



Discours

Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics

15 | 2014
Varia

Authorial Presence in French and English: “Pronoun + Verb” Patterns in Biology and Medicine Research Articles

Laura M. Hartwell and Marie-Paule Jacques



Electronic version

URL: <http://journals.openedition.org/discours/8941>

DOI: 10.4000/discours.8941

ISSN: 1963-1723

Publisher:

Laboratoire LATTICE, Presses universitaires de Caen

Brought to you by Université Toulouse 1 - Capitole



Electronic reference

Laura M. Hartwell and Marie-Paule Jacques, “Authorial Presence in French and English: “Pronoun + Verb” Patterns in Biology and Medicine Research Articles”, *Discours* [Online], 15 | 2014, Online since 19 December 2014, connection on 17 February 2021. URL: <http://journals.openedition.org/discours/8941> ; DOI: <https://doi.org/10.4000/discours.8941>



Discours est mis à disposition selon les termes de la licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.



Revue de linguistique, psycholinguistique et informatique

<http://discours.revues.org/>

Authorial Presence in French and English: “Pronoun + Verb” Patterns in Biology and Medicine Research Articles

Laura M. Hartwell

LIDILEM

Université Grenoble Alpes, UJF

Marie-Paule Jacques

LIDILEM

Université Grenoble Alpes, UJF

.....
Laura M. Hartwell et Marie-Paule Jacques, « Authorial Presence in French and English: “Pronoun + Verb” Patterns in Biology and Medicine Research Articles », *Discours* [En ligne], 15 | 2014, mis en ligne le 19 décembre 2014.

.....
URL : <http://discours.revues.org/8941>

.....
Titre du numéro : *Varia*

Coordination : Catherine Bolly et Benjamin Fagard

revues.org
CENTRE POUR L'ÉDITION ÉLECTRONIQUE OUVERTE
CENTRE FOR OPEN ELECTRONIC PUBLISHING

 discours

 Presses
universitaires
de Caen

Authorial Presence in French and English: “Pronoun + Verb” Patterns in Biology and Medicine Research Articles

Laura M. Hartwell

LIDILEM

Université Grenoble Alpes, UJF

Marie-Paule Jacques

LIDILEM

Université Grenoble Alpes, UJF

.....

Certain subjective qualities of scientific research articles are exposed when authors refer to themselves through various means including pronoun use. Drawing upon the online bilingual “Scientext” corpus, we compare personal pronoun and syntactically linked verb constructions within 180 published articles in English and French in the fields of medicine and biology. This study found that overall pronoun frequency was significantly higher ($\chi^2 = 69.45$, $df = 1$, $p < 0.001$) in English (22.6 per 10,000) than in French (14 per 10,000) and that the French *on* [one] (23.8%) was significantly more frequent ($\chi^2 = 163.35$, $df = 1$, $p < 0.001$) than the English pronoun “one” (3.8%). Furthermore, while most French verbs were limited to the present and *passé composé*, English conjugation was distributed mainly between the simple past, the simple present, and the present perfect. Both the lexis and the conjugation vary with the progression of the research article and the author roles of *researcher*, *writer*, *arguer*, and *evaluator*. This paper contributes to the discussion of the representation of objectivity in scientific discourse.

Keywords: corpora, pronoun use, research article, scientific discourse

.....

Dans cet article, nous nous intéressons de manière contrastive à la présence de l'auteur scientifique dans les articles de recherche, lorsque celle-ci se manifeste à travers l'emploi de pronoms personnels. Nous appuyant sur 180 articles scientifiques en anglais et en français sélectionnés depuis la base « Scientext » dans les domaines de la médecine et de la biologie, nous avons comparé les pronoms personnels référant à l'auteur et les verbes auxquels ces pronoms sont liés syntaxiquement, de façon globale ainsi qu'en prenant en compte la structure de l'article. Nos décomptes font apparaître une différence significative ($\chi^2 = 69,45$; $df = 1$; $p < 0,001$) d'emploi des pronoms en faveur de l'anglais : 22,6 pour 10 000 dans cette langue contre 14 pour 10 000 en français. La nature des pronoms est elle aussi significativement différente ($\chi^2 = 163,35$; $df = 1$; $p < 0,001$) : le pronom français « on » est largement plus présent (23,8 %) que le pronom anglais one (3,8 %). D'autres contrastes notables tiennent aux temps verbaux. Là où le français utilise essentiellement le présent et le passé composé, l'anglais varie entre passé simple, présent simple et présent perfect. Nous constatons des variations de lexique et de temps verbaux au fur et à mesure que progresse l'article et que l'auteur adopte tour à tour les rôles de chercheur, écrivain, argumentateur et évaluateur. En définitive, cette étude contribue à l'analyse de la représentation de l'objectivité et de la subjectivité dans le discours scientifique.

Mots clés : corpus, utilisation de pronom, article de recherche, discours scientifique

1. Introduction

1 Recent studies have revealed the argumentative qualities of scientific discourse, a form of discourse traditionally considered as being *neutral* (Fløttum, 2010; Fløttum et al., 2006a; Rinck et al., 2007; Hunston & Francis, 2000; Swales, 2004; Tutin & Grossmann, 2013). Within this context, authors may refer to themselves through various means such as lexical expressions (“the author”, “this study”), citation of previous studies, the passive voice (“this test was conducted”), but the most frequent means is currently via personal pronouns (Fløttum et al., 2007). In his cross-disciplinary study, Harwood (2005) focuses on the promotional effect or a “marking out of difference” through, among other means, personal pronoun use. He finds these promotional devices in a myriad of rhetorical steps: underscoring novelty, repeating or disputing claims, self-citation, confirming methodological soundness. Furthermore, in the natural sciences, authors traditionally minimize their presence through a formalized manner of reporting that subordinates their own voice (Hyland, 2006: 32). We posit that authorial presence evoked by pronominal use varies across languages and with the rhetorical function or role. Thus, we compared verb types syntactically linked to a pronoun that refers to the authors (“if one assumes”, “we performed”, “we and others have shown”) in both English and French across the different sections of research articles (RA) in the fields of medicine and biology. We begin with a review of the literature, specifically focusing on pronoun frequency, authorial positioning, the polyphonic nature of scientific discourse, and recommended writing style.

