Traces of self and others in research articles. A comparative pilot study of English, French and Norwegian research articles in medicine, economics and linguistics

KJERSTI RONGEN BREIVEGA
University of Bergen

TRINE DAHL
Norwegian School of Economics and Business Administration

KJERSTI FLØTTUM
University of Bergen

This article presents a pilot study which has been undertaken as preparation for a comparative research project called “Cultural identity in academic prose”. The general aim of the project is to study which aspects of scientific activity are most important for what we may call cultural identity in academic writing. Whether such identities are primarily national or discipline-specific is discussed. The project involves research articles from three disciplines – medicine, economics and linguistics – and three languages – English, French and Norwegian. The central questions are related to authorial presence and stance, to the manifestation of other researchers’ voices and to the authors’ promotion of their own research. This article takes a linguistic approach, and the pilot study focuses on the use of the following categories: first person pronouns, metatextual comments, explicit and implicit references and lexical items. The pilot study comprises 18 research articles; in the large-scale study the corpus will consist of about 500 articles. In the pilot study presented here the main finding is that the proposed categories seem to be well suited to the purposes of the large-scale study. The data also allow some preliminary hypotheses about ‘non-expressive medical researchers’, ‘shy economists’ and ‘polemic linguists’ to be formulated.
Introduction

How and to what extent do authors manifest themselves in a research article? How and to what extent are other researchers given the possibility of manifesting themselves? How does the writer promote his or her own research?

These are the central questions we have addressed in a pilot study which has been undertaken as preparation for a comparative research project called “Cultural identity in academic prose” (in Norwegian: Kulturell identitet i akademisk prosa, or KIAP). The general aim of the KIAP project is to study which aspects of scientific activity are most important for what we may call ‘cultural identity’ in academic prose. We want to investigate whether such identities are primarily national (is there for example a specific Norwegian academic identity?) or whether they are primarily linked to a specific discipline (e.g. a specific medical academic identity). The concept of cultural identity, which is a very complex one, will be delimited as the set of values and the cultural basis associated with language and discipline in academic prose. The KIAP project is based on a linguistic approach, and the central issue will be to investigate to what extent linguistic and/or textual features are indicative of national or discipline-specific scientific ‘identities’.

KIAP is a comparative study involving research articles from three disciplines – medicine, economics and linguistics – and three languages – English, French and Norwegian. The choice of these disciplines is motivated by several factors: they represent three basic scientific traditions – natural science, social science and the humanities, and it is reasonable to assume that there are discursive differences between them (Breivega 2000, 2001b). Furthermore, the project participants are familiar with these disciplines through previous research on academic discourse, acquired expert knowledge or departmental affiliation. As for the languages in question, the status of English as an academic lingua franca makes it a natural frame of reference for any contrastive study of academic discourse (see e.g. Mauranen 1993). Norwegian – like the other Nordic languages (cf. Laurén & Myking 1999) – is an interesting academic language because it is a ‘minor’ language and because so little has been done in the way of contrastive studies of academic prose which includes Norwegian material; French is by contrast a ‘major’ language which boasts a considerable scientific production. Despite this fact, French material, just like Norwegian, has only been very modestly studied in a contrastive perspective.

The selection of linguistic features or categories which may be relevant to our study is based on previous research in this domain (Vassileva 1995, 1998, 2000, 2001; Löffler-Laurian 1980; Mauranen 1993) as well as on our own observations and research (Breivega 1999, 2001a,b; Dahl 1996, 1998, 2000b; Fløttum 2000b, 2001a,b). Scientific writing has traditionally been looked upon as objective, matter-of-fact oriented and only marginally characterised by authorial presence, commitment and open argumentation. Previous research has, however, shown that academic prose is characterised by a number of pragmatic and context-bound factors and that subjective elements are clearly present (cf. e.g.
Bazerman 1988; Swales 1990; Hyland 1998a). Today, as researchers are faced with the ever-increasing competitive production of research articles, making the marketing of their own research more important, the use of subjective elements can also be seen as a way of promoting and selling their own results at the expense of other researchers’ work. This makes authorial stance, including overt positive and negative evaluations, a typical feature in research articles. However, such evaluations are not used in the same way and to the same extent within different languages and cultures or within different disciplines (cf. Berkenkotter & Huckin 1995). In short, our approach to the study of cultural identity will consist of a study of linguistic features associated with writer manifestation, writer promotion and manifestation of voices of other researchers in research articles. Through such linguistic features the writer’s voice, and hence the set of values on which his or her scientific identity is based, will emerge.

In order to prepare the ground for our large-scale study and to make sure that we focus on relevant linguistic categories which are present and used in all the three languages in question, we have carried out a small pilot study comprising 18 articles. In this pilot study, which will be presented here, we focus on some selected elements related to the three questions formulated at the beginning of this introduction. The questions involve certain lexico-semantic and functional categories which are related to argumentation and enunciation and which we consider to be relevant to our study.

Materials and method

In the KIAP project one of the aims is to establish an electronic corpus of about 500 articles. This corpus is under construction, and at the moment of writing it consists of 110 articles which are relatively evenly distributed across the three languages and the three disciplines in question. The availability of this corpus will be a great advantage to the project. The resulting opportunities for automatic searches on selected linguistic features (including collocations) will save a great deal of time. A crucial issue will be to identify linguistic features which will easily lend themselves to such a methodological approach (i.e. an automatic analytical process).

