# English and Croatian citation practices in research articles in applied linguistics: a corpus-based study

**Varga, Mirna; Gradečak-Erdeljić, Tanja**

*Download date / Datum preuzimanja:* **2024-04-25**

# English and Croatian citation practices in research articles in applied linguistics: a corpus-based study

**Mirna Varga**
Sub-department of common subjects
Faculty of Humanities and Social Sciences
L. Jägera 9
31 000 Osijek, Croatia
Email: mirna.varga@os.t-com.hr

**Tanja Gradečak-Erdeljić**
Department of English language and literature
Faculty of Humanities and Social Sciences
L. Jägera 9
31 000 Osijek, Croatia
Email: tgradeca@gmail.com

## Abstract

As a mandatory constituent of academic writing, citation allows writers to acknowledge other scholars' work and to position their research against it, showing thus both contribution to previous knowledge and research novelty (Hyland 2004). Previous research has documented not only cross-disciplinary (Hyland 2004) but also cross-cultural variations, with a general tendency of Anglo-American writers to use more citations than writers of some other cultural backgrounds (Hyland 2005; Mur Dueñas 2009). By exploring the frequency, preferred types, and reporting structures of citations in two comparable sub-corpora of research articles in applied linguistics in English and Croatian, the present study aimed to provide an insight into the patterns of cross-cultural similarities and differences in the use of academic citation. The corpus comprised 32 research articles that were sampled from the representative English- and Croatian-medium publications in applied linguistics and analyzed manually. The extracted instances of citations were categorized according to the pre-established taxonomies of the citation types (Swales, 1990) and reporting structures in academic writing (Thomson & Tribble 2001). The frequency analysis showed that the English writers used more citations as opposed to the Croatian writers, which is in line with previous cross-cultural research on the use of citations (Fløttum, Dahl & Kinn 2006; Mur Dueñas 2009). In both sub-corpora writers used more non-integral than integral citations, with the highest

frequencies reported in the Introduction section of research articles. The overall findings point to the saliency of the congruent types of reporting structures in both citation formats across the two sub-corpora, with the human subjects being most frequently used in integral citations and non-reporting being the most frequent reporting structure in non-integral citations. However, in non-integral citations English writers used non-human subjects at a significantly higher frequency than Croatian writers, which suggests that in the English citations investigated more importance is placed on research activities than human agents. Whereas similarities in the use of citations between English and Croatian writers may be accounted for by the congruent disciplinary variable, the differences seem to be related to the specifics of a wider socio-cultural background in which academic writing is embedded.

**Keywords:** citation, academic writing, research article, English, Croatian

## 1 Introduction

Citation, reporting, referencing or attributing propositional content to another source is regarded as one of the defining features of academic discourse, which along with such notions as clarity, precision, concision, logical representation of scientific ideas, normative language use represents one of its universalities (Hyland 2004; Oraić Tolić 2011). According to Shaw and Pecorari (2013:A1), citation may be defined as the "deliberate and explicit use of source material" whose importance to an academic text involves multiple aspects. In particular, it reflects writers' needs to contextualize their research against the background of previous theoretical, empirical or methodological work, to engage in a continuing scholarly debate, creating at the same time a niche for the research at hand as well as signalling its contribution to the existing state of knowledge (Swales 1990; Hyland 2004; Fløttum, Dahl & Kinn 2006). As Gilbert (1977) notes, in order to gain credence for one's work, a writer needs to cite previous research and display how the novel research builds on and furthers previously validated theories or empirical findings. By carefully selecting to cite the most relevant but also the most significant prior research that has already gained approval by a disciplinary community, a writer shows allegiance to a pre-established body of knowledge which his or her research may add to, raising thus its credibility and importance (Gilbert 1977). However, this will be achieved not by merely citing significant research for its own sake but by convincing readers how the novel research draws on and extends previous knowledge (White 2001). Therefore, in a rhetorical sense, if warranted, citation represents an important persuasive tool used to support and strenghten a writer's argumentation, with an ultimate aim of gaining acceptance by disciplinary members (Gilbert 1977; White 2001; Hyland 2005). Furthermore, reference to previous research network signals the interdependence of previous, current, and prospective scientific texts, contributing thus

154

to a socially constructed nature of scientific knowledge, which is considered to be an underlying notion of a contemporary conceptualization of academic discourse (Hyland 2004; Hyland 2009; Oraić Tolić 2011).

The multiplicity of discourse-analytic approaches[1] to the study of citation practices within the realm of EAP (Swales 1986) speak in favor of both significance but also complexity involved in their use, raising thus a need for advancing knowledge of this conventional feature of academic writing. This is particularly important for intercultural research on academic writing, most notably for novice non-native writers coming from smaller academic communities, whose rhetorical conventions in academic writing, including the use of citations may be in some aspects different from those in English and thus pose difficulties when publishing their research in English-medium journals (Mur Dueñas 2009; Carter-Thomas & Rowley-Jolivet 2013). Consequently, empirical disciplinary studies on intercultural rhetorical conventions in academic writing are increasingly important as their findings might assist non-native scholars in gaining more competence in writing academic English and thus facilitate the process of making their research internationally more visible.

When it comes to the Croatian linguistics, academic writing is a severely under-researched area so knowledge on its rhetorical conventions and even more so a comparison with those of Anglo-American academic writing is still rather scarce. Driven by the motivation to make a small step in filling this gap, the present study aims to explore the nature of citation patterns in the disciplinary writing in the two comparable corpora of English and Croatian research articles in applied linguistics with the intention of providing insights into the specifics of their use.

## 1.1 Literature review

The centrality of acknowledging other scholars' work in academic writing can be supported by the existence of various citation and format styles which, among others, prescribe the norms and offer guidelines for crediting different types of sources in particular scientific domains and avoiding plagiarism. For instance, MLA (Modern Language Asociation) style is primarily used in humanities and liberal arts while APA (American Psychological Association) style is mainly used by writers in social sciences.[2] The distinctive features of the given style guides seem to mirror the disciplinary differences that may be related

---

[1]    For a more detailed overview of further scientific approaches and scientific disciplines to the study of citations, see *e.g.* White (2004).

[2]    For more information on the differenes between the central citation styles in English see *e.g.* https://owl.english.purdue.edu/owl/resource/949/01/. For a detailed overview of different styles in citing sources in Croatian, see Oraić Tolić (2011).

to the distinctive epistemological basis of specific scientific domains. As a way of illustration, while integrating an author's name in the text is equally important (in-text or integral citation) in abovementioned disciplines, humanities writers are advised to write only an author's name and the page number of the cited source due to the importance of authorship in academic writing in humanities. By contrast, in social sciences it is necessary to add a publication date of the reported source, which reflects the importance of marking a recent status of social sciences research and thereby its potential relevance.[3]

Indeed, an assumption that citation practices may be to a certain extent disciplinary-bound, has given impetus to cross-diciplinary empirical research aimed to explore how different disciplines with the distinctive subject matters and approaches to their explorations make use of citation patterns in their written work (Hyland 2004; Hyland & Tse 2004). For instance, in his comprehensive study on the frequency and types of citations in the corpus of research articles across 8 scientific disciplines, Hyland (2004) showed that writers in soft sciences (*e.g.* sociology and marketing) used considerably more citations as compared to their peers in hard sciences, such as mechanical engineering and physics. The fact that hard sciences are characterized by a relatively clearly defined research areas with a common theoretical background and methodology places fewer demands on writers to report extensively on well-established facts. By contrast, in humanities knowledge is more dispersed, less solid and more interpretative which requires from writers to place their research in a wider framework of previous knowledge and research, resulting thus in higher frequency of citations (Hyland 2004).

In addition to disciplinary variations, prior research has also documented that the use of citations may be susceptible to specifics of distinctive academic genres (Hyland 2005). For instance, as a part of a more comprehensive study on the use of metadiscourse in university textbooks and research articles in comparable disciplines, Hyland's (2005) research showed that citations (in his terms *evidentials*) are a more prominent feature of research articles as compared to textbooks. The distinctive use of citations in the given academic genres may be accounted for by their different rhetorical purposes, textbooks being primarily concerned with passing on the established knowledge rather than indications of the originators of the claims, which is a characteristic rhetorical function of citations used in research articles (Hyland 2005).

Previous research has also demonstrated cross-cultural variations in the use of citation patterns (Taylor & Cheng 1991; Fløttum, Dahl & Kinn 2006; Okamura 2008; Mur Dueñas 2009; Hu & Wang 2014; Dontcheva-Navratilova 2016), which generally points to a greater use of citations among Anglo-American writers as compared to the writers of

---

[3]    https://owl.english.purdue.edu/media/pdf/20110928111055_949.pdf

other minor academic communities. One of the most cited reasons for a higher saliency of citations by Anglo-American community is related to its heterogeneous and more competitive nature, which prompts writers to establish more links with previous research in order to position their study and persuade readership of its significance and validity (Mur Dueñas 2009; Dontcheva-Navratilova 2016).[4] By contrast, writing in smaller national academic contexts and addressing regional readership may place different demands on authors whose argumentation may be focused more on their research rather than on rhetorical engagement with previous studies (Dontcheva-Navratilova 2016).