2. Review of the literature

2.1. Pronoun frequency across languages and disciplines

2 Pronoun use is historically, linguistically, and disciplinarily bound. In their study of the structurally compressed and inexplicit nature of academic writing in English, Biber and Gray (2010) demonstrate that during the 18th century both the pronouns “I” and “we” occurred approximately 3 to 4 times per 1,000 words in academic prose. Their corpus indicated a steady decline to about 1 occurrence per 1,000 words identified in 1965, thereafter the use of the pronoun “I” remained stable, but the “personal frame” offered by the use of “we” rose by 2005 to a new peak of 4 occurrences per 1,000 words. Cultural phenomena include the traditional French use of the first-person plural (*nous*) used as the “editorial we” (referring to a single first person) and the French indefinite pronoun *on* [one], which can be employed like the inclusive English “we”.

3 The pronouns “one”/*on* have contrasting etymological evolution as the Old English *ān* has Germanic roots from the Latin number *unus*, while the French *on* comes from the Latin *homo*, referring to a human being. In contrast to British English, the formal “one” may sound pretentious in American English. Reutner (2010) highlights the multiple functions of *on*, ranging from an undetermined reference, in guise of “we” particularly in colloquial spoken French, or referring to one or

more persons with a stylistic nuance such as discretion, modesty, irony or scorn. Carter-Thomas and Chambers (2012) found a variety of authorial roles played by first person pronouns in the introductions of economics RA in both English and French. Their study also confirmed the French avoidance of the singular first person pronoun in a formal register, contrary to the often accepted English usage, especially in relation to the author role of *writer*.

- 4 Finally, differences of pronoun use have been found across disciplines. In biology, for example, Hyland (2001) found no reference to "I" as articles tend to have multiple authors, but 24.0 tokens of "we" per 10,000 words, which he compared significantly to traditionally single-authored philosophy articles for which there were 35.6 tokens of "I" per 10,000 words, but only 1.4 tokens of "we". However, Swales and Feak (2004: 284-285) found that physicists and chemists are, "perhaps surprisingly" more likely to adopt a personal stance, through the use of the pronoun "we", than within the social sciences.

2.2. Positioning through pronoun use

- 5 Indeed, among other rhetorical devices, certain pronouns can be used either when establishing *stance*, through authorial self-reference, or when establishing *engagement*, through the use of all-inclusive pronouns ("we", "one"), which position readers as discourse participants and guide them through the text (Hyland, 2004; Hyland, 2012). Fløttum et al. (2006a) demonstrate that in medical RA the pronoun "we" is typically exclusive and hence does not include the reader. *Stance* ranges from stamping one's personal authority onto arguments, more often found in the humanities, to the more discrete disguise of one's involvement in the research, especially in the natural sciences (Hyland, 2006: 29). Sancho-Guinda and Hyland (2012: 4) summarize stance as "a continuum of evaluative meaning which varies along two axes: one epistemic and interpersonal (i.e. from feelings and attitudes to a status of knowledge) and the other linguistic (i.e. from lexis to grammar)". For example, in her study of social sciences RA in French, Tutin (2010) found that French verbs conveying a strong opinion (i.e., "to think", "to admit") were employed with an all-inclusive pronoun to engage the reader around a commonly known aspect, while verbs of intention or choice ("to show", "to conclude") were tied to exclusive self-reference of the author's work. Similarly, in his study of five recurrent French verbs related to making an assertion, Grossmann (2013) demonstrates that they are more frequently syntactically linked with the more inclusive *on* [one] than with *nous* [we], especially when the assertion refers to empirical data within the French "Scientext" subcorpus of academic texts.

2.3. Authorial voice and the KIAP project

- 6 In contrast to Hyland's primarily author-centered approach, Fløttum and colleagues (Fløttum, 2010; Fløttum et al., 2006b) highlight the polyphonic nature of scientific discourse. This polyphony, emblematic of a research community, contributes to the similarities within disciplines, despite inherent differences between given languages.

The similarities were confirmed by the analysis of the “Cultural Identity in Academic Prose” corpus (KIAP – “Kulturell Identitet i Akademisk Prosa”), a collection of RA in Norwegian, French and English in the fields of linguistics, economics and medicine (Fløttum et al., 2007; Fløttum et al., 2013). One similarity in the natural sciences is the recurrent IMRAD format (Introduction, Method, Results, and Discussion), which structures the text by defined section, also affecting the text on a deeper level. Specifically, in her study of the French indefinite pronoun *on*, Gjesdal (2013) demonstrates that the IMRAD format directs the author’s voice and the corresponding rhetorical qualities. The discursive polyphony is built upon pronouns as well as other means, including a wide range of self-reference, reference to other researchers, bibliographical references or epistemic markers evoking the need for further study.

- 7 In their study of author roles (*researcher, writer, arguer, and evaluator*), Fløttum et al. (2006b) compared the use of first person pronouns within the subcorpus of English RA as found in the KIAP corpus. The 50 medical RA of the KIAP corpus contained fewer verb types linked to “we” than found in the other domains and of those 120 different verbs, the most frequent lemmas – excluding “to have” and “to be” – were: “to find”, “to use”, “to examine”, “to observe”, and “to report” (Fløttum et al., 2006b: 209). They conclude that the *writer* and *arguer* roles were infrequent within these medical texts. Likewise, in her study of the KIAP corpus, Vold (2006) found that the epistemic “to suggest” was commonly matched with “we” within linguistics, but linked to inanimate subjects (*data, studies*) within the sciences. Finally, Carter-Thomas and Chambers (2012) found in their corpus of economics RA introductions, also drawn from the KIAP corpus, that although the *writer* role tended to predominate in relation to the use of “I”, the *researcher* role predominated in first person roles (singular and plural combined) for both languages.

2.4. Recommendations in writing style

- 8 Several guide books for writing in English or French remind us of the complexity of establishing acceptable pronoun use within a research article. Zinsser (2006: 20-21) remarks on the professorial use of “one”, the tradition of English teachers proposing the “literary we” or the impersonal “it is”, but also notes that when permitted, the use of “I” allows a more natural and engaging text. In the case of multiple authors, Glasman-Deal (2010: 11-149) encourages writers to employ “we”, often nuanced by modal verbs or *dummy subjects*, such as “this article describes”; however, when referring to people or humanity in general, the passive voice or constructions with “it is” are preferable. Henderson (2013: 147-150) compares the common advice in English to favor “an appropriate noun” instead of the ambiguous “editorial we” to the common advice in French to rely upon the *nous de modestie* [*we of modesty*].
- 9 In order to examine in this respect the actual linguistic use, technically advanced corpus linguistics, complete with specific software applications and statistical tools, now allows researchers to examine large quantities of texts in finer granular detail

(Biber et al., 1998; Sinclair et al., 2004). Our study draws upon these techniques to compare the verbal qualities linked to pronominal use within the three main sections of our bilingual corpus "Scientext".

