The Distribution of Authorial Presence in Experimental Psychology Articles*

FRANCESCA BIANCHI
Università del Salento

Abstract

Traditionally, teachers, prescriptive grammars, and writing guides have imposed the use of impersonal style in scientific writing. And indeed, in scientific papers, authorial presence is frequently hidden behind passive forms or the personification of the text or experiment. Other times, however, the author surfaces primarily by means of singular or, more frequently, plural first person pronouns or determiners. A few quantitative contributions dealing with overt authorial presence in experimental papers exist, but none of them focuses specifically on psychology. Therefore, the current study aims at expanding the existing literature by analysing the distribution of overt authorial presence in a corpus of experimental psychology articles. The corpus consists of 43 articles (298,332 running words), divided by move; each sentence in the corpus was manually tagged to identify the step it performs. Hypothetically relevant key words (*I, me, my, myself, we, us, our, ourselves, author, authors, author's*) were used as starting points for the identification of distributional patterns. The data were analysed quantitatively, in order to highlight: the distribution across files of each of the selected key words; their general distribution in the corpus; their distribution across moves and steps; the most frequent relevant moves in which they appeared per section; and their most frequent collocates per section.

Keywords: psychology; corpus; communicative function.

1. Introduction

The experimental research article is probably one of the most frequently studied academic text types. It has been analysed by a huge number of authors, from several different perspectives (see for example Bazerman 1988; Berkenkotter & Huckin 1995; Bhatia 1993; Dudley-Evans 1998; Gross *et al.* 2002; Paltridge 1997; Swales 1990). From a structural and rhetorical viewpoint, research articles tend to show a fairly limited number of recursive structural elements, often organised according to conventional textual schemata. Rhetorical strategies are acquired and perpetuated within a community by means of readings of previously published papers, imposition from reviewers (who work on the basis of their acquired knowledge), and writing guides. Traditionally, writing guides, as well as teachers, and prescriptive grammars, have imposed the use of impersonal style in scientific writing. And indeed, in scientific papers, authorial presence is frequently hidden behind passive forms or the personification of the text or experiment. On other occasions, however, the author surfaces primarily by means of singular or, more frequently, plural first person pronouns or determiners. A few quantitative contributions dealing with overt authorial presence in experimental papers exist, but none of them focuses specifically on psychology. Therefore, the current study aims at expanding the existing literature by analysing the distribution of overt authorial presence in a corpus of experimental psychology articles. The analytical perspective adopted in the current study is in line with the British tradition established by Swales (1990) and analyses the rhetorical structure of the article in terms of moves and steps, i.e. two hierarchical communicative levels in which a move (frequently but not necessarily corresponding to a section or paragraph of the article) may include one or more steps. Both moves and steps correspond to communicative functions. For a review of this tradition compared to two other major traditions of genre analysis, see Hyon (1996).

1.1 Experimental psychology as rhetorical discourse

Experimental psychology 'was the first human science to establish a specialised discourse distinguished from traditional philosophical discourse' (Bazerman 1988: 259) and this community adopted 'what they perceived to be the methods of the physical and biological sciences' and imitated 'the forms of argument developed within the natural sciences' (Bazerman 1988: 257). Their rhetorical devices are summarised in the *Publication Manual* (2001) of the American Psychological Association, which is now in its 5th edition and to which prospective authors are referred by the vast majority of psychology journals.[1] The APA *Publication Manual* is composed of 'approximately two hundred oversized pages of rules, ranging from such mechanics as spelling and punctuation through substantive issues of content and organization' (Bazerman 1988: 259). Indeed, the APA *Publication Manual* provides details of the structure the articles should have, how to name each section, what to include in each section, how to quote previous literature, and much more. This manual, however, was not meant to be a linguistic description, but rather a series of suggestions for prospective authors made by psychologists and editors.

The current study focuses on one rhetorical strategy of academic writing, namely overt authorial self-reference. The following paragraphs offer an overview of the indications provided by reading and writing guides to psychology researchers.

1.2 Authorial presence: review of guidelines in reading and writing manuals

A brief description of the possible ways in which the author or experimenter is expressed in scientific experimental articles is offered by Rossini Favretti while introducing students to the reading practice relating to psychology articles (Rossini Favretti & Bondi Paganelli 1988). She lists three strategies: explicit mention through personal pronouns (first person singular, or more frequently plural); agent hidden behind passive forms; or agent hidden behind personification structures (e.g. 'This chapter starts with…').

The APA *Publication Manual* mentioned above deals with many topics, but does not provide guidelines for the use of personal pronouns, or passive forms. Some writing guides, such as the *Guide to Grammar and Style* by Jack Lynch[2] – suggested by the online Encyclopedia of Psychology – state that

> In scientific writing […] sentences are routinely written in the passive voice; the authors are therefore given less importance, and the facts are made to speak for themselves.

On the other hand, the writing guidelines provided by the Psychology Writing Centre of the University of Washington read as follows:

> Active voice is usually clearer and more to the point. Replace "nests are built" with "birds build nests". Write "I found" instead of "it was discovered". Don't be afraid to use the passive voice if it really is clearer (as we do in this paper), but favor using subject, verb, object.[3]

Finally, some other guides to style, such as the document *How to write a paper in scientific journal style and format*[4] by the Department of Biology of Bates College, Lewiston, take into consideration the fact that the use of one or another strategy may depend on the discipline and on the different sections of a scientific article and declare that

> Some disciplines and their journals (e.g., organismal biology and ecology) have moved away from a very strict adherence to the third person construction, and permit limited use of the first person in published papers. Other disciplines, especially the biomedical fields, still prefer the third person construction. Limit your use of first person construction (i.e., " I (or we) undertook this study ….): usually it is most acceptable in the Introduction and Discussion sections, and then only to a limited extent. Use first person in the methods sparingly if at all, and avoid its use in the results.

It emerges that guidelines are rather vague and sometimes even contradictory as far as this particular topic is concerned. This may depend on the fact that actual practice reveals a mixture of techniques, that preferred trends are not easily visible through qualitative analysis, or that analysis should be highly discipline-specific. In the literature, we found three authors providing quantitative[5] contributions on overt authorial presence in experimental papers; of these, however, only one considered psychology, and analysed abstracts only. These contributions are summarised in the following paragraphs.

