

Linguists of Tomorrow

Linguists of Tomorrow:
Selected Papers from the 1st Cyprus Postgraduate Student
Conference in Theoretical and Applied Linguistics

Edited by

Kleanthes K. Grohmann,
Aljona Shelkovaya
and Dionysios Zoumpalidis

CAMBRIDGE
SCHOLARS

P U B L I S H I N G

Linguists of Tomorrow:
Selected Papers from the 1st Cyprus Postgraduate Student Conference
in Theoretical and Applied Linguistics,
Edited by Kleanthes K. Grohmann, Aljona Shelkova and Dionysios Zoumpalidis

This book first published 2012

Cambridge Scholars Publishing

12 Back Chapman Street, Newcastle upon Tyne, NE6 2XX, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Copyright © 2012 by Kleanthes K. Grohmann, Aljona Shelkova and Dionysios Zoumpalidis
and contributors

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system,
or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or
otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-4438-3876-4, ISBN (13): 978-1-4438-3876-4

TABLE OF CONTENTS

Acknowledgements	vii
Introduction	viii
Introducing <i>LoT</i> <i>Kleanthes K. Grohmann, Aljona Shelkova and Dionysios Zoumpalidis</i>	
Chapter One.....	1
Coordinate Structures in European Portuguese: Gender Variation in Written Data? <i>Joana Aguiar</i>	
Chapter Two.....	22
Are the Greek Verbs No Longer Enough? Bilingual Compound Verbs in Cypriot Greek <i>Constantina Fotiou</i>	
Chapter Three.....	51
Aspect of Embedded Verbs in L2 Acquisition: Evidence from L1- Russian Learners of Cypriot Greek <i>Sviatlana Karpava</i>	
Chapter Four.....	82
Abbreviations Dictionaries <i>Mojca Kompara</i>	
Chapter Five	105
Tracking Universals Requires Grammatical Mapping <i>Barbara Lust</i>	
Chapter Six	131
For in English Infinitives Has Nothing to Do with Case <i>Thomas McFadden</i>	

Chapter Seven.....	156
What I Say, You Say! Illustration of Syntactic Priming in Cypriot Greek	
<i>Lena Papadopoulou and Natalia Pavlou</i>	
Chapter Eight.....	186
Acquisition of /sibilant + stop/ Consonant Clusters by L1 Russian-Speaking Children	
<i>Aljona Shelkovaya</i>	
Chapter Nine.....	205
Narratives in Cypriot Greek Children with SLI	
<i>Eleni Theodorou</i>	
Chapter Ten	232
On the Functional Layer of Modified Nouns in Istro-Romanian	
<i>Iulia Zegrean</i>	
Chapter Eleven	260
The Status of Ethnic and Non-Ethnic Languages of Pontic Greeks in the North Caucasus	
<i>Dionysios Zoumpalidis</i>	
Contributors.....	280

ACKNOWLEDGEMENTS

We would like to thank first and foremost the Department of English Studies at the University of Cyprus for their financial, administrative, and organizational support without which *Linguists of Tomorrow (LoT)* would not have materialized. Thanks to UCY for hosting this event in May 2010.

On an individual level, we express our gratitude to Dr. E. Phoevos Panagiotidis, who acted as faculty liaison between the Department's Board and the student conference organizers, Aljona Shelkovaya and Dionysios Zoumpalidis. As one of the two internal speakers (the other was Kleanthes K. Grohmann), we would also like to thank him for his keynote address at *LoT*. The organizers are extremely pleased that they could win Barbara Lust, Thomas McFadden, and Peter L. Patrick as external invited speakers. We thank all of them for coming to Cyprus, engaging in lively discussions with the student presenters throughout *LoT*, and of course delivering their excellent lectures. We are further pleased that the first two also contribute their presentations to this volume, which makes it even stronger.

In addition, we thank all student participants of *LoT*: all the presenters, especially those who could not contribute to this volume, the audience, which was testimony to the overall success of a well attended conference, and the student assistants who helped making the event a full success on all levels. Thank you also to all contributors to the volume, for their cooperation and often tested patience. Last but not least, a big shout of thanks to Thekla Constantinou Nearchou, the Department's heart and soul in her capacity of administrative assistant who dealt with everything we couldn't, wouldn't, and shouldn't!

Two years later, the laborious efforts of putting this volume together have finally culminated in this collection. There was a "minor" glitch in the process which delayed the enterprise by a full year, but once we got on board with Cambridge Scholars Publishing, things went very smoothly. It was our tremendous pleasure to work together so closely and well — and in the end, so incredibly quickly and efficiently — with the CSP team. Thank you very much then for all your work, Carol Koulikourdi, Amanda Millar, and Emily Surrey.

