

USING SFG TO ANALYZE AUTHORIAL EMPHASIZERS AND PROPOSITIONAL INTENSIFIERS: A COMPARISON OF ALBANIAN AND ITALIAN ACADEMIC WRITING

Vincenzo Dheskali

Abstract

Modality expresses high probability and total degree through boosters (Halliday 1985, Holmes 1990). Through them, writers reinforce statements with the assurance of reliable knowledge (cf. Hyland 1998b). This study compares the usage of boosting emphasize (e.g. *certainly*) and intensifiers (e.g. *completely*) (cf. Quirk et al. 1985) and their orientation and manifestation (cf. Halliday & Matthiessen 2014) in Albanian and Italian student academic writings in L1 and English as an L2. I compiled an Italian and an Italian English corpus (around 3 million words each) as well as an Albanian corpus (around 2.2 million words) and Albanian English one (around 700,000 words). The corpora are comparable in terms of genre, disciplinary domain, gender and the division of ‘hard’ and ‘soft’ sciences. Since very little research has been conducted on academic writing in Albania (Toska 2015), it is essential to initiate research in this field. The results showed that boosters were significantly favored in Italian and Italian English and less favored in Albanian. Conclusively, Italians show more commitment than Albanians.

Keywords

academic writing, modality, emphasize, intensifier, student corpora

1 Introduction

Different societies and academic communities have different means of conveying ideas and presenting opinions and arguments. These means are influenced by different variables such as education (how students are instructed in using boosters), linguistic affordance (students’ mental representations of subject-action-object relations), tenor (students’ low ‘power’ roles and institutional roles in the academia) and academic culture. Italian and Albanian students pertain to two different academic cultures and have probably received different elementary or even advanced English language instruction. However, the contexts remain interrelated (cf. Hösch 2006, Gjoleka 2014), which makes it interesting to explore how this is reflected in their linguistic choices. My main objective is to find out how similar or different Italian and Albanian student writers are in terms of presenting new knowledge and persuading their readers to be successful in the academia. Thus, my paper presents a corpus-based comparison of boosters in Albanian and Italian student papers in L1 and English as an L2.

This paper applies an innovative combination of Quirk et al.’s (1985) categorization, and Halliday’s (1985) and Halliday and Matthiessen’s (2014)

larger frame of modality (modalization) and intensity (degree) (cf. Dheskali 2020). My aim is to reveal patterns of language use for the analyzed data, in relation to the theoretical concepts. Moreover, I have not been able to find any study that connects boosters (mostly treated from a metadiscourse perspective such as Hyland's (1998b, 2005)) to modalization or any other concept from Systemic Functional Grammar (henceforth SFG). This makes it necessary to establish such an innovative theoretical connection.

Within language and academic writing, there are persuasive modality and intensity devices, namely boosters, which assist writers of academic texts in presenting acceptable claims and fostering agreement from their readers. By extending the concepts of Quirk et al. (1985), I would like to add that they either indicate the writer's full commitment towards the truthfulness of the proposition (author-related emphasisers), or they indicate the complete membership of an item within a semantic category (proposition-related intensifiers).

This paper consists of five sections, including this introductory section. The second section concentrates on previous literature related to the pragmatic and semantic aspects of boosters and their subcategories, author-related emphasisers and proposition-related intensifiers. The third section focuses on methodological aspects, such as my two research questions and my corpus compilation and related descriptions of the corpora. In the fourth section, a qualitative and quantitative analysis and a discussion is provided. Finally, conclusions in relation to my two research questions and analyzed aspects are presented.

2 Literature review

2.1 The definition of boosters

Boosters are relevant elements of modalization and intensity. Hyland (2005, 2017) points out that boosters such as *it is obvious that*, *in fact* and *definitely* have the function of emphasizing strength or an author's sureness regarding the given information. Bondi (cf. 2008: 32) declares that emphasisers (Crismore's (1989) term for boosters) are mostly adverbs or adverbials which ensure an increased authority or vigor to the arguments. She extends her discussion by underlining the widely accepted necessity for a more in-depth analysis of the pragmatic and semantic features of emphasisers (cf. Biber et. al. 1999, Hyland 2000a, 2000b, Conrad & Biber 2000, as quoted in Bondi 2008: 33). Before my research, there were very few studies such as Vázquez and Giner's (2009) and Bondi's (2008) focusing exclusively on boosters. Yet, Vázquez and Giner (2009: 219) suggest that "boosting deserves particular attention if we want to fully interpret the phenomenon of academic persuasion".

While boosters represent one side of the coin, hedges (e.g. *might, possibly, I believe*) represent the other side. Hedges and boosters are both interactional resources (cf. Hyland 2005, 2017) that “are employed to either heighten [boosters] or downplay [hedges] their commitment to the proposition, its scope of applicability or its relevance” (Hyland 2005, as quoted in Mu et al. 2015: 138). Despite the interest in boosters, there is an obviously higher amount of research on hedges than boosters, particularly in academic writing (see e.g. Hyland 1998b: 352-353, Bondi 2008: 33, Lafuente-Millán 2008: 70). Therefore, I focus on boosters.

2.1.1 Emphasizers and intensifiers

My study focuses on the division of boosters into emphasizeers, which strengthen the illocutionary force of the proposition, and intensifiers, which stand for degree adverbs that either indicate degrees on a scale while modifying gradable adjectives (e.g. *extremely cautious*) or show the endpoint of a certain scale (e.g. *totally different*) (cf. Bondi 2008: 39). They noticeably differ in scope, since the scope of emphasizeers covers the whole clause or clause complex (SFG concept, generally known as ‘sentence’), while intensifiers have a generally narrower scope. However, their functions may overlap when, for instance, an emphasizeer modifies a gradable predicate (Quirk et al. 1985, as quoted in Bondi 2008: 39). This division presented by Quirk et al. (1985) and complemented by Bondi (2008) is very relevant for my study. As the division is a realization of the larger systems of modality and intensity, it is necessary to introduce these systems.