The articles constituting the KIAP corpus have been selected according to several criteria. They must be published in refereed journals. In order to keep as many factors as possible under control, we would also like the articles to be written by native speakers of the respective languages. Furthermore, the corpus as a whole as well as each subcorpus (consisting of the various combinations of discipline and language) should represent variation with regard to authorship and gender. Thus, both single and joint authorship as well as articles by female and male authors will be represented in the corpus. The first of these two variation axes is particularly crucial to the KIAP project: one central category under investigation is first person pronouns, and in order to identify possible variation patterns in the use of these pronouns it is necessary to establish a
corpus which provides the opportunity to study this phenomenon in both single and joint authorship contexts.

In addition to the criteria concerning authorship, there are some discipline-specific guiding principles. The medical articles should preferably present experimental and clinical research; we are less interested in articles discussing ‘soft’ issues like social, political, and economic factors affecting health. The economics articles should preferably not be dominated by numerical data. Finally, the linguistic articles should preferably discuss linguistic material in the author’s mother tongue. These discipline-specific criteria are, however, not absolute.

The pilot study is based on an analysis of 18 research articles, with two articles from each discipline in each language (see Appendix 1). These have been coded as follows:

<table>
<thead>
<tr>
<th>Language</th>
<th>Disciplines</th>
<th>Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Economics</td>
<td>EngEcon1, EngEcon2</td>
</tr>
<tr>
<td></td>
<td>Linguistics</td>
<td>EngLing1, EngLing2</td>
</tr>
<tr>
<td></td>
<td>Medicine</td>
<td>EngMed1, EngMed2</td>
</tr>
<tr>
<td>French</td>
<td>Economics</td>
<td>FrEcon1, FrEcon2</td>
</tr>
<tr>
<td></td>
<td>Linguistics</td>
<td>FrLing1, FrLing2</td>
</tr>
<tr>
<td></td>
<td>Medicine</td>
<td>FrMed1, FrMed2</td>
</tr>
<tr>
<td>Norwegian</td>
<td>Economics</td>
<td>NoEcon1, NoEcon2</td>
</tr>
<tr>
<td></td>
<td>Linguistics</td>
<td>NoLing1, NoLing2</td>
</tr>
<tr>
<td></td>
<td>Medicine</td>
<td>NoMed1, NoMed2</td>
</tr>
</tbody>
</table>

The electronic corpus has been constructed in parallel with the pilot analysis, and we have been working both manually and automatically with the texts. We feel that this was a good way to get to know the corpus. The quantitative results presented here are very modest, and our aim in the pilot study is primarily to determine some of the categories which will be used in the large-scale study.

The categories we have studied are related to our initial questions as outlined below.

A. HOW AND TO WHAT EXTENT DO AUTHORS MANIFEST THEMSELVES IN A RESEARCH ARTICLE?
Here we have basically limited our study to the distribution and frequency of first person pronouns and metatextual comments.

© Blackwell Publishers Ltd. 2002
First person pronouns
We have studied first person pronouns in different syntactic functions and in different forms (first person possessives are also included): 3

1a) I/we present . . . in my/our study . . .
1b) Je/nous présente/-ons . . . dans ma/notre étude . . .
1c) Jeg/vi presenterer . . . i min/vår studie . . .

Metacomments
This category is divided in two: locational and rhetorical. The first category includes expressions which point directly to the text itself or to a part of it, as in the following examples:

2a) In this article/section I propose the following hypothesis . . .
2b) Dans cet article/cette section je propose l’hypothèse suivante . . .
2c) I denne artikkelen/delen foreslår jeg følgende hypotese . . .

The second category includes expressions which refer to rhetorical strategies used in the text, such as introduce, present, summarise, conclude, exemplify, illustrate:

3a) We conclude that . . .
3b) Nous concluons que . . .
3c) Vi konkluderer at . . .

Other categories which will be crucial in the large-scale study are related to modality, but the complexity of that issue made it impossible for us to deal with it in the pilot study. 4

B. HOW AND TO WHAT EXTENT ARE OTHER RESEARCHERS GIVEN THE POSSIBILITY OF MANIFESTING THEMSELVES?
To find answers to this question we first discuss explicit references in different forms and then implicit references expressed by polyphonic constructions.

Explicit reference
Explicit references can be realised in different ways. The subcategories we have studied here are references given in notes, in parenthetical citations with name and year of publication, through introductions to the reference in the text and by quotations, illustrated in 4a–d respectively:

4a) This is shown in several studies . . . [+note number]
4b) This is shown in a recent study . . . (Smith 2000)
4c) In her study related to the external factors, Smith (2000) shows that . . .
4d) Smith (2000) puts it this way: “The external factors are important in . . .”
In addition to these categories, we have made a distinction between self-reference and reference to others in each subcategory.

