EXPRESSING STANDPOINTS IN EFL WRITTEN DISCOURSE

Leonardo Juliano Recski

lrecski@hotmail.com

1. Introduction

There are a number of reasons why the presence of the writer may be considered undesirable in formal academic writing. The most obvious one is that the absence of the writer from the text helps create a sense of impersonal objectivity. Why?

There is a bit of psychology in that: The effect of first and second person pronouns in a text is closely connected with their typical usage. In spoken language, these pronouns illustrate the involvement of the speaker in what she or he is saying.

First and second person pronouns can be used to establish direct interpersonal contact with the listeners and to appeal to their reason and emotions. When used in a written text, the pronouns then provoke similar feelings in the readers (i.e. they have a metacommunicative effect (Abdi, 2002; Crismore 1990, 2003; Crismore et al. 2003; Hyland, 2002)). The readers feel that the writer is somehow involved in the text and automatically attribute the writer of the (otherwise impersonal) text with personal, subjective emotions and attitudes – as they would attribute them to a speaker they could see or hear. In other words, when the writer uses personal pronouns, the reader immediately imagines a living person, whose ideas and standpoints are subjective by nature.

Additionally, speakers are usually concerned with the reaction of the audience and they communicate with them directly, while writers (especially of formal academic texts) have to bear in mind that their work must fulfill criteria of a large variety of

1 Pós-Graduação em Letras/Inglês - Universidade Federal de Santa Catarina
independent readers. As the interaction with readers is normally one-sided, the text must be self-explanatory and self-defensible, in other words, neutral.

Thus, a text expressing notions suited to a formal register, but utilizing a number of features common to an informal mode will most likely puzzle the readers. In a way, texts with large proportions of interpersonal devices (like I-s and you-s) resemble speech. Because speech is usually highly informal and subjective, these devices will automatically carry the informality and subjectivity into the text, even unintentionally.

Anglo-American literacy tradition seems to have adopted a specific attitude towards first person reference in writing. Readers from continental Europe background usually introduce their opinions with statements of attitude (like I think, I feel) to indicate humble subjectivity (meaning “I might not be right, but my opinion is …”) (Granger, 1998). In contrast, readers with British or American background are likely to use such statements to exercise authority and emphasis (Swales, personal communication). The difference would be best discernible if the two readers were to read the same sentence aloud. While the “continental” reader would probably read something like “In my opinion, the suggestion he/she made is a true fallacy” (meaning “but perhaps someone sees the things from a different perspective), the British or American reader would stress the introductory part: “In MY opinion, the suggestion he/she made is a true fallacy” (meaning “and there’s no doubt about it”). If put in writing, the interpretation of this sentence depends on the reader, regardless to what the author’s original intention was.

The writer only has one possibility to prevent misinterpretation – that is, to avoid ambiguous expressions. If overtly presenting his or her attitude as a personal one can incite such ambiguity, it is desirable to avoid the personal tone and present the attitude in a non-personal, neutral way which prevents misinterpretations. The example above could then be rewritten into something like “The suggestion she/he made could be a fallacy.” This way, ambiguity is avoided and the interpretation is the same for readers from both cultural backgrounds.

Having said that, we see that “tolerable” personal intrusion into a formal academic text, at least in English writing, should be limited to recounting personal experience, sometimes to references to previous parts of the text (e.g. “As I have said...”) and, in a limited number, for emphasizing personal views.
Quoting Smith (1986), Petch-Tyson (1998) points out that a limited, non-intrusive presence of the writer (and the reader) can have a non-disturbing, even refreshing effect, even in formal or scientific writing. She also suggests that the use of description of personal experiences as an illustration of the text is much less intrusive and has a more positive effect than using the pronouns to mark subjectivity of opinion. Having based her findings in the theoretical work of Lakoff (1982) she came up with proofs to the theory that the key to successful use of interpersonal features is the consistency of the style with the use of interpersonal features. She then goes on to exemplify that an informal text with consistent use of informal features has a better communicative effect than an informal text suddenly changing its discourse to a formal tenor (cf. Petch-Tyson, 1998: 116).

