

A Corpus-based Comparison of Albanian and Italian Student Academic Writings in L1 and English as L2: Introducing Hedges and Boosters within Modalization, Modulation and Degree

Vincenzo Dheskali

Sixth Semester PhD Student

Chemnitz University of Technology

vincenzo.dheskali@s2015.tu-chemnitz.de

Hedges and boosters: My working definition

Hedges (e.g. *possibly, almost, I think*) and **boosters** (e.g. *certainly, completely, demonstrate*) are numerous lexical and non-lexical devices which express different degrees of authors' direct and indirect commitment regarding the probability and usuality of the expressed proposition. They are modalization, modulation and degree devices that interweave interpersonal and ideational socio-semiotic processes on a semantic level (approximators), pragmatic level (shields), and display an interaction of both. They express different forms of manifestation, orientation, prosody of 'modality', syntactic positioning, approximation, shields and polarity across and within different cultural and linguistic contexts of student academic writings.

*I notice something which **polemicists seemed to have missed**: **the demonstrable, elementary fact that Wilde is almost always right.** (CIAOE15FBL_82)*

(see also Halliday 1985; Halliday and Matthiessen 2014; Lafuente Millàn 2008; Hyland 1998a; 2005; 2017; Prince et al. 1980; Salager-Meyer 1994; Schmied 2008; 2009; 2018).

stronger colour=clearer **hedge** and **booster**
weaker colour=less **hedge** and **booster**

Hedges and boosters: My working definition

Hedges (e.g. *possibly, almost, I think*) and **boosters** (e.g. *certainly, completely, demonstrate*) are numerous lexical and non-lexical devices which express different degrees of authors' direct and indirect commitment regarding the probability and usuality of the expressed proposition. They are modalization, modulation and degree devices that interweave interpersonal and ideational socio-semiotic processes on a semantic level (approximators), pragmatic level (shields), and display an interaction of both. They express **different** forms of manifestation, orientation, prosody of 'modality', syntactic positioning, approximation, shields and polarity **across** and **within** different **cultural** and **linguistic contexts** of student academic writings.