10 This paper examines authorial presence in scientific RA as it is embodied through the use of personal pronouns. All of the verbs syntactically linked to these pronouns were identified and analyzed through manual and automatic means in order to obtain both quantitative and qualitative data, focusing on frequency as a central criterion. The diversity of verbs and their forms are investigated in order to determine the specific role of the authors as explicitly mentioned in the RA. We expected there to be rhetorical structures inherent to the scientific RA that transcend linguistic differences of the two languages. However, we also searched for differences of verb tense and pronoun use, especially concerning the pronouns "one"/*on*, as they have contrasting etymologies and evolving registers in each language (formal for English vs. often less-formal in French). We hypothesize that these pronouns plus specific verb constructions can be categorized by author roles (*researcher*, *writer*, *arguer*, and *evaluator*) in both English and French, and that physical activities are more likely to be mentioned than mental research activities.

11 We begin by describing the methodology and the corpus, which provides a comparative overview of both languages (section 3). Then, we describe the most frequent verbs and their behavior across the three main sections of the RA (section 4). The underlying discourse qualities of these "pronoun + verb" patterns are discussed within the framework of author roles: *researcher*, *writer*, *arguer*, and *evaluator* (section 5).

3. Methodology and corpus

12 We queried the online corpus "Scientext", which includes published and unpublished research texts in both French and English (Tutin et al., 2013; Tutin, 2010; Falaise et al., 2011; Tutin & Grossmann, 2013). The English "Scientext" subcorpus is considerably larger than the equivalent French subcorpus of 90 RA, which includes part of the KIAP corpus as well as other articles in biology. We thus limited the English subcorpus to 90 RA, thereby obtaining two comparable data sources of articles published in the observational fields of biology and medicine. For the English subcorpus, only articles with at least one author working in an anglophone university were retained in order to increase the probability of analyzing standard English.

13 Data were analyzed manually and automatically by using the textometry tools available on the TXM platform (TXM, see <http://textometrie.ens-lyon.fr/?lang=en>). First, we extracted personal pronouns, their syntactically linked verbs and approximately ten words to the right and to the left of each pair. After selecting occurrences actually referring to authors, we examined data related to frequency and distribution both of pronouns and verbs throughout the RA. Using automatic extraction, we queried our corpora to collect a total of 784 English and 747 French occurrences of the constructions "I, one, we/*je, on, nous* + verb".

14 The complexity of the pronouns “one”/*on* and “we”/*nous* leads to a range of possible interpretations including combinations of the authors, readers, discourse community, people in general and textual actors other than the authors (Fløttum et al., 2007). Hence, we manually checked all the occurrences collected in order to focus on those referring uniquely to the authors themselves or to a group of people that included the authors. Here, we included expressions such as “one may assess the probability” that refer to choices of methodology, but did not include statements referring to actions by participants of a given study as in “a traumatic event is only adaptive when one can have control over the traumatic event”. Although it could be argued that the authors may also be victims of traumatic events, in this case “one” refers more generally to participants of a study rather than the research community conducting a study as evoked in the previous example. There were no cases of “I” or the equivalent French *je*, except in the acknowledgements sections, which are beyond the scope of this paper. We retained all of the occurrences of “we”, 11 of the 27 occurrences of “one” in English, 113 of the 378 occurrences of the French *on* and 362 of the 369 occurrences of the French *nous*, which resulted in a collection of 768 items in English and 475 items in French (cf. Table 1 below for their distribution).

15 Then, each phrase was automatically queried to determine the lexical verb following the pronoun. Data were sorted manually to verify that the identified verbs were syntactically linked to the pronoun and to identify modal or auxiliary verbs or expressions. The English “to be” and “to have” and the French *être* [to be] and *avoir* [to have] may function as auxiliaries or as lexical verbs. For this study, only their occurrences as lexical verbs were included in the data. For example, in the expression “one would be able to establish”, the modal verb “would”, the expression “able to” and the lexical verb “to establish” are included in the data, but not the verb “be”. Also, the expression of ability via the modal verb “can” is possible in certain cases, but for others the expression “able/unable to” must be used. For example, “we were able to assess” carries the notion of ability and not the notion of possibility found in “we could assess”. The eight cases of the verb “to be” in the expression “able/unable/not able” were not included in the data. However, the syntactically linked lexical verb was counted in the data (“to establish” in the above example) and the expression containing the expression “able to” was noted for discussion. We adopted a similar method for the 475 occurrences collected from French corpus: auxiliary verbs *être* or *avoir* were not counted in the data. The modal verb *pouvoir*, which is similar to the English “can”, was noted only for the discussion on modality. Thus, the criteria for selecting the English and French verbs in this study were their syntactic link to a pronoun referring to the authors and their lexical status.

16 While the total amounts of words for the two corpora were nearly identical (339,610 for English and 339,902 for French), the volume of words and the frequency of pronoun use vary by general section of the RA: Introduction (introduction and background), Body (methods, materials, results and any other mid-paper section)

and Conclusion (discussion and conclusion) (cf. Table 1). The English RA had considerably shorter introductions (9,376 words) than in French (24,725 words) as well as shorter conclusions (12,468 and 17,439 words respectively).

17 There was a higher density of authorial pronouns in English RA (22.6 items per 10,000 words) than in French (14 items per 10,000 words), except for the introduction sections (18.1 and 20.2 items per 10,000 words, respectively). This density in English is significantly higher than in French ($\chi^2 = 69.45$, $df = 1$, $p < 0.001$). The highest concentration of personal pronoun use was found in the English conclusions (28.9 items per 10,000 words). However, the overall frequency of authorial self-reference through pronoun use was low in both languages (0.0183 percent of the total words). We note that there was no significant difference in pronoun use across the RA sections independent of language ($\chi^2 = 4.29$, $df = 2$, $p > 0.05$), which is mirrored in the similar total frequencies ranging between 17.9 per 10,000 for the body sections and 22.1 per 10,000 for the conclusions.

