

How Personal is this Text? Researching Writer and Reader Presence in Student Writing Using Wordsmith Tools

Ruth Breeze
Universidad de Navarra, Spain

Abstract. The issue of direct writer and reader presence is an important aspect of the interpersonal dimension of the text, which has been extensively studied in the context of published academic writing. To date, only a few studies have focused on relational presence in learners' writing. Explicit author-reader relations are easily quantifiable, as corpus tools enable researchers to identify the number of times a writer speaks to readers as "I" or "we" to "you", or uses first or second-person possessive adjectives. The present paper focuses on the interpersonal dimension of a small sample of essays and reports written by a group of students at a Spanish university. The pattern which emerged in the essays diverged widely from academic prose style, particularly in the overuse of "I" and "you". However, the high frequency of "we" and "our" in the reports, with low incidence of singular pronouns, indicates that students were aware of the need to modulate their style according to target genre. It is speculated that misunderstanding of the nature of the essay task might underlie some of the inappropriate choices made. Conclusions are drawn as to the need to sensitise students to genre and register in academic and professional texts.

Keywords. Writer presence, personal pronouns, possessives, word count, L2 writing.

1. Introduction

Since the advent of computer programs designed for linguistic research, many quantitative studies have been carried out which provide insights into the nature of different written genres. Although most of this research has centred on published texts, projects like the International Corpus of Learner English and the Cambridge Learner Corpus have also taken account of student writing, including texts by non-native language users. These studies have concentrated on measurable features which reflect students' general command of English, generic awareness, disciplinary appropriateness, and so on. The issue of direct writer and reader presence is an important aspect of the interpersonal dimension of the text, which has been extensively studied in the context of published academic texts. However, to date only a few studies have focused on writer and reader presence in learners' writing. Explicit author-reader relations are easily quantifiable, as corpus tools enable researchers to identify the number of times a writer speaks to readers as "I" or "we" to "you", or uses first or second-person possessive adjectives. The present paper focuses on the interpersonal dimension of a small sample of essays and reports written by a group of students at a Spanish university. After a brief overview of the theoretical discussions and empirical research into this area, we shall discuss the way these features were researched for the purposes of the present study.

2. Researching student writing: writer and reader presence in texts

In applied linguistics, the area of authorial voice has mainly been studied through the linguistic markers of stance. Hyland (1999) argues for the importance of stance in academic prose, asserting that "controlling the level of personality in a text is central to successfully maintaining interaction with readers and building a convincing argument" (1999: 99). In the writing of practising academics and researchers, the rhetorical choices which writers make in this area are at least partly influenced by the social practices of their discipline. Much

professional writing is a multi-layered entity co-produced by the author(s) and the projected readers; along these lines, some theorists prefer to foreground the discourse community as a whole rather than the writer as an individual. In this context, writers of academic texts project themselves into their text to communicate academic virtues such as intellectual integrity, involvement or commitment, and to offer a judicious evaluation of their peers.

In Hyland's analysis, the writer's expression of a socially defined *persona* has three elements: evidentiality, affect, and relation (1999: 101). Evidentiality is the writer's expression of commitment to the truth of what he or she is setting forth, and entails an evaluation ranging from possibility or probability to categorical certainty. Although in the most obvious sense expressions of evidentiality indicate the writer's stance to what he or she is reporting, it is also interesting to note that many other functions have been attributed to the evidentiality markers in academic writing, including the need to signal solidarity with readers, parading rhetorical respect for others' work, or to manipulate commitment for rhetorical or interpersonal reasons (Hyland 1999: 100-102). The main linguistic resources used to indicate evidentiality are hedges of various kinds (it-cleft constructions with possible or probable/likely, adverbs such as possibly, modals such as might, may and could, and verbs of thinking and believing), and also emphatics (constructions with obvious, definitely, of course). By contrast, affective elements reflect the writer's own or assumed viewpoint. Affect is expressed in verbs (agree, prefer), in modals of necessity (should, have to, must), in some adverbs (hopefully, unfortunately), and in some adjectives which imply a subjective appraisal (appropriate, logical, remarkable). The third aspect of authorial voice concerns the relationship between writer and reader, which is either overtly constructed by the use of personal pronouns and possessives, or achieved by use of indirect rhetorical strategies designed to involve the reader. Relational elements construct the discursive relations with an audience, and are involved in the ways that writers choose to highlight or downplay their own presence and that of their readers. The presence or absence of relational elements in a text contribute to its level of engagement or detachment. Writers can choose to engineer a feeling of complicity, by using expressions such as "as is well known" or "it is natural to suppose that". On the other hand, they can address readers directly, by using rhetorical questions like "where does that leave Canada?". Finally, on the most basic level, writers can raise the interpersonal temperature of a text by using first or second person pronouns; conversely, they can cool the text down by recourse to passives, impersonal it-cleft constructions and grammatical metaphor.