Another significant research strand on citations has been particularly interested in accounting for their structural features as well as the meanings realized in an academic text. These approaches share a common feature of being oriented towards building taxonomies of citation forms and functions, which are further used in exploring citation patterns in different disciplines (Thomson & Tribble 2001; Charles 2006; Petrić 2007).

With regard to the surface textual features, the most basic division of citations recognizes two essential structures in which they appear in the text, *viz.* integral and non-integral (Swales 1990). As the labels suggest, in the case of the former, the name of the cited author is integrated in the sentence and is usually followed by the year of the publication in parenthesis. On the other hand, in non-integral citations, the author's name is given in the parenthesis placed outside of the sentence or at some other point in the text, which is indicated by a (superscript) number. The surface distinction between the two structural types of citations enables writers to alternate between prioritizing the author, in which case an integral citation is used, or to place more focus on the information, whereby the author's name (*i.e.* his or her role) remains in the background.

Swales' distinction between the two basic citation forms has served as the basis for more detailed elaborations of the citation patterns, which aimed to illuminate their immediate rhetorical functions in the text (Thomson & Tribble 2001; Petrić 2007; Dontcheva-Navratilova 2016). In other words, scholars were interested in finding out how writers exploit different citation patterns to meet different purposes in developing their argumentation. Thus, some common rhetorical functions include attributions of the reported content to the original author or reference, directing a reader to the source of the information, commonly signaled by an abbreviation *cf.* or a comment *see*, *etc.* (Thomson & Tribble 2001; Petrić 2007).

---

[4]    For additional cross-cultural accounts of citation conventions in academic writing, see *e.g.* Hu and Wang's (2014) discussion on how different perspectives on scientism may account for the distinctive features of the Anglophone and Chinese literacy practices, including the use of citations.

In addition to the studies of the formal features and rhetorical characteristics of citations, scholars were also interested in accounting for the use and functions of reporting verbs as one of the most salient markers of reporting structures (Thomas & Hawes 1994; Hyland 2004). Different taxonomies have attempted to account for the way writers exploit the semantic and syntactic potential of the lexical verbs in reporting information from other sources (Thomson & Ye 1991; Thomas & Hawes 1994) as well as how the choice of the verbs reveals a writer's stance towards the reported content (Hyland 2004). In addition, scholars have also been interested in the usage of verb tenses and voice, with a broad aim of elucidating potential regularities between these, which is a particularly important topic for pedagogical purposes (Swales 1990; Shaw 1992; Charles 2006).

Against the background outlined thus far, the present corpus-based analysis takes a cross-cultural perspective and aims to address the following research questions:

a) What is the overall frequency of citations in the comparable English and Croatian sub-corpora?

b) What is the ratio of integral vs. non-integral citations in the two sub-corpora?

c) How is the frequency of citations distributed across the IMRAD structure of research articles in the two sub-corpora?

d) What are the most common reporting structures of integral and non-integral citations in the sub-corpora of English and Croatian articles?

e) What are the most frequently used reporting verbs in the two sub-corpora?

## 2 Methodology

### 2.1 Corpus design

To meet the above-mentioned research objectives, the comparable corpus of research articles (RAs) in a broad field of applied linguistics was compiled based on a set of pre-established criteria. First, in order to compare the citation patterns in the congruent type of article, the selected Croatian articles had to have the label *Izvorni znanstveni rad* 'Original scientific article', while the English articles had to be classified as empirical studies. Furthermore, the articles used in compiling the corpus had to follow the conventional IMRAD structure, which allowed comparing the saliency of citations across the rhetorical sections of the targeted genre in the two sub-corpora. This was achieved in the majority of cases, with the exception of four Croatian articles that had a combined Results and Discussion section. In order to allow for the comparison of the frequency counts in the corpus as a whole, a decision was made to treat the joined

sections as the Discussion section given that the authors mainly commented on the reported results and drew conclusions based on them. In addition, it should be noted that not all the articles had the explicitly headed IMRAD sections. For example, in some Croatian articles the closing section was labeled as 'Interpretation of results' followed by a separate section 'Concluding remarks'. Given their communicative functions, such textual segments were treated as the Discussion section.

With respect to the Croatian sub-corpus, the articles that met the above stated criteria were extracted randomly from two sources of publications, which are considered to be among the most relevant for applied linguistics in Croatia. The first is the journal *Strani jezici* 'Foreign Languages', which is the only Croatian-medium journal that can be classified as being specifically oriented to applied linguistics. The second source involved the conference proceedings of the Croatian Applied Linguistics Society 'Hrvatsko društvo za primijenjenu lingvistiku', which publishes original scientific articles in the targeted discipline. The fact that the given publications were selected for the compilation of the present corpus does not entail that scientific articles in applied linguistics cannot be found in other Croatian linguistics journals but the selected two were estimated to be among the most representative ones. As for the compilation of the English sub-corpus, following previous research (Hyland 2002), the journals were chosen based on the information obtained from the Croatian scholars whose primary field of study involves applied linguistics. A short written request was sent to subject specialists in which they were asked to nominate the leading English-medium journals in their sub-domains. Based on the information obtained, two journals were selected for the present corpus, *Language Learning* and *Studies in Second Language Acquisition*.

The corpus consists of 32 research articles, totaling 195 265 running words (Table 1).[5] Each sub-corpus consists of 16 articles, whereby 8 articles were extracted from each journal. Given that the aim of the study is to explore the contemporary use of academic language, the present corpus consists of the articles published in the period between 2008 and 2017 (Bowker & Pearson 2002). In addition, as one of the research objectives is to explore the frequency counts across the IMRAD sections of RAs, each sub-corpus is further divided into four corpus sections to allow the comparison of the findings across the two sub-corpora.

---

[5] The full list of the corpus articles along with the abbreviations can be found in the section Data sources at the end of the paper. The corpus included only the text body, excluding the abstract, title, keywords, graphic material, footnotes and bibliographical data. Therefore, the analysis focused only on the citations in the main rhetorical sections of the articles.

| Sub-corpus | Word count |
|---|---|
| Croatian sub-corpus | 52 159 |
| English sub-corpus | 143 106 |
| **Total** | **195 265** |

Table 1. Corpus size

As the study aims to elicit some cross-cultural patterns in the use of the academic citation, an attempt was made to ensure that the English sub-corpus comprised articles written in native English. This variable was controlled or rather approximated by ensuring that the first authors of the articles had an English-sounding name and surname and that they were affiliated with the University departments in the English-speaking country (Yang *et al.* 2015).[6] These criteria do not ensure that the language used in the articles is native indeed but following previous research (Fløttum, Dahl & Kinn 2006; Yang *et al.* 2015), an assumption was made that even if English was not the writers' mother tongue, such cases were estimated not to be high in number to skew the corpus data (Varga 2016). Against this background, we adopt the term *English writer* to refer to the American, British, and Canadian scholars, affiliated to the University departments from the given countries and the term *Anglo-American community* to refer to the academic writing originated in them (Koutsantoni 2006; Varga 2016)

**2.2 Procedure**

Once the sub-corpora were compiled, a manual analysis was conducted to categorize all citations into two basic types, integral and non-integral, as shown in the examples below. Based on previous research (Hyland 2004), the analysis included only citations that referred to other researchers, which means that self-citations were excluded from the analysis.

Furthermore, as one of the research aims was to analyze the reporting structure of each of the two citation types, several typologies of citation patterns (Swales 1990; Hyland 2004; Thompson & Tribble 2001; Charles 2006) were explored and consulted for the compilation of the taxonomy used in the present analysis. Though space limits prohibit a more detailed comparison of the given typologies, at this point it suffices to note that, different labels notwithstanding, the core reporting structures in citations refer to (non) human subjects, passive agents, and noun phrases. Based on the previous typologies and the examination of the present corpus data, the corpus analysis yielded the following categories of the reporting structures, each exemplified by the data from the two sub-corpora:

---

6    For further discussion on the language variable in the cross-cultural research on academic writing, see Varga (2016).

**Integral citations**

    **a) Human subject + reporting verb**

(1)    ***Krashen (1985) argued*** *that ample L2 input leads to acquisition.* (Eng5, 37)

(2)    ***Guo i suradnici (2009) proveli su*** *istraživanje koje je pokazalo kako su dječaci zainteresiraniji za vozila, <…>* (Cr6, 30)
    '**Guo *et al.* (2009) conducted** the research that showed that boys are more interested in vehicles <…>'

As the examples illustrate, this type of integral citations is signaled by the presence of a cited author's name followed by a reference year in the brackets and a reporting verb.