**Implicit reference**

Researchers may, however, manifest themselves in more subtle ways than through explicit references. Their presence may be implicit, expressed by polyphonic constructions like syntactic negation with *not/ne* . . . *pas/ikke* in examples 5a–c below:

5a) The research article is *not* an interesting object of study.

b) L’article scientifique *n’est pas* un objet d’étude intéressant.

c) Den vitenskapelige artikkelen er *ikke* et interessant studieobjekt.

A polyphonic conception of meaning implies that several voices or points of view are signalled in discourse. In the examples above, two points of view are presented in the same utterance, one saying that ‘the research article is an interesting object of study’ and another saying that ‘this is not valid’. There are many constructions that may be studied in this context, but in the pilot study we have limited this aspect to syntactic negation followed directly by a sentence introduced by the connectives *but, however/mais/men*. These contrastive connectives represent a potential polemic related to the presented voices.

C. HOW DOES THE WRITER PROMOTE HIS OR HER OWN RESEARCH?

The third question which we consider in the KIAP project is a complex and comprehensive one. In fact, the categories and subcategories mentioned above will also be relevant in the search for answers to this question (see also note 4 on modality). However, in this part of the study we want to focus on how promoting is done by means of selections from the lexicon. Lexical words like *new, recent, original, useful* versus *old, traditional, futile* (i.e. words that signal evaluation in one way or another) will be of special interest here. We also want to look at the collocations where key concepts such as *hypothesis, theory, method* and *result* occur. In the pilot study we have limited our analysis to the words *new, recent(ly)* and *result(s)/finding(s)* and their French and Norwegian equivalents.

**Results and discussion**

We have studied the distribution and frequency of the different categories presented above in their local as well as global context, i.e. in the different article sections such as Introduction, Body, Conclusion (or for the IMRAD-structured articles, Introduction, Method/Material, Results and Discussion) and in the article as a whole. We have also studied potential differences between languages and between disciplines. However, since our pilot corpus is quite small, there is little point in presenting all these data here. See Appendix 2 for
the number of occurrences and the frequency of the categories in each article. Some details in the quantitative data will be commented on below.

Our main finding is that the proposed categories seem to be well suited to the purposes of the large-scale study. There are, however, some new and more clear-cut distinctions that need to be developed. Although there are clear individual differences between the articles, we have also observed that it will be necessary to consider the distribution and frequency of the categories in relation to the different parts of the articles.

We shall now look at the results and observations for the different categories, related to the three questions presented in the introduction (here marked A, B and C respectively). The presentation will be illustrated by authentic examples: elements relevant to question A appear in bold, elements related to B are underlined and elements related to C are in italics.

A1. FIRST PERSON PRONOUNS

Our results give no general indications about the use of first person pronouns. There is considerable variation between the articles. The article with the highest frequency (1 in 62 words) is NoMed1. The following example is taken from this text:

Vi benyttet medikamentet ved mye tromber før rekanalisering, lokale eller perifere tromber etter behandlingen og hvis TIMI 3-blodstrøm ikke ble oppnådd på mistanken om perifer embolisering. (NoMed1: 6)

It is rather surprising that it is the discipline of medicine which provides us with this result. We must, however, emphasise here, and for the rest of the quantitative indications, that it is quite impossible to generalise on the basis of such a small corpus. But the fact that there are such variations is interesting as a starting point for the large-scale study.

An important observation for the categories to be included in the large-scale study is that the singular/plural distinction is not sufficient to account for the French material. In the French articles it will also be necessary to look into the use of the pronoun on (English: one/you, Norwegian: man/en/ein). To some extent (for example in the medical and economics articles), it seems that on is used in the more objective Material/Method sections and that nous, or a combination nous/on is used in the more subjective Discussion section. In the following text segment, there is variation between on and nous:

Malgré la différence potentielle . . . , on tient pour acquis que . . . Les ajustements apportés . . . L’utilisation d’un indice fondé sur un système différent présente toutefois des inconvénients: nous ne pouvons tenir compte des différences . . . , et nous ne disposons pas non plus de renseignements sur les différences entre les États-Unis et le Canada . . . (FrMed2: 7)

Without going into detail here, let us just note that the reference of the on pronoun (both person and number) will represent a special challenge.
The *I* versus *we* perspective (see note 3) is also an interesting feature, but again one which will not be discussed in detail here. However, an example from English can be mentioned:

In this paper *I* argue that we should abandon a speech act theory approach . . . (EngLing2: 463)

For the French corpus, it is interesting that first person singular forms are found at all. Previous research (see e.g. Loffler-Laurian 1980) has tended to confirm the absence of such authorial presence. What is particularly interesting is the changing perspective from *je* to *on* or vice versa:


L’enjeu de ce que l’*on* appelle ici “jugement” est de rapporter le référent à une catégorie notionnelle. (FrLing1: 32)

This example is also interesting with regard to the context in which the pronouns appear: *je* is first used for self-reference, then in relation to three metacomments. In the next sentence, still in a context of metacomment, the author switches to *on*.

In one of the Norwegian linguistic articles, NoLing1, a single-author article, the frequency of first person pronouns is rather low. However, the contexts in which these pronouns are used seem to be significant. In the following example, the pronoun *jeg* is related to the expressing of a strong personal attitude:

*Så langt jeg kan forstå, er denne analysen neppe holdbar.* (NoLing1: 28)

‘As far as I can understand, this analysis is hardly tenable.’

