

CITATION PRACTICES IN A CORPUS OF ROMANIAN TEXTS ON ECONOMICS

Teodora GHIVIRIGĂ

Alexandru Ioan Cuza University, Iași, Romania

Abstract: The present paper investigates the pattern of citation practices in professional communication. The research is based on a corpus of Romanian texts on Economics. Quantitative data indicate a preference for the non-integral type of quotation, interpreted as placing more emphasis on the data than on the person of the authors cited. The existence of combined forms (footnotes and author/date) seems to indicate that the young researchers are not aware of the difference and may not be very confident in using either according to generally accepted practices in the academia.

Keywords: citation, corpus, Economics, Romanian.

1. Introduction

Citation is the most overt and most immediately obvious indication that a text is indeed academic (Swales, 2014:119). The presence of citations is a form of intertextuality that gives clear evidence that a dialogue is established between those citing and those cited (*ibid.*) and the process itself is important for both parties. Reference to pre-existing knowledge has been subject to changes over time: Bazerman (cited in Hyland, 1999:343), for example, found that the number of items in reference lists had risen steadily during the 20th century from about 1.5 per article in the *Physical Review* in 1910 to more than 25 in 1980 and, given the quantity of literature published every year, it is expected to maintain the rising trend. The manner in which names and dates about authors are incorporated in a scientific text and the amount of such information used may influence its reception and evaluation. What – or whom – is quoted and in what way this information is presented may influence a text's evaluation by peer reviewers for academic journals. It sometimes also appears to be subject to reasons other than purely scientific or informational: on investigating citation practices of doctorates in Economics two American economists find that often there are parochial loyalties and preferences that influence who gets cited and who doesn't (Stigler and Friedland, 1975: 485).

Citation practices are the result of the long established conventions of the academic community, yet they are not part of the objective, detached manifestation of scientific impersonal communication, rather they have been found to be influenced by various factors: a style that may belong to a particular publication (journal), genre, discipline and its associated epistemological position, the writers' status as a professional and stage in his career development; research also suggests the influence of national or educational rhetoric (Hewitt and Lago, 2012). Academic citation serves various ends: it contributes to the construction of disciplinary knowledge, projects a credible writer ethos (Hyland,

1999), displays knowledge as well as establishes a niche for one's own research, it enables rigorous quantitative analysis of elusive but important social phenomena such as reputation, influence, prestige, celebrity, the diffusion of knowledge, the rise and decline of schools of thought etc. (Posner, 1999). The large amount of research in this area points out to the interest and importance of citation practices as part of scientific communication. Communicating science – maybe with the exception of highly formalized sciences such as mathematics or IT – involves the use of a language, which, in Hyland's terms (1997: 9) will incorporate basic assumptions about the nature of reality; therefore any language – English as well as Romanian – carries the biases and presuppositions of certain cultural traditions. If so, the question may be asked: what values, attitudes and positions held by Romanian academics can be revealed by their citation practices?

2. Description of the corpus

6 PhD dissertations from the domain Economics have been investigated for this paper. Economics is defined as "a social science that seeks to analyze and describe the production, distribution, and consumption of wealth." (in Marc's Blaug's words as found in Encyclopedia Britannica). Economics is seen as the leading discipline among a group of sciences, the creator of doctrines and principles; however, as various economists have pointed out, its position as a social science has been constantly under epistemological questioning as it incorporates some of the instruments and methodologies of (hard) sciences, in an attempt at increasing the strength and value of its predictions. The topics of the dissertations selected for the corpus therefore come entirely from this area (doctrines and their applications); more technical topics from related disciplines have been excluded for two reasons: firstly, the texts would have contained larger amounts of the mathematical apparatus in the form of tables, formulae, graphs, and significantly less linguistic material to research, which would have been less profitable for my investigation. Secondly, previous contact with texts from this domain has suggested that, given the more theoretical nature of the topics, formulation of hypotheses and overall demonstration relies more on argumentation than on the mathematical apparatus. Selection has been made on the basis of titles, with brief consultation with the coordinators, wherever possible. The topics range from the more theoretical and philosophical ones like the concept of degrowth or the current economic crisis to the works and theories of various economists, to others concerned with more technical aspects of regional development, tourism etc. The text is more than half a million words (35,745 types, 647,059 tokens) totalling 1,384 pages, with an average length of pages of 230 per thesis. The number of titles given in reference section is 1117 titles. Tables of contents, the reference section, the annexes and the texts associated with tables and graphics (Hyland, 1999: 344) have been removed, however, comments of the quantitative data presented in the tables and graphs were retained, since they also contained citations.