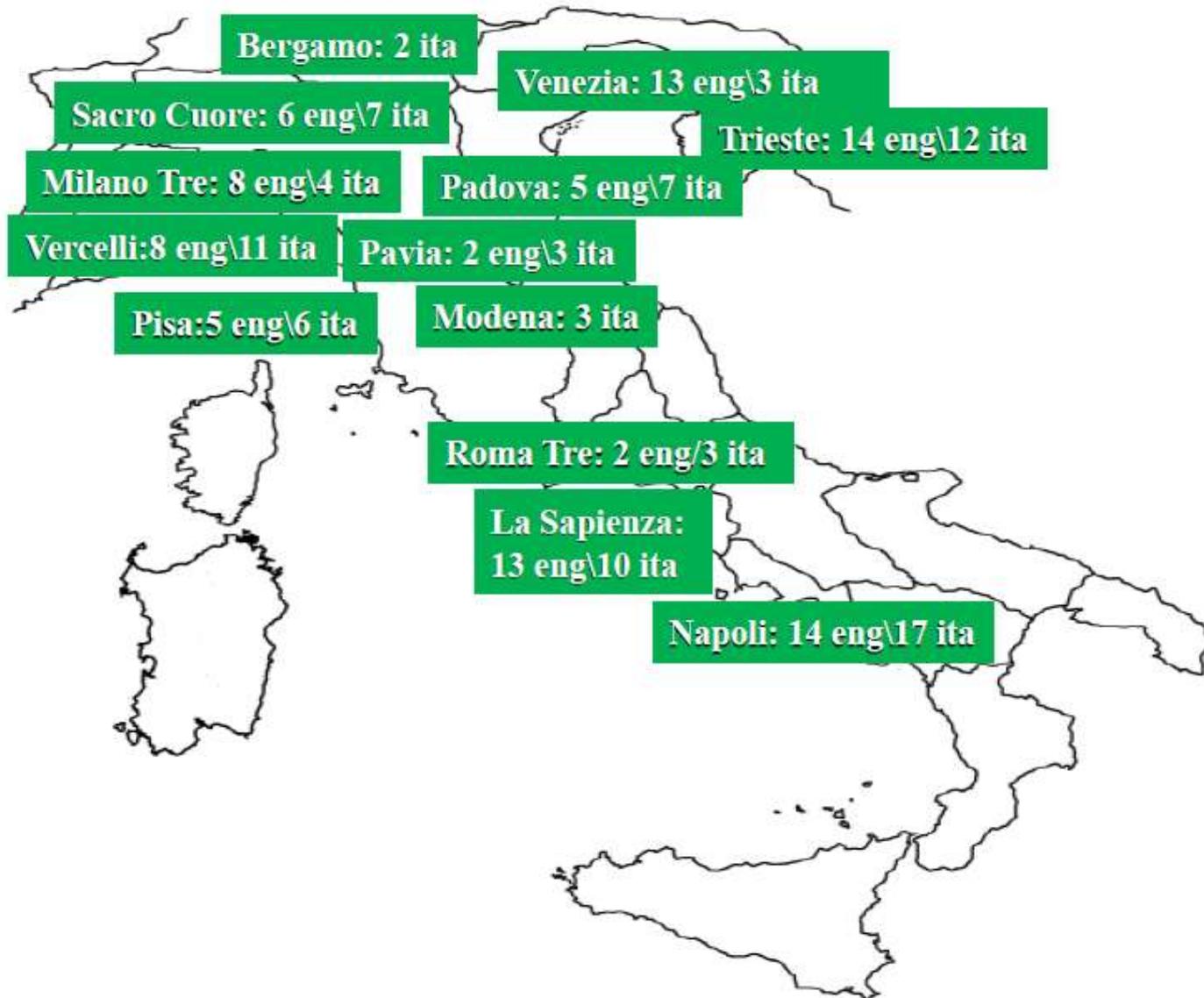
*I notice something which **polemicists seemed to have missed**: **the demonstrable, elementary fact** that Wilde is **almost always right**. (CIAOE15FBL_82)*

(see also Halliday 1985; Halliday and Matthiessen 2014; Lafuente Millàn 2008; Hyland 1998a; 2005; 2017; Prince et al. 1980; Salager-Meyer 1994; Schmied 2008; 2009; 2018).

stronger colour=clearer **hedge** and **booster**
weaker colour=less **hedge** and **booster**

Methodology

- Corpus analysis
- Qualitative and quantitative comparison with the concordance software **AntConc** (Anothony 2014)
 - 15.000 hedging and boosting instances: copied to Excel and analysed based on Bondi (2008), Halliday and Matthiessen (2014), Lafuente Millàn (2008), Prince et al. (1980) and Quirk et al. (1985) and 10 Albanian, English and Italian dictionaries. Statistical measurements through **Logistic Regression** (with Rbul, R), **Log-Likelihood** and relative frequencies x 1M words.
- Four corpora consisting of L1 and L2 writings by Albanian and Italian students



Picture 1: A map with the geographical distribution of my Italian corpora, CIAO (ita) and CIAOE (eng)