In addition, there is an important presence of first person possessives in this article:

Etter min intuitjon er døs mer marginalt enn kastes, men bedre enn for eksempel *mått*. (NoLing 1: 10)

These first person possessives occur as premodifications of nouns such as *intuisjon*, (‘intuition’), *syn* (‘view’, ‘attitude’), and *språk* (‘language’). Hence, in this linguistic article they occur in contexts where the author is presenting potentially strong views on the language (use) under study.

The following two examples, taken from NoEcon2, have the first person singular pronoun *jeg*, but this is not related to the expression of a personal view or attitude (‘I [will] start by . . . ’):

*Jeg vil starte med å se på pris- og produksjonsutviklingen på et frikonkurransemarked.* (NoEcon 2: 44)

*Jeg starter med å se på et perfekt kartell . . .* (NoEcon 2: 50)
A final comment should be made about the medical texts. The typical medical article (see Breivega 2001b), adhering to the IMRAD structure, seems to be a description of actions rather than a discussion presenting different views or attitudes. The following examples may serve as illustrations of this phenomenon:

Vi benyttet medikamentet ved mye tromber før rekanalisering, . . .
Vi behandlet andre årer hos ti pasienter.
Allerede i dag utfører vi angioplastikk i samme prosedyre som den diagnostiske undersøkelsen før de fleste pasientene.
Dette løste vi ved at vakthavende kardiolog og en sykepleier fra overvåkingsavdelingen var med. (NoMed1: 6)

These examples constitute the we-context in presenting the authors as acting subjects. The comments are descriptions of the actions which have been carried out to obtain the results presented (cf. the verb forms benyttet ‘used’, behandlet ‘treated’, utfører ‘perform’ and løste ‘solved’).

In general we consider the study of first person pronouns in different forms and syntactic functions and in various semantico-pragmatic contexts to be interesting as regards distribution and frequency. On the basis on this pilot study, we can in fact conclude that authorial presence signalled by the use of first person pronouns represents an important contribution to the determination of cultural identity in academic discourse.

A2. METACOMMENTS
Like the previous category, the analysis of the metacomments gives us no clear indications about the use of this kind of authorial presence. The only fact that can be mentioned is that metacomments do not seem to be very frequent. No clear distinction has been observed between languages or between disciplines, and rhetorical comments in particular seem to be rare. The frequency of the two categories taken together varies from 120 (EngEcon1) to 578 (FrLing2). The relatively low frequency of metacomments in the linguistic articles (except NoLing2) is rather surprising. Previous research (see for example Breivega 2001b) gives reason to expect more. On the other hand, we are not very surprised by the finding that in one of the medical articles there is just a single metacomment (NoMed2):

Treffsikkerheten for cintigrafi er vist i tabell 1. (NoMed 2: 4)

Here are some other examples:

Locational

Table 1 shows the demographic characteristics and the frequency of adverse outcomes in the study group. (EngMed1: 4)

In Section 3 we assume parameter values for technology such that multinationals do not exist . . . (EngEcon1: 212)
Les deux méthodes mentionnées ci-dessus . . . (FrMed2: 3)
‘The two methods mentioned above . . .’

I dette arbeidet vil vi beskrive våre erfaringer fra angiografilaboratoriet . . . (NoMed 1: 1)
‘In this paper we will describe . . .’

Rhetorical

Summarizing, we find that there is divergence of country endowments if $t$ is sufficiently low . . . (EngEcon1: 229)

However, as far as the early modern form of the variety is concerned, we conclude that the shape of New Zealand English, a fascinating laboratory for the study of linguistic change, can be accounted for . . . (EngLing1: 316)

Je donne ci-dessous à titre d’illustration . . . (FrLing1: 32)
‘I give the following as an example . . .’

Eg presenterer først generelle reglar, deretter meir spesielle reglar som overstyrer dei generelle reglane . . . (NoLing2: 34)
‘I present first . . .’

The French example (FrLing1: 32) is in fact typical of the rhetorical metacommentary that both a locational (ci-dessous) and a rhetorical one (donne . . . à titre d’illustration) are combined in the same sentence. As all these examples show, metatextual expressions tend to co-occur with a first person pronoun. In one of the articles, NoEcon1, the main function of first person pronouns (vi) seems to be the realisation of different metastatements. In fact, the occurrences of vi are to a surprisingly low degree related to the expression of the author’s personal view:

I første del av denne artikkelen ser vi på en stilisert markedscontekst hvor en regulert dominerende aktør møter konkurranse fra en uregulert konkurrent med markedsmakt. (NoEcon1: 138)
‘In the first section of this article we look at . . .’

Such observations will be subject to more careful analysis in the large-scale study. In any case, personal pronouns and metacommentary are instances of writer manifestation, since they are indications or traces of a person ‘behind’ the text who organises the content. It is symptomatic that definitions of metadiscourse often include modality constructions and other traces of authorial presence (see e.g. Vande Kopple 1985: 83), and previous research on meta-expressions in academic discourse has often been carried out along these lines (see e.g. Hyland 1998b). However, in the KIAP project we will treat personal pronouns, metatextual expressions and modality constructions as separate features.
justify these distinctions by referring to the importance of studying the **different kinds of presence** which they manifest.