    **b) Non-human subject + reporting verb**

This type of citation can be manifested by a:

- Possessive noun phrase (genitive *'s* or *of/by* phrase) (English)

(3)    *The correlations found in this study are higher than those reported in **Horst et al.'s** (1998) study (.31 ≤ r ≤ .36), <…>* (Eng16, 668)

- Possessive noun phrase/Genitive phrase (Croatian)

(4)    Spomenimo **Schmittovo istraživanje (1997)** koje je rezultiralo klasifikacijom SUV-a sukupno 58 različitih strategija koje <…> (Cr12, 413)
    'We should mention **Schmitt's (1997) study** that has resulted in the taxonomy of 58 different vocabulary learning strategies <…>'

    **c) Passive agent (English)**

(5)    *The second set of theories relates to CLI, **explored by** Jarvis and Pavlenko (2007).* (Eng12, 453)

    **d) Adjunct agent phrase (English)/Preposition or prepositional phrase[7] (Croatian)**

(6)    ***According to Haskell et al.'s (2003) constraint satisfaction model***, *singular nouns being neither semantically nor phonetically plural may appear in compounds (e.g. rat catcher).* (Eng11, 198)

---

[7]    In Croatian: *prijedlog* or *prijedložno-padežni izraz* (Silić & Pranjković 2005).

(7)    ***Prema Miederu*** *(1985: 119) poslovice su kratke općepoznate rečenice nastale u*
       *narodu <…>* (Cr10, 207)
       '**According to Mieder** (1985, p. 119), proverbs are short, general statements that
       originated among ordinary people <…>'

### e) Non-citation[8]

In this citation type, a cited author's name appears without the publication year in the
brackets and is followed by a reporting verb. Alternatively, this citation type may include
instances where personal or possessive pronouns or noun (phrases) occur without the
publication year.[9]

(8)    *In addition,* ***Sparks and Ganschow proposed*** *that L2 anxiety surveys would likely*
       *be measuring either students' level of language learning ability, self-perceptions*
       *of their language learning skills, or both.* (Eng14, 871)

(9)    ***Deci i Ryan razlikuju*** *pet izvora motivacije, od eksternalne preko introjicirane,*
       *identificirane i integrirane pa do intrinzične motivacije <...>* (Cr13, 5)
       '**Deci and Ryan differentiate between** five sources of motivation, ranging from
       external, introjected, identified, integrated to intrinsic motivation <...>'

### Non-integral citations

In non-integral format, the following reporting structures were distinguished:

### a) **Human subject + reporting verb** (English and Croatian)

(10)   ***L1 researchers have long known*** *that early success in L1 skills (e.g., in word*
       *decoding in first grade) predicts performance in L1 reading and spelling*
       *throughout elementary school and into high school (see, e.g., Cunningham &*
       *Stanovich, 1997; Stanovich, 2000).* (Eng14, 889)

(11)   ***Mnogi primijenjeni lingvisti*** *(cf. Ellis, 1997; Read, 2000; Schmitt, 2000)* ***ističu***
       *aktivnost dvaju procesa u usvajanju vokabulara: implicitnoga i eksplicitnoga*
       *učenja.* (Cr12, 412)
       '**Many applied linguists stress** the activity of two processes in the vocabulary
       acquisition: implicit and explicit learning.'

---

8      The term has been adopted from Thomson and Tribble (2001).
9      The occurrences which signal general reference (Charles 2006) and do not refer to the
previously cited authors or their work were excluded from the analysis, *e.g.*: *A small number of*
*studies suggest that social networks play a beneficial role in the SA experience.* (Eng5, 40)

**b) Non-human subject + reporting verb** (English and Croatian)

(12) *Specifically, **these investigations have demonstrated** the effectiveness of using acoustically varied, compared with acoustically consistent, presentation formats during training for this contrast (Lively, Logan, & Pisoni, 1993; Lively, Pisoni, Yamada, Tokura, & Yamada, 1994; <...>)* (Eng13, 188)

(13) ***Rezultati istraživanja** korelacija u tom odnosu (Müller, Andreitz i Hanfstingl, 2008) **ukazuju na** direktne i indirektne veze između institucionalnih uvjeta, potpore koju nastavnici dobivaju od kolega i uprave škole, i motivacije učenika.* (Cr13, 6)
'**Research findings** on the correlations in that relation (Müller, Andreitz & Hanfstingl, 2008) **point to** direct and indirect relationships between institutionalized conditions, the support that teachers obtain from colleagues and school management and students' motivation.'

**c) Impersonal -*It* subject or passive subject (English)/Impersonal or passive construction** (Croatian)

(14) *While **it has been noted** that literate bilingual children use sound similarities to identify cognates (Dressler et al., 2011; Kieffer & Lesaux, 2008), it is not fully understood if <...>* (Eng3, 449)

(15) *<…> pod utjecajem psiholingvistike i sociolingvistike u zadnjih **se** tridesetak godina sve više **istražuju** individualne razlike (Dornvei, 2005) <…>* (Cr8, 8)
'<…> influenced by psycholinguistics and sociolinguistics, individual differences **have been increasingly studied** over the last thirty years (Dornvei, 2005) <…>'

**d) Non-reporting[10]**

In this type of non-integral citations there is no clear signal on the reporting structure, yet the citation is marked by the presence of the reference given in the brackets (Swales 1990). In such instances writers may, for instance, paraphrase or summarize reported authors' findings or other research data without explicit signals of the report, as shown in the following instances:

(16) *When asked to produce synthetic compounds made up of two nouns in which the first nonhead noun is plural, English-speaking children (Gordon, 1985; Nicoladis, 2003; Oetting & Rice, 1993; van der Lely & Christian, 2000), native German-speaking children (Clahsen, Marcus, & Bartke, 1993), ... more often include irregular than regular plurals in compounds.* (Eng11, 195)

---

[10]    The term has been adopted by Swales (1990).

(17)    *Postoji više klasifikacija strategija učenja vokabulara (npr. O'Malley i Chamot,*
        *1990; Oxford, 1990; Pavičić Takač, 2008; Rubin i Thompson, 1994; Schmitt,*
        *1997; Wenden i Rubin, 1987).* (Cr2, 33)
        'There are many classifications of vocabulary learning strategies (*e.g.* O'Malley
        & Chamot, 1990; Oxford, 1990; Pavičić Takač, 2008; Rubin & Thompson, 1994;
        Schmitt, 1997; Wenden & Rubin, 1987).'

When it comes to counting citations, we followed the methodological procedure outlined
in Mansourizadeh and Ahmad (2011) and treated the instances of both single and multiple
references as one citation point. This is particularly important with respect to the use of
non-integral citations containing multiple references in brackets or "clusters" (Fløttum,
Dahl & Kinn 2006), as shown for instance in example (17) above.

The raw frequencies of the citation patterns were calculated for each sub-corpus as a
whole and for each RA section. In order to compare the findings across the two sub-
corpora, raw frequencies of the identified citations were normalized to a text length of
1,000 words (Biber 1988).

## 3 Results and discussion

### 3.1 Distribution of integral vs. non-integral citations

As can be seen in Table 2, the findings (in terms of normalized frequencies) point to a
higher frequency of the citation use (including both integral and non-integral citations)
in the English (n/1000 = 8.69) as opposed to the Croatian sub-corpus of research articles
(n/1000 = 7.28). Generally, the present findings conform to previous cross-cultural
research, showing a more preferred use of citations by English writers of RAs as opposed
to RA writers of other languages, such as Spanish (Mur Dueñas 2009), Norwegian or
French (Fløttum, Dahl & Kinn 2006).

| | English sub-corpus | | | Croatian sub-corpus | | |
|---|---|---|---|---|---|---|
| | Raw frequency | n/1000 | % | Raw frequency | n/1000 | % |
| Integral citations | 586 | 4.1 | 47% | 176 | 3.37 | 46% |
| Non-integral citations | 658 | 4.6 | 53% | 204 | 3.91 | 54% |
| **TOTAL** | 1244 | **8.7** | 100% | 380 | **7.28** | **100%** |

Table 2. Overall frequencies of citation use in the corpus

Despite this general difference in the overall results, the present analysis points to some similar tendencies in the use of citations across the two sub-corpora. As demonstrated in Table 2, though the differences in the frequencies between integral and non-integral citations are not particularly striking, non-integral citation is a more preferred type of citation format in both sub-corpora. This suggests that both English and Croatian writers generally place the cited authors both in the foreground and background depending on the rhetorical purposes but do not seem to particularly favor one over the other citation format. Though Hyland's (2004) research findings also pointed to a higher frequency of non-integral than integral citations in congruent disciplinary writing, the present English findings diverge from these in that the ratio between the two citation types is considerably lower. More specifically, Hyland's (2004) analysis pointed to 65% of non-integral and 35% of integral citations in RAs in applied linguistics, while the present analysis showed 53% of non-integral vs. 47% of integral citations. Given that detailed information on Hyland's corpus and methodology is rather missing, the findings may not be directly comparable. Nevertheless, the discrepancies between Hyland's and our findings may be due to the fact that in our analysis, as previously mentioned, non-integral citations were counted as one citation point which may have resulted in lower counts than would have been obtained if each name in the brackets was counted as one instance of a non-integral citation.

