

## AUTHOR IDENTITY IN ENGLISH ACADEMIC DISCOURSE: A COMPARISON OF EXPERT AND HUNGARIAN EFL STUDENT WRITING\*

KRISZTINA KÁROLY

School of English and American Studies  
Eötvös Loránd University  
Rákóczi út 5.  
H-1088 Budapest  
Hungary  
karolyk@ludens.elte.hu

**Abstract:** This paper investigates the expression of authorial identity in English published research articles and Hungarian EFL students' MA theses written in English. Based on the analysis of a computer corpus of 100 research papers in applied linguistics, it compares the frequency of one indication of author identity, author pronouns, and the rhetorical functions for which these pronouns are used in the two corpora. Frequency counts show that student writers employ a considerably higher number of author pronouns than expert writers do. Concordances of the data reveal that qualitative differences also appear in the two corpora regarding the rhetorical functions fulfilled by the personal reference used. The study also reports on the results of interviews conducted with student writers to reveal the reasons for their performance and shed light on their perceptions of pronoun use in English academic discourse. The interviews portray a considerable discrepancy between what students claim about the nature of academic discourse and what they actually do in writing. It is therefore argued that EFL students need special training in EAP courses that pays attention to the various aspects of style and takes into account the students' stereotypical cultural, linguistic, academic and educational background.

**Keywords:** English for Academic Purposes, research articles, theses, author identity, personal reference

\* I am grateful to the Hungarian National Scientific Research Fund (project number: OTKA 047017) for granting financial support for the current research project.

## 1. Introduction

It is a well established view in the literature that English has become the lingua franca of academic communication (e.g., Clyne 1987; Crystal 1995; Kaplan 1983; 1989; 1993; Medgyes–Kaplan 1992; Medgyes–László 2001; Swales 1990). Medgyes and László go even further to argue that “in a world that has granted English the privilege of *primus inter pares*, native English speakers have a distinct advantage over non-natives and, by implication, non-native speakers of English are handicapped in all areas of competence and discourse” (*op.cit.*, 261). Therefore special emphasis is generally placed in tertiary education on helping students studying or majoring in English master the prevailing conventions of English academic and formal writing, as they benefit from it not only during their university and (potentially) later academic career, but—should they opt for a different career—also in their later professional life. This applies to the situation in Hungary too, where, especially in the past 15–20 years, a significant increase could be witnessed in using English as the vehicle of scholarly and professional communication. Some have even gone as far as claiming that Hungary may in fact gradually be included among the countries where English is used as a second rather than a foreign language (Kachru 1985; Medgyes–László 2001).

Despite the stated importance of being aware of the norms of Anglo-Saxon academic and formal writing, however, there seems to be no unanimous agreement either in theory or in writing pedagogy regarding a crucial stylistic aspect of writing, namely the appropriateness of the use of *I* (and person markers in general) in written academic discourse. Disagreement may be encountered in research (Hinkel 1999; Jordan 1997) as well as in writing pedagogy (e.g., in writing manuals and text books; for a thorough analysis of the advice communicated by textbooks in the past thirty years see Chang–Swales 1999). A similar lack of consensus may be perceived among the various academic environments, too, where some professionals tend to categorically resist its legitimacy, while others consider it adequate under certain conditions. What all this boils down to is that there can be no overall consensus as to the use of personal reference because it arises out of disciplinary norms and practices.

The establishment of an effective authorial identity is of paramount importance in creating a successful argument (Hyland 2002b, 354). The notion of authorial (or writer) identity is used here as suggested by Ivanič (1997) to refer to all those aspects of writing that reflect the writer’s

relationship with his/her text; more precisely, “how authoritative s/he feels, what s/he wants to say, how s/he wants to represent him/herself in the writing, and the conflicts s/he faces between what s/he might ideally want and the constraints imposed by conventions” (Clark–Ivanič 1997, 134).

There are several forms of discursive identity construction: the various adverbial stance types as described by Biber–Finegan (1988), the anticipatory *it* (e.g., *It is interesting to note that...*) as shown by Hewings–Hewings (2002), or personal reference (Harwood 2005a;b;c; Hyland 1999; 2002a;b) to just mention a few. This study seeks to explore the latter, “the most visible expression of a writer’s presence in a text” (Hyland 2002b, 354), that is, the use of first person reference. More precisely, the article reveals how undergraduate students majoring in English with a Hungarian L1 background employ personal pronouns in expressing authorial identity compared to expert writers. The study is hoped to provide implications for the teaching of English for Academic Purposes (EAP) in contexts where, although being a dominant language of communication, English is not the official language. Besides the controversy mentioned above, another motivation for the current investigation is Hyland’s (2002b) study, which reports that as a result of the advice communicated by many writing text books and writing instructors, Hong Kong students tend to considerably underuse writer pronouns compared to expert writers.

Replicating, in part, Hyland’s (2002b) research in the Hungarian context, the main purpose of this study is therefore to compare the use of personal reference that expresses the writer’s stance, also referred to as “author pronouns” (*op.cit.*, 352; singular: *I, my, me*; plural: *we, our, us*) in English published research articles and in Hungarian students’ MA theses, both written in the field of applied linguistics. More precisely, it aims to compare the frequency of author pronouns in the two corpora and the rhetorical functions for which the pronouns are used. To shed light on some of the reasons why students perform the way they do in their texts, the study also explores what students consider as prestigious writing in the particular university context to which they belong, how they feel about the use of personal reference in academic writing in general, and what reasons motivate them in their choice of using or avoiding personal pronouns.