	Introduction	Body	Conclusion	Total
English	9,376 words 17 "we"/"one" 18.1 per 10,000	317,766 words 715 "we"/"one" 22.5 per 10,000	12,468 words 36 "we"/"one" 28.9 per 10,000	339,610 words 768 "we"/"one" 22.6 per 10,000
French	24,725 words 50 <i>nous/on</i> 20.2 per 10,000	297,738 words 392 <i>nous/on</i> 13.2 per 10,000	17,439 words 33 <i>nous/on</i> 19 per 10,000	339,902 words 475 <i>nous/on</i> 14 per 10,000
Total	34,101 words 19.6 per 10,000	615,504 words 17.9 per 10,000	29,907 words 22.1 per 10,000	679,512 words 18.3 per 10,000

Table 1. Frequency of self-reference pronouns by language and RA section

18 Frequency of verbs in either language was the first factor that conditioned our analysis, although all verbs syntactically linked to a specific pronoun were included in the data. We therefore begin by describing the nature of the most frequently occurring verbs and their patterns. We hypothesize that self-reference, including *engagement*, through personal pronoun use reveals different roles across the three main sections of the RA and that some overlap will be found in English and French. An analysis of the semantic quality of these more frequent verbs offers an insight into the scientific activities associated with pronoun use. We begin with a global overview of the verbs found in each language, before focusing on the specific activities that pertain to the author roles of *researcher*, *writer*, *arguer*, and *evaluator*.

4. Overview of verb frequency and behavior

19 In this section, we give an overview of the features of authorial presence created by personal pronouns, notably the most frequent verbs found in this context and the roles they reveal, pronoun reference, lexical verb range, verb tenses and finally modality.

4.1. Most frequent verbs syntactically linked to a pronoun referring to the authors

20 We first computed the most frequent lexical verbs syntactically linked to a pronoun referring to the authors as shown in Table 2. Frequency (given in parentheses) was the main criterion of verb analysis, with the objective of determining semantic trends and the corresponding author role. The verbs “to use”, “to show” and “to identify” were found to be among the ten most common verbs in medical and biology abstracts irrespective of their grammatical subject (Hartwell, 2013). Here, we can observe several semantic similarities among the ten most frequent verbs in both languages. These semantic trends are representative of the steps of scientific research beyond the domains of medicine and biology.

	Frequency rank	Most frequent verbs
English (768)	1	“to find” (55)
	2	“to use” (50)
	3	“to show” (29)
	4	“to identify” (27)
	5	“to demonstrate” (21)
	6	“to describe” (20)
	7	“to test” (19)
	8	“to examine” (17)
	9	“to investigate” (16)
	10	“to observe” (14)
French (475)	1	<i>observer</i> (57) [to observe]
	2	<i>constater</i> (23) [to notice]
	3	<i>retrouver</i> (17) [to find]
	4	<i>comparer</i> (13) [to compare]
	5	<i>analyser</i> (11) [to analyze]
	5	<i>étudier</i> (11) [to study]
	5	<i>noter</i> (11) [to note]
	8	<i>penser</i> (10) [to think]
	8	<i>proposer</i> (10) [to propose]
	8	<i>rapporter</i> (10) [to report]
8	<i>voir</i> (10) [to see]	

Table 2. Most frequent verb lemmas syntactically linked to a self-reference pronoun by language

21 Indeed, the majority of these frequently occurring verbs in both languages are specifically related to observational activities undertaken by the authors in their role as a *researcher*. Hence, in both languages, authors personalize, through pronoun use, their participation in objective observational research acts, such as “to find” (55), “to observe” (14), *observer* (57) [to observe], *constater* (23) [to notice]. Although the French verb *constater* and the English “to find” are generally not considered

equivalent, one may also draw connections between these two verbs within a scientific context, as both can evoke a hybrid of mental and material actions, e.g.: "We found that cells can interact with a portion of the fibronectin first type III repeat".

22 The third verb of the French list, *retrouver* [to find] was generally employed with the French pronoun *on* (10 tokens out of 17) to confirm that something that was expected actually holds true, e.g., *on retrouve bien le processus décrit dans le modèle* [as expected one does find the process described in the model]. This French pattern in the present tense is closely mirrored by two English constructions built on an inclusive pronoun "one" and an epistemic modal auxiliary verb. The first case is with the verb "to find" (1) in the conditional ("when analyzing complexes in a given network [...] one would find"). The second case, employing the less-frequent verb "expect" (8), is more recurrent and includes a range of modal verbs ("might" (1), "should" (1), "would" (4)). In all of these cases, the authors are bringing the reader into the research process. In contrast, clearly methodological research acts evoked include the English verbs "to use" (50), "to test" (19), "to examine" (17). The research action "to investigate" (16) may be compared to the French *analyser* (11) [to analyze] or *étudier* (11) [to study].

23 In contrast to the present tense, which directly involves the reader as if he or she was a direct witness of the research, we found several patterns that build upon a first person plural pronoun in an expression referring to past events:

1. "We also found that" (4)
2. "We have previously shown" (3)
3. "We have previously identified" (3)
4. "We also could have used" (3)
5. *Nous avons analysé rétrospectivement* (3) [we analyzed retrospectively]
6. *Nous avons comparé les résultats* (3) [we compared the results]

24 The English constructions "we have previously shown/identified" are a means for the authors to reintroduce their continuing research on a given subject. The two French patterns and the English "we also could have used" refer to methodological issues, while "we also found that" allows the authors to highlight supplementary results.

25 Besides the *researcher* role, verbs related to the explanatory roles of *writer* and *arguer* are also present, to a lesser degree, in both languages. We find, among the list in Table 2, several verbs related to the descriptive qualities of a *writer*: "to show" (29), "to demonstrate" (21), "to describe" (20), *noter* (11) [to note], and *rapporter* (10) [to report]. Although semantically *voir* (10) [to see] might be associated with the act of research observation, it is in fact essentially devoted to the management of the article itself with the inclusive formula *nous avons vu* [we have seen], which serves to remind the reader of a previous statement before furthering the argumentation. Finally, only French offers verbs related to the role of *arguer*, i.e., *proposer* (10) [to propose/suggest] and *penser* (10) [to think] among the top ten most frequent verbs (cf. Table 2).