## 3.2 Distribution of citations across IMRAD

When it comes to the distribution of citations across the four major rhetorical sections of RAs, the present results also show some general similarities across the two sub-corpora. As illustrated in Figure 1, in both sub-corpora citations were predominantly used in
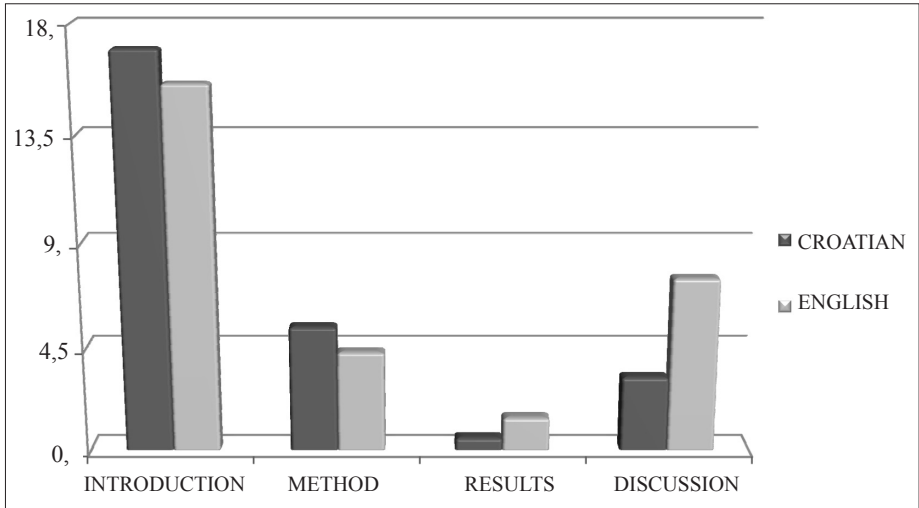


Figure 1. Normalized frequencies of citations across IMRAD in the Croatian and English sub-corpora

the Introduction followed by the Discussion and Method yet severely underused in the Results section.

These findings are generally expected given the general rhetorical functions of the respective RA sections. Thus, the centrality of the citation use in the Introduction section can be accounted for by the fact that writers need to position their research against the background of the existing knowledge as well as previous empirical findings. These rhetorical purposes call for a dense use of citations, which can take different surface features and can perform a range of different rhetorical functions (Petrić 2007). Though a detailed analysis of the rhetorical functions of citations is beyond the scope of the present paper, some broad tendencies emerging from the work on the corpus data are worth noting. For instance, in the Introduction section, writers frequently report on previous research by taking particular stance towards the reported information, as signaled by the choice of the reporting verb. Thus, in the following example, the subject initial position of the cited authors along with the choice of the reporting verb *isticati* 'highlight/stress' signals a writer's decision to foreground strong commitment to the reported proposition.

(18)   ***Dorman, Fisher i Waldrip (2006) ističu*** *da je kvaliteta interakcije koju nastavnik ostvaruje s učenicima ključna determinanta učenja* <…> (Cr13, 6)
       '**Dorman, Fisher i Waldrip (2006) highlight** that the quality of interaction that a teacher establishes with students is a key determinant of learning <…>'

Alternatively, writers may choose to summarize and interpret the reported information without providing an explicit marker of the attribution. Thus, a choice of a non-reporting non-integral citation format allows writers to condense lengthy information particularly regarding a series of congruent research that needs to be incorporated into the text and linked to the purposes of the research at hand, *e.g.*:

(19)   *Furthermore, regular plurals do occur internal to compounds in languages other than English, such as Dutch (Schreuder, Neijt, van derWeide, & Baayen, 1998) and Spanish (Lardiere, 1995).* (Eng11, 196–197)

Contrary to the Introduction, a rather low frequency of the citations in the Results section conforms to the general rhetorical purposes of this RA section that is primarily oriented to the presentation of findings without a particular need to refer to previous literature, which was shown to be more notable in the Croatian sub-corpus (Figure 1).

When it comes to the Method section, both English and Croatian writers mostly used citations to cite other scholars' methodological approaches, procedures, tools, *etc.* applied in their research, as shown in:

(20) ***The test*** *was virtually identical to the one described previously, developed by Ortega et al. (1999), with only a few changes.* (Eng2, 657)

(21) *Rezultati istraživanja prikupljeni su pomoću triju mjernih instrumenata: demografskog upitnika (Prilog 1),* **upitnika o motivaciji** *(Mihaljević Djigunović, 1998) (Prilog 2) i upitnika o strategijama učenja SILL - Strategy Inventory for Language Learning, Version 5.1 (ESL/EFL) (Oxford, 1990) (Prilog 3).* (Cr8, 12)
'The research results were obtained using three instruments: the demographic questionnaire (Appendix 1), **the questionnaire on motivation** (Mihaljević Djigunović, 1998) (Appendix 2) and the questionnaire on learning strategies SILL-Strategy Inventory for Language Learning, Version 5.1 (ESL/EFL) (Oxford, 1990) (Appendix 3).'

Finally, in the Discussion section, writers typically employed citations to compare or generally comment on their findings in relation to those of previous research, as illustrated in the following example:

(22) *Whereas* **previous research** *(Barcroft & Sommers, 2005; Sommers & Barcroft, 2007)* **indicated that** *phonetically relevant variations of the spoken form of words facilitates L2 word learning,* **the present study indicates that** *varying exemplars of referents decreases L2 vocabulary learning in absolute beginners using a time-restricted training task.* (Eng13, 205)

These rhetorical choices were shown to be particularly important to English writers who used citations significantly more frequently in this section as compared to their Croatian counterparts. Indeed, the overall results show the highest discrepancies of the citation use particularly in this section, with English findings pointing to 7.53 and Croatian to 3.2 occurrences per 1 000 words. Though a more detailed investigation on the motives for citation use would be required here to interpret the findings more plausibly, it seems that in the final sections of their RAs, Croatian writers focus primarily on their research findings and broader implications without evaluating them in relation to previous research. This may be related to the nature of the research being more locally than internationally oriented, and consequently involving fewer previous studies, which may lead to a lesser use of citations. Admittedly, this assumption stands in stark contrast with the frequency of citations in the Introduction section in the Croatian sub-corpus (Figure 1). The findings obtained are surprisingly congruent to Mur Dueñas's (2009) results in her analysis of citation use in the Spanish vs. English corpus of research articles in business management. Similarly to Spanish writers (Mur Dueñas 2009), Croatian writers tend to report rather extensively on previous literature and research in the Introduction but do not seem to engage proportionally in commenting on these in the interpretation

and implications of their research in the Discussion or concluding sections of RAs. This assumption, however, should be verified by further corpus-based investigations that might compare rhetorical functions of citations used in the Introduction and Discussion sections and explore cross-cultural specifics in the rhetorical practices in that respect.

### 3.3 Surface features of integral and non-integral citations

As stated above, one of the research aims was to explore the surface reporting structure of the two types of citations across the two sub-corpora. Following previously stated categories, the analysis yielded the results presented in Table 3.

| | CROATIAN SUB-CORPUS | | ENGLISH SUB-CORPUS | |
|---|---|---|---|---|
| **INTEGRAL CITATIONS** | Raw freq | n/1000 | Raw freq | n/1000 |
| Human subject | 94 | **1.80** | 289 | **2.01** |
| Non-human subject | 10 | 0.19 | 125 | 0.87 |
| Passive agent (only English) | – | | 32 | 0.22 |
| Adjunct | 29 | 0.55 | 2 | 0.01 |
| Non-citation | 43 | 0.82 | 138 | 0.96 |
| **NON-INTEGRAL CITATIONS** | | | | |
| Human subject | 7 | 0.13 | 51 | 0.35 |
| Non-human subject | 33 | 0.63 | 234 | 1.63 |
| Impersonal –*It* subject/passive agent (Cro. obezličenje/pasivna konstrukcija) | 29 | 0.55 | 93 | 0.64 |
| Non-reporting | 135 | **2.58** | 280 | **1.95** |

Table 3. Frequencies of reporting structures in integral and non-integral citations across the two sub-corpora

As can be seen, when it comes to integral citations, in both English and Croatian sub-corpora, **human subject + reporting verb** is by far the most frequently employed reporting structure. The preference for the subject position of the cited authors' names suggests that disciplinary writers in both languages prefer highlighting the role of the cited authors in constructing knowledge, as reported by previous research on citation practices in applied linguistics (Hyland 2004). The subject position shifts readers' attention to the cited author, contributing thus to a greater visibility of his or her work within a disciplinary context (Hyland 2005).