In a comparative study of features belonging to the three categories of evidentiality, affect and relation across eight academic disciplines, Hyland (1999) found considerable variation in the ways in which they were used. For example, relational markers were scarcely found except in philosophy texts, whereas hedges were much more frequent in arts than in the "hard" sciences. He concludes that variations in stance are an important aspect of professional academic discourse, and that authorial choices in this area are often conditioned by the conventional practices of their discourse community.

The texts under analysis in the present study are not fully fledged research papers, nor are their writers participating as yet in a professional or academic discourse community. Yet it is arguable that some of the same criteria can be applied to them, in terms of appropriateness within the genres attempted here. The best student texts, the ones which are mature, sophisticated, cohesive and coherent, are those which seem to speak with a mature authorial voice. In this, the degree of personal presence in the text is a crucial factor. Heavy use of personal pronouns in an academic or professional text creates a style that is perceived as too simple, or too similar to speech. Frequent use of "I", and even more crucially, "you", could often be associated with a text that is too "personal", that is raw or too childish for public presentation.

In fact, personal pronouns and possessives are a very common feature of the spoken language, with "I" and "you" in the top 10 words in corpora of spoken English such as the conversation subcorpus of the Cambridge and Nottingham Corpus of discourse in English (CANCODE) or the Cambridge and Nottingham Corpus of Business English (CANBEC) (O'Keeffe et al. 2007: 207). However, there is evidence that pronouns in general are very sparingly used in academic writing (Biber et al. 1999: 235), and that personal pronouns are even more unusual, as it is generally understood by mature writers of English that their role is strictly limited, other than in narrative, certain types of journalism, and texts of a directive or instructive nature (Biber 2006: 51), is limited. This may be conditioned by the fact that in many professional and academic contexts, the level of explicit personal involvement in the text is tightly controlled, since values of critical objectivity or scientific neutrality take precedence. Although some changes have been noted in recent years, there is evidence to suggest that while use of the first person singular is acceptable in disciplines such as philosophy, such forms are still avoided in "harder" disciplines such as statistics or the sciences (Hyland 1999). On the rare occasions where the first person is used, a complex range of pragmatic factors comes into play, including self-promotional issues and stance towards disciplinary gate-keepers (Harwood 2005). Moreover, it has been noted that L2 writers often tend to adopt a rather conservative stance towards such relaxations of linguistic norms, perhaps because their position of disadvantage leads them to exercise maximum care (Chang and Swales 1999).

Regarding the textual genres involved in the present study, business reports are generally supposed to be written "objectively", and overuse of personal pronouns would be felt to detract from the objective claims made in the text. As for the essay genre, though this is ostensibly an exercise in self-expression, the target style is a discourse type very much within the literate-critical-objective tradition, and as in much academic prose, the aim is to express or persuade, but through the language of "what is or may be" rather than "what I think". Writers are concerned with "producing something that will be consistent and defensible when read by people at different times and in different places" (Chafe 1982: 45). The resulting impersonal style has led some linguists to talk of the low "writer visibility" that is characteristic of English academic genres in general (Petch Tyson 1998). An impersonal style has been shown to play an important part in scientific and scholarly discourse, functioning in parallel with other types of hedging to modulate the strength of claims within the epistemological framework of the discipline concerned, as well as to protect the "face" of the writer (Meyer 1997; Harwood 2005). It is reasonable to suppose that the constraints operating in such academic texts affect a wide range of academic writing, and there is no reason to assume that "apprentice work" such as the undergraduate essay is not bound by the same conventions. Thus we might expect the better essays and reports to show evidence of a controlled interpersonal temperature and concealment of professional practice "behind a screen of empiricist impersonality" (Hyland 1999: 118).