## 2. Author identity in academic discourse

### 2.1. The role of the discourse community in shaping writer identity

A discourse community is a group of individuals who share a set of social assumptions and routines in pursuing some common objectives (Swales 1990). As Flowerdew (2000) argues, “the notion of discourse community is relevant in the study of academic literacy because it stresses the participatory, negotiable nature of learning and the fact that learning is not always based on overt teaching” (*op.cit.*, 128). It shapes the conventions of academic genres and has also been shown to play a decisive role in shaping the generic competence of young scholars (Biber et al. 2002; Flowerdew 2000; Hyland 1999). Following Clark–Ivanič (1997), by conventions we mean “abstract ‘rules of behaviour’, or prototypical ways of doing things: practices that are ratified by the social, cultural and institutional context” (*op.cit.*, 137). One needs to get used to and learn the specialized discourse of the particular discipline and employ the conventions appropriately to obtain and/or maintain community membership.

Research in academic literacy has extensively dealt with the complexity of this task in the various disciplines (e.g., Harwood 2005a;b;c; Hyland 2002a; Petch-Tyson 1998; Shaw 1991) and in the case of native speakers (for an overview see e.g., Atkinson 1999; Flowerdew 2000; Clark–Ivanič 1997), but considerably less attention has been devoted to non-native speakers, not to mention the case of EFL learners, the main focus of the current study. Newcomers in the discourse community—even in their L1—need to develop a new **identity** (Ivanič 1997) or authorial **persona** (Hyland 1999, 120) to be able to effectively and adequately represent themselves, their positions and their readers. In doing so, as Clark and Ivanič (1997, 135–6) also point out, there are several intervening factors at play: personal history, idiosyncrasy, the requirements of the discipline and its stereotypical writing conventions.

The EFL context makes this task even more difficult by having to create a new writer identity not in one’s native language but in a foreign language, where one is likely to be required to fit different cultural/discourse conventions (Connor 1996). In such a situation, several additional factors complement the process: the individual’s L1 and FL writing competence, the L1 writing conventions acquired through education and previous writing experience, and the (potentially different or conflicting) requirements and discourse conventions of the given discipline in L1 and FL. For the participants of the current investigation obtaining adequate

generic competence in EAP is further complicated by the fact that they write in a FL (English) for a predominantly Hungarian (and not English) L1 academic audience, who also have their stereotypical requirements and conventions (to be described in section 3) conditioned by the fact that they work in the Hungarian university context. When producing EFL academic discourse, all of these factors need to be known, considered and harmonized by the novice science writer.

Students strive to meet the expectations of teachers and the discourse community and conform to the conventions and requirements, as these seem to be associated with success. As the above list of factors shows, novice science writers—particularly if their L1 is not English—need to possess a **“complex set of skills** which are only partially definable in linguistic terms” (Grabe–Kaplan 1996, 171, emphasis mine). Helping them in acquiring these skills poses an extra challenge for EFL writing pedagogy. The present study is intended as one step further in understanding how one means of expression (the use of first person reference) is applied by expert and novice writers in order to aid and inform the teaching of the elusive concept of style.

## 2.2. Conceptions of style in written academic discourse

Stylistic appropriateness, due to the complexity of the notion, is difficult to obtain even in one’s L1, but it is even more so in a SL/FL. On the one hand, this difficulty originates from the fact that languages, cultures, disciplines, discourse communities, institutions, text types differ in terms of their preferences regarding what is considered to be stylistically appropriate writing (Malcolm 1999). On the other hand, it is also important to note that “all conventions are not equal: some conventions are more ‘conventional’ in the everyday sense than others” (Clark–Ivanič 1997, 138). At a given historical moment, some conventions may have a more privileged position than others, but with language dynamically changing over time these conventions may also lose their earlier prestige.

The change of conceptions regarding the dominantly impersonal tone attributed to academic writing neatly illustrates the elusive nature of certain norms. Students have long been instructed to avoid the personal tone (i.e., avoid the use of e.g., *I*) in academic writing and produce impersonal, factual discourse (i.e., use, for instance, passive) and thus increase their credibility. Drawing on Clancy and Ballard’s (1992) work, Jordan (1997, 244) summarizes the features whose combinations are claimed to

account for the formal and academic nature of scientific discourse. According to this description, “the academic writer’s *tone* is: serious, impersonal, formal *rather than* conversational, personal, colloquial” (italics in the original). Research in academic discourse and analyses of writing text books seem to reinforce this idea (Jordan 1997; Hinkel 1999; Chang–Swales 1999).

Based on a thorough analysis of writing manuals published between the 1960s and 1990s, Chang and Swales (1999, 148) compiled a list of the ten most often cited ‘rules’ of appropriate scholarly writing. In this list, the avoidance of the first person pronouns to refer to the author(s) of the text occupies the first place, which means that this feature was commented on by the largest number of handbooks and manuals. The remaining nine positions are taken by features such as the use of broad reference (i.e., anaphoric pronouns that can refer to antecedents of varying length), split infinitives, conjunctions at the beginning of the sentence, prepositions at the end of the sentence, run-on sentences, sentence fragments, contracted forms, direct questions and exclamations, i.e., phenomena that are all associated with informal spoken language.