It is a little irony to see that the learners are supposed to write argumentative essays which should be, from their definition, formal, and formal style should be avoided from their definitions and personal references; at the same time the topics of the essays include:

- Some people say that in our modern world, dominated by science, technology and industrialism, there is no longer place for dreaming and imagination. What is YOUR opinion?
- In the 19th century, Victor Hugo said: “How sad it is to think that nature is calling out but humanity refuses to pay heed.” Do YOU think it is still true nowadays?

In such cases, personal references and subjective attitudes are certainly hard to avoid. Nevertheless, the following paper attempts to answer questions concerning reader and writer visibility:

(i) How much do native speakers use writer / reader references? How do the learners compare to them?
(ii) What are the main reasons for using them by the native speakers and non-native speakers? Are there any differences in usage between them?

This paper was inspired by the above mentioned paper by Stephanie Petch-Tyson (1998). At the start, I followed a similar procedure, so I will refer to her paper further on in this section. There is no danger on my coming up with the same results – I
worked with different corpora and while leaving out some of her observations I extended others, which were merely touched in her paper.

Below I present the most important signals of writer / reader visibility as based on a list presented by Petch-Tyson:

1. First person reference (singular, plural)
2. Second person reference (specific, general)
3. Speakers’ mental processes (think, believe, etc.)
4. Monitoring of information flow (you know, I mean, etc.)
5. Emphatic particles (just, really)
6. Fuzziness (and so on, etc., ‘…’)
7. References to situation of reading/writing (this X, here, now, etc.)
8. Evaluative modifiers
9. Imperatives
10. Questions (direct / rhetorical)

We will have a look at the first three, although in a different order. As Petch-Tyson’s research showed, in comparison with the first three, features 4 through 8 play a minor role in essay writing and do not influence overall findings. Imperatives do not appear very often in a relatively small corpus – they are so rare that they can hardly be a subject of research. Direct questions in a number of student corpora have been dealt with in a paper by Tuija Virtanen (1998: 94-106), who offers a very insightful look at their distribution, including their placement and functions within the discourse. Therefore, they will not be analyzed here.

2. The data – The International Corpus of Learner English

This paper is based on findings of a portion of the ICLE. ICLE is a project started by the team of Profesor Sylviane Granger at Catholic University of Louvain, Belgium. Basically, ICLE corpus is a large computer collection of essays, written by advanced university students, who use English as a non-native language. Among others, ICLE includes contributions from French, Belgium, Dutch, Spanish, Finnish, Czech, Norwegian, Swedish, Polish, Chinese and Brazilian students and is still growing. For the sake of objectivity, the topics and length of all essays were standardized. Each
author is asked to contribute one essay 1000 words, or two essays of 500 words. The topics for the students to choose from include:

- Crime does not pay
- The prison system is outdated. No civilized society should punish their criminals: it should rehabilitate them
- Most university degrees are theoretical and do not prepare students for the real world. They are therefore of very little value.
- The role of censorship in Western society
- Marx once said that religion was the opium of masses. If he was alive at the end of the 20th century, he would replace religion with television
- All armies should consist entirely of professional soldiers: there is no value in a system of military service
- Feminists have done more harm to the cause of women than good
- In his novel Animal Farm, George Orwell wrote “All men are equal: but some are more equal than others.” How true is this today?
- Money is the root of all evil

At present, most of the material is in a raw form (untagged [i.e. with no part-of-speech analysis] and unparsed [with no syntactic encoding]). To provide a means of comparison for all the learner essays, ICLE contains a collection of texts written by British and American students (LOCNESS – Louvain Corpus of native English Essays). LOCNESS can be considered (up to a certain extent) a referential material for comparisons of underuse and overuse of features in learner essays. Basically, an “overuse” in a learner corpus denotes a feature frequency higher than in LOCNESS; similarly, “underuse” means a reduced use of a feature in comparison with the native material.