When it comes to other structures of integral citations, the findings show variations between the two sub-corpora. Thus, English writers make a significantly greater use of non-human subjects as opposed to Croatian writers. As previously noted, this category of integral citations can be expressed by a possessive noun phrase, as illustrated in the following example:

(23) *While **Dressler's (2000) findings** hinted at the potential for emergent literate children to rely on words' shared phonology to recognize cognates, very few studies have confirmed this hypothesis.* (Eng12, 450)

This construction allows writers to attribute a reporting act to an inanimate subject, while retaining the active voice, which makes the whole argumentation more visible than is the case with the impersonal or passive constructions (Dorgeloh & Wanner 2009). Indeed, the present findings show a considerably lower frequency of the given verbal forms especially in integral citation format, suggesting that disciplinary writers opt for a more personal rather than impersonal involvement with previous scholarly work (Table 3).

According to the corpus findings, Croatian writers do not show preference for this type of integral reporting structure. Instead, they make a more frequent use of adjunct agents as compared to English writers, as illustrated by the following example:

(24) ***Prema Hillu (2001****), jezik je u osnovi leksički, a jezični se obrasci slažu u rasponu –* <…> (Cr11, 146)
'**According to Hill (2001**), language is essentially lexical, and linguistic patterns are built in a range of – <…>'

As can be seen, the adjunct reporting structure allows writers to drop a reporting verb and consequently omit adopting explicit evaluative stance towards the reported information, leading to a more distant position towards it (Yang 2013). However, the extent to which the given structure is a salient citation format in the Croatian academic writing as opposed to others would require further corpus-based disciplinary research.

When it comes to non-integral citations, the findings show that non-reporting is the most preferred citation format by both English and Croatian writers. This format allows integration of reported information into the structure of one's argumentation without taking an explicit stance towards it and consequently assuming responsibility towards it, which largely matches rhetorical purposes of writers in both language groups (Yang 2013). However, the findings that merit closer attention refer to the use of inanimate subjects, which were shown to be particularly favored by English writers. As can be seen in Table 3, inanimate subjects (*e.g.* research, findings, study) were used far more frequently than animate agents in the English sub-corpus. Higher frequency of the former

in the given citation pattern indicates that in citing other scholars' work or knowledge, more importance is given to research activities (both through concrete and abstract nouns) rather than to human agents as their originators.

As can be seen from the results, while human subjects are considerably more prevalent in integral citations, which, as noted before, give a prominent status to the cited author, non-human subjects are preferred in non-integral citations. As previously mentioned, one of the reasons for the latter might be convenience provided by this structure in summarizing and underscoring shared features of congruent research, such as research objectives, as shown in the following example:

(25)   ***Although few studies*** *have addressed language play in terms of multicompetence (but see our discussion below of Belz, 2002; Belz & Reinhardt, 2004; Pomerantz & Bell, 2007), many have highlighted the resources L2 users deploy for play and humor.* (Eng1, 75)

By employing such citation clusters writers may cite multiple studies in one citation point, which, given the size of the international Anglo-American academic community and the space limitations in specific disciplinary journals' editorial policies, seems to be a suitable citation pattern frequently exploited by English writers. To what extent such clusters are more exploited than single-authored non-integral citations and how the latter stand in relation to single-authored integral citations might be a subject of further disciplinary research, which would certainly illuminate distinct ways in which writers employ the given patterns in order to meet desired rhetorical purposes.

By contrast, Croatian writers employ the given citation clusters markedly less frequently and mostly prefer non-reporting citations instead. Though motivations for the preferred status of this citation type should be further investigated, as stated earlier, writers seem to use this pattern as it allows them to convey the reported information without explicit subjective involvement (Yang 2013). This rather impersonal reporting style appears to be in line with the predominantly impersonal and non-evaluative characterization of the Croatian scientific writing style (Silić 2006). However, as previously mentioned with other citation categories, this assumption needs to be verified against further empirical findings.

### 3.4 Use of reporting verbs

The final aim of the present corpus analysis was to provide a broad insight into the choice of the reporting verbs in the two sub-corpora. In our attempt to explore the preferred choices of the reporting verbs used by the English and Croatian writers of RAs, we conducted a frequency analysis of the extracted reporting verbs identified in each sub-corpus. However, due to a relatively high ratio of the reporting verbs that occurred rather

infrequently, we present and discuss here only the frequencies of the verbs that occurred five or more than five times in the corpus as a whole.

The overall findings show a significantly greater diversity of reporting verbs used in the English as opposed to the Croatian corpus. More specifically, the English writers used a range of 126 different reporting verbs while the Croatian writers made use of 73 verbs. With respect to the English corpus, 26 reporting verbs (19%) occurred five or more than five times, while 49% of the verbs were used only once in the whole sub-corpus. By contrast, the analysis of the Croatian sub-corpus showed that 12 reporting verbs (16%) occurred at the congruent frequency, while some 65% of the verbs occurred only once.

As can be seen in Table 4, the findings demonstrate that English writers show preferences towards specific lexical verbs, with an overwhelming centrality of the reporting verb *find*, followed by *suggest, show, argue,* which were half as frequent. The remaining verbs showed a considerably less frequent use in the sub-corpus as a whole. It is interesting to note that the present findings provide support to Hyland's (2004) results which point to the centrality of exactly the same reporting verbs (*i.e. suggest*, *find*, *argue*, *show*) in the corpus of RAs in applied linguistics, suggesting their salient status in the written discourse of this discipline.

| Croatian sub-corpus | | | English sub-corpus | | |
|---|---|---|---|---|---|
| **Reporting verb** | **Raw frequency** | **n/1000** | **Reporting verb** | **Raw frequency** | **n/1000** |
| *isticati* 'highlight'/'emphasize' | 14 | **0.26** | **find** | 106 | **0.74** |
| *provesti* 'conduct' | 12 | 0.23 | suggest | 58 | 0.40 |
| *definirati* 'define' | 10 | 0.19 | show | 45 | 0.31 |
| *navoditi* 'note' | 10 | 0.19 | argue | 42 | 0.29 |
| *pokazati* 'show/'demonstrate' | 10 | 0.19 | demonstrate | 31 | 0.21 |
| *smatrati* 'consider' | 7 | 0.13 | indicate | 21 | 0.14 |
| *tvrditi* 'argue' | 7 | 0.13 | note | 17 | 0.11 |
| *ukazati* 'indicate' | 6 | 0.11 | report | 16 | 0.11 |
| *utvrditi* 'find/'establish' | 6 | 0.11 | examine | 16 | 0.11 |
| *govoriti* 'say' | 5 | 0.09 | use | 13 | 0.09 |
| *nazivati* 'label/'call' | 5 | 0.09 | report | 12 | 0.08 |
| *upozoriti* 'warn' | 5 | 0.09 | provide | 12 | 0.08 |
| | | | propose | 12 | 0.08 |
| | | | conclude | 9 | 0.06 |
| | | | investigate | 9 | 0.06 |
| | | | focus | 8 | 0.05 |

| | | | test | 7 | 0.04 |
| | | | predict | 7 | 0.04 |
| | | | ask | 7 | 0.04 |
| | | | highlight | 7 | 0.04 |
| | | | conduct | 7 | 0.04 |
| | | | consider | 7 | 0.04 |
| | | | develop | 6 | 0.04 |
| | | | describe | 5 | 0.03 |
| | | | emphasize | 5 | 0.03 |
| | | | introduce | 5 | 0.03 |

Table 4. Most frequent reporting verbs (> 5 occurrences per sub-corpus) in the Croatian and English sub-corpora

As shown in Table 4, unlike the English sub-corpus, the findings of the Croatian sub-corpus do not point to the saliency of a particular lexical verb. The Croatian writers showed rather similar preferences towards a group of lexical verbs, with the verb *isticati* 'highlight' showing a slightly higher frequency as compared to the verbs *provesti* 'conduct', *definirati* 'define', *navoditi* 'note', and *pokazati* 'show'/ 'demonstrate'. Thus, compared to previous and present findings on the reporting verbs used in the given disciplinary discourse in English, Croatian writers made rather different choices in the use of the most frequent reporting verbs, the only exception being the verb *pokazati* 'show', which was shown to be among the five most frequently used reporting verbs in both sub-corpora. However, if we look at the Croatian list of the most frequent verbs, we can see that 7 out of top 12 verbs have also been rated as the most frequently occurring verbs in the English sub-corpus, which is indicative in terms of disciplinary preferences towards the use of the congruent lexical verbs in citations. However, the English cognates of the remaining Croatian verbs were used either only once or not even once, suggesting their rather unrecognized status as the reporting verbs in the given disciplinary sub-corpus in English.

Another somehow discrepant finding between the two sub-corpora concerns a rather low incidence of the reporting verbs with a "tentative semantic component" (*e.g. suggest*, *indicate*) in the Croatian as opposed to the English sub-corpus (Thomas & Hawes 1994, 138). This is particularly striking for *suggest*, which was ranked as the second most frequent verb in the English sub-corpus. By contrast, its Croatian cognate 'sugerirati' was not only ranked among the most frequent Croatian verbs, but surprisingly was not used even once as a reporting verb in the Croatian sub-corpus.