Adherence to such rules has been emphasized in the Hungarian academic context, too, where local writing conventions seem to show a considerable difference from the Anglo-Saxon tradition (on a description of some of these rhetorical differences see Connor (1996, 144); on academic writing in particular see Árvay–Tankó 2004 and Futász 2006). Formal/impersonal writing is explicitly taught in EAP courses at Hungarian universities, as the personal/subjective tone generally rejected by English writing text books and manuals is traditionally not rejected in formal writing in Hungarian. Tolerance towards personal tone in Hungarian formal writing is promoted by the very nature of the Hungarian language, too, which tends to disfavour the use of, for instance, the passive voice in general and favours active sentence constructions instead. The frequent use of the passive sounds alien to an EFL writer with a Hungarian L1 background and therefore poses difficulties even for advanced learners.

Recent research, however, has discovered a diachronic change in the (im)personal nature of English academic writing and has demonstrated a trend to move away from impersonal to a more personal tone (Chang–Swales 1999; Hyland 2002b; Ivanič–Simpson 1992; Myers 1989). In other words, empirical investigations including linguistic analysis of texts, interviews with expert writers, manuals surveys prove that less formality

is **also** accepted now. Based on empirical evidence, Chang and Swales (1999), for example, claim that

“although disagreement still existed among the manuals published in the 1960s and 1970s, those which were published after the 1980s tended to encourage the use of first person pronouns more overtly and rigorously. These authors concurred that the use of *I* and *we* does not itself make a piece of writing less formal or less objective.” (*op.cit.*, 149)

Both Hyland (2002b, 351) and Chang–Swales (1999, 145) argue against the universal nature of earlier norms and claim that scientific discourse is not uniformly impersonal and that there is considerable interdisciplinary variation in this respect. Based on a large-scale corpus-based investigation, Hyland (2002b) concludes that “not all disciplines follow the same conventions of impersonality, and [...] in fact there is considerable scope for the negotiation of identity in academic writing” (*op.cit.*, 351).

### 2.3. The rhetorical functions of personal reference in student writing

Personal reference is used to express a wide range of rhetorical functions (a detailed discussion of these may be found in Hyland 1999; Kuo 1999; Myers 1989; 1992; Tang–John 1999). The present study builds on Hyland’s work, who argues—referring to extensive corpus-based analyses across various disciplines—that academic writers employ person markers for three main purposes: (1) to organize arguments and structure text, (2) to introduce or discuss research activities, and (3) to explicitly indicate attitude to findings or align the author with theoretical positions (Hyland 1999, 118).

In the case of Hong Kong students Hyland (2002b) found that, as a result of the stylistic advice communicated by EAP courses and writing text books, they tend to underuse author pronouns (compared to expert writers) or use them, as he puts it, “unadventurously, referring to their texts rather than their ideas” (*op.cit.*, 353). In his corpus he found that expert writers were three times more likely to use writer pronouns (i.e., *I*, *my*, *me*, *we*, *our*, *us*) than university students. Hinkel’s (1999) comparative analysis of native and non-native (Chinese) speakers also demonstrated a significant difference, with native speakers employing considerably more first person pronouns.

In terms of the rhetorical functions for which author pronouns are used in text, Hyland’s analyses showed that expert and student writers differed notably. While expert writers applied them dominantly to

present claims and bottom-line results and to link themselves with their main contribution, student writers used them to acknowledge assistance, state purpose and explain procedures. They avoided using direct reference to themselves as authors when communicating more important argumentative functions such as presenting or justifying claims. Chang and Swales's (1999) findings of a similar study also indicate that MA thesis writers felt uncomfortable about referring to themselves as *I*, except when describing their own field experiences.

The results of the above mentioned investigations seem to harmonize with an earlier remark made by Clark and Ivanič (1997, 153), namely that “the overarching idea is that writers differ in how much they feel, and appear to be, in control of the act of writing: how much they feel themselves to be not just ‘writers’ but also ‘authors’ with the authority to say something”. This constitutes a major problem for novice writers. Based on the results of previous research (e.g., Hinkel 1999; Hyland 2002b), it may be assumed that lack of feeling of authority in the discipline and/or in the academic context may be an explanation for the under- and misuse of writer pronouns in student texts. The current study is intended to reveal the validity of this assumption in the Hungarian context.

### 3. Method

This study has two main objectives. First, it explores the frequency and the main rhetorical functions of personal reference in published research articles and university students' MA theses based on text analysis. Secondly, in order to reveal some of the reasons behind the patterns of pronoun use, it investigates students' views, perceptions and practices regarding the use of these items with the help of interviews.

#### 3.1. Corpus design

The aims of the study necessitated the construction of two corpora. On the one hand, a corpus of English published research articles was constructed (henceforth referred to as the “RA Corpus”), and, on the other hand, a corpus of EFL student MA theses was built (referred to as the “Thesis Corpus”).

The RA Corpus contains altogether 50 randomly chosen articles from five prestigious journals in the field of applied linguistics. The journals



were selected on the basis of accessibility and contents. I selected journals that were electronically available in the institution I work for. Furthermore, it was also important to construct the corpus of journals which represent different areas of research in the field. The titles of the journals I finally worked with are listed in Appendix B. Ten papers were selected from each journal published between 1995 and 2005. In order to ensure the comparability of the RA and the Thesis corpora, only single-authored papers have been selected to form part of the RA Corpus. All of the articles are based on original data and report on studies conducted in applied linguistics. The RA Corpus contains just over 420,000 words and the average length of the papers it includes is 8470.96 words.