Our pilot study has also revealed that it may be fruitful to establish a third category of metacommments. This category may be called **instructions to the reader** and is related to how the article should be read. Such instructions manifest themselves by more or less deontic expressions such as *il faudra/on devra* ('it is necessary/one must'), *cela vise à préciser que...* ('it aims to specify that'), or as in the following authentic examples:

The important point to note is that ... (EngEcon1: 230)

Note that in our model ... (EngEcon2: 265)

Il est à noter que ... (FrMed1: 3).

One can discuss whether such expressions should be placed in the metacomment group; however, this issue will not be pursued further here.

**B1. EXPLICIT REFERENCE**

Perhaps the most conspicuous manifestations of the voices of other researchers are in the form of explicit references to their works. The general finding of our pilot study is that references are indeed an important part of academic discourse. References made by the author to insights gained through earlier work is a characteristic feature of the three languages and disciplines we have studied. In general, it seems that the medical articles tend to display a relatively high frequency with regard to explicit reference; however, medical researchers do not let other researchers manifest themselves in the text - they are only referred to in endnotes, by name, title and year of publication (see also Salager-Meyer 1998).

The linguistic articles show particularly great variation in the total frequency of references (for NoLing1 the frequency is 53, highest of all the articles, and for FrLing1 the frequency is 396, the lowest). In addition, variation with regard to the kind of reference is important (see below). Our pilot study indicates that there seems to be more variation between disciplines than between languages. It also seems worthwhile to study the distribution of references with regard to the different sections of the articles.

As mentioned above, explicit references can be realised in different ways;¹⁰ we have studied the following subcategories (presented with examples):

a) **References given in notes:**¹¹

Smoking is one of the strongest risk factors for adverse perinatal outcomes, (16) ... (EngMed1: 6)

Au Québec, les anomalies congénitales représentent la deuxième cause de mortalité prénatale et la sixième cause, ..., de mortalité prématurée. (FrMed1: 1)
Den største ulempen ved trombolytisk behandling er manglende reperfusjon hos bortimot 30–40% (1). (NoMed1: 1)

b) References given in parenthetical citations with name and year of publication:

In emerging markets free deposit insurance extended to banks in newly liberalised financial markets has long been cited as a source of instability in financial markets (McKinnon and Mathieson, 1981; Hanson and de Melo, 1983; DiazAlejandro, 1985; Corbo et al., 1986; Ballón, 1991; McKinnon, 1991; Velasco, 1991). (EngEcon2: 266)

Ceux-ci [des critères] permettent de comparer les distributions de revenus au vu de certains critères et notamment la taille (ATKINSON, BOURGUIGNON [1987], MAURIN [1986]). (FrEcon1: 270)

Det norske genussystemet har, som alle andre genussystem som er analysert til no, (jf. Corbett 1991), ein semantisk kjerne. (NoLing2: 35)

c) References with introductions:

Dixit and Norman (1980) further increased the generality of basic factor-proportions theory and clarified the problem of factor-price equalization (FPE), in particular by deriving a factor-price equalization set within a traditional Edgeworth box. (EngEcon1: 210)

This contention is supported by Bauer (1999), who supplies very strong lexical evidence in favour of the mixed origins of New Zealand English. (EngLing1: 308)

De nombreuses formes fonctionnelles sont estimées dans la littérature (voir LECHENE [1993] pour un survey sur ce sujet). (FrEcon1: 271)

Afin de disposer d’estimateurs convergents, on utilise une méthode proposée par CHAMBERLAIN [1984] qui consiste à se ramener à un modèle dans lequel . . . (FrEcon1: 272)

Med Corbetts Gender frå 1991 kom synet på genustilordning som grammatisk komponent inn i det lingvistiske rampelyset . . . (NoLing2: 32)

Dobrin (1997) hevdar at genustilordningsreglane står i eit hierarkisk tilhøve til kvarandre . . . (NoLing2: 32)

d) References with direct quotations:

. . . according to de Klerk (1996: 10), who writes that “the two basic formative elements of modern South African English were the largely vernacular settlements of the 1820s and the largely standard-speaking Natal settlement of the 1840s”. (EngLing1: 312).

Comme le souligne Walras, “on n’offre pas pour offrir, . . .” [Walras (1874), p.76]. (FrEcon2: 119)

© Blackwell Publishers Ltd. 2002
Hockett's definition is: “Genders are classes of nouns reflected in the behaviour of associated words” (Hockett 1958: 231). (NoLing2: 29)

As mentioned above, in addition to these categories, we have made a distinction between reference to works by the article author(s) (self-reference) and reference to works by other scholars (reference to others) in each subcategory. The following are examples of self-reference in category c, introduced reference:

One major result of this paper, that multinationals occur when countries are similar in size and relative endowments, is also found in our 1998 paper (Markusen and Venables, 1998). (EngEcon1: 213)

Nous avons montré, dans Gosselin (1996a et b, 1999a et b), que les conflits sont . . . (FrLing2: 63)

Våre tidligste erfaringer ble beskrevet i 1998 (8), og resultater fra angioplastikk hos pasienter som overlevde hjertestans ble publisert i 1999 (9). (NoMed1: 1)

It will be particularly interesting to study these self-references related to the (non-)use of first person pronouns. For example, note the author’s (Gosselin’s) use of the construction with lequel to refer to himself and his own research in the following example (compared to the use of nous in example FrLing2: 63 above):


As for references to other researchers, it will be interesting to study their lexical context in order to answer the questions about promotion (are references to others used to support one’s own research or to refute work by other researchers?).