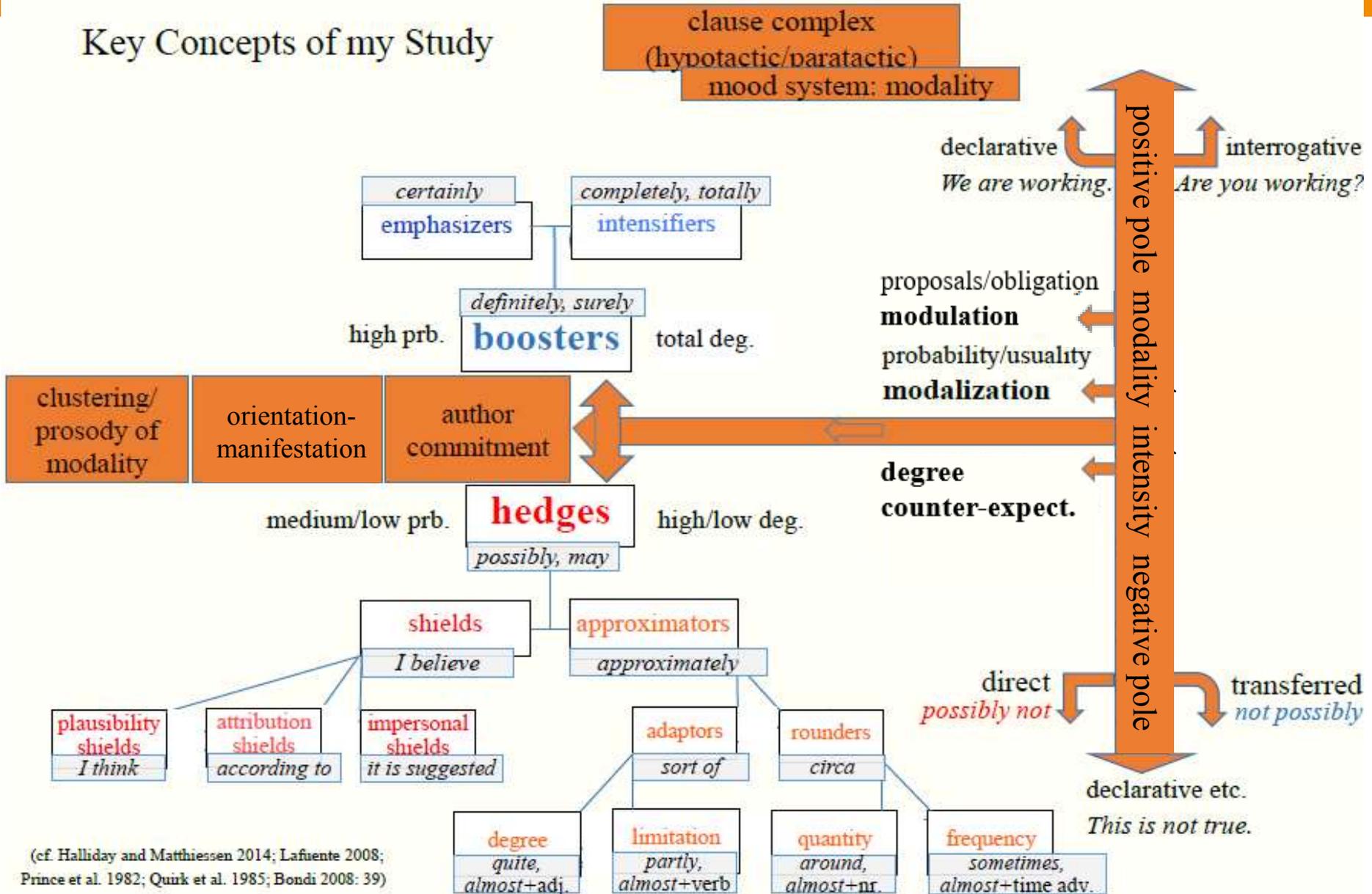
Italian Corpus (CIAO)			Italian English Corpus (CIAOE)		
type of papers	# papers	# words	type of papers	# papers	# words
PhD theses	54	2.035.346	PhD theses	57	1.984.929
MA theses	28	884.886	MA theses	26	712.014
BA theses	2	42.209	MA term papers	3	7.744
BA term papers	3	6.958	BA theses	4	67.778
males	40	1.392.173	males	45	1.325.840
females	47	1.577.226	females	45	1.446.625
total (2003-2015)	87	2.969.399	total (2003-2015)	90	2.772.465
Albanian Corpus (CAR)			Albanian English Corpus (CARE)		
type of papers	# papers	# words	type of papers	# papers	# words
PhD theses	52	2.285.498	PhD theses	9	337.184
			MA theses	10	144.225
			MA term papers	8	22.214
			BA theses	13	110.054
			BA term papers	1	2.620
males	26	1.108.837	males	28	485.807
females	26	1.176.661	females	13	130.490
total (2010-2015)	52	2.285.498	total (2009-2015)	41	616.297