4.2. Pronoun reference

- 26 In English, the pronoun “we” was predominately related to the researchers’ acts or to notions of *stance*. That is to say, it described actions or arguments of the authors and not a positioning of the reader through the use of an all-inclusive pronoun. In six cases of referring to previous studies, the authors broadened the subject to “we and others”, which was syntactically linked to the present perfect (“have shown” (3), “have found” (2), “have observed” (1)). In contrast, the English “one” (11) was used all-inclusively, as in “if one has defined”. The French *on* [one] is occasionally employed to refer exclusively to the authors: *on relevait un syndrome inflammatoire* [one noted an inflammatory syndrome]. The significantly greater ratio of the French use of the often ambiguous *on* (113) instead of *nous* (362) compared to the English ratio of “one”/“we” (11 and 757 respectively) is one of the main linguistic differences between the two corpora ($\chi^2 = 163.35$, $df = 1$, $p < 0.001$).

4.3. Lexical range

- 27 As can be seen in Table 3, the four most recurrent verbs in both languages constituted about one-fifth of the verbs examined. The number of different verbs remains greater in English (182) than in French (155), however there is no statistical difference between the two languages related to overall frequency ($\chi^2 = 1.39$, $df = 2$, $p > 0.05$). The English lexical verbs have a mean of 4.2 occurrences, while in French the mean drops to 3.1 occurrences for an overall expected frequency of 3.7 occurrences per verb type, which clearly contrasts with the most frequent verbs (*mean* = 33.9 occurrences per different verb type). The significant difference of distribution ranging from most frequent verbs (271 tokens) to only 1 to 2 tokens per verb (250 tokens) ($\chi^2 = 2638.61$, $df = 2$, $p < 0.001$), supports the notion of stabilized pronoun and specific verb pairs irrespective of the language. These findings mirror the results of Carter-Thomas and Chambers (2012), who find that 11 verbs account for more than one-third (42.5% out of 318 tokens) of all the English verbs linked to a first person subject pronoun, while 12 verbs account for more than half of the French verbs (51.8% out of 226 tokens) linked to this type of pronoun in economics RA introductions.

	4 most frequent verbs	Middle-frequency verbs	1 to 2 tokens
English (768)	4 different verbs 161 tokens 21%	65 different verbs 469 tokens 61%	113 different verbs 138 tokens 18%
French (475)	4 different verbs 110 tokens 23.5%	46 different verbs 253 tokens 53.3%	105 different verbs 112 tokens 23.6%

Table 3. Percentages of verb frequencies by language

4.4. Verb tense and aspect

28 Finally, an examination of tense, aspect, and mood of the verbs shows that, while authors of English relied essentially on three tenses (present (36.1%), simple past (42.8%), and present perfect (18%)), the present (40%) and the *passé composé* (49.5%) dominated in French. The English present perfect can be translated by diverse means into French, notably the present and the *passé composé*. There was a higher frequency of other verb tenses in French, notably the imperfect tense (3.5%), which again can be translated into a range of tenses in English, including the present simple, the past continuous, or the conditional.

29 As can be seen in Figure 1, the main difference of tense across sections and languages is the presence of the present perfect in English, with slight differences for the other tenses. In the introductions (Intro), where authors present the major premises, the present tense is more frequent in English (47%) than in French (38%), for which there is a greater presence of the *passé composé* (44%). In the central sections (Body), in which the authors describe their research steps, the percentage of the French *passé composé* increases to 60.3%, and the English simple past predominates at 43.6%, but more notable is the presence of the English present perfect (17.3%). However, in the conclusions, the English present perfect reaches 38.9%, as in: “we have (now) identified” (4) or “we have described” (2), as authors restate the main trust of their article. In both languages, but notably in French, the use of the present tense within the conclusions reminds readers of the innovative or breaking edge nature of the findings: “This is, as far as we know, the largest ever test of its predictions”.

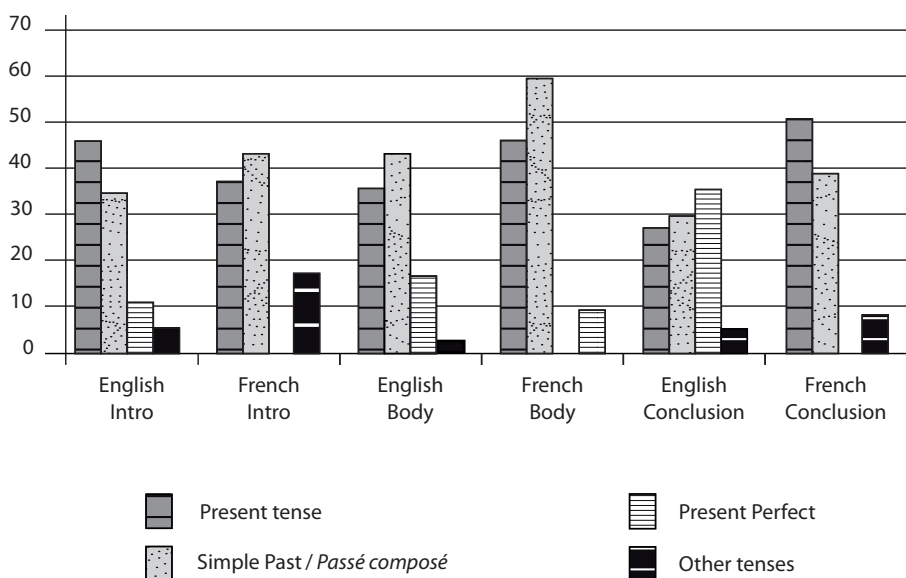


Figure 1. Verb tenses by language and RA section

30 Our findings concerning “to find” (35 tokens of a past tense (63.8%), 10 tokens of the present perfect (18%) and 10 tokens of the present tense (18.2%)) contrast slightly with a greater percentage of the past tense within the KIAP corpus, in which the distribution of tenses for “find” (54) in the medical RA subcorpus was 47 tokens of the past tense (87%), 5 tokens of the present perfect (9.3%) and 2 of the present tense (3.7%) (Fløttum et al., 2006b). This may be due to the presence of biology RA in our corpus.

4.5. Modality

31 A positive or negative form of the present “can” (21) was the most frequent modal verb in the English corpus, but the most frequent modal-verb string was: “we would expect” (6). Similar to the English modal verb “can” or “may” (1), a notion of ability or potentiality is expressed through the use of the French modal verb *pouvoir* (15 tokens in all, 10 with *nous* and 5 with *on*). However, authors also relied upon the marker of modality “able to/unable to” (5), including: “we may not have been able to identify any other members”. Although the modality of obligation “must” (2) (“we must await”) was found, authors also conveyed this through the use of “have to” (6). Interestingly, only one case is paired with “we”, while there are five occurrences with “one”, such as: “one has to analyze”, “one does not have to re-compute”, and “one has to screen”. These three segments broaden the subject to the reader and the scientific community, who are thereby encouraged to adopt the research methodology. The almost exclusive use of the subjunctive with the modal *pouvoir* (5 out of 7 occurrences) is also striking as this association reinforces the modality each one bears, creating a combination of possibility and uncertainty.