2.1.2 Orientation and manifestation

“[I]n English there are two axes of variation by which modal assessments can be construed: subjective versus objective and explicit versus implicit” (Martin & White 2005: 130). These two axes of variation include different means of conveying probability within modality. Writers may choose to project their statements as facts and mark objectivity (e.g. *it is likely*) or subjectivity (e.g. *I strongly believe*). Subjective and objective expressions of author’s stance are grouped within ‘orientation’ whereas author’s explicitness and implicitness are grouped within ‘manifestation’. All categories of ‘orientation’ and ‘manifestation’ are interrelated, as it will be shown in the following list.

1a) Subjective implicit: commonly a finite (modal auxiliary) within the clause and mood: *Miri who **could** be the main actor, his friends **may** be influenced.*

1b) Subjective explicit: normally a projecting mental clause which is supported by an idea clause: ***I guess, I don't think** it was inappropriate for Marlon, **I suppose** Lucrezia was disappointed by these events.*

2a) Objective implicit: clause, mood – usually carried by a mood adjunct, in particular a modal adverb: *The process will **probably** be successful. This **certainly** influences the audience. Trump will **possibly** win the elections.*

2b) Objective explicit: usually expressed through a relational clause containing a factual carrier, clause and modal attribute and nominal group: ***It is certain** that he would never be rude to someone, **It is probable** that, **It isn't possible** that he would have done this.*

Other intermediate elements between explicit and implicit levels are subjective forms, such as *in my opinion*, and objective forms: *in all probability* (cf. Halliday & Matthiessen 2014: 688).

Point 2a) reflects Halliday and Matthiessen's (2014) inclusion of adverbs as exclusively objective-implicit forms. My main argument is that the same adverb can shift between objective-explicitness and implicitness. This discussion will be extended in the qualitative analysis. Including orientation and manifestation in a study on boosters is an innovative combination that has, to my knowledge, not been applied in any previous study. It makes the study of boosters more complete.

2.2 Modality and intensity: Modalization and degree

The system of polarity represents ways of building semantic space between the positive and negative (cf. Halliday & Matthiessen 2014: 162). In between stands modality (see Section 1), which is further divided into modalization and modulation. Modulation “refers to the semantic category of proposals” (Halliday & Matthiessen 2014: 691-692). Since boosters predominately occur within modalization, this will be the main focus of my study. Within the ‘indicative’ modalization, there are intermediate possibilities such as degrees of probability: *possibly/probably/certainly* that convey either ‘yes’ or ‘no’ and contain various levels of likelihood (cf. Halliday 1985: 335, Halliday & Matthiessen 2014: 177-692, Dheskali 2017: 139). Probability can be expressed through low (*possibly*), medium (*probably*) and high (*certainly*) values. Meanwhile, intensity stands on a similar level to modality and is subdivided into categories of adjuncts expressing degree (high, low or total) (cf. Halliday & Matthiessen 2014: 189). In my paper, I investigate high probability and total degree, as they are the most prominent boosting phenomena in my corpora.

By respecting the delicacy principle (higher system-lower level-realization or example) of SFG, Figure 1 below connects the previously mentioned concepts with author-related emphasers and proposition-related intensifiers. This shows

how these two booster types connect with the larger levels of modalization, modulation and degree. The concepts are borrowed from Halliday and Matthiessen (2014) and Quirk et al. (1985).

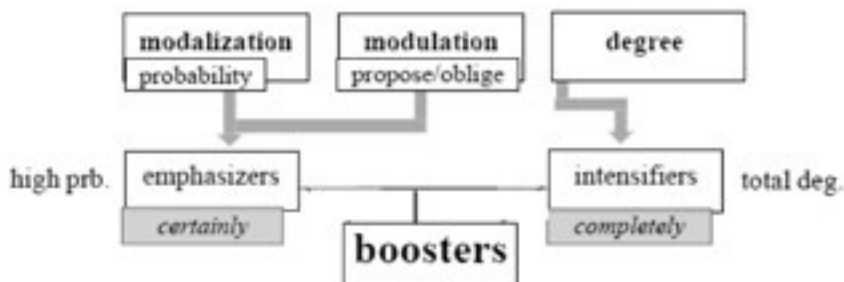


Figure 1: A summary of the concepts treated in this paper

3 Methodology

3.1 Research questions

Two major research questions have been developed in order to compare the attitudes of Albanian and Italian students in their academic writing:

1. What are the pragmatic, semantic and lexico-grammatical differences in the use of proposition-related intensifiers and author-related emphasizees within and across the corpora?

2. What are the differences in the frequency of proposition-related intensifiers and author-related emphasizees within and across the corpora?

The first research question seeks to analyze proposition-related intensifiers and author-related emphasizees in qualitative terms. The second research question compares the same aspects from a quantitative point of view.

3.2 Corpus compilation

Corpus linguistics, as McEnery and Hardie (2012: i) explain, is “the study of language data on a large scale, computer-aided analysis of very extensive collections of transcribed utterances or written texts”. Additionally, they suggest that a corpus should be representative and large enough for analysis and machine-readable. My corpora fulfill these two criteria as the following paragraphs will reveal.

My corpora are named CAR (Corpus of Albanian Research), CARE (Corpus of Albanian Research in English), CIAO (Corpus of Italian Academic Outputs)

and CIAOE (Corpus of Italian Academic Outputs in English). The papers are written by students of the Albanian Universities of Tirana and Vlora, and 13 Italian universities.

One of the most striking features of the corpora is the similar number of words in the Italian corpora, which have around three million words each (see Table 1). CAR also has a large number of around 2.2 million words. There is also a similar number of papers and words for males and females in CIAO, CIAOE and CAR. All the different academic levels (BA, MA, PhD) of students are present. Schmied (2011: 16) points out how the compatibility of the corpora has been a major challenge of comparative analyses on academic writing. This is particularly noticeable in the CARE corpus, which has only about 600,000 words.