In terms of form of citation, the situation is surprising and rather confusing: of the 6 texts, 1 uses only footnotes, 3 use only the Harvard style (name and date), while the rest use a combination of the two, with sources of cited material presented in both author-

date manner and in footnotes, sometimes even for the same source; moreover, in three of the theses, the footnotes alternately contain both reference and the writer's comments on the texts. Also information from sources is presented as direct quotations (marked by quotation marks) to short quotes and terms quoted, to paraphrasing/summarizing from one author to generalization from multiple sources (Hyland, 2000); no block quotation has been identified in any of the texts.

3. Current state of research

A citation is defined as a research report which has a specific reference point that is clearly identifiable. (Charles, 2006:314) While the definition does not seem to be problematic or subject to debate, its typology has been regarded from various angles and classifications of citations have been suggested with various criteria in mind: the syntactic patterns in which they occur (Swales, 1990; Charles, 2006); the writer's position towards the content cited (Hyland, 2004); and in terms of their rhetorical function (Thompson & Tribble, 2001, Petrič, 2007) or syntactic role (Hyland 1999). In terms of rhetorical function for instance, Petrič (2007) distinguishes among: attribution, exemplification, further reference, statement of use, application, evaluation, establishing links between sources, comparison of one's own findings or interpretation with other sources and other. In terms of their syntactic function/role within the sentence, citations have been classified as: a. author as SUBJECT; b. author as AGENT; c. author as ADJUNCT and d. author as part of a noun phrase consisting of the author's name and a possessive or an agentive structure, and e. author {OTHER} (Swales, 2014). In terms of form, a proposed taxonomy (Hyland, 2000, in Petrič 2012:104) divides them into quote, block quotation, summary, and generalization from multiple sources. The most frequently used and quoted is that proposed by Swales (1990: 148), who classifies them according to the criterion of level of textual integration into integral and non-integral citations, where integral citations are those where the name of the cited author occurs in the citing sentence while non-integral forms make reference to the author in parenthesis or by superscript numbers. Thompson (2001) further develops Swales' typology taking into account a combination of formal linguistic criteria, such as the syntactic position of the citation within a clause, and its function, such as whether the citation identifies the origin of an idea or is used as an example (in Petrič, 2007). Flottum et al. (2006) propose an extended version of Swales' classification into: a. non-integral reference (with numbers [3]); b. partly integral reference, e.g. x is y (*Author*, 2015); c. semi-integral reference, e.g. Jones (2015) has observed that...; d. fully-integral reference, e.g. *Author* (2015) claims that "... / claims: "...[54]; and e. a combination of patterns. However, given the extensive size of the texts in the corpus and the different genre they belong to, which is less concentrated and compressed than the research article mostly discussed in these papers, the models presented above were considered inapplicable to the corpus under investigation.

4. Methodology and findings

I have chosen the Swales model with an adaptation due to the nature of the textual material; given the type of discourse required by the genre, the PhD, which is far