We dedicate this volume to the memory of Dr. Pavlos Pavlou, former colleague (Grohmann), supervisor (Zoumpalidis), teacher (Shelkovaya), and much more, who passed away too early.

INTRODUCTION

INTRODUCING *LoT*

KLEANTHES K. GROHMANN,
ALJONA SHELKOVAYA
AND DIONYSIOS ZOUMPALIDIS

We are very happy to present this volume as testimony of a very fruitful and highly successful two-day postgraduate student conference organized at the University of Cyprus in Nicosia, 7–8 May 2010. The full title of the event was *Linguists of Tomorrow: The 1st Cyprus Postgraduate Student Conference in Theoretical and Applied Linguistics* — or *LoT* for short. Our expressed hope is that the numerical addition was not just a token of idealistic optimism, but that the *LoT* concept will indeed be picked up in subsequent years by future post-graduate students in the Department. This said, with the two student organizers and editors slowly but surely on their way out, it really would have to be organized by the “next generation”.

This present collection is a selection of written up presentations from *LoT*. The conference program featured a total of 24 talks by postgraduate students from 14 different countries plus three invited presentations from abroad (Barbara Lust from Cornell, Thomas McFadden from Tromsø, and Peter L. Patrick from Essex) and two invited faculty presentations from the UCY Department of English Studies (E. Phoevos Panagiotidis and Kleantes K. Grohmann). The three invited presentations not included here dealt with ‘The linguistic rights of asylum speakers’ (Patrick), ‘Greek verb-making morphology’ (Panagiotidis), and ‘The Gen-CHILD project: A first view from the acquisition of clitics in Cypriot Greek’ (Grohmann).

The nine student and two faculty papers which are included here touch on different topics in theoretical and applied linguistics. We decided to arrange them in alphabetical order for this volume, but for the purposes of our overview, we decided to group them thematically. There is, however, considerable overlap — is syntactic processing part of syntax or psycholinguistics, for example, or should psycholinguistics itself be classified on

the “theoretical” or the “applied” side of linguistics? In any case, one way of cutting the *LoT* pie could be along the following lines.

Language acquisition and development is a perfect interface between theoretical and applied linguistics. One of the *LoT* keynote speakers, **Barbara Lust**, made this clear in her lecture, written up here as ‘Tracking Universals Requires Grammatical Mapping’ (Chapter Five). Arguing that different types of evidence from language acquisition foreshadow the existence of linguistic universals, she discusses both early and continuous access of functional categories and ‘structure-dependence’ constraining anaphora as well as continuous sensitivity to general structural properties of the grammar acquired. The evidence reveals continuous constraint on children’s early linguistic representations which is consistent with versions of principles and parameters hypothesized in a linguistic theory of Universal Grammar as well as several typological universals. She sketches motivation for a theoretical paradigm, Grammatical Mapping, articulating ways in which it differs from other current approaches to a comprehensive theory of language acquisition. Lust provides several examples of empirical research instantiating test of three essential hypotheses deriving from this paradigm. Finally, she considers implications for a theory of Universal Grammar as a model of the faculty of language raised by this approach to a comprehensive theory of language acquisition.

Three other chapters deal directly with language acquisition and development. Concerning first language acquisition, **Aljona Shelkova** contributes a paper entitled ‘Acquisition of /sibilant + stop/ Consonant Clusters by L1 Russian-Speaking Children’ (Chapter Eight). In this study, she investigates the influence of first language phonotaxis on the reduction patterns of /sibilant + stop/ clusters in L1 Russian-speaking children. The emphasis of the data analysis is children’s reduction patterns of some clusters of falling sonority as compared with those of /s/ clusters.

Continuing the theme of first language acquisition, but considering language-impaired populations, **Eleni Theodorou** investigates ‘Narratives in Cypriot Greek Children with SLI’ (Chapter Nine), providing a first description of the narrative ability of Greek Cypriot children with Specific Language Impairment and compared with typically developing children. She aims to identify special areas of difficulty in respect to narration for language-impaired children.

Moving on to second languages, **Sviatlana Karpava** contributes her work on ‘Aspect of Embedded Verbs in L2 Acquisition: Evidence from L1-Russian Learners of Cypriot Greek’ (Chapter Three). This paper aims to contribute to the investigation of child and adult second language acquisition, the role of UG, and the possible cross-linguistic transfer in

L2A. It focuses on the acquisition of embedded aspect in Cypriot Greek by learners from different L1 backgrounds (Russian and Georgian), which are different from the CG aspectual system.

The above-mentioned dual status of psycholinguistics comes out in the next chapter, which involves psycholinguistic investigations of (theoretical aspects of) syntax and (first language) acquisition. In ‘What I Say, You Say! Illustration of Syntactic Priming in Cypriot Greek’ (Chapter Seven), **Lena Papadopoulou and Natalia Pavlou** investigate error patterns in children’s *wh*-question production as outlined in two experiments, the Syntactic Priming Experiment in Cypriot Greek from the first author’s Ph.D. dissertation and the “Guess What Game” from the second author’s M.R.A. thesis.