Italian (CIAO)			Italian English (CIAOE)		
genre	# papers	# words	genre	# papers	# words
PhD thesis	54	2,035,346	PhD thesis	57	1,984,929
MA thesis	28	884,886	MA thesis	26	712,014
BA thesis	2	42,209	MA term paper	3	7,744
BA term paper	3	6,958	BA thesis	4	67,778
male	40	1,392,173	male	45	1,325,840
female	47	1,577,226	female	45	1,446,625
total (2003-2015)	87	2,969,399	total (2003-2015)	90	2,772,465
Albanian (CAR)			Albanian English (CARE)		
genre	# papers	# words	genre	# papers	# words
PhD thesis	52	2,285,498	PhD thesis	9	337,184
			MA thesis	10	144,225
			MA term paper	8	22,214
			BA thesis	13	110,054
			BA term paper	1	2,620
male	26	1,108,837	male	28	485,807
female	26	1,176,661	female	13	130,490
total (2010-2015)	52	2,285,498	total (2009-2015)	41	616,297

Table 1: Genres and gender in the Albanian and Italian corpora including their respective number of papers and word totals according to AntConc

In terms of disciplinary domains, the corpora contain papers from both, ‘hard’ and ‘soft’ sciences (Languages and Literature, Social Studies, Medicine, Chemistry, Physics, Economics, Mathematics and Computer Sciences) since these two opposite divisions of disciplinary domains have different conventions and language styles (see e.g. Vázquez & Giner 2009).

Beyond the description of the corpora, it needs to be clarified that in this paper, sememes across languages such as PROVE (which includes the English form and corresponding Albanian and Italian equivalents) are capitalized. The examples (from my corpora or previous literature) and the specific forms in Albanian, English or Italian (e.g. *probably*) are indicated in italics. This will be seen in the following qualitative and quantitative analysis subsections.

4 Analysis and discussion

4.1 Qualitative analysis

4.1.1 Intensifiers

The boosting category of intensifiers is the main focus of this subsection. From all selected boosters (see Appendix 1), SIGNIFICANTLY showed some of the most interesting differences based on the context. In Example 1, *significantly* was classified as a proposition-related intensifier because the student of Medicine is merely referring to an experimental context where certain elements influence the behavior or results of other elements. He is presenting the high intensity of the caused impact. This impact is expected to be evaluated in an objective manner by the author, in relation to the results discovered. The objectivity of *significantly* can partly be seen in the use of *it was observed*. If the results of a certain test show a statistically high variation between the results from different variables, the word *significantly* can be used. Therefore, this usage of *significantly* is propositional since the author does not appear to subjectively judge this event. The same propositional use was found in the less ambiguous Example 2. In this example, the scope of *significantly* has a limited extension on *different* and compares the difference of various *ranges*. In such cases, the narrow scope helps in the interpretation of *significantly* (or other boosters) as proposition-related.

- (1) *Në përfundim u pa se llaku i fluorit (NaF) frenonte ndjeshëm prodhimin e acidit laktik (183). (CAR12MPM_34)*
In the end, it was observed that the fluor spray (NaF) was significantly slowing down the production of lactic acid (183). (transl. by the author)
- (2) *On the contrary the BeCp*2 profile is significantly different, more clearly in the lower energy range [...]. (CIAOE14FPC_21)*

This subsection showed two examples of the intensifying function of *significantly*. The next subsection will focus on emphasizeers as well as their interesting overlap with intensifiers. As it will be shown, for statements which

were more in the form of opinions rather than showing statistical significance, *significantly* was often classified as an author-related emphaziser.

4.1.2 Emphasizers

Author-related emphazisers convey modality and stance by adding strength to the modified predicate (cf. Bondi 2008: 39). The way they overlap with proposition-related intensifiers is a relevant aspect examined in my paper. The first example below includes *significantly* as an emphaziser because it is fronted (sentence-initial), it has a wider scope extending on the following information and it has the special Italian English usage of the lexeme with the meaning ‘special meaning’, ‘meaningfully’ (according to Cambridge Online Dictionary, 2019). These aspects enable a closer connection to *significantly* with the author by presenting it as author-related. In Example 4 from CARE, *definitely* means ‘without doubt’, ‘absolutely’ (according to Grande Dizionario Hoepli Inglese). Through this adverb, the BA student of Languages and Literature is boosting her commitment towards the truthfulness of the proposition. She appears to be trying to persuade the reader and avoid possible doubts about her claim of *the internet* being *the fastest-growing tool*. The claim has a persuasive nature deriving from general knowledge, which supports the use of *definitely*. As a result, *definitely* was classified as an authorial emphaziser, in contrast to *definitely* in Example 5. In this example from CIAOE, *definitely* conveys the meaning of ‘completely’, ‘without exclusions’ (according to Grande Dizionario Hoepli Inglese). The MA student of Languages and Literature has used *differentiate definitely one’s class from that class* to convey a total intensity or completeness of the difference between two classes. If we replaced *definitely* with *completely*, the meaning of the clause complex would remain the same. It is surprising to see how a strong booster such as *definitely* can also affect the propositional level and express intensity.

- (3) *Most significantly at the beginning of this scene we are shown the ruins of a destroyed (or still to be completed) church.* (CIAOE13MML_78)
- (4) *The internet is definitely the fastest-growing communication tool [...]* (CARE13FBL_24)
- (5) *[...] in order also to differentiate definitely one’s class from that class of nobility not able to follow the same rate of consumption [...]* (CIAOE12FPS_13)

This striking qualitative finding for *definitely* (found in many similar examples) confirms the predictions of Quirk et al. (1985) and Bondi (2008: 39), which were mentioned in Subsection 2.1.1.

4.1.3 Orientation and manifestation

The interaction between manifestation and orientation is a crucial point in defining boosters. Halliday and Matthiessen (2014) offer a valuable introduction to these ‘axes of variation’. However, there is one related aspect that needs further elaboration. While the authors presented adverbs as exclusively objective-implicit forms (see Subsection 2.1.3), I argue that there is an overlap between the objective-explicit and objective-implicit categories within modalizing adverbs. The positioning of boosters appears to play a crucial role in their manifestation. If, for instance, a booster such as *clearly*, is placed in-between the clause complex or next to a verb, as in Example 6, it would represent an objective-implicit expression of the author. *Clearly* expresses its asseverative obviousness, but it is not the main focus of the clause complex. Meanwhile, in Example 7, *obviously* (another adverb expressing asseverative obviousness according to SFG) is separated by commas, as a sort of pause unit of the clause complex. My assumption is that in such cases, novice writers appear to present information more explicitly. Therefore, such examples were often ranked as objective-explicit. In this gradual increase of emphasis and explicitness, Example 8 contains *obviously* as a lexeme placed right at the beginning of the clause complex, simultaneously acting as a marked theme (purposive) of it. This usage (position) appears to enable the novice writer to ‘warn’ the readers that all the following information is clear and obvious. It transforms lexemes such as *obviously* from objective-implicit into objective-explicit. Even though not every case leads to such a transformation of normally objective-implicit forms, this appears to be a common phenomenon in academic writing.