lengthier than the research article and allows more room for elaborations, connections between ideas, etc., the corpus offered instances when an economist's name – or his theory/idea/concept/term – is just mentioned generally, without any specific date, while sometimes only the title is mentioned. I considered that since the writers decide to introduce it in the discourse by name, a dialogue with the respective author is initiated. Therefore, following Hyland (1999:345) and distancing myself from Thompson (2001) I introduced such cases in the final count. I followed Thompson (2001) in not counting the repeated use of the same name that had already been labelled, when it was recurrent within the same paragraph, with the mention that if this was done in a different section, related to a different topic/concern, it was counted. I labelled such instances a "mention" and I describe it very similar to Charles (2006:315) as a "general reference", a category that includes "instances where, although there is no citational name or number, a marker indicates that the clause introduces a report on the work of other researchers." To these I added just instances when an author was mentioned, sometimes even remotely related to the discipline (Aristotle, Virginia Woolf, Einstein, whose quote, incidentally is given "Nu tot ce contează poate fi numărat, și nu tot ceea ce poate fi numărat contează" [1] txt), as I took into account the fact that the writers explicitly wish to include them in their argument and therefore to establish a textual relation with them. Although they are similar to Thompson's non-citations (2001, in Petric 2007:240) counted as integral citations, (where the author's name is not followed by other data, such as publication year), I preferred to label them differently and classify them other than integral or non-integral, as the lack of elaboration on the writer's part seems to indicate a different relation between author, content of cited material and the writer himself. Two other reasons influenced my choice of the Swales model. One is that it has been most consistently interpreted in relation to the writer's position regarding the source cited and it was therefore expected to allow an interpretation for my data. Another reason is that since it has been widely used in studies on citation in various languages (French, Spanish, Norwegian), with various types of writers (novice vs experts), various genres – RA to MA and Phd dissertations in various disciplines (social sciences vs sciences), various L1 writers (native vs non-native English of Spanish, French, Norwegian, Chinese, etc.), and since I found only one similar study for Romanian (Blaj, 2006), it could provide some ground for comparison.

The texts have been manually annotated for types of citations, with some cases where a citation could be arguably attributed to two classes. The cases in which a legitimate reference is made to textual objects and which attributed scientific (ideas, concepts, terminology) material to other sources (for example 'Gricean strictures', 'Davidson's argument') were also counted (see Hyland 1999:345). Where more than one author or a group of authors are cited, usually between brackets in non-integral citations, each author or group of authors has been counted as an individual instance (Petric, 2007:66). Direct quotations were identified on the basis of conventional signals, i.e., quotation marks, author's name, and page numbers (Petric, 2012:105). The so-called "scare quotes" are stretches of text marked by scare quotation marks (i.e., inverted commas placed around a term or expression without an accompanying citation, typically used in order to question a term or to refer ironically to it and thus help the writer distance himself from it were not taken into account. Nor were instances where terminology (Petric, 2012:110) was placed between quotation marks, indicating the writers' concern not to take credit for phrasing which they did not own. It is interesting to

note, however, that such behaviour is reported in novice writers and has been interpreted as a mark of insecurity related to whether it is necessary to signal such terms in all cases, while common practice is to do so at first mention only (ibid.) e.g: intrarea pe piață a unor brokeri “low-cost” [1]ⁱⁱ, [indicele] cunoscut sub denumirea de “Dow Jones” [1]. No specialized texts places a term such as the widely used Dow Jones industrial average index between quotations marks.

A rather unusual instance of quotations marks which is surprisingly frequent in the corpus (but has previously been identified by the author in a small(er) corpus of research articles on Economics in Romanian dating from ten years ago or more, when such texts were still published) is when writers incorporate a figurative meaning – or what is thought to be one – in their texts (e.g. se constată o “cronicizare” a acestei stări de lăncezeală [1]ⁱⁱⁱ, cel mai sigur instrument pentru a “cheltui” atunci când sectorul privat... [1]^{iv}, “cantitatea de educație” primită [3]^v, canalul de distribuție ce „iriga” piața [9]^{vi}) Such instances were not counted as citations. A similar practice is also described by Petrić (2012:110), who associates it to novice writers (she found such instances in MA dissertations). Another methodological remark to be made here is that, as other researchers have noted (Flottum et al., 2006, for instance), while most citation instances are straightforward, there are examples which are difficult to ascribe to one class (only). Following the KIAP model, they have been included in a special residual group. There have also been cases when the same sentence can contain several types of citation, e.g. *Aries* (integral, date further down on the page), *făcând referire la această distrugere a identității de clasă, îl evocă pe Alvin Toffler* (mention) (no indication of title, date or any other information) [5]^{vii}. Each has been interpreted and ascribed to one group accordingly.