5. Discussion of “pronoun + verb” by author role

32 In accordance with community practices, authors assume a range of roles. Fløttum et al. (2006a) have defined four such categories of which the boundaries retain a certain degree of subjectivity: *researcher*, *writer*, *arguer*, and *evaluator*. According to this school of thought, the process of researching may be embodied in verbs such as “to analyze”, “to assume”, “to consider”, “to examine”, “to find”, and “to study”. Actions related to the verbal or graphic explanatory role of writing range from “to describe”, “to illustrate”, “to present”, “to begin by” to “to focus on”. Third, taking a stand can be manifested through the verbs “to argue”, “to claim”, “to dispute” and “to reject”. Finally, although less frequent in the KIAP corpus, the role of *evaluator* conveys evaluative or rare emotional reactions through, for example, “to feel” or “to be skeptical”. We examine here our corpus in light of these four categories with specific attention to the different RA sections.

5.1. Researcher role

33 In the hard sciences, the role of researcher can combine both material and cognitive processes, of which observation might be considered a hybrid of both aspects.

Apart from one exception, the English "use" (50) embodies the material aspects of scientific research, as in: "we used acetylene because of the higher explosive potential". In French, the semantically equivalent verb linked to a pronoun, *utiliser* (8) is considerably less frequent. The procedural qualities of "use" are highlighted by its frequent pairing with the modal verb "can" that express capacity: "we can use" (2), "we could also have used" (4), and "we could only/therefore use" (2). The material aspects related to the verb "test" (19) also invite modality: "we were (not) able to test" (3), "we could not test" (1). However, it is also employed in the mental process of reflection as in the coined expression: "we have tested the hypothesis" (1). In comparison to "test", "to investigate" (16) often introduces the main trust of the study: "In this study, we investigate the effects of dipole depth and orientation on source localization". The link between the two verbs can be noted in: "To test this, we investigate the relationship between airway inflammation".

34 A range of verbs encompass the fundamental notion of research acts. Verbs related to research methodology are found in this category, notably:

- "to identify" (27)/*identifier* (3);
- "to develop" (8)/*établir* (6);
- "to compare" (7)/*comparer* (13).

35 These verbs are rarely in a present tense. For example, "to compare" is only found in the simple past form; 7 out of the 8 occurrences of the lemma "to develop" are in the present perfect ("we have developed a clustering algorithm"); and "to identify", found only once in the present tense, is repeated in the expression: "we were unable to identify" (3). Finally, the researchers glean results through observational acts, such as:

- "to find" (55)/*retrouver* (17);
- "to observe" (14)/*observer* (57).

36 The processes accomplished before the actual writing are rarely found in the present tense:

- "to examine" (17)/*analyser* (11);
- "to study" (8)/*étudier* (11).

37 In English, members of the research community are credited for their contribution by the expression: "we and others have shown" (3). This expression reinforces the hypothesis that "we" tends to be exclusive in scientific English (cf. Fløttum et al., 2006a). The French *observer* was used in both the *passé composé* and present, but was combined with distinct pronouns depending on the tense: *nous avons observé* [we observed], but *on observe* [one observes]. The French *observer* is linked in 41 out of 57 occurrences to the pronoun *on*, thereby including the reader in the process and making the observation more general, impersonal, and less challengeable. Semantically close to "to observe", *voir* (10) [to see] is often used to guide the reader through the reading process as in: *nous avons vu que les inflexions observées* [we have seen that the observed inflections]; thus it falls into the category of *writer* role.

5.2. Writer role

38 The majority of “pronoun + verb” patterns related to the role of writer are devoted to guiding the reader through the argumentative progression of the RA. One of the central acts is informing the reader through a variety of tasks:

- “to describe” (20)/*décrire* (3);
- “to report” (12)/*rapporter* (10);
- “to show” (29)/*montrer* (6);
- “to demonstrate” (21)/*démontrer* (2).

39 Second, especially in French, we find verbs related to the act of “noting”:

- “to notice” (2)/*constater* (23);
- “to note” (2)/*noter* (11).

40 However, *constater* (23) [to notice] occurs 19 times with the pronoun *on*, either for generally known information as in: *on constate depuis de nombreuses années* [one has noticed for several years]; or in statements specifically pertaining to the study: *on constate aussi une mydriase transitoire de l’œil gauche* [one also notices a temporary mydriasis of the left eye]. *Voir* (10) serves also to remind the reader of what has already been discussed (*nous avons vu* (7) [we have seen]) or to instruct the reader to refer to graphics. The English “to see” (7) is essential for this latter writing technique, as in: “Looking at Appendix 2, we see unadjusted sample sizes”.

5.3. Arguer

41 While the writer role deals with describing the factual information, the arguer defends a position. The French verbs related to taking a stance, *proposer* (10) [to propose] and *penser* (10) [to think], are both found in the top ten most frequent verbs. Direct English verbal equivalents, such as “to propose” (3) and “to think” (0), are rare. We found French verbs linked with modality as in: *on peut penser* (3) [literally, “one can think”]. This use of the pronoun *on* positions the readers, engaging them in the process of reflection. This pattern is semantically mirrored by the English expression “we (do not) know” (10), which can also be considered to engage the reader. However, the English “know” is employed without calling upon a form of modality.

42 In English, authors are also likely “to conclude” (10), “to suggest” (8) or “to speculate” (4). Irrespective of the RA section, these verbs are found primarily in the present tense. An exception to this is a simple past tense of “to conclude”, which was found throughout the RA sections, as in this example from an introduction: “we concluded that bacteria engulfed by macrophages can reach the joint”.

5.4. Evaluator

43 Finally, authors also employ verbs to evaluate a concept, thereby revealing the subjective nature of the RA. An example of this subjective perspective is found in the French confirmation: *une intuition que nous avons souhaité confirmer par une*

recherche épidémiologique simple [an intuition that we wished to confirm with a simple epidemiological study]. In English, both “to believe” (6) and “to feel” (5) can be used to take an unambiguous stance, e.g., “we believe it is time to have a more discriminatory approach” or “we feel that this novel pharmacological approach has revealed a clear dissociation”.