The issue of writer visibility and the need for learners to adjust a predominantly spoken command of English to the requisites of formal writing have been the focus of various pieces of research into L2 writing. Ringbom (1998) found that in the essays by Spanish L2 writers of English in his corpus, "we" occurred at a frequency of 98/10,000 words, while in the corpus of essays by NS it occurred at a frequency of only 34/10,000 words. "Our" had frequencies of 60/10,000 for Spanish students and 25/10,000 for NS. "I" was found at a frequency of 36/10,000 (Spanish students) and 25/10,000 (NS). "You" occurred at a frequency of 33/10,000 (Spanish students) and 8/10,000 (NS). In their greater writer/reader visibility, the essays by Spanish students were similar to those by students from other European language backgrounds. In broad terms, the essays by Spanish students in Ringbom's study used the first

and second person singular less than many European language groups, but used the first person plural more than most.

Other studies have also found much higher incidences of first and second person pronouns in texts by L2 writers of various European language backgrounds than in texts by L1 writers of English (Petch-Tyson 1998). Although Petch-Tyson's study does not include Spanish students, her conclusions are interesting, and merit consideration here. She found that the differences between NS and NNS writers concerning personal pronoun use were qualitative as well as quantitative. So, for example, NS writers used first person singular mainly to recount personal experiences, while NNS (particularly Dutch and French) writers used the first person to express opinions, or to talk about themselves as writers within the context of the text. In addition to this, she surmises that one of the problems inherent in NNS use of personal pronouns may be their inconsistency. Some L1 writers used them to create an informal tone, which appeared to be a deliberate strategy, maintained throughout the text. However, the L2 writers used them sporadically, in a way that jarred with attempts to keep to a more formal style. Mention should also be made here of Hyland's thesis (2002) that some L2 writers, specifically those he encountered in Hong Kong, were liable to underuse personal pronouns, perhaps as a form of overcompensation, or because they were uncomfortable with the notion of personal authority. He suggests that the problem might be that L2 writers lack awareness of the rhetorical ends to which NS writers actually use these features. Despite their different findings, both Hyland and Petch-Tyson point to personal pronouns as an area of difficulty for NNS writers. These features may be over- or underused by different groups of learners, and there seems to be agreement that both their use and their omission may be problematic.

3. Materials and results

Against the background described above, it seemed likely that the use of personal pronouns and possessives in L2 learner texts might yield interesting results. Research was carried out on 30 TOEFL-style essays and 30 reports written as homework assignments by students enrolled on the upper-intermediate-level 6-credit course "Inglés empresarial" at the University of Navarra. Students had written the essays and reports after standard classroom preparation including study of a model text in each case. In order to determine the number of personal pronouns and possessive adjectives present in the essays and reports, samples comprising the first 180 words of each essay and the first 220 words of each report were prepared in text format, after correction of spelling and typing errors that might distort the results of automatic word counts. Then WordSmith was used to generate a count of the number of instances in each sample of the words: I, my, we, our, you and your. Table 1 shows the findings for each essay, and Table 2 displays the findings from the reports.

Table 1. Personal pronouns and possessives in the essays.

Essay	I	my	we	our	you	your
1	2	0	0	0	4	0
2	0	1	4	3	0	0
3	2	0	0	0	0	0
4	5	1	1	1	0	0
5	1	1	0	0	1	2
6	1	0	0	0	3	1
7	1	0	0	0	0	0
8	0	0	0	0	0	0
9	1	1	0	0	0	0
10	0	0	0	0	1	0

11	6	0	0	0	0	0
12	1	1	3	0	0	0
13	0	0	1	0	0	0
14	0	0	0	0	1	0
15	2	0	1	0	0	0
16	0	1	1	0	0	0
17	1	0	0	0	0	0
18	1	0	0	0	0	0
19	0	0	0	0	0	0
20	0	0	8	1	0	0
21	1	0	1	0	0	0
22	2	0	4	1	0	1
23	0	0	1	2	3	1
24	0	0	0	0	0	0
25	0	0	0	0	0	0
26	2	1	0	0	9	0
27	0	1	2	1	0	0
28	4	0	0	0	0	0
29	2	1	0	0	0	0
30	0	0	0	0	1	0
Total	35	9	27	9	23	5

Table 2. Personal pronouns and possessives in the reports.