According to Thomas and Hawes (1994), *suggest* belongs to the group of Tentativity verbs, a sub-group of Discourse verbs, which along with *indicate*, *hypothesize*, *propose,*

*etc.* signal writers' cautious and tentative stance in reporting claims, which in a pragmatic sense of a word is concerned with the use of hedges in academic writing (Hyland 1998), as shown in:

(26)  ***Murphy (2000) suggested*** *that one reason children might omit regular plurals from English compounds could stem from the fact that the plural [-s] morpheme consistently is found at the end rather than in the middle of words.* (Eng11, 197)

Based on the findings, which point to a relatively high ratio of Tentativity reporting verbs overall (*cf.* frequency of *suggest* and *indicate*) in the English sub-corpus, it may be suggested that English writers tend to take a rather hedged stance in reporting other scholars' work as opposed to Croatian writers. Though further research should verify the status of the reporting verbs carrying a hedging potential in the Croatian disciplinary discourse and academic writing in general, a low incidence of the given semantic verb class may be broadly in line with the overall lower frequency of hedging markers in the Croatian academic writing as opposed to that in English (for further discussion see Varga 2016).

Despite different preferences towards the use of individual lexical verbs across the two sub-corpora, the findings point to congruent tendencies of English and Croatian writers towards the use of the same semantic verb classes (Hyland 2004; Yang 2013). Thus, the findings show that English writers used mostly discourse-based reporting verbs (*e.g. suggest*, *argue*, *demonstrate*), accounting for 52% of the instances. On the other hand, 36% of the most frequently used reporting verbs referred to research activities (*e.g. find*, *show*), while the verbs denoting cognition acts (*e.g. conclude*, *predict*, *consider*) were used less frequently by comparison (12%).

Similarly, in the Croatian sub-corpus, the discourse verbs, such as *isticati* 'highlight', *definirati* 'define', *navoditi* 'note' accounted for 67% of the occurrences while the verbs denoting research activities, such as *provesti* 'conduct', *pokazati* 'show'/ 'demonstrate', *ukazati* 'indicate' were used in 25% of the cases and only one referring to cognition, *i.e. smatrati* 'consider' showed 8% of the occurrences. Both English and Croatian findings conform to the results of previous research on the distribution of semantic verb classes used in research article writing in applied linguistics (Hyland 2004; Yang 2013). For instance, Hyland's (2004) findings show the congruent ranking of the use of semantic categories of the reporting verbs in the equivalent disciplinary corpus, with the discourse-based verbs showing the highest frequencies (59%), followed by the reporting verbs denoting research (30.5%) and cognition (10.5%). Overall, the fact that in citing other scholars' work both English and Croatian disciplinary writers refer more to communicative-based activities rather than those referring to different aspects of research may be accounted

for by a more discursive and interpretative nature of research in soft sciences in general, and applied linguistics in particular (Hyland 2004).

## 4 Concluding remarks

The aim of the present research was to investigate the use of citations in the research articles written by Croatian and English applied linguists and thus contribute to a greater understanding of a particular aspect of rhetorical practices by two writing cultures. The general findings show that citations were more frequently used in the English as opposed to the Croatian sub-corpus, which conforms to the general tendency of Anglo-American academic writing towards the use of more citations as opposed to academic writing styles of other language communities (Hyland 2005). As previously mentioned, in accounting for the distinctive rhetorical practices, prior research mainly underscores a large size of the international Anglophone academic community and its wide readership as opposed to those in local academic contexts (Mur Dueñas 2009; Dontcheva-Navratilova 2016). Generally, it may be argued that the present findings conform to these trends, in terms of the incomparable discrepancy between the sizes and scope of audience the English and Croatian applied linguists address. In addition, the subject matter and the type of the research presented in the corpus articles may also be a factor affecting the present findings. Unlike the English articles that addressed English or some other world languages, some of the research presented by Croatian writers was more relevant to the domain of applied linguistics at the local rather than international level, which, given the amount of prior disciplinary research could have affected the amount of citations they integrated into their writings. However, the extent to which the present findings reflect disciplinary-specific rhetorical preferences or those that may be related to the characteristic feature of the Croatian academic writing needs to tested against far more extensive corpus research.

While different socio-cultural backgrounds might broadly account for the differences in the overall frequencies of citations across the two sub-corpora, the congruent disciplinary and rhetorical variables may be underlying the similarities between the findings. First, in both sub-corpora, the findings point to the significantly greatest density of citations in the Introduction sections, followed by the Discussion section, as the two most argumentative and interactive RA sections. By contrast, the use of citations in the middle sections of research articles is considerably smaller by comparison given their predominantly descriptive characterization. The congruent distribution of citations across the IMRAD structure broadly reflects the major rhetorical purposes of each RA section, which are considered to be universally generic features of academic writing irrespective of the language background. However, the most striking finding in the use of citations across the two-sub-corpora relates to their considerably higher frequency in the Discussion

section of the English as opposed to the Croatian articles. Higher frequency of citations in the Discussion section of the English research articles suggests that English writers engage more in interpreting their findings while referring to previous related research. By contrast, Croatian peers seem to be more focused on the discussion and implications of their own research without as much need to draw parallels to prior studies as their English peers. However, these broad assumptions should be tested against more detailed cross-cultural research on rhetorical moves of each RA section as well as rhetorical functions of citations that have not been analyzed separately in the present study.

An additional similarity in the analysis relates to the cross-cultural preference towards the use of non-integral as opposed to integral citations. In both sub-corpora the discrepancy in the findings is surprisingly similarly small, which suggests that both groups of writers exploit almost equally the potential of both citation types. When it comes to the reporting structures, in both sub-corpora **human subject + reporting verb** is the most frequently employed type of integral citations, which reflects disciplinary preferences towards emphasizing scholars' contribution to disciplinary knowledge (Hyland 2004). In other words, the use of the given reporting structure of citations reflects broader epistemological foundations of soft sciences that place importance on human involvement in establishing disciplinary knowledge, which is viewed as a shared process, accounting for more visible authorship in reporting on other people's work (Hyland 2004).

With respect to the use of non-integral citations, both English and Croatian writers used most frequently **non-reporting citations** in which the only reporting signal is the writer's name in the brackets. It may be assumed that both groups of writers readily employ this citation pattern as it allows them to freely summarize or interpret the reported content without breaking the argument with any explicit form of acknowledgment of the cited author (Hyland 2004; Okamura 2008). Yet, given the high rate of this reporting structure in both sub-corpora, more investigation is needed in order to clarify what type of content is mostly reported in this way and at which points and for what reasons priority is given to this rather than the other citation pattern.

One of the most notable distinctions in the preferred choices of reporting structures regards the use of non-human subjects, which were shown to be a more salient reporting structure of English than Croatian, both integral and non-integral citations. As previously discussed, by opting for this citation form, writers give a more prominent role to research activities rather than their agents whose visibility is thus diminished. The high incidence of non-human subjects in the English sub-corpus may be related to their overall status of formulaic scientific language in English academic prose, reflecting its general characterization of being more object- rather than agent-oriented (Biber *et al.* 1999; Dorgeloh & Wanner 2009). Though human subjects were used more frequently

than non-human subjects in the present English sub-corpus, a relatively high frequency of the latter in the subject position suggests that foregrounding scholars' work may have almost as important role as human agents in the given disciplinary writing. Within the context of the present study, this finding is even more important given a considerably lower frequency of non-human subjects by the Croatian writers who, on the other hand, used more adjuncts in reporting other scholars' work. It seems that the use of adjuncts that allow omission of a reporting verb, conforms well to the conventionally impersonal and predominantly informative (rather than evaluative) characterization of the Croatian scientific style in which a lack of expressivity and suppression of writer's explicit involvement in conveying message may still be regarded as one of its implicit norms (Silić 2006).

Finally, the corpus analysis of the reporting verbs showed that despite differences in the saliency of individual lexical verbs, writers in both sub-corpora showed preference towards the congruent semantic verb class in citations, viz. discourse-based verbs, followed by research-based verb class, while the use of cognition verbs was rather limited. A finding that merits attention and thereby further investigation is a markedly less frequent use of the reporting verbs with hedging potential by Croatian than English writers, which may be a part of a broad tendency of the Anglo-American academic discourse towards a more frequent use of hedges as opposed to other academic writing cultures (Hyland 2005). Nevertheless, the overall findings on the equivalent ranking of the salient reporting verbs reflects wider disciplinary characteristics, in particular a rather discursive nature of applied linguistics, in which, as in other soft sciences that deal with less solid subjects of enquiry than hard sciences, knowledge construction relies heavily on interpretation and reiteration involving human reasoning (Becher & Trowler 2001; Hyland 2004).