The main difference between ICLE and other English corpora is that ICLE was designed primarily to compare the English language produced by non-native speakers (in other words, their interlanguage) with that of native speakers. It can give an insight into the actual production of language and help us understand the mistakes and discoursive features produced by learners of various mother tongue backgrounds.

As a result a better understanding of peculiarities and non-standard features of learner English can be achieved, which, in turn, may help improve the teaching of
English as a foreign language in the countries concerned. The main value of ICLE lies in its use of advanced learner English as a research material. People who contributed essays typically reached a fluent command of English, almost free from basic mistakes. Still, the transfer from their mother tongues is discernible and the non-native English differs from the native one in many aspects. Therefore, ICLE may be used to help identify blind spots of English language teaching and highlight weaknesses in teaching methodology, be they specific of learners’ national background or not.

2.1 Material used in this study

The portion of the ICLE employed in the present rendering consists of the following subcorpora:

<table>
<thead>
<tr>
<th>Subcorpus</th>
<th>Length (tokens)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Czech</td>
<td>221,583</td>
</tr>
<tr>
<td>Brazilian</td>
<td>33,754</td>
</tr>
<tr>
<td>French</td>
<td>134,236</td>
</tr>
<tr>
<td>Spanish</td>
<td>177,903</td>
</tr>
<tr>
<td>Dutch</td>
<td>130,208</td>
</tr>
<tr>
<td>Finnish</td>
<td>129,045</td>
</tr>
<tr>
<td>Native</td>
<td>182,318</td>
</tr>
</tbody>
</table>

2.2 Basic procedure

The data in Table 1 were retrieved using WordSmith’s Wordlist (Scott, 1996) on raw, non-tagged corpora. In each corpus, the percentages of separate frequencies (e.g. for words like *me, my, mine* and *myself*) were then added up to give the overall scores for each person.

3. Findings

Table 1 and corresponding Graph 1 present frequencies (in percentages for easy comparison among corpora of different lengths) for first and second person pronouns, both for singular and plural.
Table 1. Percentages of first and second person pronouns in ICLE

<table>
<thead>
<tr>
<th></th>
<th>Czech</th>
<th>Brazilian</th>
<th>French</th>
<th>Spanish</th>
<th>Dutch</th>
<th>Finnish</th>
<th>Native</th>
</tr>
</thead>
<tbody>
<tr>
<td>I, me, my,</td>
<td>1.28</td>
<td>0.79</td>
<td>0.55</td>
<td>0.49</td>
<td>0.58</td>
<td>0.70</td>
<td>0.29</td>
</tr>
<tr>
<td>mine, myself</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>you, your(s),</td>
<td>0.47</td>
<td>0.42</td>
<td>0.38</td>
<td>0.29</td>
<td>0.47</td>
<td>0.37</td>
<td>0.14</td>
</tr>
<tr>
<td>yourself/-ves</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>we, our(s),</td>
<td>1.21</td>
<td>0.97</td>
<td>1.24</td>
<td>1.15</td>
<td>0.52</td>
<td>0.93</td>
<td>0.45</td>
</tr>
<tr>
<td>ourselves</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>2.76</td>
<td>2.18</td>
<td>2.17</td>
<td>1.93</td>
<td>1.57</td>
<td>2.00</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Graph 1. Percentages of first and second person pronouns in ICLE

The data can fully support the findings of Ringbon (1998), who, concerned with the vocabulary frequencies, noticed the learners’ overuse of all the pronouns. The proportions of overuse differ across the corpora, though: Czech learners use four times the number of first person singular pronouns used by natives, and while the Dutch use only slightly more of *we/our/…*, the French and the Czech use it three times more than the native speakers. As for the use of the first and second person pronouns, the closest to the native speakers are the Dutch – but this closeness is still a 200 per cent of total overuse.