B2. IMPLICIT REFERENCE
We restricted the pilot analysis to syntactic negation with not/ne pas/ikk(j)e followed by the contrastive or concessive connectives but, however/mais/men. This is of course a very limited selection, and we did not expect to find many occurrences. However, our small corpus did turn up some tokens, which implies that it may be interesting to look for this polemic construction in the large-scale study. Here are some examples:

Ils [ces résultats] montrent en effet que ce n’est pas tant la présence de l’enfant . . . qui réduit l’aisance financière, mais que celles-ci étaient déjà moindre auparavant. (FrEcon1: 290)

Det har ikke vært behov for å innkalle ekstra personell under prosedyren, men det har vært nyttig med flere enn to til stede. (NoMed1: 6)
We also observed a relatively high frequency of negations alone. This result lends further support to the hypothesis according to which polyphonic constructions indicate not only an exchange of views but an exchange which is polemic. This is of interest to our general question about the existence of cultural identity in academic prose. Finally, we found many other traces of polyphony which require further analysis in the same perspective, such as connectives other than the ones identified here.

C. LEXICAL PROMOTION
The findings in this part of our pilot study are rather meagre, and we have therefore chosen not to present them quantitatively. Some comments are, however, relevant.

We selected the lexemes new, recent(ly) and result(s)/finding(s) and their French and Norwegian equivalents. The material displayed occurrences of all of these. We also found support for our hypothesis that it would be fruitful to study key words like method, data/materials, results etc. in a textual perspective. We need to examine their use and context with a view to identifying semantic isotopies, both positive and negative ones. In addition, we must be careful to distinguish between the good/bad dimension on the one hand and the dimension involving author(s)/other researchers on the other. An occurrence like ‘interesting result’ can refer to the findings reported in the article itself or to findings obtained by other researchers; in the same way, a ‘new method’ can refer to the method used by the author or to a method used by other researchers. We also have to make other distinctions regarding the relation between the expression used and its semantic reference. In the KIAP project our main concern will be expressions used to describe the actual research process as well as the writing process itself, rather than expressions used to describe various aspects of the object of study. Hence, the following example where the author mentions a new verb (det...nye verbet) in Norwegian is not very interesting to our study:

Vi finner også en viss evidens ved det relativt nye verbet synse. (NoLing1: 19)

We end this section with some further examples from our material which clearly indicate the importance of the lexical dimension of the study:

Our findings in non-smoking mothers suggest that the positive associations previously reported among first births might simply reflect inadequate adjustment for confounding variables. (EngMed1: 6)

In addition to the new results we derive, we hope that combining these themes in a single framework provides a useful synthesis of a number of strands in modern international economics. (EngEcon1: 212–3)

This result provides a link between Krugman’s work on economic geography (Krugman, 1991) and more conventional trade models. (EngEcon1: 231)
En combinant ces données . . . , nous croyons ainsi obtenir des résultats plus précis . . . (FrMed1: 2)

Le résultat n’est guère surprenant pour le diplôme, mais moins immédiat en ce qui concerne . . . (FrEcon1: 286)

Ce nouvel argument sera un élément de la détermination du processus de consommation. (FrEcon2: 129)

Helt nylig er det foreslått å benytte et paraplyfilter (Angioguard) for å hindre perifer embolisering, det samme som nylig er tatt i bruk ved angioplastikk av halskar. (NoMed1: 6)

Likevel synes ikke disse fordelene å være avgjørende for et godt resultat. (NoMed2: 6)

Examples like these are directly relevant to our question C, related to the promotion of one’s research, but also to question A, related to the manner in which authors manifest themselves. A final example illustrates the use of interesting lexical items other than the ones analysed in the pilot study:

L’idée traditionnelle . . . se trouve donc ici réfutée . . . (FrLing1: 36)

Conclusion

In this article we have presented the main goals and research questions of the KIAP project, as well as our very preliminary findings based on a small pilot study of 18 texts. With this pilot study as our basis, in addition to related research in the field of academic discourse, we hope to be able to carry out a large-scale study with the intention of shedding some light on different cultural identities in academic prose. The main emphasis of the analysis will be on the individuals (i.e. the writer(s) and other researchers) ‘behind’ the research activity and their specific attitudes, opinions and values, as realised by specific linguistic expressions.

Some of the hypotheses which we would like to investigate can be formulated as follows:

- Medical researchers are non-expressive writers who do not let other researchers be heard in their texts.
- Economists are shy writers who avoid promoting themselves by means of personal pronouns in connection with stating personal opinions.
- Linguists are polemic writers who involve several other researchers in the text, often to refute their points of view.

