>
<i>convaincu</i>	<i>incontestable</i>	<i>persuadé</i>	<i>vrai</i>
<i>crucial</i>	<i>indéniable</i>	<i>pertinent</i>	<i>vraisemblable</i>

Table A6: French adjectives selected

<i>affirmation</i>	<i>devoir</i>	<i>intuition</i>	<i>présomption</i>
<i>argument</i>	<i>difficulté</i>	<i>jugement</i>	<i>prétention</i>
<i>avantage</i>	<i>doute</i>	<i>légitimité</i>	<i>preuve</i>
<i>avis</i>	<i>erreur</i>	<i>lucidité</i>	<i>problème</i>
<i>besoin</i>	<i>espérance</i>	<i>mérite</i>	<i>regret</i>
<i>certitude</i>	<i>espoir</i>	<i>nécessité</i>	<i>revendication</i>
<i>chance</i>	<i>exploit</i>	<i>obligation</i>	<i>risque</i>
<i>complexité</i>	<i>facilité</i>	<i>occasion</i>	<i>sagesse</i>
<i>conclusion</i>	<i>hypothèse</i>	<i>opportunité</i>	<i>sentiment</i>
<i>conscience</i>	<i>illusion</i>	<i>particularité</i>	<i>stratégie</i>
<i>constat</i>	<i>impératif</i>	<i>pertinence</i>	<i>supposition</i>
<i>constatation</i>	<i>importance</i>	<i>peur</i>	<i>tentation</i>
<i>crainte</i>	<i>impression</i>	<i>possibilité</i>	<i>thèse</i>
<i>défaut</i>	<i>injonction</i>	<i>postulat</i>	<i>tort</i>
<i>défi</i>	<i>intérêt</i>	<i>préoccupation</i>	<i>utilité</i>

Table A7: French nouns selected

<i>accepter</i>	<i>déclarer</i>	<i>menacer</i>	<i>regretter</i>
<i>admettre</i>	<i>décréter</i>	<i>mériter</i>	<i>ressentir</i>
<i>affirmer</i>	<i>défendre</i>	<i>montrer</i>	<i>rétorquer</i>
<i>apparaître</i>	<i>démontrer</i>	<i>nécessiter</i>	<i>revendiquer</i>
<i>apprécier</i>	<i>douter</i>	<i>nier</i>	<i>risquer</i>
<i>argumenter</i>	<i>échouer</i>	<i>objecter</i>	<i>savoir</i>
<i>assurer</i>	<i>espérer</i>	<i>opposer</i>	<i>se contenter</i>
<i>attester</i>	<i>estimer</i>	<i>oser</i>	<i>se douter</i>
<i>avouer</i>	<i>établir</i>	<i>oublier</i>	<i>sembler</i>
<i>concéder</i>	<i>étonner</i>	<i>paraître</i>	<i>sentir</i>
<i>conclure</i>	<i>exiger</i>	<i>penser</i>	<i>sous-entendre</i>
<i>confirmer</i>	<i>hésiter</i>	<i>postuler</i>	<i>soutenir</i>
<i>conjecturer</i>	<i>imaginer</i>	<i>préférer</i>	<i>suggérer</i>
<i>considérer</i>	<i>importer</i>	<i>présumer</i>	<i>supposer</i>
<i>constater</i>	<i>inférer</i>	<i>prétendre</i>	<i>témoigner</i>
<i>convaincre</i>	<i>intéresser</i>	<i>proposer</i>	<i>trouver</i>
<i>convenir</i>	<i>juger</i>	<i>prouver</i>	<i>valoir</i>
<i>craindre</i>	<i>manifester</i>	<i>réaliser</i>	<i>veiller</i>
<i>croire</i>	<i>manquer</i>	<i>reconnaître</i>	<i>voir</i>

Table A8: French verbs selected