3.1 First person singular pronouns (I, my, me, mine, myself)
Of all first person singular pronouns, the word I takes the largest portion in all the corpora. The words myself and mine are hardly used and the distribution of my and me is in close connection with the frequency of I. Therefore, the main (and only) objective of this subsection will be a fairly detailed analysis of the use of the pronoun I.

To get a better insight into the use of the pronoun I, Petch-Tyson decided to manually analyze the corpora for its concordances. She observed that in about a half of the occurrences the natives connect the pronoun with verbs in the past tense, often used for recounting personal experiences. The only learner corpus in which the writers seemed to do the same was the Finnish one; the Finns also expressed their feelings and attitudes in real life. On the contrary, in the Swedish, Dutch and French corpora such occurrences were scarce – the learners used the pronouns prevalently to make personal appearance within the text and say what they think.

Using her observations as a starting point for the direction of my research, I have taken on a different approach. My concern was as objective frequency analysis of the actual verbs used in connection with the pronoun I, not a subjective observation. Therefore I retrieved the data that say what kinds of verbs follow the pronoun I (looking at the categories auxiliary / lexical and past tense / present tense) and subsequently found what the actual verbs are. The aim was to see whether Petch-Tyson’s observations, i.e. that the native and Finnish students used past tense more than all the other students, corresponds with statistical results.

3.1.1 Procedures

For Table 2 tagged corpora were needed\(^2\). My original intention, i.e. to retrieve data in WordSmith’s Concord and Wordlist, failed. Due to program bugs the search algorithms in WordSmith are non-standard and unreliable; the wildcards work on a slightly different, but significant way than in other programs, especially when combined with the option “Characters within words” (in Settings / Text characteristics /characters within words).

---

\(^2\) The corpora were automatically tagged using the TOSCA-ICLE Tagger. The TOSCA tagger uses 256 tags, which means that a very refined analysis can be carried out. For example, the tagger distinguishes 22 word classes, many of which are subcategorized to give a total of 256 lexico-grammatical tags, 78 of which are for verb types alone and 15 punctuation and pause tags. For further details see Aarts et al. (1997).
After much experimenting it turned out that the easiest way to obtain the required data is to use MicroConcord, which nicely searched all the tagged texts. The main disadvantage of the program was its search limitations, i.e. the number of concordances is restricted to 1500 lines. It did not matter in this case as all the occurrences of specified verb types were below the limit in all corpora.

The search word was **I_PRON**(pers,sing) (so that all other meanings of *I* (like numbers) be eliminated), the context horizon was set to 0,3 (i.e. 0 words to the left and 3 words to the right – the program was to look for all verbs following the pronoun within a three-word distance, even if they are preceded by an adverb (like in *I also think*)), and the context words search strings to appear within that horizon were set to *_VB(aux,*,pres* / *_VB(lex,*,pres* / *_VB(aux,*,past* and *_VB(lex,*, past* respectively.

MicroConcord does not display frequencies in per cents, but supplies exact number of occurrences (as long as it is lower than about 1500). By relating these numbers to the total number of the pronoun *I* in the corpus I finally achieved the desired frequencies of verb types as presented in Table 2.

### 3.1.2 Findings

Apparently, most verbs are in present tense and hence can hardly express past experiences. Moreover, a manual examination disclosed the fact that 30 to 50 per cent of occurrences of past tense auxiliaries are connected to the word *would* alone, mostly when used in phrases like “*I would like to say that…*”. That means that a great number of past tense verbs is not used for recalling past experience, but again, for appearances of the author within the text as well as for subjective commentaries.