- (6) *This phenomenon **can be clearly** seen in the LCC data.* (CIAOE11FPL_7)
- (7) *The first step towards a national identification was, **obviously**, the reinvigoration of the language.* (CIAOE12FPL_47)
- (8) ***Obviously**, we have $\psi(u(cid:48)) \in C 1(j0, Rj)$.* (CIAOE14FPMI_17)

Through this qualitative analysis, I was able to clarify the concepts that are relevant for this paper. The following quantitative analysis will present an additional interpretation of the results of the statistical tests and other results.

4.2 Quantitative analysis

4.2.1 The variables of corpus and author commitment in relation to boosters

One of the major features of logistic regressions is their applicability with depended binary variables. Since this paper derives from my larger PhD project on modalization, modulation and intensity, the binary variables of hedges or boosters were included. However, the focus here will be solely on boosters. Out of circa 15,000 tokens, only 8,088 tokens of hedges and boosters were used for the statistical test. Other tokens had to be excluded when preparing the data. To prepare the data for the logistic regression through Rbrul 3.1.1. (Johnson 2017), elements such as alpha clauses and subjective-explicit forms were excluded. Since these variables were unevenly distributed and had zero occurrences for many words, excluding them was the only way to ensure accurate results. Such problems were particularly prominent for the CARE corpus. Therefore, the CARE corpus had to be excluded. It contained many levels (categories) with non-measurable total frequencies of less than five, many empty cells, various types of interaction between the variables, for instance, the variable of orientation-manifestation with the one of author's commitment. Excluding them was the only way to ensure accurate results of the logistic regression through Rbrul 3.1.1. The variables included in this paper are the corpora (indicated as corpus) and author commitment (including author-related and proposition-related categories). It is necessary to explain the statistical measurements used and their benefits. One effective way to interpret logistic regression coefficients is by using odds ratios. The ratio of two odds, $(p1 / (1-p1)) / (p2 / (1-p2))$, is referred to as the odds ratio (cf. Gelman & Hill 2007: 82). Tables 3 and 4 contain the logodds ($\log [p/(1-p)]$), which are the logarithm of the odds ratio of hedges. The odds indicate the probabilities of occurrence/success or non-occurrence of boosters. Higher positive values would mean that boosters were favored in a certain corpus and so on. The benefits of using the logodds are that they present infinite +/- values (and the effect size) for boosters and can be easily updated with new data. N indicates the total number of tokens for boosters and hedges. The proportion shows the degree or percentage of boosters. In Table 3, for instance, 62.6 per cent (0.626) out of 2,906 tokens from CIAO are boosters. As it is known, the p value indicates the level of significance between the different strengths presented by these variables. As is known, p-values indicate the level of significance between the different strengths presented by these variables (see also Chatterjee & Ali 2006, Gelman & Hill 2007).

After having explained how the data was prepared for the test and which measurements were used, the discussion can move to the actual results obtained. The Somers. Dxy, which is related to the overall effect size, had a medium-low value of 0.332. The R square (R²) value, which determines how much the independent variables explain or affect the dependent variables, was relatively low (only 0.115). Since only 11.5 per cent of the variation is explained, it means that other variables might explain the usage of boosters.

For the variable of corpus, the results from Table 2 clearly show that boosters were significantly favored in CIAOE, followed by CIAO. Out of 2,906 hedges and boosters in CIAO, 62.6 per cent of them were boosters, with the highest positive logodds value of 0.254 across the corpora. The results from CIAO and CIAOE indicate that Italian students appear as fully committed to their claims in their L1 and L2 writing. This goes along with the (cultural stereotypes) claims by the Italian university teachers interviewed, who predicted such linguistic behavior by their students. The students especially preferred total intensity items such as ENTIRELY and TOTALLY. Appearing confident does not necessarily mean being confident, but it surely shows a preference for presenting high probability values (e.g. *certainly*) and total degree (e.g. *entirely*). It does also not exclude the fact that Albanian and Italian students used a considerable number of hedges, indicating partial commitment in writing. However, boosters were less favored in CAR, which demonstrates that Albanian students prefer to appear more tentative in their writing. In terms of tenor (the communicative situation with the audience) the lower vertical power role and institutional role as students or novice writers seemed to influence Albanians (rather than Italians), who appeared quite tentative as expected in such an academic hierarchy that includes more 'powerful' and experienced supervisors, teachers and other experts (cf. Argyle & Henderson 1984, Pulcini & Furiassi 2004, Matthiessen 2010). The results are partly influenced by the widely known difficulty of selecting equivalents. Another implication can be drawn on the basis of linguistic affordances (how the action expressed by the lexeme, especially verb, can be mentally imagined (cf. Aronin & Singleton 2010, Singleton 2013), which may often be different for Italians and Albanians. The different mental representations for these two language groups in terms of subject-action-object relations may contribute to the interpretation of this result. As Panajoti (2015: 177) confirms, "Albanian students have for long been encouraged to use either impersonal forms of addressing or the plural *we* to approach their own study". While analyzing similar doctoral dissertations by Albanian students, she discovered a strong tendency for using impersonal forms by both female and male candidates in various disciplines when treating their own works, which she interpreted as an unwillingness to "appropriate" their works

(Panajoti 2015: 183). These findings, the relatively high number of impersonal hedges in my corpora and the more natural connection of impersonal *it* forms with hedges rather than boosters may have contributed to the more frequent use of hedges by Albanians. This may not have applied for Italians.