Report	I	my	we	our	you	your
1	0	0	1	0	0	0
2	0	0	4	0	0	0
3	0	0	0	0	0	0
4	0	0	0	1	0	0
5	1	0	4	0	0	0
6	0	0	4	0	0	0
7	0	0	5	1	0	0
8	0	0	2	0	0	0
9	0	1	0	0	0	0
10	0	0	1	0	0	0
11	0	0	4	0	0	0
12	0	0	4	0	0	0
13	0	0	8	2	0	0
14	0	0	2	0	0	0
15	0	0	1	5	0	0
16	0	0	1	0	0	0
17	0	0	0	0	0	0
18	0	0	5	0	0	0
19	1	0	5	6	0	0
20	0	0	7	3	0	0
21	3	0	0	4	0	0
22	0	0	2	0	0	0
23	0	0	0	0	0	0
24	0	0	0	0	0	0
25	0	0	6	4	0	0
26	0	0	0	0	1	0
27	0	0	4	1	0	0
28	1	2	0	0	0	0
29	1	0	5	4	0	0
30	0	0	0	0	0	0
Total	7	3	75	30	2	0

4. Discussion

In general, the pattern that emerges is that the novice L2 writers in this study used personal pronouns and possessives in ways that did not entirely fit with the target genres. None the less, students evidently perceived a generic difference between the essay and the report task, because their use of pronouns differed widely between the two.

Table 1 shows that the most popular relation word in the essays was "I", which occurred a total of 35 times, followed by "we" (27 instances) and "you" (23 instances). The possessive adjectives were all less frequent (9 occurrences of "my" and "our", and 5 of "your"). Given that formal prose is often characterised by an absence of relation markers of this kind, it seems that the students in this group as a whole used personal pronouns, particularly "I", rather more than might be appropriate. In this, there was some degree of similarity to the findings by Ringbom (1998) described above, in which various groups of L2 writers were found to use personal pronouns more often than L1 writers of English. In Ringbom's study, the Spanish writers used "I" 36 times per 10,000 words, while the L1 writers of English used "I" only 25 times. Here, the students used "I" 35 times in a corpus consisting of only 5,400 words, thus exceeding not only Ringbom's L1 writers, but also his findings for Spanish speakers. On the other hand, in the present study, "we" was used 27 times in 5,400 words. Ringbom's sample of Spanish writers used "we" more than the present group (98/10,000 words), while his L1 writers used "we" at a frequency of 34/10,000, that is, less than the students in our sample. It therefore seems that the student essays in this study showed a marked tendency to over-use "I", and a rather more moderate tendency to over-use "we", unlike the Spanish writers in Ringbom's study, whose over-use of "we" was greater than their over-use of "I".

Although "you" was used less frequently than the first person pronouns, the fact that it was used at all could be regarded as considerably more aberrant. In Ringbom's study, "you" was found at a frequency of 33/10,000 in Spanish writers' essays and 8/10,000 in L1 writers' essays. Here, in the essays, "you" occurred 23/5,400, thus easily exceeding both of Ringbom's groups. Excessive use of "you" is a particularly striking phenomenon, and merits further discussion. "You" is rarely used in academic English, partly because it is not usually necessary or appropriate to address the reader directly, and partly because the impersonal use of "you" to mean "one", which is very common in spoken English, is not considered acceptable in writing. These students appeared to use "you" mainly in the latter sense:

The first idea is, if you have the idea of start a business, is because you want to make profit. If you have the opposite idea, probably you will be ruined in a short period of time. (Essay 26)

Students such as the writer of essay 26 evidently had not been sensitised to the absence of impersonal "you" from academic registers.

One final point about pronouns and possessives in the essays is that their use varied greatly among the students in this sample. Some, such as student 26, wove almost their entire essay around a discussion of what "you" should do. Others, such as student 11, were highly visible writers, accepting the potential invitation in the essay title to express their own opinion, making explicit their own intention as writers, and not hiding behind the conventions of persuasive argument:

I disagree a little with this statement. Although I think that the main aim of any business is to make a profit, I also consider that they must respect some aims such as environmental, social or human conditions.

To begin with, I would like to write about environmental matters that a business must take into account in production decision looking at the neoclassical economic theory. (Essay 11)

Students such as this writer matched well with the findings by Petch-Tyson (1998) outlined above; like Petch-Tyson's French and Dutch students, student 11 used "I" to talk about his opinions ("I disagree"), and to talk about himself as a writer in that particular text ("I would like to write").

As far as the reports were concerned, the pattern was very different from that found with the essays. The only personal pronoun used with any frequency was "we" (75/6,600 words), and the only possessive used frequently was "our" (30/6,600). The contrast with the essays is interesting, because it shows that the same students were generally able to modulate their style or register in accord with the task, and that they had some awareness of the genre conventions governing the report. The students hardly used "I" or "you" at all in the reports, but used "we" more often in the reports than in the essays. One possible explanation for this phenomenon is that the report was correctly perceived as a type of collective task involving others, whereas the essay prompt seemed to be asking for a "personal opinion", thus triggering a highly personal reaction in terms of "I think", rather than the less overtly personal style of persuasive prose that would be more usual in an academic essay.