<table>
<thead>
<tr>
<th></th>
<th>Present tense</th>
<th>Past Tense</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>auxiliary</td>
<td>lexical</td>
</tr>
<tr>
<td>Czech</td>
<td>29</td>
<td>44</td>
</tr>
<tr>
<td>Brazilian</td>
<td>30</td>
<td>49</td>
</tr>
<tr>
<td>French</td>
<td>32</td>
<td>51</td>
</tr>
<tr>
<td>Spanish</td>
<td>29</td>
<td>51</td>
</tr>
<tr>
<td>Dutch</td>
<td>28</td>
<td>38</td>
</tr>
<tr>
<td>Finnish</td>
<td>29</td>
<td>50</td>
</tr>
<tr>
<td>Native</td>
<td>24</td>
<td>46</td>
</tr>
</tbody>
</table>

**Table 2.** Percentages of verb types used by ICLE learners after the pronoun *I*
What is most puzzling about the figures above is that the results do not only disapprove of the results of Petch-Tyson’s analysis, but are even more contradictory to them. The Finns are not the learners who use more past tense verbs after *I*, they are actually the ones who use it the least! Therefore, results of the statistical inquiry cannot support Petch-Tyson’s observances. Neither does the inspection of the concordances reveal remarkable differences in the use of *I* and other learners.

This becomes even more obvious when looking at the actual verbs employed after *I* as presented in the next section.

### 3.2 Verbs following the pronoun “I”

For a better understanding of the results in Table 2 a closer look at the actual verbs was needed. A search on words can disclose factual differences between the native speakers and the learners in the use itself, not only in its amount.

#### 3.2.1 Procedures

For the retrieval non-tagged, raw corpora were used. They were run through Wordlist with the option of clusters activated, set to two words (*Settings / Min & Max. Frequencies / Clusters – Size: 2, activated*). This way, Wordlist counted the number of occurrences of all bigrams in the corpora (instead of single words). These numbers were then used for computing amounts of each of the verbs in relation to the total number of the pronoun *I* in the particular corpus (as seen in Table 3).

#### 3.2.2 Findings

As we can see in Table 3, apart from *think, mean, guess, hope, feel, know and believe* the words are all auxiliaries. Nevertheless, in the learner corpora, the lexical verbs together occur more often than auxiliaries. It is also possible to notice that the learners’ pet word is *think*, generally used for stating personal opinion.

The natives balance out the use of *think* by utilizing the argumentative words *believe* and *feel*. The learners seem to be unaware of these words; *believe* only appears in the Brazilian, Spanish and Finnish and *feel* makes no appearance in the learners’ top
ten at all. In the words of Granger (1998: 129), think is the learners’ “cover-all” verb for expressing their own view.

Petch-Tyson makes an interesting observation when she notices that the pronoun I appears in “chains” in the corpora. The following excerpts from the Spanish corpus are self explanatory:

**Table 3. Top ten verbs following I in the ICLE subcorpora (per 100.000 words)**

<table>
<thead>
<tr>
<th></th>
<th>Czech</th>
<th>Brazilian</th>
<th>French</th>
<th>Spanish</th>
<th>Dutch</th>
<th>Finnish</th>
<th>Native</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I think</td>
<td>I think</td>
<td>91 I think</td>
<td>5 I think</td>
<td>51 I think</td>
<td>64 I have</td>
<td>18 I have</td>
</tr>
<tr>
<td>2</td>
<td>I am</td>
<td>I am</td>
<td>47 I am</td>
<td>2 I have</td>
<td>31 I have</td>
<td>36 I think</td>
<td>16 I think</td>
</tr>
<tr>
<td>3</td>
<td>I do</td>
<td>I have</td>
<td>42 I have</td>
<td>2 I do</td>
<td>25 I am</td>
<td>35 I</td>
<td>15 I</td>
</tr>
<tr>
<td>4</td>
<td>I have</td>
<td>I can</td>
<td>27 I do</td>
<td>25 I would</td>
<td>1 I have</td>
<td>20 I would</td>
<td>26 I feel</td>
</tr>
<tr>
<td>5</td>
<td>I would</td>
<td>I</td>
<td>18 I have</td>
<td>25 I don’t</td>
<td>1 I would</td>
<td>20 I would</td>
<td>26 I feel</td>
</tr>
<tr>
<td>6</td>
<td>I was</td>
<td>I do</td>
<td>19 I do</td>
<td>1 I had</td>
<td>15 I believe</td>
<td>22 I am</td>
<td>11 I am</td>
</tr>
<tr>
<td>7</td>
<td>I can</td>
<td>I was</td>
<td>18 I mean</td>
<td>12 I mean</td>
<td>1 I will</td>
<td>14 I don’t</td>
<td>14 I will</td>
</tr>
<tr>
<td>8</td>
<td>I know</td>
<td>I would</td>
<td>18 I don’t</td>
<td>10 I will</td>
<td>1 I don’t</td>
<td>12 I will</td>
<td>11 I was</td>
</tr>
<tr>
<td>9</td>
<td>I will</td>
<td>I could</td>
<td>15 I hope</td>
<td>8 I believe</td>
<td>7 I can</td>
<td>11 I was</td>
<td>9 I do</td>
</tr>
<tr>
<td>1</td>
<td>I mean</td>
<td>I know</td>
<td>15 I can</td>
<td>5 I know</td>
<td>7 I could</td>
<td>11 I guess</td>
<td>7 I know</td>
</tr>
</tbody>
</table>

In my case, for example when I studied in secondary school, I had time to … Sometimes I ask myself why I am studying and I ask myself why I did not … of the human being. Anytime, if I am in mess, I just hope not be here … … children and technology. When I was a child, I used to read short stories … … obvious, if I know that in a few months I will not be there, I will not mind … Could people survive? Personally, I think I that could. When I just imagine … Before I expose these problems, I must say that I am very much against the … When I analyze my own experience, my life, I understand that I was wrong …

This means that the distribution of the pronoun I is not uniform throughout the corpora; rather it clusters in separate essays in succeeding sentences. The overall scores are then influenced by the inclinations of individual authors to employ them rather than by a general strategy of all learners with the same mother tongue background.

I tried to provide evidence for this claim. The most convincing is a visualization of the distribution of the pronoun I in the entire Brazilian corpus. Figure 1 shows the dispersion plot for the pronoun I across individual learner essays. From the left to right we can see the name of the file, the total number of words it has and the total number of hits that were found for I in that particular essay. We can see that while some essays are peppered with the pronoun, others use I only once or twice (the plot area in the figure indicates the length of the essays, each vertical line representing one instance of the pronoun). In the essay named icle30.txt (the first from the top to the bottom) it is possible to observe the chain of I-s permeating most of it. Notice, also, that on in this
particular essay, the frequency of the pronoun *I* (25) represents more than 10 per cent of the total number of words the whole essay (236 words)!

The situation could be compared to counting the county’s average salary: a small amount of millionaires increase the overall mean average of salaries, while most people remain quite poor. Likewise, a relatively small group of “*I*-pronoun millionaires” influences the total number for the whole corpus. High frequency of the word thus indicates that there are a large number of writers who are prone to overusing *I* instead of all the users employing the pronoun too extensively.

### 3.3 First and second person plural pronouns

The excerpts below demonstrate that the pronouns *we* and *our* may be used for references to the general public (1) and (2), to the citizens of a country (3) and (4), to mankind or people in general (5) and (6) – generally, large groups. Learners also use them as general referents, similar to the pronoun *one* or the general *you* (7) and (8). More restrictedly, the pronouns are employed to denote a specific group, e.g. students of a university (9) and (10), or to refer to the author of the text (11).

---

**Figure 1.** Dispersion plot of the pronoun *I* in the Brazilian subcorpus (an extract)

---

3 Please note that these are only observations and presumptions based on manual inspection of the concordances, not statistically based results. For a more profound research a word-to-word analysis of all corpora would be needed. The scope of this paper discouraged me to pursue such a detailed scrutiny.
(1) Despite the futility of life, we still have to choose to live life to the full, and accept our condition with dignity.