The Thesis Corpus is taken from the thesis sub-corpus of the Hungarian Corpus of Learner English<sup>1</sup> and contains altogether 50 MA theses written (between 2004–2007) by students majoring in English language and literature and specializing in applied linguistics at Eötvös Loránd University, Budapest, Hungary. All the theses were written in English by advanced EFL learners whose mother tongue is Hungarian. The theses report on empirical studies conducted in the field of English applied linguistics. The Thesis Corpus contains just over 870,000 words and the average length of the theses is 17684.10 words.

The theses written at Eötvös Loránd University are to follow the relevant departmental guidelines, which attempt to approximate as closely as possible the norms and requirements of professional English language research articles. Hence the motivation for the current investigation and for regarding the two genres (the RAs and the theses) as comparable. The thesis guidelines of the Department of English Applied Linguistics (School of English and American Studies, Eötvös Loránd University) offer information both regarding form (e.g., structure of the thesis, documentation of sources) and content (e.g., review of the literature, research questions, designing and piloting instruments, doing analysis and writing up results). The theses therefore comprise a comparable corpus with published research articles, the only major difference lying in (1) the average length of the papers (which is about twice as long as journal articles) and (2) the audience. The readers of the theses are the students' university tutors and not the professional audience of journal articles. This entails

<sup>1</sup> The corpus has been built at the School of English and American Studies, Eötvös Loránd University, in cooperation with the Institute for Linguistics of the Hungarian Academy of Sciences.

a different power/status relationship, which—as the findings will also demonstrate—may have an effect on the way personal reference is used.

### 3.2. Data coding and analysis

Articles and theses were scanned to produce electronic corpora. The two corpora were then examined using the *WordSmith Tools* text analysis and concordance programme (Scott 1996) to determine the frequency of personal reference expressing the author's stance (singular: *I, my, me*, and plural: *we, our, us*) and examine their context (using the KWIC function of the programme, i.e., “key word in context” concordance). All target items revealed by the programme were manually examined by myself and a fellow researcher working independently to ensure they all expressed the writer's stance, that is, they could be regarded as genuine author pronouns expressing exclusive first person uses. Reliability of analysis was checked comparing the coding of the first 100,000 words of each corpus. This activity produced an inter-coder reliability of 0.91 (Kappa), indicating a high degree of agreement.

### 3.3. Interviews

To explore university students' views and practices regarding the use of author pronouns and shed light on their reasons for opting to use or avoid these, the analysis of texts was complemented with interviews. Altogether 10 thesis writers (in the last year of their studies) and 10 students from first-year undergraduate academic writing courses were interviewed. Thesis writers, having completed all of their MA courses may be considered as experienced writers in the university context. The students in the first-year academic writing course are novice EFL writers, without much practice or experience in academic writing in English. Still, as participants of the course, they are engaged in the process of developing their EFL writing proficiency and forming some idea (at least in theory) of the kind of writing that will be expected of them at the university. The aim of this course is to familiarize students with the conventions of English academic writing and help them master the writing skills necessary for success during their university studies.

The interview protocol was piloted before use and the necessary modifications were made to ensure it elicits the kind of data that is required

for the purposes of the investigation. The list of altogether six questions primarily focused on (1) what students regard as prestigious writing in their immediate academic/university context, (2) how they feel about the use of first person pronouns, and (3) when they opt for using personal pronouns (for the interview questions see Appendix A). The interviews were conducted in Hungarian and each took about 15 minutes. The interviews were tape-recorded and later on transcribed for analysis.

## 4. Results and discussion

### 4.1. Frequency of author pronouns

Both corpora were electronically searched for the pronouns *I*, *my*, *me*, *we*, *our*, *us* and each case of genuine author pronoun was counted. The results of the analysis confirm the idea that scientific writing is not as impersonal as it is often claimed to be. The mean frequency calculated for all author pronouns per research paper in the current RA Corpus (Table 1, mean frequency: 33.8) harmonizes with Hyland's (2002b, 353) results for research papers written in the field of applied linguistics, where an average frequency of 32.3 pronouns was found. Interestingly, however, the distribution of singular and plural pronouns in the present corpora differs to some extent: the RA Corpus contains a considerably higher number of singular pronouns on average (21.36) than plural ones (12.44).

Table 1  
Mean frequency of author pronouns per research paper

Corpus	Average number of words	All author pronouns	Singular pronouns ( <i>I</i> , <i>me</i> , <i>my</i> )	Plural pronouns ( <i>we</i> , <i>our</i> , <i>us</i> )
RA Corpus	8470.96	33.80	21.36	12.44
Thesis Corpus	17684.10	102.40	85.70	16.70

Contrary to Hyland's findings, however, which showed that expert writers were more likely to use author pronouns in their texts than Hong Kong students, Hungarian student writers used author pronouns much more frequently than expert writers. As Table 1 shows, the mean frequency of all author pronouns in the Thesis Corpus is 102.40 (more than three times higher than in the RA Corpus). The distribution of singular and

plural pronouns is also interesting. Student writers seem to opt for using singular pronouns in the vast majority of the cases (85.70) and not plural ones (16.70).

As the theses are on average twice as long as the journal articles, we could expect the longer texts to include more metadiscoursal uses of personal pronouns (Swales 1990). To neutralize differences resulting from text length, the frequency of author pronouns was recalculated (per 10,000 words, Table 2). These counts reinforce the results of the previous analysis: in the Thesis Corpus the frequency of all author pronouns is considerably higher (57.85) than in the RA Corpus (39.88). The latter shows a similar frequency to what Hyland's (2002a: 354) overall figures indicated: in his study 41.2 author pronouns appear per 10,000 words in published scientific writing.