- 44 Authors also take a strong stance on the future of their research, employing a form of “to wish” (4): “we wish to release the information to the public domain”; or “to hope” (1) evoking future studies: “we now hope to extend this approach”. However, other expressions containing the verbs “to believe” or “to feel” did not convey a strong opinion (“we believe this is the first reported case”, “we feel this may be a particular problem”). Finally, the reader is engaged in this subjective evaluation through the French *on*: *si l’on veut comprendre* (2) [if one wants to understand].

6. Concluding remarks

- 45 Personal pronoun use in English contrasts with French discourse in both type and frequency. Proportionally, the French *on* was frequent (23.8%), while the English “we” was vastly more present than the pronoun “one” (3.2%). Furthermore, although the English texts contained relatively more self-reference through pronoun use, both languages showed a greater proportion within introductions and conclusions compared to the central sections. Both languages also contained expressions built with “able to”, “can”, “may” or *pouvoir* [can] that firmly link conclusions to the given data.

- 46 Finally, the English pronoun “we” was related here to notions of *stance* or actions of the authors, while the pronoun “one” engages the reader through the use of an all-inclusive pronoun. The multiple facets of *on* (cf. Reutner, 2010) are exemplified in the extensive entry in the *Trésor de la langue française* (see <http://atilf.atilf.fr/>), ranging from contexts in which the indefinite third personal pronoun *on* embodies the first person singular or plural as well as the second person singular or plural, to other contexts where *on* refers to no particular person. Further studies might categorize the contexts of *on* in the scientific French RA, as the data confirmed their presence.

- 47 This study found little direct correspondence between the four most frequent verbs in English and French, although these verbs constituted approximately one-fifth of all verbs in both languages. Many of the higher frequency verbs were related to the author role of researcher and particularly linked to essential observation (“to find”, “to observe”/*observer*), but also to explanatory functions of the writer (“to show”, “to describe”, *noter*, *constater*), as well as activities related to research, in which the researcher reveals the existence of data, but does not modify them. Other verbs relate more specifically to actions taken by the researcher: “to identify”, “to investigate”, *comparer*, *analyser*, and *étudier*. While these lexical choices may appear objective, they also embody research choices that reveal the authors’ commitment as driving forces of the research, which inherently implies a subjective quality.

48 In English, authors relied more frequently upon a “pronoun + verb” pattern. The most frequent English verbs appear to cover a wider range of material research acts, including “to use” and “to test”, often in the simple past or present perfect. In contrast, the most common French verbs include certain evaluative acts related to results, such as *penser* (10) [to think] and *proposer* (8) [to propose]. There was an equal number of English equivalents to these two verbs (“to conclude” (10), “to suggest” (8)). However, the greater overall number of items in English make these verbs proportionally less frequent. In both languages, these evaluative acts were conjugated more frequently in the present tense, thereby emphasizing the up-to-date relevancy of the research.

49 This research is limited to the discourse of the published RA. In contrast, in their comparison of RA and conference presentations, Carter-Thomas and Rowley-Jolivet (2001) found that RA contained frequent instances of extraposition in both the active (“we feel it is these difference that”) and passive voice (“it has been found that”). However, within the oral presentation of the same subject, they found that the speaker supplied the same information using “we” (“we’ve demonstrated”, “we think that probably”). Citing supporting information from the *Longman Grammar of Spoken and Written English*, they suggest that instead of hedging devices, these lexical bundles are stable collocations of academic prose. Hence, further studies are needed to compare the current findings to other contexts, such as oral presentations.

50 The orientation of research in the natural sciences is largely conditioned on public or private financing. Gross and Chesley (2012) suggest that industrial financing in biomedical research impacts the language of RA, notably that hedging decreases in correlation with increased financing. They suggest that authors with substantial funding employ more persuasive language than those without high-stakes funding. Hence, further studies might examine how authorial presence manifests itself in high-stakes research that is subject to significant financing and profits or to public debate. This last question is highlighted by Fløttum’s (2010) analysis of the Intergovernmental Panel on Climate Change’s 2007 “Summary for policymakers” written by a 40-members panel in which there are no personal pronouns referring to the authors.

51 We have sought to demonstrate empirically, through a corpus study, the differences and similarities of authorial presence through pronoun use in French and English RA in medicine and biology. Beyond surface differences of personal pronoun choice (e.g., “we” vs. “one”), stance vs. engagement, material vs. mental acts as played out in author roles, this study underscores the interest of corpus studies to evaluate discursive functions and the subjective qualities of scientific research.

Acknowledgements

52 This study was made possible by a grant from the Grenoble 1 Research “Pôle Sciences Humaines et Sociales”. We are also grateful to Agnès Tutin for her constructive comments and Didier Piau for his statistical formulae.