Kuo (1998) analysed all personal pronouns in 36 scientific journal articles from three hard science fields, namely computer science, electronic engineering, and physics. With regard to pronouns that might indicate authorial presence, this author found that first-person plural pronouns and the corresponding determiner (*we*/*us*/*our*) were the most frequent ones in the corpus. On the other hand, first-person singular pronouns (*I*/*me*/*my*) did not occur at all; this absence was at least partly explained by the fact that even in single-authored papers, the writer referred to him/herself as *we*. Kuo also analysed distribution of *we* across discourse functions, but this was done on a sample of 9 research articles only. The analysis showed that exclusive *we*/*us*/*our* appeared in sentences performing one of the following functions: explaining what was done (the vast majority of cases for *we* and *us*, and highly frequent for *our*); showing results or findings (highly frequent for both *we* and *our*); showing commitment or contribution to research (the majority of cases for *our*; rather frequent also for *we*); proposing a theory or approach; stating a goal or purpose; justifying a proposition; hedging a proposition or claim; giving a reason or indicating necessity; comparing approaches or viewpoints; expressing wish or expectation. Some of these functions, such as explaining what was done, seemed to be distributed across all the sections of a journal article. Distribution of functions across sections, however, was not analysed systematically by this author.

Hyland (2001; 2003) focused on the use of self-citation and exclusive first person pronouns in two subsequent papers, both based on a corpus of 240 research articles in eight disciplines from both hard and soft fields (mechanical engineering, electrical engineering, marketing, philosophy, sociology, applied linguistics, physics, and microbiology). Unfortunately, psychology did not appear in this corpus. The 2001 paper focuses primarily on the key words *I*, *me*, *my*, *we*, *us*, *our*, *this writer*, *the research team* (and other similar expressions) used for self-mention, and on their discourse functions. In Hyland's data the vast majority of occurrences were pronouns (81%). As far as frequency of self-mention forms per discipline is concerned, each discipline seemed to prefer one or the other type of pronoun, regardless of the traditional distinction between hard and soft sciences. However, one general trend is visible in Hyland's tables, in almost all disciplines and with both pronouns/determiners: the subject pronoun is generally more frequent than the determiner, which in its turn is more frequent than the object pronoun. The author concluded saying that authorial presence surfaces at those points in the article where the authors are best able to promote themselves and their individual contributions; their intrusion is aimed at strengthening their credibility and their role in the research and helps them gain credit for their claims. Furthermore, in different disciplines, different degrees of authorial presence are allowed. In the 2003 paper, Hyland counted and analysed frequency and forms of self-reference, making a distinction between articles and abstracts. Furthermore, he analysed the distribution of authorial presence across discourse functions, on a limited sample of 10 research articles from each discipline. Self-reference was generally lower in abstracts than in papers, in almost all disciplines. Singular first person pronouns were virtually inexistent in all the hard sciences considered,

while in the soft sciences both singular and plural pronouns were frequent. A few occurrences of other forms of reference were found in the papers, scattered across all disciplines except philosophy, but none was found in the abstracts. Finally, four main purposes were identified for self-reference: stating a goal or outlining the structure of the paper; explaining a procedure; stating results or making a claim; elaborating an argument. The frequency 'hierarchy' of these functions depended on the discipline and was different in papers and abstracts.

Lastly, Martín Martín (2003) analysed the distribution and frequency of occurrence of first person pronouns in phonetics and psychology. His attention was focused on abstracts only, and his aim was to compare English and Spanish rhetorical practice. His data showed that first person pronouns, though present in the corpus, are not a relevant feature of abstracts in these disciplines. Almost 30% of the occurrences in English indicated 'the author as describer of the research', while 16.6% of occurrences were of 'the author as fully committed claim maker'. The other identified functions for exclusive first person pronouns were: 'the author as experiment conductor'; 'the author as opinion holder', and 'the author as cautious claim maker'. Interestingly, several of the occurrences of the author as fully-committed claim maker appeared in the Results and Discussion sections of the abstract.

The aim of the current study, therefore, was to expand the existing literature by analysing the distribution of overt authorial presence in a corpus of experimental psychology articles and taking advantage of corpus concordancing tools[6] and quantitative methods of analysis.

2. Materials and methods

The current study is based on an annotated corpus of 40 experimental psychology articles taken from 15 different journals, for a total of 232,244 running words. This is a subset of the *Psychology Corpus* described in Bianchi & Pazzaglia (2007, subsection 1.4.1)[7]. In the corpus considered, slightly more than 80% of the journals are publications of the American Psychological Association (APA). The other articles all come from psychology journals published by Hogrefe & Huber.

Following Gledhill's and Bowker and Pearson's approaches (Bowker & Pearson 2002; Gledhill 1995, 1996, 2000), this corpus is divided into folders, each folder corresponding to a different move typical of experimental articles. Folders do not strictly correspond to the different sections of articles (structural units), but rather to moves (primary-level functional or communicative units) which may or may not overlap with the actual sections of articles. Moves were selected *a priori* following the work of various authors including Gopnik (1972) and Gläser (1995) and the preliminary analysis of a limited number of randomly selected articles. The moves appearing in the corpus are the following: Titles; Notes; Thanks; Abstract; Introduction; Literature Review; Method; Results; Discussion; Conclusion. For the sake of comparison with similar analyses carried out in the literature, only the following moves were considered and discussed in the current study: Abstract; Introduction; Literature Review; Method; Results; Discussion; Conclusion[8].

Furthermore, steps, or communicative functions, were manually annotated within each move. The steps identified during the annotation phase are listed in

Table 1, along with the tags used. Any of the steps listed in the table could theoretically be found in any of the moves listed above.

For the purpose of the current study, the analyses were carried out starting from concordances of key words that might be indicative of explicit authorial presence. The following key words were taken into consideration: *I, me, my, myself, we, us, our, ourselves, author, authors, author's*. Concordance lines were retrieved with *Wordsmith Tools* for each key word and for each of the moves. The data retrieved were tabulated in Excel and integrated with information about the file from which the concordance line was taken, the step/communicative function in which the key word appears, the syntactic role of the key word (subject; direct object; other), and, when necessary, other information (such as inclusive or exclusive function, in the case of key words *we/us/our*).