*Table 2*  
Frequency of author pronouns per 10,000 words

Corpus	Number of words	All author pronouns	Singular pronouns ( <i>I, me, my</i> )	Plural pronouns ( <i>we, our, us</i> )
RA Corpus	10,000	39.88	25.20	14.68
Thesis Corpus	10,000	57.85	48.42	9.43

Since Hungarian university students have generally been instructed in EAP courses to avoid the personal tone and make their text objective and factual, these findings are surprising. It is important to note, though, that students face a number of cultural, rhetorical, linguistic difficulties when trying to acquire appropriate stylistic competence in EFL academic writing (e.g., the generally more subjective nature of formal writing in Hungarian, the resistance of the Hungarian language to using passive structures, the traditionally more authoritarian educational system, etc.). This huge difference of tone between the theses (MA students' final and most significant academic assignment before graduation) and published research articles (the "model" to be followed by English majors), calls for a reconsideration of the ways in which the notion of style is dealt with in EAP courses designed for students with a non-English (Hungarian) L1.

#### 4.2. The rhetorical functions expressed by author pronouns

To explore the rhetorical functions for which author pronouns are used, the immediate linguistic contexts of the author pronouns were investigated using KWIC concordances. A concordance of the data showed that there are remarkable differences for example in the verbs most commonly co-occurring with author pronouns in the two corpora. It needs to be noted here though that rhetorical functions may not only be distinguished by reference to the verb as there are also instances where the pronoun does not control a verb (e.g., *contrary to our findings*). The differences in the verbs commonly co-occurring with author pronouns in the two corpora may indicate (a) distinct representations of writer presence (i.e., different perceptions of authority) and (b) differing perceptions of the principal functions for which these pronouns may be appropriately used. In what follows, these two will be described in more detail.

(a) As argued earlier in this paper, the way one formulates his/her statements is crucial to be accepted by the scientific/academic community. As regards the representation of writer presence, a considerably more confident and factual tone (writer stance) may be observed in the RA Corpus when defining aims, expressing opinion, opposition or formulating criticism or appraisal. Some of the most frequently occurring verbs in the RA Corpus appear in sequences such as *We propose that...*, *We shall present a revised taxonomy...*, *In this paper I have demonstrated...*, *I have assumed in my conclusion that...*; all of which reflect a confident writer stance. The Thesis Corpus, on the other hand, contains a large number of sequences which show less confidence and more uncertainty on the part of the writer regarding the claims or decisions made: e.g., *I tried to get insight into...*, *I will attempt to...*, *this phenomenon, I think, is not so pervasive as...*, *I would like to analyze...*

(b) The two corpora also indicate a somewhat different distribution of the functions for which writers use author pronouns in their texts. As mentioned earlier, previous research (Hyland 1999, 118) has shown that person markers fulfil three main functions in the papers of academic writers: they (1) organize arguments and structure text, (2) introduce or discuss research activities, and (3) explicitly indicate attitude to findings or align the author with theoretical positions. Frequency lists of the verbs co-occurring with person markers suggest that expert writers dominantly use them for different purposes and in different sections of the paper than student writers.

In the RA Corpus, author pronouns are mostly (in 81% of all cases) used to organize arguments and structure text (function 1) and to explicitly indicate attitude to findings or align author with theoretical positions (function 3). More precisely, author pronouns dominate when the writer intends to

- specify aims (e.g., *In this paper **we** investigate the...*, *In this study **I** will focus on...*, *Here **we** attempt to...*, *This is the issue **I** address in this paper.*);
- structure text (e.g., *Let **us** now take a look at...*, ***We** will first briefly review...*, *In this section **I** discuss...*, ***I** turn to this issue now...*, *Now **I** turn to explore these general observations in more detail...*);
- state claim, indicate the author's attitude to theoretical positions or findings of other scholars (e.g., ***We** also propose that...*, ***I** believe the assumption...*, ***We** have extensively argued that...*, *...at this level of proficiency **I** would be surprised if this were the case...*, *In the present study, **I** take a conservative approach and...*);
- report on the author's own findings (e.g., ***My** results support the claim **I** made earlier that...*, *...**we** find many obligation bundles with a second person subject...*, ***I** would characterize this set of data as...*, ***I** hope that **I** have also demonstrated that they are consistently functional.*); and
- formulate theoretical or practical implications (e.g., ***I** conclude by identifying...*, *...contributed a great deal to **our** understanding of...*, *...lead **us** to reconsider earlier assumptions...*, *...such awareness would, **I** believe, have lasting, positive consequences...*).

Consequently, the sections where author pronouns appear in the highest frequency are the Review of the Literature (17%), the Method (18%), the Results (32%) and the Conclusion sections (28%) of the papers. Merely 5% of the total number of author pronouns appears in the Introduction sections of the RA Corpus.

In the Thesis Corpus, on the other hand, their main functions are (in 74% of all cases) to organize arguments and structure text (function 1) and to introduce or discuss research activities (function 2). Taking a closer look at the data, the corpus provides evidence of a high number of author pronouns used for the following purposes:

- to specify aims (e.g., *In this thesis **I** would like to analyze...*, *In my research **I** focus on...*, ***I** will investigate...*);

- to structure text (especially to refer to the way the thesis is structured: e.g., *In the following **I** am going to look at...*, *In what follows **I** am going to present...*, *First, **I** will present the method, then...*, *Finally, **we** shall outline...*);
- to describe research design and procedures (e.g., ***I** conducted a quantitative study to...*, ***We** designed a questionnaire to...*, ***I** asked 25 students...*); and
- to report on the author's own findings (e.g., *Based on the findings of my research **I** think that...*, *The results of the analysis confirm what **I** assumed earlier...*, *Based on these figures I see a marked difference between...*).