Two chapters deal with syntax proper. **Thomas McFadden**, the other keynote speaker contributing to this volume, claims that ‘*For* in English Infinitives Has Nothing to Do with Case’ (Chapter Six). He argues that the distribution of *for* can be more accurately described in terms of restrictions on overt and covert complementizers, which shows a similar distribution. He presents evidence that these restrictions must be characterized at least partly in prosodic terms. This chapter is a first attempt at making these precise and motivating them. McFadden then proposes that the differences of detail in the restrictions on the two complementizers can be accounted for if central aspects a finer left periphery in English are adopted, but at the same time, he argues against certain peripheral aspects of his account which depend on Case theory.

Iulia Zegrean contributes a host of novel and intricate data ‘On the Functional Layer of Modified Nouns in Istro-Romanian’ (Chapter Ten). In this paper, she aims to provide an accurate description of data concerning definiteness and modified nouns in an underdescribed dialect. Doing so, she proposes a tentative analysis following recent approaches to nominal expressions within generative studies. This is a first attempt to formalize an aspect of the nominal grammar of Istro-Romanian..

‘Are the Greek Verbs no Longer Enough? Bilingual Compound Verbs in Cypriot Greek (Chapter One) is the title of **Constantina Fotiou**’s paper. Greek Cypriots purportedly use English loans widely, and even code-switch. The data of this study come from natural speech recordings, enriched by non-recorded speech from participant observation, focusing on bilingual compound verbs.

In ‘The Status of Ethnic and Non-Ethnic Languages of Pontic Greeks in the North Caucasus’ (Chapter Eleven), **Dionysios Zoumpalidis** examines what languages are employed by Pontic Greeks residing in the Stavropol region, more specifically, in the village of Kyrpichnyi which is

located outside the town of Essentuki. The aim of this study is to identify the current position of the ethnic and non-ethnic languages within the Pontic Greek community, and to examine the reasons why and how each language obtained a low or high status in the community in question.

Continuing such ethno- or sociolinguistic concerns, ‘Coordinate Structures in European Portuguese: Gender Variation in Written Data?’ (Chapter One) is the question **Joana Aguiar** asks. The objective of this paper is to observe the frequency of prototypical coordinate structures in written texts and how these may vary according to the writer’s gender and education.

Applying linguistic insights to dictionary studies, **Mojca Kompara**’s paper deals with ‘Abbreviations Dictionaries’ (Chapter Four) against the background of abbreviations being a growing phenomenon which is dealt mainly by Slovene orthographic dictionaries. The aim is to present how abbreviations are included in foreign monolingual dictionaries, Slovene monolingual and bilingual dictionaries, specialized abbreviations’ dictionaries, online dictionaries, and data bases.

And without further ado, we’re herewith happy to present the volume as a whole.

CHAPTER ONE

COORDINATE STRUCTURES IN EUROPEAN PORTUGUESE: GENDER VARIATION IN WRITTEN DATA?^{*}

JOANA AGUIAR

1. Introduction

The objective of this paper is to observe the frequency of prototypical coordinate structures in written texts and how these may vary according to the writer's gender and, to a lesser extent, education.

In recent years the sociolinguistic analysis has been applied to other fields besides phonology, such as morphology and syntax (Winford, 1996; Kroch, 2001; Henry, 2002; Mondorf, 2002; Pintzuk, 2003; Macaulay, 2005; Kortmann, 2006; Cheshire, 2005; among others). Some of the most studied morphosyntactic phenomena include multiple negation; was/were variation; use or lack of third person singular form; inadequate use of present forms instead of Past Simple tense; inadequate use of *done* instead of *do*; and use of relative pronoun *what* (Cheshire, 2009).

Considering syntax as the linguistic object of sociolinguistic analysis is not pacific though. Some of the methodological and conceptual obstacles mentioned (Winford, 1996; Cheshire et al., 2005; Cheshire, 2009) are: the

^{*} This work has been made possible thanks to the financial support of a PhD grant from Fundação para a Ciência e Tecnologia (PhD scholarship SFRH-BD-64349-2009). I take this opportunity to thank my advisor, Prof. Doutora Pilar Barbosa, and my co-advisor, Prof. Doutora Conceição Paiva, for their helpful observations and scientific guidance. I would like to thank the audience at *Linguists of Tomorrow* (University of Cyprus, 2010) for their fruitful comments and questions and to the reviewers of this paper for the pertinent observations and corrections. Further investigation concerning this subject was also presented at *19th International Postgraduate Linguistics Conference* (University of Manchester, 2010). Any remaining errors are my responsibility.