Values (normalized/10,000 words)	
Integral (I)	8.03
Non-integral (NI)	20.09
Mention (M)	4.66

Table 1. Types of citation in the corpus (normalized values for 10,000 words)

The figures for the three types of citation are provided in table format and the values have been normalized per 10,000 words. As shown in the table, the most frequently used type of citation in the corpus is the non-integral type NI (20.09/10,000) with integral citations (I) the second (8.03/10,000) and the mention (M) a much smaller ratio (4.66/10,000). Data in Hyland’s study (1999) show that softer disciplines (such as Sociology, Philosophy, Marketing, Applied Linguistics) tend to employ significantly more citations than hard sciences (Engineering, Electronic engineering, Physics). Direct quotation is relatively common in the writing of social scientists but rather infrequent or entirely absent from the writing of scientists (e.g., Hyland, 2000:105). The data that can be related to the findings (61% NI, 25% I and 14% for mentions), are those for Marketing (NI 70.3% vs I 29.7%), Sociology (NI 64.6% vs I 35.4%) or Philosophy (NI 35.4% vs I 64.6%). Comparing the values I found (NI 61% vs I 25% vs M 14%), the closest values would be those of Sociology, and the relation of Economics to other social sciences is obvious, although not always graciously acknowledged by economists, who aspire to the status of hard, sciences; however, the figures are not comparable to any of the sciences in Hyland’s study. On the other hand, Mansourizadeh and Ahmad (2011) note a lack of

non-integral citations in the article of novice writers and a definite preference for integral untransformed citations: this would indicate that the status of the writers in the corpus is closer to that of expert rather than novice. Values for each paper differ, and taking this as an indicator, they may reveal various degrees of expertise in using citations. In Blaj (2006), which contains the only figures existing for Romanian, the microcorpus consisting of texts from the social studies show the same preference for NI (with a ratio of NI 52 to I 35), but there is no indication whether these are absolute figures or percents.

5. Conclusions

With respect to the preference for integral vs non-integral types, the choice of integral/ non-integral citation is a complex product of a number of factors including citation conventions, genre, discipline and individual study type (Charles, 2006). A writer can choose either to place focus on the researcher, the cited text, or the piece of research by including the citation in a sentence (integral), or to emphasize the finding, the data, the concept, and de-emphasize the researcher and the product by placing the citation outside the sentence (non-integral). Previous studies seem to agree that “the use of integral citation with a human subject leads to prominence of the cited author” (Weissberg & Buker, 1990:316) and that generally “softer disciplines tend to employ more citations, with engineering and physics well below the average, although the frequencies for molecular biology appear to differ considerably from this general picture.” (Hyland, 1999:346)

The figures seem to indicate a distinct preference for the non-integral type of citation. Evidence from the corpus suggests that PhD researchers tend to place the emphasis on the data (of which much is statistical), on the information, rather than the person: in this the texts are closer to hard sciences than to social sciences. As could be seen from the examples provided, when integral citations are used, there is often a considerable distance between the name and the date, which makes it difficult to trace and follow the author; moreover, such instances often look like a combination of the two, as if even when they are used the preference would go to NI. The existence of combined forms (footnotes and author/date) seems to indicate that the young researchers are not aware of the difference between the two and not very confident in using either according to generally accepted practices in the academia.

The use of combined forms of quotation can be seen as disclosing insecurity in addressing the issue, probably related to the attempt at avoiding any possible charges of plagiarism. The use of terminology between quotes (not only scare quotes) and/or of phrases associated by the PhD students – or considered to be so – with connotative meaning may indicate that to a certain extent they feel uncomfortable about the process of text production and the degree of freedom and creativity authors can have in their relation with language. However, the presence of integral citations also suggests that they are willing to engage in a dialogue with other scientists and place them in a prominent position as creators of knowledge.

The authors of the texts investigated here have, it should be mentioned, an ambivalent status: they are at the same time students (PhD programs are a form of education) and researchers (they are to present their ideas not only to their coordinator, but to the scientific community in the form of articles for journals during their PhD program). This results in a paradoxical situation: their knowledge of the mechanics of

text production for publication is taken for granted and their various degrees of exposure to scientific prose, as found in the monographs, research articles and other sources they consult in the process of preparation, is not uniform or in agreement with the general conventions. From these facts a number of pedagogical conclusions can be also derived, which do not, however, make the topic of this paper.