(2) Generally, the authors do not withhold the intense emotions experienced by the characters; this graphic and realistic font provides us with an unbiased view of emotions that we thought were unique to our white culture.

(3) For example, they say that Britain is a small island which has stood alone for centuries. We fought in both world wars and survived and although we fought with France etc, we were not part of the community.

(4) The defense of our country is the ultimate goal of our armed forces. If there is no trust amongst our troops, then how can we, as civilians, trust in them?

(5) If we, the people, allow big business to pollute and rape our land, we will all be lesser in the end. We must find better ways to remove and transport our natural resources. Along with this effort, we must push to make large corporations, such as Exxon accountable for their effects on the environment.

(6) Aids is the most common disease of our world today.

(7) When reading the play, we tend to impose our own moral standards and judgements on the characters.

(8) To summarise then, our feelings of sympathy and possible admiration for them are never evoked.

(9) If we had an honor code at Louvain, I would feel an obligation to myself and to my university to turn cheaters in.

(10) What makes a good story apart from the confrontation and conflict, an element that we have been told since the beginnings of our literary education must be present in order for any good work to exist?

(11) Thus, as we have said, he came from a Bourgeois family, he was an only child but was psychologically neglected by his father and mother.

Referring back to Table 1, it is possible to observe the learners’ tendency (with the exception of the Dutch) to use the pronouns we and our more often than the native speakers.

Another interesting phenomenon disclosed by Table 1 and Graph 1 is that even though the frequencies of first person pronouns vary greatly, the number of second
person pronouns is much more balanced in the learner corpora, being about two to three times higher than in the native corpus.

Inspection of the examples containing the pronoun you shows that a vast majority is used to function as a general referent; the reader is being addressed scarcely.

To illustrate this, the following examples were selected from the Dutch corpus, completely at random:

(12) The way you act on television is a very important factor when you are trying to sell yourself.
(13) You always have to wonder whether your partner loves you or just your money. And who are your real friends? You always have to doubt about this sorts of things.
(14) You also have to be more frightened on the street. Because people are able to see whether you are rich or not. For all these reasons, I believe, it's much better to stay simple. If you ask me, you'll enjoy life much more!!!!
(15) I ask her why she did this, and she tells me that, by enlisting for a number of years, you can earn a "tuition waiver".
(16) They are forced to believe that you deserve to be punished if you do not keep the Sabbath or if you forget to pray five times a day (= Salat).
(17) When you go into the army, you are totally independent, you have to do everything on your own. This way, you get a pretty big sense of responsibility.

The fact that the native speakers employ both we and you much less than the learners is a bit of a surprise. Similarly to the learners, the pronouns employed by the native speakers function as general referents in their essays, but phrases making use of them are less common among them. How do the native speakers express the same as the learners, then?

In order to check that, I examined the frequencies of the pronoun one, which could, in many cases, replace the instances of we and you. The survey, though, revealed that the native students do not use the pronoun one more often than the learners.

The conclusion might be drawn that either the topics of their essays do not require the use of general referents, or that the native speakers employ completely different strategies, which avoid personal pronouns at all, e.g. the use of passives or – ing participles.
4. Conclusion and pedagogical implications

It would be unwise to simply prohibit students from using first person pronouns. Much more important is to teach them to use alternative constructions which can express the same idea, and, above all, lead them to maintain consistency in their style at all levels – lexical, syntactic, stylistic and, of course, at the level of the content itself.

Clever, sparse use of author appearances can help refresh a dry academic tone, making it more involving and interesting – particularly when expressing personal experience, not “just” standpoints. In contrast, texts burdened with extensive writer presence are virtually drowned in their subjectivity. When that happens, the original content is covered by the persona of the author, even though he or she does not intend to do so. Finding the right balance should be one of the main aims of English stylistic exercises – at least I should think so.

REFERENCES