Table 1: Albanian and Italian corpora including respective word totals according to AntConc

Italian Corpus (CIAO)			Italian English Corpus (CIAOE)		
sections	# papers	# words	sections	# papers	# words
Lang. & Lit.	40	1.402.214	Lang. & Lit.	41	1.197.841
Social Studies	8	483.368	Social Studies	8	462.368
Medicine	9	86.084	Medicine	10	116.077
Chemistry	8	233.434	Chemistry	8	199.209
Physics	6	172.326	Physics	8	257.065
Economics	8	386.148	Economics	8	267.383
Math. & Inform.	8	205.825	Math. & Inform.	7	264.778
total	87	2.969.399	total	90	2.772.465
Albanian Corpus (CAR)			Albanian English Corpus (CARE)		
sections	# papers	# words	sections	# papers	# words
Lang. & Lit.	8	605.556	Lang. & Lit.	33	293.899
Social Studies	8	483.872	Social Studies	3	127.883
Medicine (+M.Law)	3 (+1)	103.037	Chemistry	1	69.513
Chemistry	8	214.619	Physics	1	14.979
Biotechno. & Physics	8	334.607	Economics	2	75.173
Economics	8	361.906	Informatics	1	34.850
Math. & Inform.	8	181.901			
total	52	2.285.498	total	41	616.297

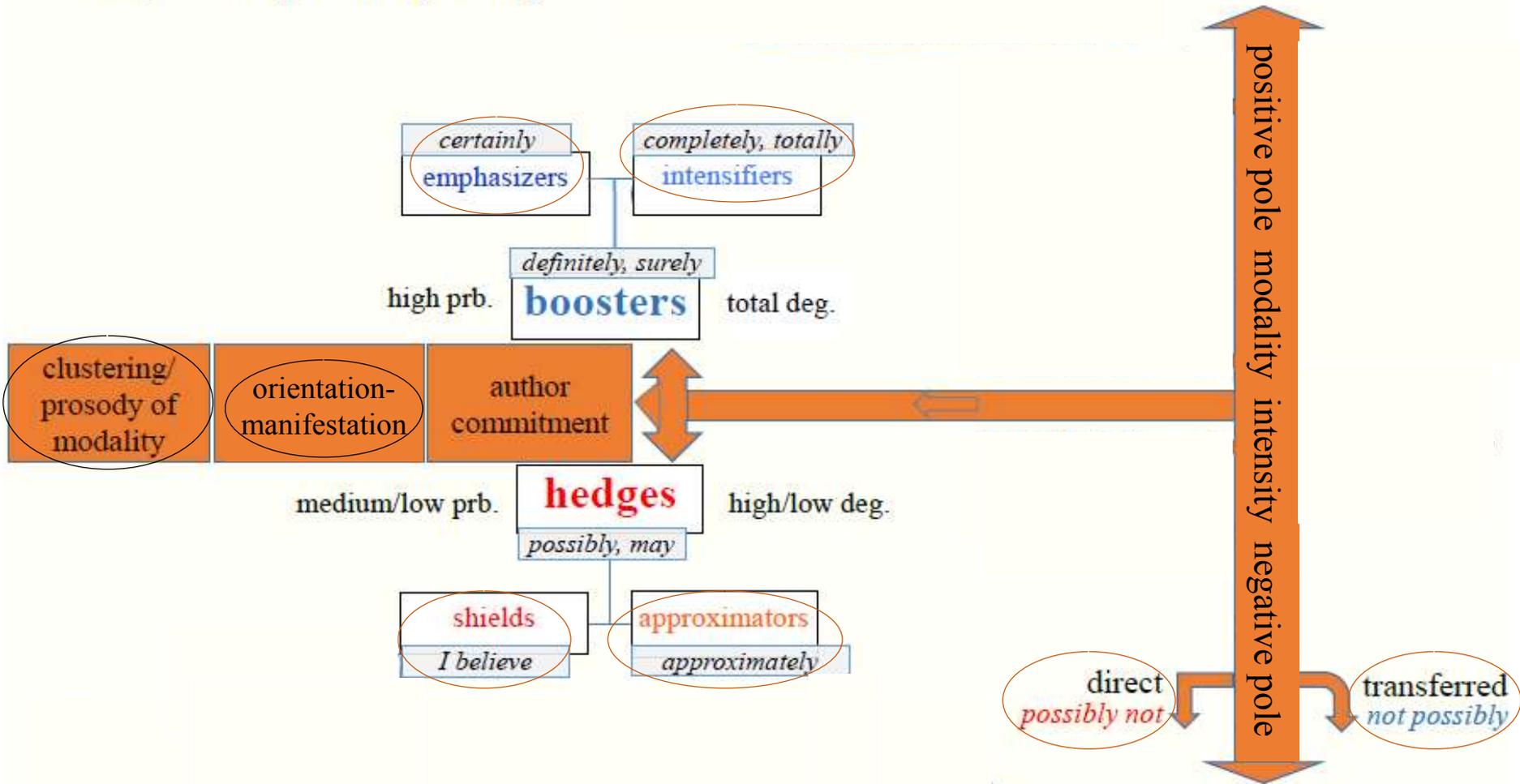
Table 2: Albanian and Italian corpora according to sections (genres), texts and words