## 4.1 Limitations and recommendations for further research and pedagogic work

The present corpus results should be interpreted within certain limitations, which involve different aspects of the analysis. The first relates to establishing the criteria for comparison of cross-cultural findings or *tertia comparationis* required in intercultural research of this kind (Connor & Moreno 2005). More specifically, despite the criteria presented earlier, the compilation of the present corpus consisted of the articles extracted on a rather random basis, disregarding the cross-cultural match in terms of sub-disciplines, type of research, *etc*. Given the small size of the Croatian linguistic academic community, this variable would be hard to control in terms of building a corpus of a representative size. An additional limitation refers to the size of the corpus, which included only 16 articles per language group, and which therefore allows for the interpretation of the findings in very generic terms (Hyland 2004).

There are a number of possible avenues for further research stemming from the present analysis. As previously noted, the study did not account for the rhetorical functions of citations in each RA section, an analysis of which would provide a deeper insight into the motivations for their use. A rhetorical analysis of this kind, supplemented by interviews with subject-specialists, would provide a more profound insight into the disciplinary rhetorical practices of the citation use in the given pair of cross-cultural academic writing. In addition, the present study provided only a very general perspective on the reporting structure of citation forms as well as the use of reporting verbs. More cross-cultural research is needed, however, on examining other aspects of their use. One of it may be the examination of correlations between a particular reporting structure and a semantic class of reporting verbs, which might reveal some regularity in their use (Charles 2006). An additional related and important research area that has not been covered by the present research refers to the analysis of the verb tenses and reporting verbs. The practical implications of findings of such analysis, in particular those that would uncover cross-cultural differences, would be particularly beneficial to non-native disciplinary scholars in writing their research in academic English.

In sum, despite its restricted scope, it is believed that the present research may serve as a starting point for further cross-cultural enquiries into the use of academic citations as an inherent aspect of academic writing. The practical implications of the findings of this or similar research, particularly those pointing to the areas where rhetorical conventions between the two language groups differ (Sanderson 2008), might provide a solid basis for the course materials of much needed academic writing courses, particularly at the doctoral study programs in Croatia. In the context of English as a global language of research and publication, such practically-oriented yet corpus-informed courses in disciplinary writing are becoming increasingly important in non-native English academic settings.

## References

Becher, Tony & Paul R. Trowler. 2001. *Academic tribes and territories: intellectual inquiry and the cultures of disciplines*. Buckingham Philadelphia: Society for Research in Higher Education and Open University Press.

Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Essex: Pearson Education.

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

Carter-Thomas Shirley & Elizabeth Rowley-Jolivet. 2013. Citation from a cross-linguistic perspective: the case of French researchers publishing in English. *ESP Across Cultures* 10, 111–125.

Charles, Maggie. 2006. Phraseological patterns in reporting clauses used in citation: a corpus-based study of theses in two disciplines. *English for Specific Purposes* 25, 310–331.

Connor, Ulla & Anna I. Moreno. 2005. Tertium comparationis: a vital component in contrastive rhetoric research. In *Directions in applied linguistics: essays in honor of Robert B. Kaplan.* Paul Bruthiaux, Dwight Atkinson, William G. Eggington, William Grabe & Vaidehi Ramanathan, eds. Clevedon, UK: Multilingual Matters. 153–164.

Dontcheva-Navratilova, Olga. 2016. Rhetorical functions of citations in linguistics research articles: a contrastive (English-Czech) study. *Discourse and Interaction* 9 (2), 51–74.

Dorgeloh, Heidrun & Anja Wanner. 2009. Formulaic argumentation in scientific discourse. In *Formulaic language: Volume 2. Acquisition, loss, psychological reality, and functional explanations*. Roberta Corrigan, Edith A. Moravcsik, Hamid Ouali & Kathleen M. Wheatley, eds. Amsterdam: John Benjamins. 524–544.

Fløttum Kjersti, Trine Dahl & Torodd Kinn. 2006. *Academic voices: across languages and disciplines.* Amsterdam: John Benjamins.

Gilbert, G. Nigel. 1976. Referencing as persuasion. *Social Studies of Science* 7 (1), 113–122.

Hu, Guangwei & Guihua Wang. 2014. Disciplinary and ethnolinguistic influences on citation in research articles. *Journal of English for Academic Purposes* 14, 14–28.

Hyland, Ken. 1998. *Hedging in scientific research articles.* Amsterdam: John Benjamins.

Hyland, Ken. 2004. *Disciplinary discourses: social interactions in academic writing.* Ann Arbor: University of Michigan Press.

Hyland, Ken. 2005. *Metadiscourse: exploring interaction in writing*. London, New York: Continuum.

Hyland, Ken. 2009. *Academic discourse: English in a global context*. London, New York: Continuum.

Hyland, Ken & Polly Tse. 2004. Metadiscourse in academic writing: a reappraisal. *Applied Linguistics* 25 (2), 156–177.

Koutsantoni, Dimitra. 2006. Rhetorical strategies in engineering research articles and research theses: advanced academic literacy and relations of power. *Journal of English for Academic Purposes* 5, 19–36.

Mansourizadeh, Kobra & Ummul Ahmad. 2011. Citation practices among non-native expert and novice scientific writers. *Journal of English for Academic Purposes* 10, 152–161.

Mur Dueñas, Pilar. 2009. Citation in business management research articles. In *Cross-linguistic and cross-cultural perspectives on academic discourse*. Eija Suomela-Salmi & Fred Dervin, eds. Amsterdam: John Benjamins. 49–60.

Oraić Tolić, Dubravka. 2011. *Akademsko pismo: strategije i tehnike klasične retorike za suvremene studentice i studente*. [Academic writing: strategies and techniques of classical rhetoric for contemporary students]. Zagreb: Naklada Ljevak.

Okamura, Akiko. 2008. Use of citation forms in academic texts by writers in the L1 & L2 context. *The Economic Journal of Takasaki City University of Economics* 51 (1), 29–44.

Petrić, Bojana. 2007. Rhetorical function of citations in high- and low-rated master's theses. *Journal of English for Academic Purposes* 6, 238–253.

Sanderson, Tamsin. 2008. *Corpus, culture, discourse*. Tübingen: G. Narr.

Shaw, Philip. 1992. Reasons for the correlation of voice, tense and sentence function in reporting verbs. *Applied Linguistics* 13 (3), 302–319.

Shaw, Philip & Diane Pecorari. 2013. Source use in academic writing: an introduction to the special issue. *Journal of English for Academic Purposes* 12, A1–A3.

Silić, Josip & Ivo Pranjković. 2005. *Gramatika hrvatskoga jezika: za gimnazije i visoka učilišta*. [Grammar of Croatian language: for secondary schools and institutions of higher education]. Zagreb: Školska knjiga.

Silić, Josip. 2006. *Funkcionalni stilovi hrvatskoga jezika*. [Functional styles of the Croatian language]. Zagreb: Disput.

Swales, John. 1986. Citation analysis and discourse analysis. *Applied Linguistics* 7, 39–56.

Swales, John. 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Thomas, Sarah & Thomas P. Hawes. 1994. Reporting verbs in medical journal articles. *English for Specific Purposes* 13 (2), 129–148.

Thompson, Geoff & Yiyun Ye. 1991. Evaluation in reporting verbs used in academic papers. *Applied Linguistics* 12 (4), 356–382.

Thompson, Paul & Chris Tribble. 2001. Looking at citations: using corpora in English for academic purposes. *Language Learning & Technology* 5 (3), 91–105.

Varga, Mirna. 2016. *Epistemic modality in academic discourse in Croatian and English language*. [PhD dissertation]. Osijek: Filozofski fakultet u Osijeku.

White, D. Howard. 2004. Citation analysis and discourse analysis revisited. *Applied linguistics* 25 (1), 89–116.

Yang, Linxiu. 2013. Evaluative functions of reporting evidentials in English research articles of applied linguistics. *Open Journal of Modern Linguistics* 3 (2), 119–126.

Yang, An, Shu-yuan Zheng & Guang-chun Ge. 2015. Epistemic modality in English-medium medical research articles: a systemic functional perspective. *English for Specific Purposes* 38, 1–10.

## Data sources

Croatian sub-corpus

Cr1   Aladrović Slovaček, Katarina. 2011. Razvoj diskursne kompetencije u ranome učenju hrvatskoga jezika [The development of discourse competence in early learning of Croatian language]. In *Diskurs i dijalog: teorije, metode i primjene, Zbornik radova s 24. međunarodnog skupa HDPL-a*. Vladimir Karabalić, Melita Aleksa Varga & Leonard Pon, eds. Zagreb, Osijek: Hrvatsko društvo za primijenjenu lingvistiku, Filozofski fakultet Sveučilišta Josipa Jurja Strossmayera. 323–338.

Cr2   Andraka, Marija & Ankica Jurković. 2016. Potiču li udžbenici za rano učenje stranog jezika razvoj strategija učenja vokabulara? [Do EFL coursebooks for young learners promote the use of vocabulary learning strategies?]. *Strani jezici* 44, 29–53.

Cr3   Bagarić Medve, Vesna & Silvija Berkec. 2016. Komunikacijska jezična kompetencija studenata na početku studija njemačkoga jezika i književnosti. [Communicative language competence of students at the beginning of the study of German language and literature]. *Strani jezici* 45, 4–22.