Despite my consideration of the frequency, semantic, pragmatic and syntactic aspects when selecting equivalents, it is possible that I missed one matching equivalent or overlooked a frequent lexeme in one language simply because it does not have an equivalent in English. The inclusion of only epistemic adverbs, lexical verbs and introductory phrases, represented by 20 frequent English lexemes and a much larger number of their equivalents in Albanian and Italian does not fully support my implications on the usage of hedges and boosters. Overall, the variable of corpus was the second most influential variable with a significant p-value of $0.000...33 < 0.001$.

corpus	logodds	n	proportion
CIAO	0.254	2,906	0.626
CIAOE	0.158	2,737	0.625
CAR	-0.412	2,445	0.464
p-value			4.25e-33

Table 2: The variable of corpus influencing the usage of boosters (and hedges)

Table 3 below indicates the results of the variable of author commitment. On the one hand, boosters were significantly favored in the author-related category. The relative contribution of the author-related level demonstrates that the majority of authorial forms were boosters. This is shown by the high factor weight of 0.582 as well as the positive logodds value of 0.332 and the very high proportion score of 0.695 (69.5%). Author commitment proves that Albanian and Italian novice writers prefer to show complete commitment and establish a direct connection with their propositions while boosting. On the other hand, boosters were significantly disfavored on the propositional level. Finally, author commitment was the variable with the second highest relative strength due to the significant p-value of $0.000...36 < 0.001$.

author commitment	logodds	n	proportion
author-related	0.332	3,372	0.695
proposition-related	-0.332	4,716	0.492
p-value			8.92e-36

Table 3: The variable of author's commitment influencing the usage of boosters (and hedges)

While the logistic regression is the largest test I used to compare the data, there are several detailed comparisons of single lexemes in CARE and CIAOE that only the following log-likelihood test can display.

4.2.2 A pairwise comparison between CARE and CIAOE in their occurrences of hedges and boosters

This subsection includes a pairwise comparison between the English corpora of log likelihood (indicated as LL) and effect size (oddsratio, indicated as OR) of the most interesting boosters, out of a larger set of boosters (see Appendix 1). For further clarification on the exact equivalents across languages, Appendix 2 can be consulted. In the comparison of the English corpora, PROVE, CLEARLY and DEFINITELY show a significant difference in favor of CARE when compared to CIAOE (see Table 4 below). This is indicated by the LL values which are all much higher than the critical value of 3.84. Moreover, the OR values for CLEARLY, DEFINITELY and PROVE are more than 1.0 and therefore, they can be considered quite high. SIGNIFICANTLY, SURELY and CERTAINLY were significantly more frequent in CIAOE than in CARE. However, their LL values (indicated with -) are smaller than, for instance, the ones of CLEARLY.

booster	CARE	%	CIAOE	%	LL	OR
PROVE	86	0.01	234	0.01	+14.62	1.65
SIGNIFICANTLY	42	0.01	342	0.01	-15.36	0.55
SURELY	8	0.00	85	0.00	-6.85	0.42
CERTAINLY	17	0.00	132	0.00	-5.16	0.58
CLEARLY	109	0.02	181	0.01	+60.28	2.71
DEFINITELY	18	0.00	34	0.00	+7.93	2.38

Table 4: A comparison of boosters in CARE and CIAOE in terms of log-likelihood and size effects

As shown in Table 4, Albanians employed boosters such as PROVE significantly more often than Italians whereas Italians employed others such as SURELY more often. This indicates that the (idiosyncratic) preference for certain individual words (and non-use of others) is a strong factor that influences the linguistic choices of these students. In this sense, Albanians preferred to emphasize evidence by using PROVE significantly more often than Italians. They generally show a stronger tendency to use objective-implicit and subjective-explicit forms in their writings. However, this discussion is restricted to the specific items displayed in Table 4 which were selected as the most interesting ones. The expected over-confidence of Italians (as Italian university

teachers claimed) does not prevail in the use of the lexical verbs (PROVE) or epistemic adverbs (CLEARLY and DEFINITELY). As for the quite low OR numbers for SIGNIFICANTLY, SURELY and CERTAINLY, they can be related to the small interrelation between CARE and CIAOE. Finally, these boosters did not only show external differences with other boosters but also internal ones in relation to their emphasizing and intensifying functions. This will be further explained in the next subsection.

4.2.3 The quantitative results of the same lexeme as an emphasizer and intensifier

As shown in Subsection 4.1.2 above, the categories of author-related emphasizers and proposition-related intensifiers may often overlap (in terms of language use), depending on the lexeme. For this small-scale study, the interesting case of SIGNIFICANTLY was chosen.

Before going into detail on the frequencies of SIGNIFICANTLY, it is necessary to explain the details regarding the choice of equivalents across languages for this sememe. These details include semantic meanings as well as syntactic and pragmatic features of the selected boosters. I partly followed Aijmer and Simon-Vandenberg (2004: 1785), who presented pragmatic markers (including boosters) as having “wide meanings and different contextually based functions”. Therefore, these functions can be conveyed through different translations. *Significantly* in Italian can be expressed in many ways, such as the closer *sensibilmente* (*significantly*) and the peripheral ones (in terms of surface features) *in (una) maniera sensibile* (*in a significant (sensitive) manner*), *in (un) modo sensibile* (*in a significant way*), *in (una) maniera significativa* (*in a significant manner*), and *in (un) modo significativo* (*in a significant way*). Albanian showed rich semantic fields like Italian, with forms such as *në mënyrë (shumë) të ndjeshme/konsiderueshme* (*in a (very) significant/considerable manner*), apart from the closer cognate *ndjeshëm*. Another suggested Italian form for *significantly* was *notevolmente*, which was not included due to its close similarity to *noticeably*. The chosen lexeme, *sensibilmente*, was explained in Grande Dizionario Hoepli Inglese as related to the senses or deriving from the senses. However, this lexeme has a second meaning and can function as a booster, which makes it similar to *significantly*. In my data, its frequent boosting function was found and all its instances that were related to senses and perception were excluded. Admittedly, the usage of *significativo* in Italian may be considered as an anglicism. The inclusion of only the forms *in (un) modo significativo* and *in (una) maniera significativa* showed all the possible used forms that express the same pragmatic and semantic meaning as the English form *significantly*.