Key Concepts of my Study



(cf. Halliday and Matthiessen 2014; Lafuente 2008; Prince et al. 1982; Quirk et al. 1985; Bondi 2008: 39)

Key Concepts of my Study



(cf. Halliday and Matthiessen 2014; Lafuente 2008; Prince et al. 1982; Quirk et al. 1985; Bondi 2008: 39)

Orientation (objective, subjective) and Manifestation (explicit, implicit)

- 1a). Subjective explicit: projecting mental clause, supported by an idea clause: *I guess, it is believed that, I don't think it was inappropriate for Federico, I suppose Lucrezia was disappointed by these events.*
- 1b). Subjective implicit: commonly a finite (modal auxiliary) within the clause and mood: *Miri who could be the main actor, his friends may be influenced.*
- 2a). Objective implicit: clause, mood - usually carried by a mood adjunct, esp. a modal adverb: *certainly*. / Placed in between the sentence e.g. in a neutral position: *Trump will possibly lose the elections.*
- 2b). Objective explicit: usually expressed through a relational clause containing a factual carrier, clause etc.: *It is certain, it is probable that* (see Halliday and Matthiessen 2014: 688; Martin and White 2005: 130-131). / Emphasized modal adverb e.g. placed initially: *Obviously, it was successful.*, parenthetically.

Values	<u>Transferred negative</u>
High:	<i>certainly</i> = not possibly
Medium:	<i>probably</i> = not probably
Low:	<i>possibly</i> = not certainly (Halliday and Matthiessen 2014: 180).

*my theoretical additions

Research questions

1. What are the qualitative differences in using hedges and boosters and their semantic and pragmatic subcategories between the Albanian and Italian L1 and L2 corpora?
2. What choices do students make in the usage of hedges and boosters: word order, orientation, manifestation, and prosody of modalization?
3. What are the quantitative differences in using hedges and boosters, and their semantic and pragmatic subcategories between the Albanian and Italian L1 and L2 corpora?
4. What are the quantitative differences in the usage of hedges and boosters: word order, orientation, manifestation, and prosody of modalization?

Research question 1:

What are the qualitative differences in using hedges and boosters and their semantic and pragmatic subcategories between the Albanian and Italian L1 and L2 corpora?