Quantitative analysis of the data shows that author pronouns appear in the highest frequency in the Introduction (11%), Method (37%), and Results sections (39%) of the theses. The Review of the Literature and the Conclusion sections show a considerably lower percentage, with a mere 6% and 7%, respectively. This shows a similar tendency to Hyland's (2002b) findings, according to which, as he put it, the vast majority of all uses performed "these relatively innocuous and text internal roles, which commit the reader to little, and carry only a weak sense of identity" (*op.cit.*, 354). Students tend to avoid directly referring to themselves as authors when expressing higher level argumentative functions (e.g., stating claims or expressing/negotiating disagreement).

The high number of author pronouns in the Introduction sections of the Thesis Corpus may in part be due to the student writers' practice of including here information that does not typically (or at least not so often) form part of the introductory sections of research articles. Thesis writers very often make note of their motivation (personal or professional) for conducting the research reported on, for example:

"The idea of choosing Civilization as a topic for my thesis originates from a surprising recognition during my English studies at [...]. Being an [...] major student for five years now I have encountered several rather shocking discoveries related to both my own and other fellow students' general knowledge about the United Kingdom."

or:

"As a freelance translator I mainly work with British television programmes and encounter a number of difficulties arising from the special characteristics of subtitling."

or:

“This field of study Intercultural learning became the most interesting area to me during my years of study as [...]. Moreover, I have a special personal interest in intercultural matters, as I was an exchange student to the United States where I realised that it was extremely important to be able to navigate between cultures in order to make ourselves understood.”

Furthermore, besides describing the rationale for and the purposes of the investigation, several of the Introductions state the research questions and/or the hypotheses explicitly, which also very often make use of author pronouns:

“Therefore in my study I aim to answer the following research questions.”

or:

“If we accept this as a working hypothesis, based on a quantitative and a qualitative analysis we can test it, and draw conclusions regarding how the differences in the two languages influence the production of appropriate translations.”

Even though student writers use a high number of personal pronouns, this does not seem to make a successful contribution to their discursive identity. The data reinforce the assumption that student and expert writers have a somewhat different perception of authority and of the possible (and desired) uses of first person pronouns. In what follows I will turn to describing what exactly student writers’ perceptions of pronoun use are, based on the interviews conducted in the Hungarian university context.

#### **4.3. Student writers’ perceptions of pronoun use in English academic writing**

The interviews brought surprising results in the light of the findings of the text analyses which, compared to the RA corpus, showed a considerable over- and misuse of author pronouns by students in their theses. Their texts do not seem to be in congruence with what they claim in the interviews. Regarding their perceptions of prestigious writing at the university they—including both first-year students and thesis writers—almost unanimously said that style in English academic writing has to be ‘formal’ and ‘objective’, and that ‘personal/subjective views, opinions are to be avoided’. Some of the other adjectives they frequently mentioned



were 'impersonal', 'neutral', 'dry', 'distanced', 'factual', and 'sophisticated'. They also mentioned the importance of using complex sentence structure (as opposed to simple sentences) and passive (rather than active) constructions to minimize the impression of subjectivity and thus increase credibility. These features echo the advice communicated by the majority of writing manuals.

Although students were not explicitly asked about this, with the exception of one student, they all gave voice to the difficulties or problems they face when trying to express themselves within the confines of this style. Some, especially first-year students, mentioned that they find this style of composing very 'alien' and 'strange' and that they feel these norms in fact 'limit their self-expression', or even take away their 'freedom as writers' or 'character'. They find it hard to 'find the balance', as they put it, between being able to express their views and original ideas (one expectation often stated about successful writing at the university) and keeping to the objective and impersonal tone that they are instructed to use (another expectation communicated by several of their teachers as well as some writing manuals).

As regards students' feelings concerning the use of personal pronouns and *I* in particular, their opinions were somewhat divided. The majority of the participants (11 out of the 20) categorically stated that *I* is inappropriate and is thus to be avoided in academic writing whenever possible. Out of the remaining 9 participants, 7 considered its use appropriate (and even necessary) in some cases, and two students said that it is completely acceptable if not employed too often (i.e., its frequent repetition does not make the text 'monotonous' and 'awkward sounding', as they put it). Those who argued that in some cases it is appropriate said that to them it would sound 'unnatural' to completely avoid *I* when, for example, announcing the aim of the study (especially the research questions) or describing the procedures of research. One of them claimed that using *I* is sometimes necessary to be able to 'draw the line between what others say and what I think'.

The students' main (stated) motives for using or avoiding personal pronouns (mostly *I*) seem to reflect what they have been instructed to do by their tutors. They claim to use *I* for the following functions (cited directly from the data): 'to explain the motivation for doing the given piece of research', 'to show that something is my own idea', 'to state the aim of my study', 'to describe what I did in my research and how', 'to explain my decisions', 'to express my opinion', 'to relate my experiences',

‘to acknowledge help’ (either in conducting the research or in writing the study). All of these reasons seem to reflect what Ivanič (1997) and Clark–Ivanič (1997) refer to as “identity” or “voice” (“self-representation”: Ivanič–Camps 2001), and also indicate that—at certain points in their papers—students consider it important and justifiable to be present as authors.