Name of step	Tag	Description
Background	<backg> </backg>	reporting past research
Conclusions	<conc> </conc>	final remarks – conclusions
Discussion	<discussion> </discussion>	author's comments about findings, material, procedure, etc.
Findings	<findings> </findings>	description of results
General statement	<gen> </gen>	author's general considerations
Limits	<limits> </limits>	limits of current experiment
Material	<mat> </mat>	description of material
Method	<method> </method>	in the case of descriptions of subjects, material and procedure are not distinct
Need for extension	<further> </further>	need for further data, analysis, etc.
Object	<obj> </obj>	aim of the study
Procedure	<proc> </proc>	description of procedure
Rationale	<rationale> </rationale>	when reference is made to a theory and not to empirical past research
Relevance	<relevance> </relevance>	why are these previous studies relevant to the current experiment?
Subjects	<subj> </subj>	description of subjects
Niche	<niche> </niche>	describing/explaining one's research niche
Metatextual reference	<metatextual> </metatextual>	Metatextual reference
Hypothesis	<hypothesis> </hypothesis>	Statement of hypothesis
Quote	<quote> </quote>	Citation of other authors or texts (eg. questionnaire items)
Criteria	<criteria> </criteria>	Description of criteria/logic used to create material or analyse data
Acknowledgments	<genack> </genack>	Acknowledgments
Note to the text	<textnote> </textnote>	Note to the text

Table 1. Steps: Annotation scheme

The data were analysed quantitatively, in order to highlight: a) the distribution of each of the selected key words across files; b) their general distribution in the corpus; c) their distribution across moves and steps; d) the most frequent relevant steps in which they appeared per move; and e) their most frequent collocates per move. Some of the analyses were carried out considering all of the selected keywords, while others were carried out only on first person plural pronouns and determiners, as this was the only category of authorial self-reference with a statistically significant number of occurrences. Any irrelevant occurrence of a keyword, i.e. occurrences that were not indicative of authorial presence, were excluded from our calculations. Among the concordance lines excluded there were, for example, all those cases in which the word *author/s* referred to the author of a test or paper quoted for reference, rather than the author of the paper under analysis (e.g. *According to the test authors, the MC scale has satisfactory internal consistency*), or cases where the key word appeared within a quote (e.g. "*The teacher wants us to try new things*": quote from an experimental subject's response). The following sections illustrate and discuss the results of the analyses we performed.

3. Results and discussion

3.1 Distribution of authorial presence across files

Analysis of the distribution of the relevant instances of the selected key words provided interesting insights into rhetorical standards of the psychology community and the (lack of) impact of traditional writing guidelines. The key word *author* appeared in 10 different articles, mostly in the abstract; *I/my* appeared in 3 different articles; and *we/us/our* appeared in all the remaining articles. Interestingly, those articles where the word *author* was used as self reference in the Abstract then shifted to using *we/our* in the sections that followed. All the articles in the corpus contain some form of overt authorial presence, despite traditional writing guidelines, including those of APA, which advocate impersonal style. It would be interesting to see whether any connection exists between these authorial choices and the two publishing houses.

3.2 General distribution

First of all we analysed the distribution of overt authorial presence in the whole corpus. To this end, wordlists were generated for each folder/move using *Wordsmith Tools*, concordance lines were run for each of the selected key words, and irrelevant occurrences were deleted. This provided a first means of comparison with the previous studies.

The results of this general-level analysis are summarised in Table 2: the first column indicates whether the author resorted to first person singular, first person plural, or impersonal type of explicit reference; the second column shows the type of word representing overt authorial presence; the third column provides the raw occurrences of that word; the fourth column shows the relative percentage with respect to the given type of reference; the fifth column offers the percentage with respect to the entire set of overt references in the corpus; finally, the last column shows the syntactic position in which the key word appeared (subject/object/other).

Type of reference	Key word	Hits	Relative % (type of ref.)	General % (overt references)	S / O / Other position
1st person singular pron.	I	5	83	0.58	S (100%)
	Me	0	0	0	
	My	1	17	0.12	S (100%)
	Myself	0	0	0	
1st person plural pron.	We	599	69	68.77	S (100%)
	Us	17	2	1.95	O (99%); Other (1%)
	Our	240	29	27.55	S (50%); O (6%); Other (44%)
	Ourselves	0	0	0	
Impersonal ref.	Author	0	0	0	
	Authors	9	100	1.03	S (67%); O (11%); Other (22%)
	Author's	0	0	0	
	Authors'	0	0	0	

Table 2: General distribution of overt referential key words

As the number of hits shows, first person plural pronoun reference clearly dominated. Even in cases of impersonal overt reference the plural form (*authors*) was the only one used. Indeed, dominance of plural reference was expected, given that psychology articles are typically characterized by joined authorship. Furthermore also Kuo (1998) and Hyland (2003) have shown neat prevalence of plural reference in most academic fields. However, a few examples of reference by singular pronoun also occurred in the Psychology Corpus. This is in keeping with Hyland's (2003) study, where a high number of instances of singular personal pronouns were found in soft science papers.

Abstracts almost exclusively included references by means of the key word *author* (9 instances) and one case only of reference by personal pronoun (*we*). This finding is in keeping with Martín Martín (2003), who found that the use of personal pronouns was not a widespread practice in psychology abstracts.

Let us now take into consideration personal reference through pronouns only. If we explore the details of their syntactic roles, subject roles (*we*/*I*) outnumbered determiners (*our*/*my*), which in turn outnumbered object roles (*us*/*we*). This is true for both first person singular and first person plural reference. This type of distribution (*S* > *DET* > *O*) appeared in four moves, namely: Literature review, Introduction, Method, and Results. So far, the results match those in Hyland (2003). However, in our corpus, the Discussion and Conclusion moves behaved differently, showing a virtually equal number of subject and object pronouns. Abstracts could not be considered in this analysis, as they included only one instance of overt authorial presence through personal pronoun, in the form of the key word *we*.