1.1 New introduced categories:

plausibility shields (e.g. *I think*)

attribution shields (e.g. *Sumra claims*) (Prince et al. 1980; Kaltenböck, Mihatsch & Schneider 2010)

impersonal shields (e.g. *it is suggested*),

multiple shields (e.g. *one of them is probably Ludovico di Breme, as Ungaretti himself suggests* [...]) ((transl.V.DH from CIA009MPL_7))

Clear inclusion / exclusion criteria

Frequently excluded instances for the lemma *prove*: *the desire of proving, failure to prove, has the burden to prove, wish /want /aim to prove, we should prove, let us prove, we have to prove.*

Frequently included instances: *we are now in position to prove, it suffices to prove and we start by proving.*

stronger colour=clearer hedge and booster
weaker colour=less hedge and booster
*my theoretical additions

Research question 1:

What are the qualitative differences in using hedges and boosters and their semantic and pragmatic subcategories between the Albanian and Italian L1 and L2 corpora?

1.2 Overlap: The dual functions of different lexemes as hedges (*suggest, believe, not entirely, not clearly*) and boosters (*strongly suggest, strongly/really/truly believe*)

Subcategories: an overlap between the propositional and authorial categories within the same hedge and booster. Reflected in the quantitative analysis.

Similar to Varttala (cf. 2001: 11-12) and Skelton (cf. 1988: 38) and different from Prince et al. (cf. 1980: 6-20).

S (1) *The internet is **definitely** the fastest-growing communication tool [...]* (CARE13FBL_24).

S (2) *[...] 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).

stronger colour=clearer **booster**
weaker colour=less **booster**

Research question 2:

What choices do students make in the usage of hedges and boosters: word order, orientation, manifestation, and prosody of modalization?

2.1 The interaction and overlap between modulation and modalization (ideational and interpersonal) within the same lexeme functioning in an area of uncertainty and negotiation.
Modulation: Hedges.

Modulation (+ideational) proposing/recommending an idea, possible plan, or action (COD, QEP, FGJSSH)

S (4) *It is **suggested** to the teachers that these habits should be avoided in the classroom and proper use of nonverbal communication be used in order to have great impact on the teaching learning process.* (CARE15MML_7)

Modalization (+interpersonal) communicating an idea or feeling without stating it directly or giving proof (COD)

S (5) *This **would perhaps suggest** that, for example, that 'Derived VC' is **actually** an extreme form [...].* (Woodhall, 2003: 15; cited in Vázquez and Giner 2009: 229)

hedge and **booster**

Research question 2:

What choices do students make in the usage of hedges and boosters: word order, orientation, manifestation, and prosody of modalization?

2.2 The non-applicability or the partial applicability of the phenomenon of transferred negation (e.g. *not clearly=vaguely*) in Albanian and Italian (in contrast to English)

S (6) *Prima della loro nascita esistevano alcune iniziative ma **non chiaramente** definibili WISE.* (CIA008FPE_4)

*Before they were created, some initiatives existed, but they were **not as clearly** [vaguely] definable as WISE.* (transl.V.DH)

S (7) *La questione, che **non riguarda certamente** solo la metodologia ma che interessa trasversalmente ogni disciplina umana perché [...].* (CIA010MPS_54)

*The matter that **certainly does not** only have to do with the methodology but it also transversely interests every human discipline because [...].* (transl.V.DH)

hedge and booster

Research question 3:

What are the quantitative differences in using hedges and boosters, and their semantic and pragmatic subcategories between the Albanian and Italian L1 and L2 corpora?

corpus	logodds	n	proportion
CAR	0.412	2445	0.536
CIAOE	-0.158	2737	0.375
CIAO	-0.254	2906	0.374

Table 3: The variable of corpus influencing the usage of hedges (and boosters).

Logistic Regression model (hedges) $\ln(B/(1-B))$ fitted for H, their training L2 writings: boost
 $\ln(\text{odds}) = \ln(B/(1-B))$, in the algorithm of the odds ratio of H, infinite +/- values, easily updated
 Albanian data, 1 appear more tentative: hedge.

Proportion: percentage of H,

N values: total values of H and B (see also Gelman and Hill 2007; Chatterjee and Ali 2006)

Exclusion of CARE and other categories, empty cells, interaction

p: probab. of occurrence

p-1: probab. of non-occurrence

Research question 3:

What are the quantitative differences in using hedges and boosters, and their semantic and pragmatic subcategories between the Albanian and Italian L1 and L2 corpora?