The considerable higher frequency of the booster SIGNIFICANTLY as a proposition-related intensifier can be clearly seen in all the corpora that are shown in Figure 2. In CIAO, the uses as an intensifier are slightly less frequent. This is probably the result of the higher use of *sensibilmente* (*significantly*) to indicate a ‘special meaning’, which appears also to be authorial. More than two thirds of the usages of SIGNIFICANTLY boosted the propositional level rather than the author-proposition relation. This means that SIGNIFICANTLY was often found in comparative, superlative and quantity-oriented contexts. It needs to be emphasized that not all instances of SIGNIFICANTLY found in quantity-oriented contexts that express the semantic meaning of ‘statistically significant’ (as in *significantly higher/increase*) are inevitably intensifiers. It also needs to be mentioned that not every instance of SIGNIFICANTLY expressing a ‘significant manner’ (as in *significantly improve*) is necessarily ranked as an author-related emphazier. Despite many exceptions, the relation between the semantic categories of SIGNIFICANTLY and its authorial or propositional relation is very strong.

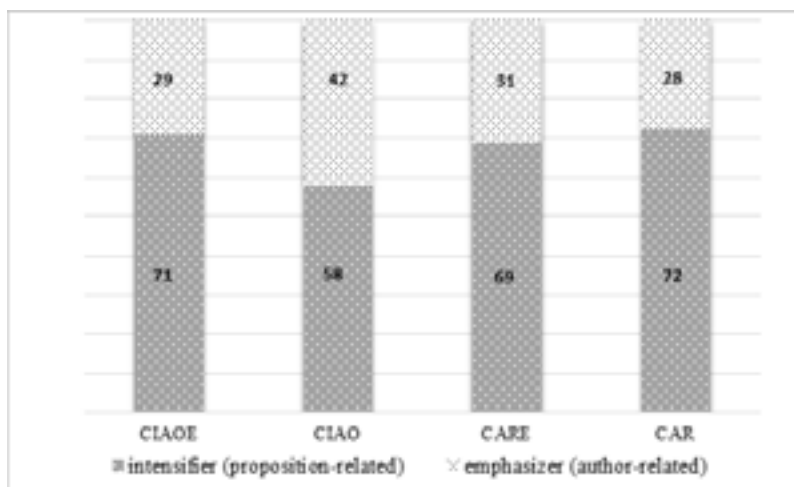


Figure 2: The percentages of SIGNIFICANTLY as an author-related emphazier and proposition-related intensifier across the four corpora

Through the case of SIGNIFICANTLY, the frequent overlap in language use between intensity and emphasis within the boosting category was revealed. This was the last relevant aspect, which leads us to the concluding remarks on the findings.

5 Conclusions

This paper focused on pragmatic, semantic and lexico-grammatical aspects of boosters in Albanian and Italian student academic writing in L1 and English as an L2. In this final section, the main findings, including the answers to the two research questions, will be summarized. The first research question was related to the semantic, pragmatic and lexico-grammatical aspects of boosters. From a qualitative perspective, there was a context-dependent variation in the proposition-oriented (semantic) and author-oriented (pragmatic) way as conveyed by the same boosters (SIGNIFICANTLY, DEFINITELY). The variation confirms the general claims of Quirk et al. (1985) and Bondi (2008) that the functions as intensifier and emphazier often overlap for the same lexeme. It is quite surprising that even a strong booster such as DEFINITELY can affect the propositional level and express intensity. In terms of orientation and manifestation, there was an important elaboration within SFG. It consisted in finding that objective-explicitness does not only consist of forms such as *it is clear, it is obvious*, and objective implicitness is not only bound to adverbs. The assumption was that (boosting) adverbs placed right at the beginning of the clause complex, as marked themes, are often explicit and mark author's commitment more directly.

The second research question was: What are the differences in the frequency of proposition-related intensifiers and author-related emphaziers within and across the corpora? Generally, Italians appeared more confident and fully committed in their L1 and L2 writing by boosting more than Albanians in their L1, who preferred to appear more tentative (by hedging). These differences are statistically significant as the logistic regression showed. This cross-national difference is partly explained with reference to education (i.e. Albanian students hedge more, partly due to the way they are often taught to use impersonal forms or somehow appear tentative in their writings, cf. Panajoti 2015: 177) and linguistic affordance (i.e. different mental representations of subject-action). In relation to linguistic affordance, Albanians use, for example, more impersonal forms in my corpora. Such forms may be more related to hedges than boosters. It can also be explained with reference to tenor, i.e. Albanians could have adapted more to their lower vertical 'power' role and institutional role as novices and students, in comparison to more 'powerful' and experienced teachers, supervisors and other readers, in an interactional academic writing context (cf. Pulcini & Furiassi 2004, Matthiessen 2010). Another possible explanation is related to academic culture, i.e. leading Italian students to 'appear' fully committed, as interviewed Italian teachers predicted. At the macro level, the L1-L2 comparison for Italian

students showed a strong similarity between L1 and English writings in terms of perceptions of audience and writer as well as metadiscoursal conventions. This finding is different from what Hyland (2003) and Connor (1996) predicted. At the micro level, there were L1-L2 differences hinting at a preference for individual lexemes and non-use of others. Albanians, for instance, emphasized significance and preferred to prove credibility to their claims through a significantly more frequent use of *significantly* and *prove*, whereas Italians showed certainty through *surely* and *certainly*.

The small size of the CARE corpus is a limitation of this paper. Despite the relevance of dependency and disciplinary domains in a discussion of boosters, as Bondi's (2008) study also showed, these aspects are beyond the focus of this paper. Further studies could focus on these relevant aspects within academic writing.