inclusion of a sociolinguistic approach in a syntactic theory, the constitution and definition of the variables and their semantic equivalence, the scarce frequency of occurrence of some phenomena in the generalization of the results, and the interference of individual discourse style.¹

Although coordinate structures are a well-studied topic in European Portuguese (Peres, 1997; Colaço, 2005; Matos, 2005; Peres & Mascarenhas, 2006; Chaves, 2007; Mória, 2008; among many others), there is still a lack of sociolinguistic investigation focusing the frequency and distribution of these structures. Nevertheless, considering that gender differences were found in the frequency of coordinate conjunctions in oral samples in English (Macaulay, 2005), the research questions that conducted this paper are the following (i) can the same differences be found in written samples? And in other language? Furthermore, (ii) is gender encoded in terms of syntactic choices or preferences? Or can other variables, such as education, clarify the (possible) variation observed in written samples?

In order to discuss the role of gender in the occurrence of coordinate structures, this paper is organized as follows. Section 2 will review some background to some aspects of the variables ‘gender’ and ‘education’ in a number of processes of variation and change; and in section 3 a small introduction will be given to coordinate structures in European Portuguese (henceforth EP). In section 4 the methodology used will be explained. The results and discussion will be presented in sections 5 and 6, respectively. Section 7 summarizes the main outcomes.

2. Some background

2.1. Gender

Gender has been always considered an important variable (Eckert, 1989; Labov, 1972, 1990, 2000; Chambers, 2003:ch.3; Coates, 2004; Holmes & Meyerhoff, 2004; Cheshire, 2002, 2005; among many others) and many studies reveal that men and women do not express themselves in the same way. Besides the distinct lexical choices - for example, the range of lexical items to distinguish colors is broader in women’s discourse, as well as the use of *empty* adjectives (Lakoff, 1973), there seem to be many linguistic strategies used differently by men and women: women tend to integrate more apologies, and empathy expressions, such as *you know* or *sort of*

¹ The influence of the individual style in sociolinguistic analysis is explored by Tannen (1984), Biber & Finegan (1994), Rickford & Eckert (2001), Schilling-Estes (2002), among many others.

(Holmes, 1995); and girls are likely to integrate more personal and affective information in the discourse, whereas boys tend to interpret the questions literally, giving factual answers (Cheshire & Williams, 2002; Cheshire, 2005). Also Mulac (1998) refers that men use more references to quality, as well as judgmental adjectives and elliptical sentences when describing a picture, whereas women use more intensive adverbs, references to emotions, uncertainty verbs, such as *see* and sentence initial adverbs. Moreover, there is also evidence from syntax: regarding the position of adverbs in English and in oral samples, Mondorf (2002) concluded that men tend to place adverbial clauses in the beginning of the sentence whereas women place them in the end. Considering the use of pragmatic expressions in oral conversations, according to Erman (1992), women use the expressions *you know*, or *you see* to establish connections between consecutive arguments more often than men do. Nevertheless, these results are dependent on whether the informants are in same-sex or mixed-sex conversation groups.

The differences are also visible in writing, as pointed out by several studies (Aries & Johnson, 1983; Tannen, 1990; Herring & Paolillo, 2006): men tend to write about facts and objects and women prefer to write about feelings and relations. Also, according to Johnstone (1993), in storytelling men tend to convey more details about time, place and objects, whereas women use more personal names and reported speech (*ibidem*, 72). Moreover, according to Biber et al., 1998; Palander-Collin, 1999; Argamon et al., 2003; Cheshire, 2005; among others, texts written by male informants seem to be more *informational* (Biber, 1995: 141-151), i.e., they present more determiners, prepositional phrases as post-nominal modifiers, and cardinal numbers while the texts written by female informants contain more *involvedness features* (*idem, ibidem*), presenting a higher frequency of singular person pronouns. Nevertheless, as stated in Argamon et al. (2003: 11), the values for *involvedness* and *informational* features vary according to the gender (fiction/non-fiction) of the text: pronouns are more frequent in fiction texts and determiners in non-fiction texts. Also, a study based on weblogs entries (Herring & Paolillo, 2006) reveals that traditional *female stylistic features* are more common in personal journal entries and traditional *male stylistic features* in filter entries (on politics mainly), notwithstanding the gender of the writer, which demonstrates the paramount importance of genre rather than gender in written texts.

In terms of syntactic structures, Jespersen (1922) considers that male and female speech present considerable differences in terms of hypotactic and paratactic structures employed, stating that men prefer long complex

structures and women coordinate structures, comparing the first with Chinese boxes and the later with a string of pearls.² Notwithstanding the importance of Jespersen's conclusions, it is important to emphasize that at the time women's access to culture, education and market place was very restricted, which may have misled the observations, as the author mentions. Furthermore, more recent studies (Trudgill, 1974; Labov, 1990; Cheshire, 2002) reveal that social status and local identity are variables that influence gender variation. Also, Holmes (1998) highlights the importance of the context. According to her, men have better performances in public contexts and women in informal contexts as men search for social enhancement and women for the maintenance of relations.