Furthermore, if we analyse the noun phrases where referential determiners appear, authorial presence surfaces primarily in thematic/subject position. In fact, in this corpus, the number of *we* plus the number of *our* in subject position amounted to 83.5% of total cases of first person plural reference; and the only instance of *my* appeared in a subject noun phrase. In line with this, 67% of the hits of *author* were subjects.

To complete this series of analyses, we looked at the distribution of *our* in terms of Subject (S), Direct Object (DO) or Other Indirect (OI) role, across sections (Table 3). Results confirmed similarity between the Literature Review, Introduction, Method, and Results moves, where Other Indirect role preceded Direct Object, which in turn preceded Subject (OI > DO > S). Once more, the Discussion and Conclusion moves behaved differently, Subjects outnumbering Direct Objects, outnumbering in turn Other Indirect roles (S>DO>OI)[9]. Finally, no instances of *our* appeared in the Abstract section.

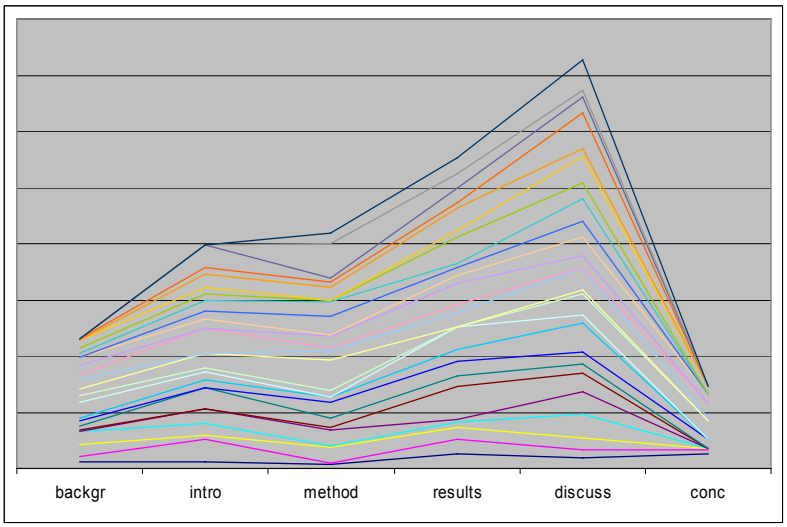
OUR	Abstr.	Lit. rev.	Intro.	Method	Results	Discuss.	Conc.	Notes	Tot.	%
Subj.		3	7	4	14	72	20	1	121	50.42%
Obj.		1	1	2	5	3	1		13	5.42%
Ind.		7	11	9	25	47	7		106	44.16%
TOT	0	11	19	15	44	122	28	1	240	100%

Table 3: Distribution of key word *our* in terms of S/DO/OI role, across sections

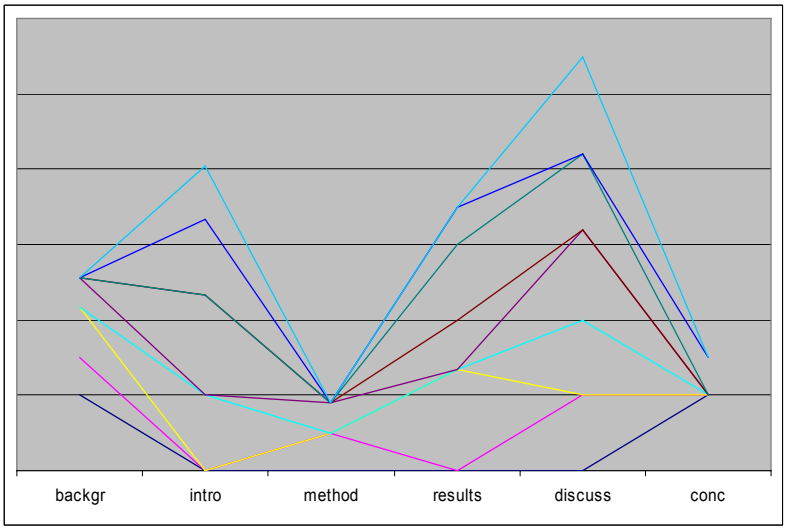
3.3 Distribution of authorial presence across sections

To establish the incidence of authorial presence across sections, we decided to concentrate on the key words *we/us/our* only, as these represent the overwhelming majority of occurrences of overt reference. Only cases of *we/our/us* of the exclusive type were considered.

Graphs 1 and 2 show the distribution of overt authorial presence by first-person plural pronoun/determiner across moves. To produce these graphs, percentage values were used, calculated on the basis of the total number of occurrences of the three key words. In the graphs, each line corresponds to a different file in the corpus. Moves are reported on the *x* axis. The Abstract move is not present, because it never included occurrences of the key words taken into consideration. We must recall, however, that this move includes a few instances of reference using the key word *authors*. For a clearer graphical representation, two separate graphs were produced: one for those files where instances of the key words appeared in all sections, and one for those files where presence of the key words was discontinuous across sections.



Graph 1. Distribution of key words *we/us/our* across sections: files with constant presence of these key words



Graph 2. Distribution of key words *we/us/our* across sections: files with discontinuous presence of these key words

In both graphs, the Discussion section clearly appears as the top rising point of inflection, i.e. the section with the highest number (%) of the key words considered. This is true for all files, except 5 (85% of the files). In the vast majority of the files, the Introduction and Results sections are also points of inflection, though lower than the Discussion section. On the other hand, the Literature Review, Method and Conclusions sections are falling points of inflection, i.e. sections with the lowest number (%) of the key words taken into consideration.

These results may tentatively be explained by the content of each move: in the Literature Review move, focus is on previous literature rather than on the current study and its author/s; authorial presence, then, increases in the Introduction, as this move presents the current study and/or the article and is the place where authors frequently present their working hypotheses; in the Method move focus shifts to participants, materials and procedures, which are supposed to be described in an objective, detached type of language; attention returns to the current experiment in the Results section, where data are presented, and above all in the Discussion section, where the author/s can openly comment the results and their hypotheses; finally, the Conclusion section is probably poor in authorial presence markers because of the need to try and generalize results. Confirmation of these preliminary explanations will be searched for in an analysis of the distribution of steps across the different sections.