Appendices

Appendix 1 - A complete list of all hedges and boosters used for the logistic regression.

hedges	APPROXIMATELY, ALMOST, NEARLY, KIND OF, POSSIBLY, PROBABLY, SUGGEST, SEEM, BELIEVE, IN MY OPINION, FROM MY POINT OF VIEW, FROM MY PERSPECTIVE
boosters	ENTIRELY, COMPLETELY, TOTALLY, SIGNIFICANTLY, CLEARLY, OBVIOUSLY, OF COURSE, DEFINITELY, SURELY, CERTAINLY, WITHOUT DOUBT, UNDOUBTEDLY, DEMONSTRATE, PROVE

Appendix 2 - The specific boosters and their Albanian and Italian counterparts that were compared in terms of LL and frequencies

	compared boosters	
English lexeme	Albanian equivalent / translation	Italian equivalent / translation
<i>prove, establish facts, demonstrate, show, validate, confirm, find, verify, certify, attest, testify, sustain</i>	<i>të provosh, të faktosh, të vërtetosh, të dëshmosh</i>	<i>provare, comprovare, attestare, dare prova/e, fornire prova/e, evidenziare, dimostrare (close to demonstrate), certificare, documentare, testimoniare</i>
<i>definitely, certainly, absolutely, clearly, decidedly, doubtless, obviously, unequivocally, undeniably, unquestionably, without doubt, surely</i>	<i>absolutisht, në mënyrë absolute, ne menyre absolute (not found), pa asnjë përjashtim, plotësisht, krejt</i>	<i>assolutamente, in (una) maniera assoluta, in (un) modo assoluto, in assoluto, in senso assoluto, decisamente, con decisione, risolutamente, indiscutibilmente, fermamente, sicuramente, con certezza, in via definitiva, definitivamente, certo, prontamente</i>
<i>clearly, without any doubt, apparently, certainly, definitely, distinctly, evidently, obviously, audibly, incontestably, incontrovertibly</i>	<i>qartësisht, në mënyrë të/ të tregosh qartë, me qartësi qartas, qartazi</i>	<i>chiaramente, chiaro (incl. in (un) modo chiaro), in (una) maniera chiara, senza alcun dubbio, distintamente, ovviamente, inequivocabilmente</i>
<i>surely, without doubt, absolutely, assuredly, certainly, indeed, inevitably</i>	<i>sigurisht, me siguri</i>	<i>sicuramente, di sicuro, senza esitazione, con sicurezza, con padronanza, certamente, di certo</i>
<i>certainly, without doubt, absolutely, assuredly, exactly, of course, unquestionably, surely</i>	<i>natyrisht, siç është e natyrshme, sigurisht, pa dyshim</i>	<i>certamente, di certo, certissimamente (spoken), con (assoluta) certezza, in tutta certezza, con sicurezza, di sicuro, indubbiamente, innegabilmente, certo</i>
<i>significantly, considerably, decidedly, extremely, greatly, much, notably, quite, very, very much, incomparably</i>	<i>ndjeshëm, ndjeshem, ndjeshme, në mënyrë (shumë) të ndjeshme/konsiderueshme, mjaft i madh, i dukshëm, i rëndësishëm</i>	<i>sensibilmente, in (una) maniera sensibile, in (un) modo sensibile, in (una) maniera significativa, in (un) modo significativo, notevolmente, considerabilmente, decisamente, nettamente, alquanto</i>

Notes:

- ¹ The criteria of selection were high frequency in my corpora, pragmatic similarity, semantic similarity, previous literature, e.g. Fraser (cf. 2010: 22ff), Hyland (1998a: 154-177), Toska (2012), Trajkova (cf. 2015: 145ff) and suggestions by experts (Memushaj, Kabashi, personal communication, 6-7 December 2016).
- ² The lexemes indicated in bold are the lexemes selected on the basis of the researcher's perception/competence. The elements that are not in bold represent the ones added from the previously mentioned sources above. The elements underlined are additional equivalents, which were not included for this paper.
- ³ It is clear that a detailed contextual analysis was conducted for 15,000 instances, including each single element in this table and from Appendix 1. Only the boosting instances were considered for the indicated frequencies in this paper. If, for instance, *provare* refers to feelings, it cannot be considered an equivalent of the booster *prove* and it is excluded from the analysis.

References

- Aijmer, K. and Simon-Vandenberg, A. M. (2004) 'A model and a methodology for the study of pragmatic markers: The semantic field of expectation.' *Journal of Pragmatics* 36, 1781-1805.
- Anthony, L. (2014) *AntConc* (Version 3.4.3) [Computer Software]. Tokyo, Japan: Waseda University. Available from <http://www.laurenceanthony.net/>.
- Argyle, M. and Henderson, M. (1984) *The Anatomy of Relationships*. London: Heinemann, Harmondsworth: Penguin.
- Aronin, L. and Singleton, D. (2010) 'Affordances and the diversity of multilingualism.' *International Journal of the Sociology of Language* 125. *Special Issue: The Diversity of Multilingualism*. 105-129.
- Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (1999) *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.
- Bondi, M. (2008) 'Emphatics in academic discourse: Integrating corpus and discourse tools in the study of cross-disciplinary variation.' In: Ädel, A. and Reppen, R. (eds) *Corpora and Discourse: The Challenges of Different Settings*. Cambridge: Cambridge University Press. 31-55.
- Cambridge Dictionary*. (2019) Cambridge: Cambridge University Press. Retrieved on 10 May 2020 <<https://dictionary.cambridge.org/de/>>.
- Chatterjee, S. and Ali, S. H. (2006) *Regression Analysis by Example*. 4th ed. New Jersey: John Wiley & Sons.
- Connor, U. (1996) *Contrastive Rhetoric*. Cambridge: Cambridge University Press.
- Conrad, S. and Biber, D. (2000) 'Adverbial marking of stance in speech and writing.' In: Hunston, S. and Thompson, G. (eds) *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press. 1-27.
- Crismore, A. (1989) *Talking with Readers: Metadiscourse as Rhetorical Act*. Frankfurt: Peter Lang.
- Dhaskali, V. (2017) 'Hedging and boosting in Albanian, British and Italian online news articles in Europe.' In: Schmied, J. and van der Bom, I. (eds) *Working with Media Texts: Deconstructing and Constructing Crises in Europe*. Göttingen: Cuvillier Verlag. 135-164.