Gender and age differences were also found in the distribution of coordination structures in oral samples in English (Macaulay, 2005: 88-90): the percentage of occurrence of these constructions is higher in adult women and middle-class boys. Also, the percentage of occurrence of coordinate structures is different in adolescent/adult and men/women speech: adults use more *but*, and the values for *and* and *so* are higher in men when compared to women's. Bearing in mind that Macaulay's study was carried out in oral data and considering that written and oral registers may have distinct functional and linguistic features (Biber, 1995: 412-413), an analysis of written texts may reveal a different organization of ideas with more recurrence of complex structures (Biber, 1988, 1995:311) and probably a different distribution of coordinate structures.

² The full quote is provided below:

If we compare long periods as constructed by men and by women, we shall in the former find many more instances of intricate or involute structures with clause within clause, a relative clause in the middle of a conditional clause or vice versa, with subordination and sub-subordination, while the typical form of long, feminine periods is that of co-ordination, one sentence or clause being added to another on the same plane and the gradation between the respective ideas being marked not grammatically, but emotionally, by stress and intonation, and in writing by underlining. In learned terminology we may say that men are fond of hypotaxis and women of parataxis. Or we may use the simile that a male period is often like a set of Chinese boxes, one within another, while a feminine period is like a set of pearls joined together on a string of *ands* and similar words. (Jespersen 1922: 251-252)

2.2. Education

As mentioned in the previous section, informant's access to formal education may influence the conclusions drawn when observing certain linguistic variables. Considering that the education level is very often related to the career opportunities, education as a social factor is recurrently analyzed along with social status. For instance, in Eckert's study on phonologic variation in Belten High (2000) the level of education of the informants' parents is correlated with their profession/occupation and, consequently, their social status. In fact, according to Labov (1972: 112-115), the *socioeconomic index* should be calculated considering three aspects: education, profession and family income, and for Wolfram & Fasold (1997: 94-98) and Trudgill (1974) the level of education is one of five aspects which should be considered if we intend to observe the variable social stratification.

Focusing on syntax, previous works on oral data considering Basil Bernstein's definition of *public language*,³ namely Macaulay (1991, 2005), showed that middle-class informants tend to use a broader variety and combination of constructions. The analysis also showed that this tendency was not very significant, except in the use of non-restrictive relative clauses (more frequent in middle-class speakers) and the use of more discourse markers and *dislocated syntax*⁴ (more frequent in working-class informants).

Considering education apart from social status, Scherre & Naro (1992), in an analysis of noun/verb agreement in oral samples from Brazilian Portuguese spoken in the city of Rio de Janeiro, found that the level of education of the informants is directly proportional to the percentage of noun/verb agreement, *i.e.*, the informants with lower levels of education tend to have lower percentages of noun/verb agreement when compared with the results of the informants with higher levels of education. Nevertheless, the authors also point out the importance of the linguistic variables, such as the "discourse and clausal level parallel processing effects" (*ibidem*).

³ Bernstein (1971: 42) defines *public language* as containing: "(1) short, grammatically simple, often unfinished sentences, a poor syntactical construction with a verbal form stressing the active mood; (2) simple and repetitive use of conjunctions (*so, then, and, because*)", among others linguistic features.

⁴ Dislocated syntax constructions are "constructions that have a highlighting or intensifying effect", such as demonstrative focusing, clefting, noun phrase prepositions, left dislocation and right dislocation (Macaulay, 2005:93).

3. Coordinate Structures

In European Portuguese the structures of coordination, their constitution and boundaries have long been topic for discussion (Peres, 1997; Colaço, 2005; Matos, 2005; Peres & Mascarenhas, 2006; Chaves, 2007; Mória, 2008 among many others).

Like in other European languages, such as English, French or Spanish, coordinating constructions in Portuguese may be structured with coordinators or with no overt conjunctions, also referred as *asyndeton*. In Portuguese the most frequent coordinating conjunctions, therefore known as *prototypical* or *canonical*, are: *e* ‘and’ to establish additive connections; *mas* ‘but’ to establish contrast/adversative connections; and *ou* ‘or’ to establish alternative connections. Although these coordinating conjunctions, especially *e*, are used to establish other semantic nexus, in this paper only the prototypical coordinating constructions will be analyzed, as exemplified in the examples (i), (ii) and (iii):

(i) additive connection with *e*:

- (1) *O homem ideal não ousa sair do etéreo mundo dos livros e ainda bem que não o faz.*

‘The ideal man doesn’t dare to get out of the ethereal world of the books **and** fortunately he doesn’t.’