Acknowledgment. The present paper is in partial fulfillment of the requirements for the postdoctoral project POSDRU/159/1.5/S/133652 *Profilul autorului de text de specialitate – utilizarea formelor autoreferențiale și referința la alți autori în discursul științific (economic) în limba română.*

References

1. Anthony, L. 2014. AntConc (Version 3.4.3) [Computer Software]. Tokyo, Japan: Waseda University. Available from <http://www.laurenceanthony.net/>
2. Blaj, L. 2006. "Citation practices and authorial identity in Academic Writing on food and foodways". In *Romanian Journal of English Studies*, 3.
3. Blaug, M. s.a. Economics, in Encyclopedia Britannica, at <http://www.britannica.com/>, retrieved February 2015.
4. Charles, M. 2006. "Phraseological patterns in reporting clauses used in citation: A corpus-based study of theses in two disciplines". In *English for Specific Purposes* 25, pp. 310–331.
5. Encyclopedia Britannica,
6. Fløttum, Kjersti, T. Dahl, T. Kinn. 2006. *Academic Voices: Across languages and disciplines*, John Benjamins Publishing Company.
7. Hewitt, E., Ángel Felices Lago. 2010. Academic style and format of doctoral theses: The case of the disappearing discussion chapter. In *Ibérica* 19 (2010), pp. 119-140
8. Hyland, K. 2004. *Disciplinary discourses: Social interactions in academic writing*. Ann Arbor: University of Michigan Press.
9. Hyland, K. 2000. *Disciplinary discourses: Social interactions in academic writing*. Harlow, UK: Longman.
10. Hyland, K. 1999. "Academic Attribution: Citation and the Construction of Disciplinary Knowledge". In *Applied Linguistics*, 20/3, pp. 341-367.
11. Hyland, K. 1997. "Scientific Claims and Community Values: Articulating an Academic Culture". In *Language & Communication*, Vol. 17, No. 1, pp. 19-31.
12. Mansourizadeh, A. K., Ahmad, U. K. 2011. "Citation practices among non-native expert and novice scientific writers". In *Journal of English for Academic Purposes*, 10(3), pp. 152-161.
13. Mur Dueñas, P. 2012. "Getting research published internationally in English: An ethnographic account of a team of Finance Spanish scholars' struggles". In *Ibérica* 24, pp. 139-156.
14. Petrić, B. 2007. „Rhetorical functions of citations in high- and low-rated master's theses". In *Journal of English for Academic Purposes* 6, pp. 238–253.
15. Posner, R. 1999. "The Theory and Practice of Citations Analysis, with Special Reference to Law and Economics". In J. M. Olin, *Program in Law and Economics Working Paper* No. 83, pp 1-29.
16. Stigler, J. G., C. Friedland. 1975. "The Citation Practices of Doctorates in Economics". In *Journal of Political Economy*, Vol. 3, 83, pp 477-507.
17. Swales, J. M. 2014. "Variation in Citational Practice in a Corpus of Student Biology Papers: From Parenthetical Plonking to Intertextual Storytelling". In *Written Communication*, Vol. 31(1), pp. 118–141.

18. Swales, J. M. 1990. *Genre Analysis: English in Academic and Research Settings*, Cambridge: Cambridge University Press.
19. Thompson, P. 2005. "Points of focus and position: Intertextual reference in PhD theses". In *Journal of English for Academic Purposes*, 4, pp. 307–323.
20. Thompson, P., C. Tribble. 2001. "Looking at Citations: Using Corpora in English for Academic Purposes". In *Language Learning & Technology*, Vol. 5, No. 3, September, pp. 91-105.
21. Weissberg, R., S. Buker. 1990. *Writing Up Research*. Prentice Hall.

-
- ⁱ "Not everything that matters can be counted, and not everything that can be counted matters", which made me ironically aware of the risks of my own enterprise.
- ⁱⁱ the presence on the market of "low-cost" brokers [1]; [the index] known as the "Dow-Jones" industrial average index [1]
- ⁱⁱⁱ this state of lassitude tends to become "chronic"... [1]
- ^{iv} the safest instrument to "spend" when the private sector... [1]
- ^v "the amount of education" provided...[3]
- ^{vi} the distribution channel that "irrigates" the market...[9]
- ^{vii} Aries [economist] [integral citation, containing publication year, provided further down the page], referring to this destruction of class identity, invokes Alvin Toffler [mention, no indication of title, date or any other information] [5]