3.4 Distribution of authorial presence across steps

Distribution of authorial presence across steps was assessed considering all types of authorial reference. Table 4 shows the distribution of authorial presence across steps, in decreasing order of frequency, regardless of move. In the third column, percentages are reported, shortened to the second decimal place. In the last column, examples from the corpus are provided (in shortened form, because of space limitations), with the indication of the move in which they were found. Here, the key word has been capitalized.

Step	Hits	%	Example from the corpus
Discussion	203	23.23	<discussion>*OUR effects were quite consistent for estimates of the causes …*</*discussion*> [Conclusions]
Method	157	17.96	<method>*The picture-recall test allowed US to analyze effects of valence and arousal separately;*</method> [Discussion]
Object	107	12.24	<obj>*OUR questions of interest related to the variables added after the academic skills' variable.*</*obj*> [Results]
Hypothesis	101	11.56	<hypothesis>*OUR hypothesis was that sad subjects would perceive situationally caused events as more likely…* </*hypothesis*> [Results]
Findings	72	8.24	<findings>*For m. corrugator activity, WE found a significant Defensiveness x Emotion effect:*</findings> [Results]
Procedure	68	7.78	<proc>*In Session 2, WE assessed behaviors over the preceding 2 weeks.*</*proc*> [Method]
Limits	44	5.03	<limits> *Another shortcoming of the study involves OUR use of measures for which only limited validity and reliability…*</limits> [Discussion]

Step	Hits	%	Example from the corpus
Material	25	2.86	<mat>OUR *protocol, however, diverged in two ways from the CTS*.</mat> [Method]
Niche	23	2.63	<niche>*What is different about OUR study when compared with others is that we focus on signif…*</niche> [Introduction]
Metatextual reference	17	1.95	<metatextual>*Before turning to the details of the longitudinal study, WE will briefly summarize the literature that explores our main…*</metatextual> [Introduction]
Background	14	1.6	<backg>*Our predictions were based on appraisal theories of emotion,…*</backg> [Conclusion]
Subjects	7	0.8	<subj>*We recruited 305 participants between the ages of 13 and 17*</subj> [Method]
Conclusion	7	0.8	<conc>*WE suggest that the study of LGB identity may enrich pres…*</conc> [Conclusion]
General	6	0.69	<gen>*These more directive methods, which WE refer to as curriculum centered, typically involve structured…*</gen> [Introduction]
Note to the text	6	0.69	<textnote> …10 periods) *were not chosen as baseline in this report because WE decided to contrast a rest period with the induction period, …*</textnote> [Notes]
Criteria	5	0.57	<criteria>*Specifically, six or more symptoms of either inattention or hyperactivity/impulsivity had to be endorsed (as "pretty much present" or "very much present") on the DBD Rating Scales to meet OUR criteria for ADHD*.</criteria> [Method]
General acknowledgment	4	0.46	<genak>*WE thank Maria Boardman, Susan Bunton, Phillis George, …*</genak> [Notes]
Relevance	4	0.46	<relevance>*This study builds on previous work by the first AUTHOR (Levendosky & Graham-Bermann, 2001) in which an ecol…*</relevance> [Literature review]
Need for extension	3	0.34	<further>*OUR assessment of peer relationships could be improved by …*</further> [Discussion]
Rationale	1	0.11	<rationale>*Recently, WE reported data demonstrating that young infants both discriminate among faces on…*</rationale> [Literature review]

Table 4: Steps including authorial self-reference, in order of frequency

Overt authorial presence was found in all steps, except Quote which, as expected, never included instances of authorial self-reference. However, in order to compare the steps considered in this analysis to those that emerged in the previous studies ,'a naming conversion table' is needed, as the names and descriptions used by the different authors cannot be easily and univocally matched (Table 5). Indeed some doubts still remain, expressed by a question mark next to this matching. Martin's labels were not considered, as they were really difficult to compare.

Kuo's	Hyland's	Current
Explaining what was done	Explaining a procedure	Material / Method / Subjects / Procedure
Proposing a theory, approach, etc.		Hypothesis (?)
Stating a goal or purpose	Stating a goal or purpose	Object
Showing results or findings	Stating results or claim	Findings
Justifying a position	Elaborating an argument	Discussion / relevance
Hedging a proposition or claim	Elaborating an argument	Discussion
Showing commitment or contribution to research	Elaborating an argument	Discussion / conclusion / relevance
Comparing approaches, viewpoints, etc	Elaborating an argument	Discussion
Giving a reason or indicating necessity	Elaborating an argument	Discussion / limits
Expressing wish or expectations		Hypothesis (?)
	Expressing self-benefits	Discussion

Table 5: Steps: naming conversion table

The picture emerging from this table led us to group some steps together: the Material, Procedure, Method, and Subjects steps were grouped into 'Procedure+'; the Discussion, Conclusion, and Relevance steps were grouped into 'Discussion+'. This, in turn, led us to a general review of the initial steps and to a grouping of the Background, Rationale, and Criteria steps as 'Background+'. Recalculation of the results of the current analysis provided the ranking in Table 6, in which only the most frequent functions are reported.

Step	%
Discussion +	29.8
Procedure+	29.4
Object	12.2
Hypothesis	11.6
Findings	8.2

Table 6: General distribution of overt authorial presence: most frequent steps, after grouping

Interestingly, all the functions in which authorial self-reference emerged in the previous literature are among the top five in our corpus. However, in terms of ranking no similarity seems to exist between the data in our corpus and those considered by the other authors. This confirms Hyland's (2003) result that the frequency 'hierarchy' of these functions depends on the discipline.

If we look at overt authorial distribution within the different moves, the pictures in Tables 7a-7f emerge. Moves are listed in decreasing order of frequency. Abstracts have not been included in these tables, given the few occurrences of the authorial key words used in this study; however, this move will be briefly discussed after the comments on the other moves.