On the other hand, they also tend to avoid personal pronouns, as they claim, in order to sound ‘objective and thus reliable’, ‘factual’, ‘formal’, and ‘not subjective’. There was a very interesting point made by a good number (6) of the participants of the interviews: they said that they often find it hard to stay on topic and remain always fully relevant when reporting on a piece of research. They are frequently tempted to share experiences (positive or negative ones) and views that may not be linked organically to the project or are not crucially important from the point of view of the outcomes of the study. Still, as students either learnt from them (personally and/or professionally) or just simply ‘want to show off’, they argue, with how well they have done the job, these experiences or views form part of their texts. Relating these, by nature, entails a more frequent use of personal pronouns. Let me note here that this ‘temptation’ (as one of the students called it) is understandable as conducting research for an MA thesis is a major task and forms a significant component of the students’ learning process and training at the university. This may therefore explain a considerable part of the overuse of personal pronouns found in the Thesis Corpus. The views communicated by students harmonize with the assumption put forward in Hvítfeldt’s (1992) study, according to which some writing traditions are more personalized than those of English, as a result of the greater importance played by everyday experiences in the construction of one’s “idea of truth” (*op.cit.*, 33).

## 5. Conclusions

This study aimed to compare the use of author pronouns in English published research articles and Hungarian EFL students’ MA theses written in English. Frequency counts show that student writers employ a considerably higher number of personal pronouns than expert writers. Concordances of the data reveal that besides the differences of quantity, qualitative differences also appear between the two corpora in terms of the rhetorical functions of author pronouns performed in them. While expert writers employ author pronouns to express higher level argumen-

tative functions (e.g., to state claims, express opposition, report on findings, formulate theoretical implications), students use them dominantly to refer to their text and research activity. The results of the interviews also show that students struggle with acquiring and practicing the skills necessary for success in English academic writing.

The interviews brought conflicting results compared to the outcomes of the text analyses. The interviews suggest that students are in fact (theoretically) aware of the conventions of style (specifically of those related to personal pronoun use) in English academic writing, even if it is not reflected in their papers. What all this boils down to is that while students seem to know the “rules”, they are unable to apply them adequately. Some of their difficulties are cultural/rhetorical, others originate from their somewhat categorical and (partly) outdated perceptions/intuitions regarding the formality of English academic discourse.

Therefore EAP writing courses (especially in the EFL context) need to pay special attention to this—very sensitive and complex—aspect of composition: besides training students to avoid the overuse of pronouns in their texts, they need guidance in mastering the rhetorical functions for which personal pronouns may be adequately used. A significant outcome of the analyses for future research is that in accounting for students’ use of personal pronouns, linguistic and cultural factors may be at least as important (if not more so) as the students’ apprentice status or the explicit teaching they receive.

As the notion of appropriate style is hard to grasp even in one’s L1, it is important that besides knowing the conventions of Anglo-Saxon scientific writing, course instructors familiarize themselves with the students’ perceptions of formal/academic style in their L1 and highlight English preferences in comparison to those. By raising students’ awareness of the differences between the two languages (e.g., note some of the guidelines put forward in Hyland 2002a), their sensitivity towards contrastive rhetorical features may increase and they may be able to apply the FL norms more consciously and successfully.

Furthermore, in contexts where the educational system is traditionally more authoritarian, like in Hungary, where teachers have been highly respected by their students, and the words of expert members of the academic community have been rarely assessed or criticized, special emphasis needs to be laid in EAP courses on helping students (1) acquire the necessary critical thinking skills and (2) the adequate linguistic and pragmatic means and skills of expressing their opinion, or even opposition (as also

shown by Clark–Ivanič 1997, 134, 153). These skills are crucially important for students to successfully integrate/engage in scientific discourse. EAP courses in the EFL context can only achieve their goals if they take into consideration the students' stereotypical cultural, linguistic, academic and educational background.

#### **Appendix A: Questions of the student interviews**

1. How would you characterize the style of English academic writing? Please mention the characteristics you consider important.
2. What do you think your tutors consider as successful writing at the university? Think of their requirements concerning seminar papers and theses.
3. Do you think the use of first person pronouns (e.g., *I*, *my*, *me*, *we*, *our*, or *us*) is appropriate in academic writing? If yes/no, why?
4. Do you use these pronouns when you write an academic paper? If yes, for what purpose, or in which section(s) of your paper?
5. When do you think their use is acceptable?
6. When do you think their use is unacceptable?

#### **Appendix B: Journals constituting the RA Corpus**

1. Applied Linguistics
2. English for Specific Purposes
3. Journal of Pragmatics
4. Journal of Second Language Writing
5. English Language Teaching Journal

#### **References**

- Árvay, Anett – Gyula Tankó 2004. A contrastive analysis of English and Hungarian theoretical research article introductions. In: *International Review of Applied Linguistics in Language Teaching* 42: 71–100.
- Atkinson, Dwight 1999. Language and science. In: *Annual Review of Applied Linguistics* 19: 193–214.
- Biber, Douglas – Susan Conrad – Randi Reppen – Pat Byrd – Marie Helt 2002. Speaking and writing in the university: A multidimensional comparison. In: *TESOL Quarterly* 36: 9–48.
- Biber, Douglas – Edward Finegan 1988. Adverbial stance types in English. In: *Discourse Processes* 11: 1–34.