References

- BIBER, D., CONRAD, S. & REPPEN, R. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- BIBER, D. & GRAY, B. 2010. Challenging Stereotypes about Academic Writing: Complexity, Elaboration, Explicitness. *Journal of English for Academic Purposes* 9: 2-20.
- CARTER-THOMAS, S. & CHAMBERS, A. 2012. From Text to Corpus: A Contrastive Analysis of First Person Pronouns in Economics Article Introductions in English and French. In A. BOULTON, S. CARTER-THOMAS & E. ROWLEY-JOLIVET (eds.), *Corpus-Informed Research and Learning in ESP: Issues and Applications*. Amsterdam – Philadelphia: J. Benjamins: 17-44.
- CARTER-THOMAS, S. & ROWLEY-JOLIVET, E. 2001. Syntactic Differences in Oral and Written Scientific Discourse: The Role of Information Structure. *ASp* 31-33: 19-37. Available online: <http://asp.revues.org/1752>.
- FALAISE, A., TUTIN, A. & KRAIF, O. 2011. Exploitation d'un corpus arboré pour non spécialistes par des requêtes guidées et des requêtes sémantiques. In *Actes de la 18^e conférence sur le Traitement automatique des langues naturelles (TALN 2011) (Montpellier, 27 juin-1^{er} juillet 2011)*. Available online: <http://pro.aiakide.net/publis/2011TALNPaper-Falaise-Tutin-Kraif.pdf>.
- FLØTTUM, K. 2010. A Linguistic and Discursive View on Climate Change Discourse. *ASp* 58: 19-37. Available online: <http://asp.revues.org/1793>.
- FLØTTUM, K., DAHL, T. & KINN, T. 2006a. *Academic Voices: Across Languages and Disciplines*. Amsterdam – Philadelphia: J. Benjamins.
- FLØTTUM, K. et al. 2007. Cultural Identities and Academic Voices. In K. FLØTTUM (ed.), *Language and Discipline Perspectives on Academic Discourse*. Newcastle upon Tyne: Cambridge Scholars Publishing: 14-39.
- FLØTTUM, K. et al. 2013. KIAP – Reflections on a Complex Corpus. *Bergen Language and Linguistics Studies* 3 (1): 137-150. Available online: <https://bells.uib.no/bells/article/download/367/381>.
- FLØTTUM, K., KINN, T. & DAHL, T. 2006b. "We Now Report on..." Versus "Let Us Now See How...": Author Roles and Interaction with Readers in Research Articles. In K. HYLAND & M. BONDI PAGANELLI (eds.), *Academic Discourse across Disciplines*. Berlin – New York: P. Lang: 203-224.
- GJESDAL, A.M. 2013. The Influence of Genre Constraints on Author Representation in Medical Research Articles. The French Indefinite Pronoun *On* in IMRAD Research Articles. *Discours* 12. Available online: <http://discours.revues.org/8770>.
- GLASMAN-DEAL, H. 2010. *Science Research Writing: For Non-Native Speakers of English*. Singapore – London: World Scientific.
- GROSS, A.G. & CHESLEY, P. 2012. Hedging, Stance and Voice in Medical Research Articles. In K. HYLAND & C. SANCHO-GUINDA (eds.), *Stance and Voice in Written Academic Genres*. Basingstoke: Palgrave Macmillan: 85-100.
- GROSSMANN, F. 2013. Les verbes de constat dans l'écrit scientifique. In A. TUTIN & F. GROSSMANN (eds.), *L'écrit scientifique: du lexique au discours*. Rennes: Presses universitaires de Rennes: 85-100.

- HARTWELL, L.M. 2013. Corpus-Informed Descriptions: English Verbs and Their Collocates in Science Abstracts. *Études en didactique des langues* 20: 79-94.
- HARWOOD, N. 2005. 'Nowhere Has Anyone Attempted... In This Article I Aim to Do Just That'. A Corpus-Based Study of Self-promotional *I* and *We* in Academic Writing across Four Disciplines. *Journal of Pragmatics* 37 (8): 1207-1231.
- HENDERSON, A. 2013. Le positionnement à travers la mention de l'auteur dans les écrits universitaires anglais d'étudiants français de premier cycle: implications pédagogiques. In A. TUTIN & F. GROSSMANN (eds.), *L'écrit scientifique: du lexique au discours*. Rennes: Presses universitaires de Rennes: 145-162.
- HUNSTON, S. & FRANCIS, G. 2000. *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English*. Amsterdam – Philadelphia: J. Benjamins.
- HYLAND, K. 2001. Humble Servants of the Discipline? Self-Mention in Research Articles. *English for Specific Purposes* 20 (3): 207-226.
- HYLAND, K. 2004. Patterns of Engagement: Dialogic Features and L2 Undergraduate Writing. In L.J. RAVELLI & R.A. ELLIS (eds.), *Analysing Academic Writing: Contextualized Frameworks*. London – New York: Continuum: 5-23.
- HYLAND, K. 2006. Disciplinary Differences: Language Variation in Academic Discourses. In K. HYLAND & M. BONDI PAGANELLI (eds.), *Academic Discourse across Disciplines*. Berlin – New York: P. Lang: 17-45.
- HYLAND, K. 2012. *Disciplinary Identities: Individuality and Community in Academic Discourse*. Cambridge: Cambridge University Press.
- REUTNER, U. 2010. *De nobis ipsis silemus?* Les marques de personne dans l'article scientifique. *Lidil* 41: 79-102. Available online: <http://lidil.revues.org/3013>.
- RINCK, F., BOCH, F. & GROSSMANN, F. 2007. Quelques lieux de variation du positionnement énonciatif dans l'article de recherche. In P. LAMBERT et al. (eds.), *Variations au cœur et aux marges de la sociolinguistique*. Paris: L'Harmattan: 285-296.
- SANCHO-GUINDA, C. & HYLAND, K. 2012. Introduction: A Context-Sensitive Approach to Stance and Voice. In K. HYLAND & C. SANCHO-GUINDA (eds.), *Stance and Voice in Written Academic Genres*. Basingstoke: Palgrave Macmillan: 1-11.
- SCIENTEXT. A French and English Corpus of Scientific Texts. Available online: <http://scientext.msh-alpes.fr/scientext-site-en/spip.php?article9>.
- SINCLAIR, J.M., JONES, S. & DALEY, R. 2004. *English Collocation Studies: The OSTI Report*. London – New York: Continuum.
- SWALES, J.M. 2004. *Genre Analysis: English in Academic and Research Settings* [1990]. Cambridge: Cambridge University Press.
- SWALES, J.M. & FEAK, C.B. 2004. *Academic Writing for Graduate Students: Essential Tasks and Skills*. Ann Arbor: University of Michigan Press [2nd edition].
- TRÉSOR DE LA LANGUE FRANÇAISE INFORMATISÉ 2004. Paris: CNRS Éditions. Available online: <http://atilf.atilf.fr/>.
- TUTIN, A. 2010. *Sens et combinatoire lexicale: de la langue au discours*. Unpublished report. Grenoble: Université Stendhal.

- TUTIN, A. et al. 2013. Autour du projet Scientext: étude des marques linguistiques du positionnement de l'auteur dans les écrits scientifiques. *Texte et corpus* 4: 333-349. Available online: http://www.licorn-ubs.com/jlc6/ACTES/Tutin_etal_JLCo9.pdf.
- TUTIN, A. & GROSSMANN, F. (eds.) 2013. *L'écrit scientifique: du lexique au discours*. Rennes: Presses universitaires de Rennes.
- TXM PLATFORM. Available at: <http://textometrie.ens-lyon.fr/?lang=en>.
- VOLD, E.T. 2006. The Choice and Use of Epistemic Modality Markers in Linguistics and Medical Research Articles. In K. HYLAND & M. BONDI PAGANELLI (eds.), *Academic Discourse across Disciplines*. Berlin – New York: P. Lang: 225-249.
- ZINSSER, W.K. 2006. *On Writing Well: The Classic Guide to Writing Nonfiction* [1976]. New York: HarperCollins.