Literature Review	%
Object+	31.4
Hypothesis	26.5
Procedure+	18.0
Discuss+	15.7
Background+	6.0
Niche	2.4

Table 7a. Steps with overt authorial presence: Literature review move

Introduction	%
Object+	29.7
Hypothesis	26.9
Procedure+	25.5
Niche	4.9
Background+	4.1
Discuss+	4.1
Metatextual reference	3.4
Gen	0.7
Findings	0.7

Table 7b. Steps with overt authorial presence: Introduction move

Method	%
Procedure+	82.1
Discuss+	14.2
Hypothesis	2.8
Background+	0.9

Table 7c. Steps with overt authorial presence: Method move

Results	%
Procedure+	46.6
Discuss+	16.2
Findings	13.9
Object	9.5
Hypothesis	7.2
Metatextual reference	4.9
Background+	0.9
General	0.4
Limits	0.4

Table 7d. Steps with overt authorial presence: Results move

Discussion	%
discussion+	42.8
Limits	14.8
Findings	14.0
Hypothesis	6.2
Niche	5.4
procedure+	6.0
Object	6.0
background+	2.5
General	1.4
Further	0.9

Table 7e. Steps with overt authorial presence: Discussion move

Conclusion	%
discussion+	51.9
Limits	13.5
Findings	11.8
Hypothesis	9.5
Object	5.7
Further	1.9
General	1.9
procedure+	1.9
Niche	1.9

Table 7f. Steps with overt authorial presence: Conclusion move

Each move is characterized by a specific ranking of steps, but similarities can be seen between groups of moves. Furthermore, as we will see, overt authorial presence tends to appear in steps that seem to be particularly significant for and expected in the given move.

The Literature review move shows several similarities with the Introduction move. In particular, all the steps present in the Background move authorial presence list are also present in the Introduction move, the first three (Object+; Hypothesis; Procedure+) exactly in the same order, the other three (Discussion+;

Background+; Niche) in reverse order. The Introduction list also includes three other steps with overt authorial presence, namely Metatextual reference, General and Findings. These, like the ones higher up in the ranking, are typical steps for an Introduction, where one would expect to see a brief description of the contents and structure of the whole paper (Metatextual reference) and maybe also a brief summary of the results of the study (Findings).

The Method move shows several similarities with the Results move. As before, all the steps in the Method move (Procedure+; Discussion+; Hypothesis; Background+) are also present in the Results one. Furthermore, the Results move includes several other steps with overt authorial presence. This mirrors the wide variety of communicative functions that characterize the Result section of an article. Here, too, authorial presence tends to appear in steps that seem to be particularly significant for and expected in the given move, namely description of results (Findings) and description of tables and graphs (Metatextual reference).

Finally, the Discussion and Conclusion moves are among the richest in terms of variety of steps with authorial presence. Their lists share almost all the same steps, the top four (Discussion+; Limits; Findings; Hypothesis) even in the same order. A major difference is the presence of step Background+ in the Discussion move. Indeed, reference to the literature can be expected when discussing results.

The Abstract section represents a move on its own. Indeed, we have already mentioned that very few instances of overt authorial presence were found in this section, and nearly all of them were characterized by the key word *authors*, rather than *we/us/our*. These few occurrences were scattered through five steps: object, procedure+; discussion; findings; and hypothesis. Given the low number of hits, it is clearly pointless to rank steps in order of frequency. However, as was the case with the other moves, most of these steps (object, hypothesis, procedure, and findings) seem to be highly typical of an abstract.

These results are only partially in keeping with our tentative explanation of general distribution across moves (Section 3.3). On the one hand, the hypothesis that low presence of authorial reference in the Literature review move might be due to focus on previous literature rather than on the current study and its author/s is supported by the very few occurrences of self reference in the Background step. Furthermore, an increase in authorial presence in the Introduction due to topic shift to the current study and the author's working hypotheses is indeed supported by the wide number of steps and their ranking within this move. Finally, the wide difference in authorial presence that characterizes the Discussion and Conclusion moves (the highest number vs. one of the lowest numbers) can only be explained in terms of topic shift, as our analysis of steps shows very little difference between these two sections. On the other hand, the hypothesis that participants, materials and procedures – which are the main topics of the Methods move – are described in an objective, detached type of language, is contrasted by the fact that a high number of occurrences of authorial self-reference were found in the step Procedure+ (top of the list in the Method and Results moves, and among the top three in the Literature review and Introduction moves). Furthermore, increased presence in the Results move seem connected to the wide variety of steps that characterize this move, which accompanies topic shift to the presentation of results.

3.5 Collocates per section

Finally, we decided to attempt an analysis of collocates, to see whether this would give us some further clues to the discussion of our quantitative findings. Collocates (N+1/2/3) were assessed only for the key words *we* and *our* – the most frequent key words of all – and are summarized in Table 8. In each column, the first row of each key word shows the most frequent collocate. The Conclusion move does not appear in this list because it did not include relevant occurrences of either key word.

Key word	Abstract	Literature review	Introduction	Method	Results	Discussion
WE	studied	expected	hypothesized	used	found	found
		hypothesized	examined	assessed	conducted	expected
		examined	believed	had	used	hypothesized
			expected	tested	entered	may
					performed	
					calculated	
					computed	
					ran	
					predicted	
OUR		data	hypothesis	analyses	hypothesis	findings
		study	study	criteria	analysis	predictions
		research	aim	hypothesis	data	study
		work	groups	study	research questions	model
		view	sample	goal	expectations	account
		knowledge	procedure	raters	model	data
			intent	results	criteria	version
			expectations	laboratory	focus	effects
			prediction	variable	design	results
			knowledge	protocol	sample	
			model		participants	
			belief		predictions	
			focus		tests	
			attempts		measures	
			research			

Table 8: Collocates (N+1) of key words *we* and *our*

In the Psychology Corpus, *we* collocated in N+1/2/3 position with verbs only, as expected, all of them in the active form. Interestingly, these verbs are connected to a limited number of communicative functions: Hypothesis; Procedure; Object; and Findings. Collocations with words indicating Hypothesis are scattered in the following moves: Literature review (collocates: *expected; hypothesized*); Introduction (collocates: *hypothesized*; *believed*; *expected*); Results (only collocate: *predicted*); and Discussion (collocates: *expected*; *hypothesized*; *may*). *We* collocates with the idea of Procedure primarily in the Results move (collocates: *conducted*; *used*; *entered*; *performed*; *calculated*; *computed*; *ran*), but also distributed in the following moves: Literature review and Introduction (only collocate: *examined*); and Method (only collocate: *used*). Collocates referring to function Object are found in the Method (collocates: *assessed*; *had*; *tested*), and Abstract (only collocate: *studied*) moves. Finally, *we* collocated with the idea of Findings in the Results and Discussion moves (only collocate: *found*).