The results which have been obtained so far do not indicate any clear language differences, possibly apart from the very low frequency of first person singular
pronouns in French. This will of course be investigated further in the large-scale study. However, even at this point it seems that our initial hypothesis (Fløttum 2001c) – that cultural identity is more likely to be related to discipline than to language – will find support in our material (see the hypotheses above).

To conclude, we would like to emphasise the linguistic approach undertaken in KIAP. This we hope will be a fruitful contribution to a research field where previous research has been predominantly conducted along diachronic, pedagogical, literary and rhetorical lines (cf. Bazerman 1988; Swales 1990; Hertzberg 1995; Ventola & Mauranen 1996; Flyum 1996). The aim of the project is two-fold: we want to present quantitatively valid results with regard to the frequency of use and distribution of different categories, and we want to present results based on the qualitative study of selected categories, such as the use and context of first person pronouns, metacomments, various types of explicit and implicit reference, various lexical items and modality constructions.

Notes

1. The project was partly financed by the Meltzer foundation for five months in 2001 and will be financed by the Research Council of Norway during 2002–2004. Kjersti Fløttum is the project leader of KIAP.
2. Another important goal in the KIAP project will be to establish an electronic corpus consisting of about 500 research articles. A third goal will be to identify linguistic features which can be used in the generation of acceptable summaries of argumentative research articles (see Dahl 2000a,b).
3. We consider the study of the use of first person pronouns rather than noun phrases, third person pronouns or impersonal constructions – in different syntactic functions and in different forms – to be interesting regarding their distribution and frequency. In our study, however, the degree of authorial presence signalled by the use of first person pronouns is interesting in itself, as a contribution to the determination of cultural identity in academic discourse. For example, previous research indicates that French authors are less visible on the surface of the text than English authors (cf. Vassileva 1998). The I versus we perspective, as Vassileva puts it, is another interesting contrastive feature.
4. Modality is a very complex concept which can be defined and delimited in several ways. In some linguistic traditions the concept of modality is more or less exclusively associated with expressions of possibility, necessity, permission and obligation (i.e. epistemic and deontic modality). Other approaches take a broader and almost all-inclusive view of modality which would seem to cover “the set of linguistic phenomena which signal the presence of man in language” (Herslund 1989: 7; see also Löffler-Laurian 1980; Palmer 1986; Hunston & Thompson 2000). Our working definition of modality will represent a stance somewhere in between these extremes, covering linguistic categories which signal the speaker’s attitude towards the propositional content. This will also allow us to include in our study attitudes other than the strictly epistemic and deontic. However, we wish to restrict modality to expressions signalling some kind of attitude, thus excluding authorial presence per se (e.g. the mere use of first person pronouns). We will consider the following two main types of modality: epistemic modality (i.e. attitudes on the scale ‘certain’ vs. ‘uncertain’, ‘true’ vs. ‘false’) and evaluative modality (i.e. attitudes on the scale ‘good’ vs. ‘bad’). As regards epistemic modality constructions, we want to focus on...
especially on the phenomenon of hedging, which is very common in academic writing (see Zuck & Zuck 1986; Hyland 1998a).

5. The concept of polyphony originates from Bakhtine (1970), and a genuinely linguistic version of the theory of polyphony was proposed by Ducrot (1984). The study of this kind of implicit presence of other researchers will be based on more recent work within the NOS-H project “Linguistic and literary polyphony” by Nølke (1994, 2001) and Fløttum (1999, 2000a).

6. In the KIAP project we will carry out a basically lexico-semantic analysis of lexical words used in the research articles, taking our cue from the works of Dahl (1998, 2000b) on the lexical patterns in research articles (see Hoey 1991). These lexical patterns will be subjected to a more thorough semantic analysis based on Rastier’s interpretative theory of semantics (Rastier 1987, 1997; Malrieu & Rastier in press). The content of the lexical words chosen in order to convey a message or an information package can be specified in semantic features (sèmes). Such semantic features can be realised through different lexical choices. When specific semantic features are repeated, semantic chains or isotopies will be formed. We especially wish to identify positive isotopies (phenomena which are described positively) and negative isotopies (phenomena which are described negatively).

7. As for the division of the articles into different sections, we see clear discipline differences: the medical articles are organised according to a strict IMRAD pattern, while the linguistic articles are divided into sections with thematic subtitles. The economics articles seem to be situated somewhere between, without a strict IMRAD structure but sometimes with subheadings which are related to this structure.

8. Page references for the medical articles refer to the electronic version of these articles.

9. NoMed2 has only one metacomment; this result seems to be rather exceptional and is therefore not taken into consideration when it comes to the comparative aspect.

10. Our investigation of reference practices in the research articles will primarily be based on Rosier (1999), a monograph on citations and reported speech in French, on Tuomarla (1999), a study of the same phenomena in French scientific prose, and on various other studies carried out within the international research network CIDIT for studies on citation and reported speech strategies.

11. It is important to distinguish notes which give only bibliographical references (mostly in medical articles) from notes which contain information or comments of other kinds.

12. These hypotheses are of course very general and rather categorical. However, there are clear indications in the material we have analysed so far, as well as in previous research and observations, which point in the directions suggested.