(Source: <http://avenidacentral.blogspot.com/2009/04/capitalismo-social.html>)

(ii) contrast connection with *mas*:

- (2) *O problema não reside na avaliação em si, porque deve ela mesma existir, **mas** antes na sua desvirtuação.*

‘The problem is not the evaluation itself, because it should exist, **but** rather its distortion.’

(Source: <http://oprofano.blogspot.com/2009/04/avaliacoes-epoca-de-amiguismos-e-afins.html>)

(iii) alternative connection with *ou*:

- (3) *Poderia dizer que a culpa é do ensino pois as aulas não são apelativas o suficiente para prender atenção do aluno **ou** então porque o professor não sabe fazer valer a sua autoridade.*

‘I could say that the responsibility is of the education system because the classes are not appealing to the students **or** rather because the teacher doesn’t know how to exercise his/her authority.’

(Source: <http://kulcinskaia.blogs.sapo.pt/169126.html>)

Results on frequency data considering the type of coordinative conjunction can be found in Schiffrin (1987) and Macaulay (2005: 39-42) for English, Ohori (2004) for English and German, and will be explored in section 7.

Due to time constraints, the syntactic representation of the coordinate structures and the status of the coordinands will not be discussed in this paper.

4. Methodology

In order to analyze the distribution of the coordinate structures in written samples, 24 texts from 12 informants were collected. The informants are stratified according to gender and education (12th grade and Bachelor/Master holders), as shown in Table 1-1. All the informants are European Portuguese native speakers between 20-45 years old.

Table 1-1 Number of informants and texts according to education level and gender

Education	12 th Grade		Bachelor/Master		Total
	<i>Female</i>	<i>Male</i>	<i>Female</i>	<i>Male</i>	
Gender					
Informants	3	3	3	3	12
No. of texts	6	6	6	6	24

As mentioned in section 2.1, several studies (Aries & Johnson, 1983; Tannen, 1990; Herring & Paolillo, 2006) indicate that women’s texts tend to be mainly about relations and men’s texts about objects. To overcome this possible bias, only argumentative texts were analyzed. The themes are roughly the same: civil rights, government decisions, justice and laws, education and health. The texts were taken from blogs, randomly visited. In order to identify the gender, education level and age of the writer, the author of the text was contacted. We also requested the authorization to divulge the texts for academic purposes.

Besides controlling the social variables and the type of text, only texts between 400-500 words were selected. In this paper, only the occurrence of prototypical clausal coordination will be under analysis.

In the total, 315 coordinate structures were encoded considering gender and education of the informant and the following factors: (i) Logical-Semantic relation: addition, alternative, contrast; (ii) Conjunction: *e and*, *mas but*, *ou or*, asyndeton; (iii) Level of coordination: main-clause coordination⁵ or lower level coordination⁶ (Huddleston & Pullum, 2006: 204).

In relation to statistical analysis, all the structures were encoded and ran in statistical software SPSS 17.0 (Statistical Package for the Social Sciences). The frequencies and crosstabs results were obtained with descriptive statistics. Also, the Chi-Square test ($p \leq 0.05$) was used to confirm the significance of the results.

5. Results

In this section the frequency results of the prototypical coordination in argumentative texts that constitute our *corpus* will be presented, as well as the results of the distribution of this type of construction according to gender and education.

5.1. Coordinate Structures

Results indicate that *e* ‘and’ is the most common used coordinate conjunction (45.7%), followed by the non-overt conjunction (26.7%) and *mas* ‘but’ (14.6%).

Results also show that the conjunctions *e*, *mas*, *ou* and the asyndeton occur in 299 out of the 315 structures connected, which constitutes almost 95% of all the occurrences. Due to the lower values for the other conjunctions, only the structures linked by *e*, *ou*, *mas* and asyndeton will be considered in the results below presented and in the subsequent discussion (cf. table 1-3).

Looking at semantic relations, the results are very similar to the values for conjunctions, as expected. Thus, addition nexus represent 74.3% of all

⁵ *Começa nos gabinetes e propaga-se pelas escolas.*
 ‘It starts in the offices and it spreads to the schools.’
 (Source: avenidacentral.blogspot.com/2008/05/avenida-marginal-chatinho.html).

⁶ *Admito que não pensei muito e que provavelmente vai haver quem encontre uma razão mais plausível do que a minha.*
 ‘I admit I haven’t thought about it much and that there will be someone who will find a more plausible reason than mine.’
 (Source: kulcinskaia.blogs.sapo.pt/169126.html).

structures. Contrast and alternative relations together represent only 25.7% (cf. table 1-4).