Cr4   Barac, Maja & Nataša Košuta. 2011. Stereotipi – prepreka ili pomoć u stjecanju interkulturalne komunikacijske kompetencije? [Do stereotypes help or hinder the acquisition of intercultural communicative competence?]. *Strani jezici* 40, 133–144.

Cr5   Carović, Ines & Nikolina, Sokolić. 2011. Prijenos sportskog događaja – elementirazličitih govorničkih strategija. [The broadcast of sporting event – elements of different types of discourse]. In *Diskurs i dijalog: teorije, metode i primjene, Zbornik radova s 24. međunarodnog skupa HDPL-a*. Vladimir Karabalić, Melita Aleksa Varga & Leonard Pon, eds. Zagreb, Osijek: Hrvatsko društvo za primijenjenu lingvistiku, Filozofski fakultet Sveučilišta Josipa Jurja Strossmayera. 159–174.

Cr6   Erdeljac, Vlasta & Maja Lakuš. 2012. Leksičke veze u mentalnom leksikonu dvogodišnjaka i trogodišnjaka. [Lexical relations in the mental lexicon of two-year-old and three-year-old children]. In *Jezik kao informacija*. Anita Peti-Stantić & Mateusz-Milan Stanojević, eds. Zagreb: HDPL/CALS, Srednja Europa. 51–53.

Cr7   Horga, Damir, Vlasta Erdeljac & Ana Vidović. 2011. Govorne pogreške u elektroničkim medijima, [Speech errors in electronic media]. In *Diskurs i dijalog: teorije, metode i primjene, Zbornik radova s 24. međunarodnog skupa HDPL-a.* Vladimir Karabalić, Melita Aleksa Varga & Leonard Pon, eds. Zagreb, Osijek: Hrvatsko društvo za primijenjenu lingvistiku, Filozofski fakultet Sveučilišta Josipa Jurja Strossmayera. 175–201.

Cr8   Jakovac, Mihaela & Željka Kamenov. 2012. Uloga prethodnog iskustva, motivacije i strategija učenja u uspjehu iz engleskog jezika struke. [Prior learning experience, motivation and learning strategies as success factors in ESP learning]. *Strani jezici* 41, 7–32.

Cr9   Košuta, Nataša & Sanja Vičević. 2011. Komunikacijske strategije u ranom učenju stranoga jezika. [Communication strategies in early foreign language learning]. In *Diskurs i dijalog: teorije, metode i primjene, Zbornik radova s 24. međunarodnog skupa HDPL-a*. Vladimir Karabalić, Melita Aleksa Varga & Leonard Pon, eds. Zagreb, Osijek: Hrvatsko društvo za primijenjenu lingvistiku, Filozofski fakultet Sveučilišta Josipa Jurja Strossmayera. 339–356.

Cr10  Molnar, Draženka & Dubravka Vidaković Erdeljić. 2010. Konceptualne metafore i metonimije u prevođenju izvornih i modificiranih poslovica. [Conceptual metaphors and metonymies in translation of original and modified proverbs]. In *Prostor i vrijeme u jeziku: Jezik u prostoru i vremenu*. Mario Brdar, Marija Omazić, Višnja Pavičić Takač, Tanja Gradečak-Erdeljić & Gabrijela Buljan, eds. Zagreb, Osijek: Hrvatsko društvo za primijenjenu lingvistiku; Filozofski fakultet Sveučilišta Josipa Jurja Strossmayera. 207–217.

Cr11  Omrčen, Danja. 2011. Gramatičko znanje engleskoga kao stranoga jezika u studenata i studentica kineziologije. [Grammatical knowledge of English as a foreign language in male and female students of kinesiology]. *Strani jezici* 40, 145–157.

Cr12  Pavičić Takač, Višnja & Ana Umiljanović. 2008. Odnos između znanja vokabulara i strategija učenja vokabulara u engleskom kao stranom jeziku. [Relationship between vocabulary knowledge and vocabulary learning strategies in English as a foreign language]. *Strani jezici* 37 (4), 411–422.

Cr13  Pernjek, Jasminka & Ivana Matić. 2015. Zadovoljstvo poslom nastavnika njemačkoga jezika u Hrvatskoj. [Job satisfaction of German language teachers in Croatia]. *Strani jezici* 44, 4–28.

Cr14  Truck-Biljan, Ninočka. 2012. Dva u jedan, jedan u dva: vjerovanja učenika i učitelja o učenju dva strana jezika. [Two in one, one in two: student and teacher beliefs about learning two foreign languages]. *Strani jezici* 41, 153–170.

Cr15  Šimunić, Zrinka. 2012. Prevođenje kao konceptualno rekonstituiranje i informacijsko restrukturiranje izvornog teksta. [Translation as a conceptual reconstitution and information restructuring of the original text]. In *Jezik kao informacija*. Anita Peti-Stantić & Mateusz-Milan Stanojević, eds. Zagreb: HDPL/ CALS, Srednja Europa. 93–110.

Cr16  Vančura, Alma & Diana Tomić. 2011. Informativni složenac: analiza retoričkih vrsta unutar informativnih emisija hrvatskih televizija [Information Stew- Analysis of Rhetoric Genres in the News of Croatian Televisions]. In *Diskurs i dijalog:*

*teorije, metode i primjene, Zbornik radova s 24. međunarodnog skupa HDPL-a*. Vladimir Karabalić, Melita Aleksa Varga & Leonard Pon, eds. Zagreb, Osijek: Hrvatsko društvo za primijenjenu lingvistiku, Filozofski fakultet Sveučilišta Josipa Jurja Strossmayera. 193–205.

English sub-corpus

Eng1   Bell, Nancy, Skalicky, Stephen & Tom Salsbury. 2014. Multicompetence in L2 language play: A longitudinal case study. *Language Learning* 64 (1), 72–102.

Eng2   Bowden, Harriet W. 2016. Assessing second-language oral proficiency for research: the Spanish Elicited Imitation Task. *Studies in Second Language Acquisition* 38 (4), 647–675.

Eng3   Carrol, Gareth, Kathy Conklin & Henrik Gyllstad. 2016. Found in translation: the influence of the L1 on the reading of idioms in a L2. *Studies in Second Language Acquisition* 38 (3), 403–443.

Eng4   Cox, Jessica G. 2017. Explicit instruction, bilingualism, and the older adult learner. *Studies in Second Language Acquisition* 39 (1), 29–58.

Eng5   Dewey, Dan P., Bown, Jennifer, Baker, Wendy, Martinsen, Rob A., Gold, Carrie & Dennis Eggett. 2014. Language use in six study abroad programs: an exploratory analysis of possible predictors. *Language Learning* 64 (1), 36–71.

Eng6   Grey, Sarah, John N. Williams & Patrick Rebuschat. 2014. Incidental exposure and L3 learning of morphosyntax. *Studies in Second Language Acquisition* 36 (4), 611–645.

Eng7   Jacobs, April, Fricke, Melinda & Judith F. Kroll. 2015. Cross-language activation begins during speech planning and extends into second language speech. *Language Learning* 66 (1), 324–353.

Eng8   Lambert, Craig, Judit Kormos & Danny Minn. 2017. Task repetition and second language speech processing. *Studies in Second Language Acquisition* 39 (1), 167–196.

Eng9   Lee, Andrew H. & Roy Lyster. 2016. Effects of different types of corrective feedback on receptive skills in a second language: a speech perception training study. *Language Learning* 66 (4), 809–833.

Eng10 Miller, A. Kate. 2015. Intermediate traces and intermediate learners: evidence for the use of intermediate structure during sentence processing in second language French. *Studies in Second Language Acquisition* 37 (3), 487–516.

Eng11 Murphy, Victoria A. & Jennifer Hayes. 2010. Processing English compounds in the first and second language: the influence of the middle morpheme. *Language Learning* 60 (1), 194–220.

Eng12  Simpson Baird, Ashley, Natalia Palacios & Amanda Kibler. 2016. The cognate and false cognate knowledge of young emergent bilinguals. *Language Learning* 66 (2), 448–470.

Eng13  Sommers, Mitchell & Joe Barcroft. 2013. Effects of referent token variability on L2 vocabulary learning. *Language Learning* 63 (2), 186–210.

Eng14  Sparks, Richard L. & Jon Patton. 2013. Relationship of L1 skills and L2 aptitude to L2 anxiety on the foreign language classroom anxiety scale. *Language Learning* 63 (4), 870–895.

Eng15  Stringer, David. 2015. Embedded wh-questions in L2 English in India: inversion as a main clause phenomenon. *Studies in Second Language Acquisition* 37 (1), 101–133.

Eng16  Webb, Stuart & Anna C.-S. Chang. 2015. How does prior word knowledge affect vocabulary learning progress in an extensive reading program? *Studies in Second Language Acquisition* 37 (4), 651–675.