On the other hand, N+1 collocates of the key word *our* are associated to a wider number of communicative functions, and are sometimes difficult to classify. As expected, they are all nouns. The key word *our* frequently collocates with words indicating Hypothesis. This is true in particular for the Introduction (collocates: *hypothesis*; *intent*; *expectations*; *prediction*), and Results (collocates: *hypothesis*; *research questions*; *expectations*; *predictions*) moves. However, one collocate was also found in the Method (only collocate: *hypothesis*), and Discussion (only collocate: *predictions*) moves. Other collocates lead us back to the Procedure + function, with the majority of words indicating either Subjects or Method. These collocates were found in the following moves: Introduction (collocates: *groups*; *sample*; *procedure*; *attempts*); Method (collocates: *analyses*; *raters*; *laboratory*; *variable*; *protocol*); Results (collocates: *analysis*; *model*; *design*; *sample*; *participants*; *tests*; *measures*); and Discussion (only collocate: *model*). Collocates of *our* also frequently express the function Findings; these are scattered primarily in the Discussion move (collocates: *findings*; *account*; *data*; *effects*; *results*), but also in the Method move (only collocate: *results*), and in the Introduction and Results moves (only collocate: *data*). Less frequently collocates refer to the Object, Discussion, or Criteria functions. Collocates indicating the Object of the study were found in the following moves: Introduction (collocates: *aim*; *focus*); Method (only collocate: *goal*); and Results (only collocate: *focus*). Collocates indicating Discussion were found in the Literature review (collocates: *view*; *knowledge*), and Introduction (collocates: *knowledge*; *belief*) moves. Finally, one collocate only was found referring to Criteria (only collocate: *criteria*), in the Method and Results moves. The remaining collocates (*study*; *research; work; version*) are too general to be matched to a specific function.

Partial coincidence can then be seen between the functions highlighted by the collocates of the two key words. In fact, the Hypothesis, Findings and Object functions are common to both, while the Procedure function of collocates of *we* does not coincide with the Procedure+ function of collocates of *our*, as the former refers to the description of the procedure itself, while the latter includes mainly description of subjects or of method. Analysis of syntactic roles, however, only partially confirmed the hypothesis that the wider number of functions connected to collocates of *our* might be due to those cases when the noun phrase did

not appear in subject position. In fact, most, but not all, instances of *our+collocate* in subject position refer to the same function as collocates of *we*; others, instead, refer to other functions.
Furthermore, we checked whether direct correspondence existed between the step performed by a sentence including authorial reference and the communicative function suggested by the collocate. In several cases correspondence existed, as in the following examples: <findings>*Furthermore, WE found patterns of relationship qualities that distinguished…*</findings>; <hypothesis>*In addition, WE expected a positive relationship between defensiveness…*</hypothesis>. However many instances were also found where sentence and collocate inspired different communicative functions. In the following sentence <conc>*In summary, WE replicated the pattern of results of Study 2…*</conc>, for example, collocate *replicate* describes a Method or Procedure, and the communicative function of the whole sentence (Conclusion) is performed by phrase *In summary*, in initial position. In the sentence <limits>*WE also attempted to call the siblings of each subject but…*</limits>, instead, collocate *attempt* suggests a Discussion, or Method, or Procedure function, and the Limits step was derived from the whole sentence, and marked by the conjunction *but*. These instances also rule out the existence of a direct connection between step/collocate-function, and subject role of the key word.

4. Conclusions

This study – based on a corpus of 43 articles (298,332 running words), divided by move (each move roughly corresponding to a different section of the article) and tagged for steps – has shown that, despite the indications in the APA manual and the general indications of style guides, the authors frequently surface in experimental psychology articles. Indeed, every article in the corpus contained some form of overt authorial presence. The preferred referential expression was the use of first person plural reference (*we/our/us*), although a few instances of reference by the word *authors* (always in the plural) or *I/my* were also found. Furthermore, authorial references mostly appeared in subject position.
Overt authorial presence was seen in all moves. The highest number of hits was found in the Results and Discussion moves, while the lowest was in the Abstract, Method and Conclusions moves. In terms of steps, key words indicating authorial presence were found in all steps except Quote. The highest number of instances belonged to the Procedure+, Discussion+, Method, Object, Hypothesis and Findings steps, in that order. Within the different moves, analysis of the distribution of authorial presence across steps showed a different ranking of steps in each move. Similarities were found between the Literature review and Introduction moves, the Method and Results moves, and the Discussion and Conclusions moves, while Abstract emerged as a move with its own characteristics. Furthermore, each move also included steps directly connected to the contents of the move.
Analysis of collocates was also carried out. Collocates replicated the most frequent steps (Hypothesis, Findings+, Procedure, Object) plus a few other functions. However, no direct correspondence was found between collocate and sen-

tence in terms of communicative function, not even when the key word was in subject position.

Comparison of results with previous findings was also attempted, despite methodological differences in the definition of steps. Generally speaking, the findings of this study are in keeping with those described in the literature, even though the other studies analysed articles from fields other than psychology. Comparison seems to confirm that, for several disciplines including psychology, the rhetorical habits observable in experimental papers tend to depart from the strategies suggested by writing guides and publication manuals, at least as far as authorial presence is concerned. In particular, our data seem perfectly in keeping with considerations by Bazerman (1988: 275), who declared that although in the 'last twenty years, a major style change in the [APA] psychological journal has [...] started to take place [as a] result of the rising influence of a cognitive psychology based on the computer model', the 'new style has not yet affected the *Publication Manual* in any significant way'. Comparison between the analysis of authorial presence across moves and steps have shown a highly articulated picture which could benefit from further analysis.