3.1 ~~Proud and confident Albanians, as generally known and Albanian university teachers also predicted~~ **lower vertical power role, institutional role and personal features** of students in an academic context including more 'powerful' and experienced supervisors, teachers and other experts. (see also Argyle and Henderson 1984; Matthiessen 2010; Pulcini & Furiassi 2004) = **more hedges by Albanians**

Italians: confident, sometimes even over-confident, as Italian teachers predicted.

Research question 4:

What are the quantitative differences in the usage of hedges and boosters: word order, orientation, manifestation, and prosody of modalization?

4.2 Orientation - manifestation: hedges - significantly favored in the objective-implicit combination; boosters - frequently employed in the objective-explicit one.

orient-manifest	logodds	n	proportion
object.-implicit	0.138	6320	0.437
object.-explicit	-0.138	1768	0.376

Table 4: The single variable of orientation - manifestation influencing the usage of hedges (and boosters).

Objective-**implicitness** - **tentative** nature of **hedging** itself - the low institutional and power roles of students as actors in the academic discourse.

Objective-**explicitness** - **boosters** confirms the claims of the interviewed university teachers from Albania and Italy.

Research questions 3, 4

3. What are the quantitative differences in using hedges and boosters, and their semantic and pragmatic subcategories (a. commitment) between the Albanian and Italian L1 and L2 corpora?
4. What are the quantitative differences in the usage of hedges and boosters: word order, orientation, manifestation, and prosody of modalization?

References

- Anthony, L., (2014). *AntConc* (Version 3.4.4) [Computer Software]. Waseda University. Tokyo, Japan. Retrieved from <http://www.laurenceanthony.net/>. Accessed on 10.09.17.
- Agalliu, F., Angoni, E., Demiraj, Sh., Dhrimo, A., Hysa, E., Lafe, E. & Likaj, E. (2002). *Gramatika e Gjuhës Shqipe*. 1st Edition. Tiranë: Akademia e Shkencave të Shqipërisë.
- Argyle, M., & Henderson, M. (1984). *The Anatomy of Relationships*. London: Heinemann, Harmondsworth: Penguin.
- Akbas, E. & Hardman, J. (2018). Strengthening or weakening claims in academic knowledge construction: A comparative study of hedges and boosters in postgraduate academic writing. In: *Educational Sciences: Theory & Practice*, 18. 831-859.
- Bondi, M. (2008). Emphatics in academic discourse: Integrating corpus and discourse tools in the study of cross-disciplinary variation. In Ädel, A., Reppen, R. (Eds.). *Corpora and Discourse: The Challenges of Different Settings*, 31-55.
- Chatterjee, S. & Ali, S. H. (2006). *Regression Analysis by Example*. 4th Edition. New Jersey: John Wiley and Sons, Inc.
- Gelman, A. & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.
- Halliday, M. A. K. (1985). *An Introduction to Functional Grammar*. 1st edition. London: Edward Arnold.
- Halliday, M. A. K. & Matthiessen, C. M. I. M. (2014). *An Introduction to Functional Grammar*. 4th Edition. London: Routledge.
- 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.

References

- 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.
- Kaltenböck, G., Mihatsch, W. & Schneider, S. (Eds.). *New Approaches to Hedging*. Studies in Pragmatics, 9. North America: Emerald Group Publishing Limited. 267-288.
- Lafuente Millán, E. (2008). Epistemic and approximative meaning revisited: The use of hedges, boosters and approximators when writing research indifferent disciplines. In Burgess, S. & Martín-Martín, P. (Eds.). *English as an Additional Language in Research Publication and Communication*. Bern: Peter Lang. 65-82.
- Martin, J. R. & 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 from https://www.youtube.com/watch?v=U15qHWJcfT4&feature=player_embedded. Accessed on 10.11.2017.
- Prince, E. F., Frader, J. & Bosk, C. (1980). On hedging in physician-physician discourse. Presented at the *AAAL Symposium on Applied Linguistics in Medicine*. University of Pennsylvania, Pennsylvania.
- Pulcini, V. & Furianassi, C. (2004). Spoken interaction and discourse markers in a corpus of learner English. In Haarman, L., Morley, J. & Partington, A. (Eds.). *Corpora and Discourse*. Bern: Peter Lang. Reprinted in Prat Zagrebelsky, M. T. (Ed.). (2004). *Computer Learner Corpora. Theoretical Issues and Empirical Case Studies of Italian Advanced EFL Learners' Interlanguage*. Alessandria: Edizioni Dell' Orso s.r.l.
- Quirk, R., Greenbaum, R., Leech, G. & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.