Table 1-2 Frequency and percentage of conjunctions

Conjunctions	Frequency	%
<i>e</i> ‘and’	144	45.7%
asyndeton	84	26.7%
<i>mas</i> ‘but’	46	14.6%
<i>ou</i> ‘or’	25	7.9%
other/correlative pairs ⁷	16	5.1%
Total	315	100.0%

⁷ Examples of ‘other conjunctions’ include *e/ou* ‘and/or’, as in:

Os tais simpatizantes da causa que não se assumem como tal e/ou que não se querem comprometer com o “partido”.

‘The so-called sympathizers of the cause that don’t want to stand up for the party.’

(Source: <http://oprofano.blogspot.com/2009/07/muito-ou-pouco-independentes.html>)

Examples of correlative pairs found in our *corpus* include:

(i) *não... nem* ‘neither... nor’:

Não são preconceituosos, nem coisa que pareça

‘They are neither racist, nor anything alike.’

(Source: <http://avenidacentral.blogspot.com/2008/10/casamento-subprime-reservo-me-partir-de.html>)

(ii) *ou... ou* ‘either... or’:

Assim, ou declaram solenemente que não votam (...) ou encolhem os ombros e dizem “o que se há-de fazer?!”

‘Thus, or they solemnly declare that they don’t vote (...) or they shrug the shoulders and say “what can we do?!”’

(Source: <http://anacamarra.blogspot.com/2008/10/povo-programado.html>)

(iii) *não só... mas também* ‘not only... but also’:

não só por serem mais (...) mas também porque têm hoje em dia facilidades que não existiam há 20 ou 30 anos atrás.

‘not only because they are more, but also because today they have ways that didn’t exist 20 or 30 years ago.’

(Source: <http://kulcinskaia.blogs.sapo.pt/169126.html>)

Table 1-3 Frequency and percentage of *and*, *but*, *or* and asyndeton

Conjunctions	Frequency	%
<i>e</i> 'and'	144	48.2
asyndeton	84	28.1
<i>mas</i> 'but'	46	15.4
<i>ou</i> 'or'	25	8.4
Total	299	100.0

Table 1-4 Frequency and percentage of the semantic relations

Semantic Relation	Frequency	%
addition	222	74.3
contrast	47	15.7
alternative	30	10.0
Total	299	100.0

In relation to the level of coordination (cf. table 1-5), 63.9% of the structures are coordinated within the structure of a main clause (iv), and only 36.1% are main clause coordination (v).

(iv) lower-level coordination:

- (4) *Há um elefante no meio da sala que muitas pessoas do Sim teimam em ignorar e que muitas pessoas do Não querem transformar numa espécie de fim da civilização ocidental.*
 'There is an elephant in the middle of the room that many people voting Yes pretend to ignore and that many people voting No want to transform in the end of the western civilization.'
 (Source: <http://melcomcicuta.blogs.sapo.pt/316056.html>)

(v) main clause coordination:

- (5) *Em Portugal há poucos médicos portugueses mas há milhares de médicos e enfermeiros que nós contratamos.*
 'In Portugal, there are few Portuguese doctors but there are thousands of doctors and nurses that we hire.'
 (Source: <http://tpm.blogs.sapo.pt/157901.html>)

Table 1-5 Frequency and percentage of level of coordinate structures

Level	Frequency	%
lower level	191	63.9
main clause	108	36.1
Total	299	100.0

5.2. Coordinate Structures and Gender

The association gender with semantic relation (cf. table 1-6) reveals that addition connections are more frequent in the women's texts (79%) than in the men's (67.5%). On the other hand, alternative connections occur more often in the men's texts (15%) than in the women's (7%).

Table 1-6 Association gender/semantic relation

		Logical-Semantic relation			Total
		addition	alternative	contrast	
Gender	Female	139	12	25	176
		79.0%	6.8%	14.2%	100%
	Male	83	18	22	123
		67.5%	14.6%	17.9%	100%
		222	30	47	299
Total		74.2%	10.0%	15.7%	100%

In terms of correlation with the type of conjunction and the gender of the informant, the intersection of gender and type of conjunction reveals significant differences ($\chi^2(3)=20.449$, $p=.000$), namely in the occurrence of *ou* which occurs 4% in the women's texts and 14.6% in the men's texts; and in the occurrence of asyndetic coordination (36% in the women's texts and 17% in the men's). On the other hand, the correlation of gender with level of coordinate structures revealed no significant differences ($\chi^2(1)=0.20$, $p=.882$).

5.3. Coordinate Structures and Education

The correlation of education level and the semantic relation shows that the percentage of additive connections is lower (68.0% versus 82.3%) in the informants with 12th grade. On the other hand, alternative and contrast connections are inferior in the informants with Bachelor/Master level (8.5% and 9.2% versus 11.2% and 20.7%, respectively).