To conclude, despite any shortcomings this study might have, we believe that this paper may be of interest for the linguistic community, as it provides insight into a little analysed rhetorical practice of psychology researchers. By focusing on one single discipline, by considering all the main sections of psychology research articles, by performing detailed quantitative analysis on a full tagged corpus, and by analysing the distribution of overt authorial presence across both moves and sections, we hope to have covered a niche and provided hints for future research in the field.

NOTES

*A preliminary version of this study was presented at the ICAME 2006 Conference, 24-28 May 2006, Helsinki.

1 Interestingly, APA norms have been adopted as standards for publication also by journals of other disciplines, including linguistics.

2 http://andromeda.rutgers.edu/-jlynch/Writing/p.html

3 http://web.psych.washington.edu/writingcenter/writingguides/pdf/style.pdf

4 http://abacus.bates.edu/-ganderso/biology/resources/writing/HTWgeneral.html

5 Qualitative contributions on personal reference in academic writings are not taken into consideration here, though some interesting papers exist in the literature; among them: Harwood (2005), who proposes a functional taxonomy of self-promotional *I* and *we* and shows that scientific articles in the hard sciences carry a self-promotional flavour with the help of personal pronouns; Vladimirou (2007) who argues that expert writers in the 'academic discipline of linguistics' are surprisingly inconsistent in their use of personal reference.

6 For a discussion of the benefits of an integration of ESP and corpus approaches, see Flowerdew 2005.

7 The *Psychology Corpus* includes 67 empirical studies in psychology taken from 20 different international journals, with a total amount of 462,772 words. It exists in two parallel versions: an annotated one and a non-annotated one. When this study was carried out, the annotated version was undergoing revisions by the author; therefore only the part of the corpus that had been completely reviewed was used.

8 In the analyses carried out on the most general level, also the Note move was considered in so far as it contains further information about the experiment described in the paper. However, when getting to more specific types of analyses, this move was ignored, given the specificity of its steps and the low number of authorial references.

9 This trend was confirmed by the fact the only *my* appears in the discussion section and is in subject position.

REFERENCES

Bazerman, Charles (1988), *Shaping Written Knowledge. The genre and activity of the experimental article in science*. London: The University of Wisconsin Press.

Berkenkotter, Carol & Huckin, Thomas N. (1995), *Genre Knowledge in Disciplinary Communication: Cognition/Culture/Power*, Hillsdale - New Jersey: Lawrence Erlbaum Associates.

Bhatia, Vijay K. (1993), *Analysing Genre. Language Use in Professional Settings*, London & New York: Longman.

Bianchi, Francesca & Pazzaglia, Roberto (2007), "Student writing of research articles in a foreign language: metacognition and corpora". In: Facchinetti, Roberta (ed.), *Corpus Linguistics 25 Years on*. Amsterdam: Rodopi, pp. 261-289.

Bowker, Lynne & Pearson, Jennifer (2002), *Working with specialized language. A practical guide to using corpora*, London & New York: Routledge.

Dudley-Evans, Tony & St John, Maggie Jo (1998), *Developments in English for Specific Purposes. A multi-disciplinary approach*. Cambridge Language Teaching Library, Cambridge: Cambridge University Press.

Flowerdew, Lynne (2005), "An integration of corpus-based and genre-

based approaches to text analysis in EAP/ESP: countering criticisms against corpus-based methodologies", *English for Specific Purposes*, 24: 321-332.

Gläser, Rosemarie (1995), *Linguistic features and genre profiles of scientific English*, Frankfurt-am-Main: Peter Lang GmbH.

Gledhill, Christopher (1995), "Collocation and genre analysis: the phraseology of grammatical items in cancer research abstracts and articles", *Zeitschrift für Anglistik und Amerikanistik*, 43: 11-29.

Gledhill, Christopher (1996), "Science as collocation: phraseology in cancer research articles". In: Botley, Simon, Glass, Julia, McEnery, Tony & Wilson, Andrew (eds.) *Proceedings of the teaching and language corpora 1996*. Lancaster: UCREL: 108-126.

Gledhill, Christopher (2000), "The discourse function of collocation in research article introductions", *English for Specific Purposes*, 19 (2): 115-135.

Gopnik, Myrna (1972), *Linguistic structures in scientific texts*, The Hague: Mouton.

Gross, Alan G., Harmon, Joseph E., Reidy, Michael (2002), *Communicating Science. The Scientific Article from the 17th Century to the Present*, Oxford: Oxford University Press.

Harwood, Nigel (2005), "'Nowhere has anyone attempted... In this article I am 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.

Hyland, Kan (2001), "Humble servants of the discipline? Self-mention in research articles", *English for Specific Purposes*, 20: 207-226.

Hyland, Ken (2003), "Self-citation and self-reference: Credibility and promotion in academic publication", *Journal of the American Society for Information Science and Technology*, 54 (3): 251-259.

Hyon, Sunny (1996), "Genre in three traditions: implications for ESL", *TESOL Quarterly*, 30 (4): 693-722.

Kuo, Chich-Hua (1999), "The use of personal pronouns: Role relationships in scientific journal articles", *English for Specific Purposes* 18: 121-138.

Martín Martín, Pedro (2003), "Personal attribution in English and Spanish scientific texts", *Bells*, 12, retrieved on 10 Feb. 2006 from www.publicacions.ub.es/revistes/bells12/PDF/art09.pdf

Paltridge, Brian (1997), *Genre, Frames and Writing in Research Settings*, Amsterdam & Philadelphia: John Benjamins' Publishing Company.

Publication Manual of the American Psychological Association (2001), Washington, D.C.: American Psychological Association.

Rossini Faretti, Rema & Bondi Paganelli, Marina. (1988), *Il testo psicologico. Aspetti della traduzione e della lettura in lingua inglese*, Bologna: Pitagora Editrice.

Swales, John (1990), *Genre analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.

Vladimirou, Dimitra (2007), "'I suggest that we need more research'. Personal reference in linguistics journal articles". In: Gabrielatos, Costas, Slessor, Richard & Unger J.W. (eds), *Papers from the Lancaster University Postgraduate Conference in Linguistics and Language Teaching, 2006, Vol. 1: Papers from LAEL PG 2006*, Lancaster: Lancaster University.