One striking point with pedagogical implications arises out of a comparison of what the students wrote with the two texts which the students had been given as models. Here, it is also noteworthy that the model report given to the students as a guide was written almost entirely in the passive, with only one, albeit prominent, instance of "us". Apart from this one instance, no personal pronouns or possessives were used in the model. Similarly, the TOEFL-style essay given to the students as a model contained no personal pronouns or possessives. It seems that most students did not pick up and emulate this particular feature of the style of these models, either because they did not notice it, or because they did not regard it as significant. On the basis of this, it is possible on the one hand to conclude that the level of writer/reader presence in the students' texts was not due to imitation of any model, and on the other, to speculate that students generally lacked sensitivity to this aspect of genre.

5. Conclusion

It was evident that the apprentice writers in the present study opted for a much greater writer-reader presence in the text than would generally be considered appropriate, or than was in fact the case in the models they had been given. Both the essays and the reports used the first person rather more than would be usual in academic texts. But it was their heavy use of the pronoun "you" in the essays that was particularly discordant, since "you" is rarely used in academic essays in English. This phenomenon again points to a lack of awareness of genre conventions and a general problem with formal register.

In general, the analysis of the essays and reports in terms of appropriate relation leads to the conclusion that these students lacked sensitivity to these genres and were unaware of many of the conventions operating when one writes formal essays or reports in English. It was evident that many of their inappropriate linguistic choices were the result of following the norms of spoken English rather than trying to notice and imitate the formal written language. In conclusion, it would certainly be worthwhile including discussion of writer-reader presence in future academic writing programmes. Students need to be sensitised to the conventions governing use of personal pronouns and possessives in formal academic and professional writing, and equipped with alternative linguistic means for conveying authorial stance and voice that are more appropriate to these genres.

As far as the use of computer tools is concerned, this paper has shown that programs such as Wordsmith serve as a useful quantitative complement to analysis of student writing. Small-

scale studies such as this prove particularly useful in that they are based on texts that are homogeneous as far as genre and topic are concerned, which means that their findings provide detailed insights into the issues involved in writing very specific types of text.

6. References

- Biber, D. 2006. *University Language. A Corpus-Based Study of Spoken and Written Registers*. Amsterdam/Philadelphia: John Benjamins.
- Biber, D., S. Johansson, G. Leech, S. Conrad and E. Finegan. 1999. *The Longman Grammar of Spoken and Written English*. London: Longman.
- Chafe, W. 1982. "Integration and involvement in speaking, writing and oral literature". D. Tannen, ed. *Spoken and Written Language: Exploring Orality and Literacy*. Norwood (NJ): Ablex. 35-53.
- Chang, Y. Y. and J. Swales. 1999. "Informal elements in English academic writing: threats or opportunities for advanced non-native speakers?" C. Candlin and K. Hyland, eds. *Writing: Texts, Processes and Practices*. London: Longman. 145-167.
- 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: 1207-1231.
- Hyland, K. 1999. "Disciplinary discourses: writer stance in research articles". C. Candlin and K. Hyland, eds. *Writing: Texts, Processes and Practices*. Harlow: Longman. 99-121.
- Hyland, K. 2002. "Options of identity in academic writing". *English Language Teaching Journal*, 56 (4): 351-358.
- Meyer, P. G. 1997. "Hedging strategies in written academic discourse: strengthening the argument by weakening the claim". R. Markkanen and H. Schröder, eds. *Hedging and Discourse: Approaches to the Analysis of a Pragmatic Phenomenon in Academic Texts*. New York: Walter de Gruyter. 21-41.
- O'Keeffe, A., M. McCarthy and R. Carter. 2007. *From Corpus to Classroom*. Cambridge: Cambridge University Press.
- Petch-Tyson, S. 1998. "Writer/reader visibility in EFL written discourse". S. Granger, ed. *Learner English on Computer*. Harlow: Addison, Wesley, Longman. 107-118.
- Ringbom, H. 1998. "Vocabulary frequencies in advanced learner English: a cross-linguistic approach". S. Granger, ed. *Learner English on Computer*. Harlow: Addison, Wesley, Longman. 41-52.