References

- Salager-Meyer, F. (1998). Language is Not a Physical Object. Françoise Salager-Meyer responds to Peter Crompton's "Hedging in academic writing: Some theoretical aspects" in *English for Specific Purposes*, 16(4). 295-302.
- Schmied, J. (2008). Hedges in specialised vs. popular academic interaction: A case study of medical texts. *Discourse and Interaction 1.2.*, Brno: Masaryk University. 85-98.
- Schmied, J. (2009). Discourse approaches to specialised and popular academic English: Comparing stance, hedges and modality. In M. Dynel (Ed.), *Advances in Discourse Approaches*. Newcastle upon Tyne: Cambridge Scholars Publishing. 76-91.
- Schmied, J. (2018). A global view on writing research articles for international journals. In Schmied, J., Hofmann, M. & Esimaje, A. (Eds.). *Academic Writing for Africa: The Journal Article*. Göttingen: Cuvillier. 1-18.
- Skelton, J. (1988). The care and maintenance of hedges. *ELT Journal*, 42(1). 37-43.
- Stefanllari, I. (2007). *English-Albanian Dictionary*. Kent: Bay Foreign Language Books. (Original work published 1999).
- Toska, B. (2015). Structural, textual and commentary linkers in Albanian phd dissertations: A Pilot Study. In Schmied, J. (Eds.). *Academic Writing for South Eastern Europe: Practical and Theoretical Perspectives*. Göttingen: Cuvillier Verlag. 163-177.
- Varttala, T. A. (2001). *Hedging in scientifically oriented discourse: exploring variation according to discipline and intended audience*. (Unpublished doctoral thesis). University of Tampere, Finland.
- Vázquez, I. & Giner, D. (2009). Writing with conviction: the use of boosters in modelling persuasion in academic discourses. *Revista Alicantina de Estudios Ingleses*, 22. 219-237.

A Corpus-based Comparison of Albanian and Italian Student Academic Writings in L1 and English as L2: Introducing Hedges and Boosters within Modalization, Modulation and Degree

Vincenzo Dheskali

Sixth Semester PhD Student

Chemnitz University of Technology

vincenzo.dheskali@s2015.tu-chemnitz.de

To appropriately convey plausible reasoning, refer to other sources, refer to our results, and use various realizations of modality and its metaphorical extension by balancing the right amount of conviction with tentativeness, it is at the heart of effective writing within the academic context. (cf. Halliday and Matthiessen 2014; Hyland 1998a; 1998b)

Those aspects should be considered by researchers focusing on hedges and the less researched boosters (Bondi 2008; Akbas and Hardman 2018) within the underexplored Albanian context (Toska 2015) and Italian context as well as university teachers in their classrooms.

Main variables and their categories

Main variables	their categories/levels
word order	fronted, thematic, neutral, medial, final (afterthought), parenthetic, in parentheses
a. commitment	proposition-related (approximator for H, intensifier for B) or author-related (shield for H, emphasizer for B)
corpus	CAR, CARE (not for Log. Regress.), CIAO, CIAOE
orient-manifest	subjective-explicit, subjective-implicit, objective-explicit, objective-implicit

Table 5: The main variables related to hedges and boosters and their different levels or categories.

current predictors	p. value
word order	6.21e-64
a. commitment	8.92e-36
corpus	4.25e-33
orient-manifest	1.80e-04

Table 6: A general comparison of the significance of all variables influencing the usage of hedges (and boosters).