- Candlin, Christopher N. – Ken Hyland (eds) 1999. *Writing: Texts, processes and practices*. Longman, Harlow.
- Chang, Yu-Ying – John M. Swales 1999. Informal elements in English academic writing: Threats or opportunities for advanced non-native speakers? In: Candlin – Hyland (1999, 145–67).
- Clancy, John – Brigid Ballard 1992. *How to write essays*. Melbourne, Longman Cheshire.
- Clark, Romy – Roz Ivanič 1997. *The politics of writing*. Routledge, London.
- Clyne, Michael 1987. Cultural differences in the organization of academic texts. In: *Journal of Pragmatics* 11: 211–47.
- Connor, Ulla 1996. *Contrastive rhetoric*. Cambridge University Press, Cambridge.
- Crystal, David 1995. *The Cambridge encyclopedia of the English language*. Cambridge University Press, Cambridge.
- Flowerdew, John 2000. Discourse community, legitimate peripheral participation, and the non-native-English-speaking scholar. In: *TESOL Quarterly* 34: 127–50.
- Futász, Réka 2006. Analysis of theoretical research article introductions written by undergraduate students: A genre-based approach. In: *Acta Linguistica Hungarica* 53: 97–116.
- Grabe, William – Robert B. Kaplan 1996. *Theory and practice of writing*. Longman, London.
- Harwood, Nigel 2005a. “I hoped to counteract the memory problem, but I made no impact whatsoever”: Discussing methods in computing science using *I*. In: *English for Specific Purposes* 24: 243–67.
- Harwood, Nigel 2005b. “Nowhere has anyone attempted... In this article I aim to do just that”: A corpus-based study of self-promotional *I* and *WE* in academic writing across four disciplines. In: *Journal of Pragmatics* 37: 1207–31.
- Harwood, Nigel 2005c. “We do not seem to have a theory... The theory I present here attempts to fill this gap”: Inclusive and exclusive pronouns in academic writing. In: *Applied Linguistics* 26: 343–75.
- Hewings, Martin – Ann Hewings 2002. “It is interesting to note that...”: A comparative study of anticipatory ‘it’ in student and published writing. In: *English for Specific Purposes* 21: 367–83.
- Hinkel, Eli 1999. Objectivity and credibility in L1 and L2 academic writing. In: Eli Hinkel (ed.): *Culture in second language teaching and learning*, 90–108. Cambridge University Press, Cambridge.
- Hvitfeldt, Christina 1992. Oral orientation in ESL academic writing. In: *College ESL* 2: 29–39.
- Hyland, Ken 1999. Disciplinary discourses: Writer stance in research articles. In: Candlin – Hyland (1999, 99–121).
- Hyland, Ken 2002a. Authority and invisibility: Authorial identity in academic writing. In: *Journal of Pragmatics* 34: 1091–112.
- Hyland, Ken 2002b. Options of identity in academic writing. In: *ELT Journal* 56: 351–58.

- Ivanič, Roz 1997. Writing and identity: The discorsal construction of identity in academic writing. John Benjamins, Amsterdam & Philadelphia.
- Ivanič, Roz – David Camps 2001. I am how I sound: Voice as self-representation in L2 writing. In: *Journal of Second Language Writing* 10: 3–33.
- Ivanič, Roz – John Simpson 1992. Who's who in academic writing? In: Norman Fairclough (ed.): *Critical language awareness*, 141–73. Longman, London.
- Jordan, Robert R. 1997. *English for academic purposes. A guide and resource book for teachers*. Cambridge University Press, Cambridge.
- Kachru, Braj B. 1985. Standards, codification and sociolinguistic realism: The English language in the outer circle. In: Radolph Quirk – Henry G. Widdowson (eds): *English in the world*, 11–30. Cambridge University Press and British Council, Cambridge.
- Kaplan, Robert B. 1983. Language and science policies in new nations. In: *Science* 221: 4614.
- Kaplan, Robert B. 1989. English as language of science. In: *Vox* 2: 49–53.
- Kaplan, Robert B. 1993. The hegemony of English in science and technology. In: *Journal of Multilingual and Multicultural Development* 14: 151–72.
- Kuo, Cheh-Hua 1999. The use of personal pronouns: Role relationships in scientific journal articles. In: *English for Specific Purposes* 18: 121–38.
- Malcolm, Ian G. 1999. Writing as an intercultural process. In: Candlin – Hyland (1999, 122–42).
- Medgyes, Péter – Robert B. Kaplan 1992. Discourse in a foreign language: The example of Hungarian scholars. In: *International Journal of the Sociology of Language* 98: 67–100.
- Medgyes, Péter – Mónika László 2001. The foreign language competence of Hungarian scholars: Ten years later. In: Ammon Ulrich (ed.): *The dominance of English as a language of science. Effects on other languages and language communities*, 261–86. Mouton de Gruyter, Berlin & New York.
- Myers, Greg 1989. The pragmatics of politeness in scientific articles. In: *Applied Linguistics* 10: 1–35.
- Myers, Greg 1992. “In this paper we report...”: Speech acts and scientific facts. In: *Journal of Pragmatics* 17: 295–313.
- Petch-Tyson, Stephanie 1998. Writer/reader visibility in EFL written discourse. In: Sylviane Granger (ed.): *Learner English on computer*, 107–18. Addison Wesley Longman, Harlow.
- Scott, Mike 1996. *WordSmith tools*. Oxford University Press, Oxford.
- Shaw, Philip 1991. Science research students' composing processes. In: *English for Specific Purposes* 10: 189–206.
- Swales, John M. 1990. *Genre analysis: English in academic and research settings*. Cambridge University Press, Cambridge.
- Tang, Ramona – Suguanthi John 1999. The ‘I’ in identity: Exploring writer identity in student academic writing through the first person pronoun. In: *English for Specific Purposes* 18: S23–S49.