- Dhaskali, V. (2020) *A corpus-based comparison of Albanian and Italian student writing in L1 and English as L2: Hedges and boosters as modalization by degree*. Doctoral thesis. Chemnitz: Technische Universität. Retrieved on 10 January 2021 <<https://nbn-resolving.org/urn:nbn:de:bsz:ch1-qucosa2-384460>>.
- Fraser, B. (2010) 'Pragmatic competence: The case of hedging.' In: Kaltenböck, G., Mihatsch, W. and Schneider, S. (eds) *New Approaches to Hedging. Studies in Pragmatics 9*. North America: Emerald Group Publishing Limited. 15-34.
- Gelman, A. and Hill, J. (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.
- Gjoleka, M. (2014) 'Il contributo degli studiosi italiani per la lingua e la cultura albanese.' Paper presented at the international conference "La lingua italiana in Albania tra passato e futuro", organized by the Department of Italian Language, Faculty of Foreign Languages, University of Tirana.
- Halliday, M. A. K. (1985) *An Introduction to Functional Grammar*. London: Edward Arnold.
- Halliday, M. A. K. and Matthiessen, C. M. I. M. (2014) *An Introduction to Functional Grammar*. 4th ed. London: Routledge.
- Holmes, J. (1990) 'Hedges and boosters in women's and men's speech.' *Language & Communication* 10(3), 185-205.
- Hösch, E. (2006) *Storia dei Balcani*. Bologna: Il Mulino.
- Hyland, K. (1998a) *Hedging in Scientific Research Articles*. Amsterdam: John Benjamins.
- Hyland, K. (1998b) 'Boosting, hedging and the negotiation of academic knowledge.' *TEXT* 18(3), 349-382.
- Hyland, K. (2000a) *Disciplinary Discourses: Social Interactions in Academic Writing*. London: Longman.
- Hyland, K. (2000b) 'Hedges, boosters and lexical invisibility: Noticing modifiers in academic texts.' *Language Awareness* 9(4), 179-197.
- Hyland, K. (2003) *Second Language Writing*. Cambridge: Cambridge University Press.
- Hyland, K. (2005) *Metadiscourse: Exploring Interaction in Writing*. London: Continuum.
- Hyland, K. (2017) 'Metadiscourse: What is it and where is it going?' *Journal of Pragmatics* 113, 16-29.
- Johnson, D. E. (2017) Rbrul. (Version 3.1.1.) [Computer Software]. York, UK: University of York. Available from <<http://www.danielezrajohnson.com/Rbrul.R.>>.
- Lafuente-Millán, E. (2008) 'Epistemic and approximative meaning revisited: The use of hedges, boosters and approximators when writing research in different disciplines.' In: Burgess, S. and Martín-Martín, P. (eds) *English as an Additional Language in Research Publication and Communication*. Bern: Peter Lang. 65-82.
- Martin, J. R. and White, P. R. R. (2005) *The Language of Evaluation: Appraisal in English*. London: Palgrave Macmillan.
- Matthiessen, C. M. I. M. (2010) 'Language evolving: Notes towards a semiotic history of humanity.' *The International Systemic Functional Congress* [Webinar]. Vancouver: University of British Columbia. Retrieved on 10 November 2017 <<https://www.youtube.com/watch?v=U15qHWJcfT4>>.
- McEnery, T. and Hardie, A. (2012) *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.
- Mu, C., Jun Zhang, L., Ehrlich, J. and Hong, H. (2015) 'The use of metadiscourse for knowledge construction in Chinese and English research articles.' *Journal of English for Academic Purposes* 20, 135-148.

- Panajoti, A. (2015) 'Authorial identity in PhD theses in Albania(n).' In: Schmied, J. (ed.) *Academic Writing for South Eastern Europe: Practical and Theoretical Perspectives*. Göttingen: Cuvillier Verlag. 177-187.
- Picchi, F. (ed.) (n.d.) *Grande Dizionario Hoepli Inglese*. Hoepli Editore. La Repubblica. it. Available from <http://dizionario.repubblica.it/inglese.php>.
- Pulcini, V. and Furianassi, C. (2004) 'Spoken interaction and discourse markers in a corpus of learner English.' In: Haarman, L., Morley, J. and Partington, A. (eds) *Corpora and Discourse*. Bern: Peter Lang. 107-123.
- Quirk, R., Greenbaum, R., Leech, G. and Svartvik, J. (1985) *A Comprehensive Grammar of the English Language*. London: Longman.
- Schmied, J. (2011) 'Academic writing in Europe: A survey of approaches and problems.' In: Schmied, J. (ed.) *Academic Writing in Europe: Empirical Perspectives*. Göttingen: Cuvillier. 1-22.
- Singleton, D. (2013) 'Affective dimensions of second language ultimate attainment.' In: Gabrys-Barker, D. and Bielska, J. (eds) *The Affective Dimension in Second Language Acquisition*. Bristol: Multilingual Matters. 29-34.
- Toska, B. (2012) *Studim përqasës i konektorëve në tekstet argumentuese të gjuhës angleze dhe asaj shqipe*. Doctoral thesis. Tirana: University of Tirana, Albania. Retrieved on 3 May 2018 <<http://www.doktoratura.unitir.edu.al>>.
- Toska, B. (2015) 'Structural, textual and commentary linkers in Albanian PhD dissertations: A pilot study.' In: Schmied, J. (ed.) *Academic Writing for South Eastern Europe: Practical and Theoretical Perspectives*. Göttingen: Cuvillier Verlag. 163-177.
- Trajkova, Z. (2015) 'Expressing commitment and detachment in non-native English BA graduate papers.' In: Schmied, J. (ed.) *Academic Writing for South Eastern Europe: Practical and Theoretical Perspectives*. Göttingen: Cuvillier Verlag. 143-163.
- Vázquez, I. and Giner, D. (2009) 'Writing with conviction: The use of boosters in modelling persuasion in academic discourses.' *Revista Alicantina de Estudios Ingleses* 22, 219-237.

Vincenzo Dheskali is research assistant at the University of Chemnitz, Germany. His main tasks include coordinating the PhD group at the Department of English, organizing International Corpora, leading the Eye Tracking Experiments and coordinating the project "Truths" (an exchange project funded by the DAAD). He investigates academic writing and journalistic writing and has defended his PhD in Corpus Linguistics. He speaks English (CPE), Italian (C1), German (C1) and Albanian (mother tongue) fluently, and has a basic knowledge of Spanish (B1).

Address: Vincenzo Dheskali, Department of English and Linguistics, Faculty of Humanities, University of Chemnitz, Reichenhainer Straße 39, Room 217, 09107 Chemnitz, Germany. [e-mail: vincenzo.dheskali@gmail.com]