13. However, even in single-author linguistics articles in Norwegian the use of the singular form jeg/eg is rare.

References


© Blackwell Publishers Ltd. 2002


*Travaux du cercle linguistique de Copenhagen*, vol. XXIII. Copenhagen. 
7–15.

E.B. Johnsen, *Virkelighetens forvaltere. Norsk sakprosa. Første bok*. Oslo: 
Universitetsforlaget. 187–205.


the construction of discourse*. Oxford University Press.

Hyland, K. (1998a) *Hedging in scientific research articles*. Amsterdam: John 
Benjamins.


fagspråkstudiar*. [Nordica Bergensia 20].


(http://www.msh-paris.fr/texto/)

study*. Frankfurt: P. Lang.


France.

France.


research, knowledge and society*. Copenhagen Business School. 495– 
504.

Cambridge University Press.

du discours rapporté direct*. Academia Scientarium Fennica, Helsinki. 
Humaniora 308.

*College Composition and Communication* 36.1: 82–93.

written discourse in English, Bulgarian and Russian. *International Journal 
TRACES OF SELF AND OTHERS IN RESEARCH ARTICLES


[Received 17/2/02; revised 20/5/02]

Kjersti Fløttum
Department of Romance Studies
University of Bergen
Øysteinsgate 1
N-5007 Bergen
Norway
e-mail: kjersti.flottum@roman.uib.no

Appendix 1 – Pilot study material

(The medical articles were accessed on the internet, so no page numbers are given for them here.)

EngEcon1

EngEcon2

EngLing1

EngLing2

© Blackwell Publishers Ltd. 2002


Appendix 2 – KIAP pilot study

<table>
<thead>
<tr>
<th></th>
<th>Total words</th>
<th>1st p. pronouns</th>
<th>French on</th>
<th>Metacomments</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>total/freq.</td>
<td>total/freq.</td>
<td>total/freq.</td>
<td>total/freq.</td>
</tr>
<tr>
<td>EngEcon1-j</td>
<td>8773</td>
<td>88/100</td>
<td></td>
<td>73/120</td>
<td>31/286</td>
</tr>
<tr>
<td>EngEcon2-s</td>
<td>7280</td>
<td>28/260</td>
<td></td>
<td>19/383</td>
<td>43/169</td>
</tr>
<tr>
<td>EngLing1-j</td>
<td>8427</td>
<td>126/67</td>
<td></td>
<td>26/324</td>
<td>52/162</td>
</tr>
<tr>
<td>EngLing2-s</td>
<td>10056</td>
<td>57/176</td>
<td></td>
<td>23/437</td>
<td>76/132</td>
</tr>
<tr>
<td>EngMed1-j</td>
<td>2257</td>
<td>16/141</td>
<td></td>
<td>8/282</td>
<td>39/58</td>
</tr>
<tr>
<td>EngMed2-j</td>
<td>2593</td>
<td>2/1297</td>
<td></td>
<td>9/288</td>
<td>23/133</td>
</tr>
<tr>
<td>FrEcon1-s</td>
<td>7750</td>
<td>2/388</td>
<td>74/105</td>
<td>23/337</td>
<td>48/161</td>
</tr>
<tr>
<td>FrEcon2-j</td>
<td>4600</td>
<td>37/124</td>
<td>4/1150</td>
<td>16/288</td>
<td>68/68</td>
</tr>
<tr>
<td>FrLing1-s</td>
<td>9100</td>
<td>18/506</td>
<td>114/80</td>
<td>27/337</td>
<td>23/396</td>
</tr>
<tr>
<td>FrLing2-s</td>
<td>5200</td>
<td>22/236</td>
<td>31/168</td>
<td>9/578</td>
<td>53/98</td>
</tr>
<tr>
<td>FrMed1-j</td>
<td>2550</td>
<td>14/182</td>
<td>8/318</td>
<td>5/510</td>
<td>30/85</td>
</tr>
<tr>
<td>FrMed2-j</td>
<td>3500</td>
<td>9/389</td>
<td>21/167</td>
<td>11/318</td>
<td>29/121</td>
</tr>
<tr>
<td>NoEcon1-j</td>
<td>5157</td>
<td>55/94</td>
<td></td>
<td>38/136</td>
<td>32/161</td>
</tr>
<tr>
<td>NoEcon2-s</td>
<td>7100</td>
<td>14/507</td>
<td></td>
<td>37/192</td>
<td>52/137</td>
</tr>
<tr>
<td>NoLing1-s</td>
<td>7700</td>
<td>54/143</td>
<td></td>
<td>28/275</td>
<td>144/53</td>
</tr>
<tr>
<td>NoLing2-s</td>
<td>8000</td>
<td>85/94</td>
<td></td>
<td>63/127</td>
<td>49/163</td>
</tr>
<tr>
<td>NoMed1-j</td>
<td>1850</td>
<td>30/62</td>
<td></td>
<td>8/231</td>
<td>22/84</td>
</tr>
<tr>
<td>NoMed2-j</td>
<td>2500</td>
<td>24/104</td>
<td></td>
<td>1/2500</td>
<td>46/54</td>
</tr>
</tbody>
</table>

s = single author  j = joint authors  freq. = average of 1 occurrence in x no. of words