Table 1-7 Association of the education level and the semantic relation ($\chi^2(2)=8.739$, $p=.013$)

		Logical-Semantic relation			Total
		addition	alternative	contrast	
Education	12 th grade	115	19	35	169
		68.0%	11.2%	20.7%	100%
	Bachelor/Master	107	11	12	130
		82.3%	8.5%	9.2%	100%
	Total	222	30	47	299
		74.2%	10.0%	15.7%	100%

These results are similar when education level with type of conjunction are intersected (cf. table 1-8). The most significant results can be seen in the distribution of the conjunction *e* ‘and’ – 40.8% in the 12th grade informants and 57.7% in the informants with a Bachelor/Master- and in the distribution of asyndetic coordination – 31.4% in the 12th grade informants and 23.8% in the informants with a Bachelor/Master.

Table 1-8 Crosstab education *type of conjunction ($\chi^2(3)=10.153$, $p=.017$)

		Conjunction				Total
		<i>e</i>	<i>ou</i>	<i>mas</i>	asyndeton	
Education	12 th grade	69	14	33	53	169
		40.8%	8.3%	19.5%	31.4%	100%
	Bachelor/Master	75	11	13	31	130
		57.7%	8.5%	10.0%	23.8%	100%
	Total	144	25	46	84	299
		48.2%	8.4%	15.4%	28.1%	100%

The intersection education with level of coordinate structures revealed no significant differences ($\chi^2(1)=3.734$, $p=.053$).

6. Discussion

In the previous section, the results obtained for the distribution of type of conjunction, semantic relation, and level of coordination according to the gender and education of the informants were presented. In this section we will discuss the results considering some results for English and German languages and the double crosstabs of the variables.

The ratio of *e* to *ou* in our *corpus* is very similar to the results stated in Ohori (2004: 61-62) for English but distant from the values stated in Ohori (2004: 61-62) for German and in Schiffrin (1987: 128) for English (cf. Table 1-9). On the other hand the result for the ratio *e* to *mas* in our *corpus* is very similar to the values stated in Schiffrin (1987: 128) and Macaulay (2006: 40) for English, as seen in Table 1-10.

Table 1-9 Ratio of *e* to *mas* based on our corpus, Schiffrin (1987) and Ohori (2004)

Ratio <i>e</i> : <i>ou</i> ('and :or')			
European Portuguese	English		German
5.8 : 1	18.9 : 1 Schiffrin (1987)	6.2 : 1 Ohori (2004)	17.1 : 1 Ohori (2004)

Table 1-10 Ratio of *but* to *and* based on our corpus, Schiffrin (1987) and Macaulay (2006)

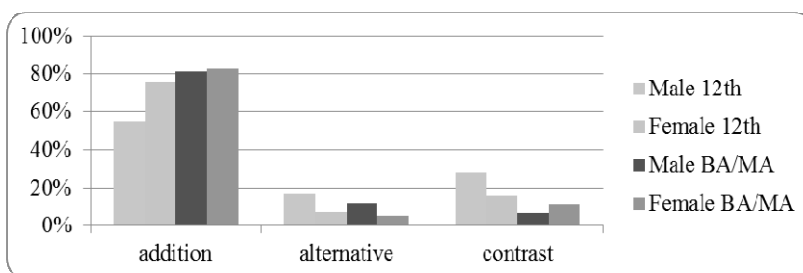
Ratio <i>e</i> : <i>mas</i> ('and : but')		
European Portuguese	English	
3.1 : 1	2.3 : 1 Schiffrin (1987)	3 : 1 Macaulay (2006)

These differences may be due to the type of text analyzed and to the methodology used to calculate the frequency of conjunctions in the *corpora*. First, both in our *corpus* and in Schiffrin's analysis only constructions of clausal coordination were considered. Nevertheless, only *and* conjunctions with the canonical semantic-value of addition were selected. For that reason, no time, condition or cause nexus coordinated by *and* were included in our analysis, aspect that can explain the lower values for the ratio of *and* to *or* in our *corpus* when compared to Schiffrin's (1987). Second, the German *data* (Ohori, 2004: 41-66) consist of ten storytelling texts collected by the author, and the study in Schiffrin (1987) was based on sociolinguistic interviews, register which seems to increase the frequency of the conjunction *and*. Also, Halliday (1987) considers that *grammatical intricacy*, i.e., the quantity of clauses in a sentence, is more common in oral language, whereas in written language we find more *lexical density*. Third, the lower values for the ratio of *and* to *or* in English (Ohori, *ibidem*) and in Portuguese may be related to the text type. For European Portuguese only argumentative texts were analyzed and Ohori (*ibidem*) collected three formal conversations from the *Corpus of Spoken*

Professional American English. Thus, because some of samples of this *corpus* were composed by faculty council meetings and White House press conferences, the informants may have used other linking devices rather than *and*. Finally, the ratio of *but* to *and* is very similar in our *corpus*, in Schifffrin (1987) and in Macaulay (2006), even though Macaulay's and Schifffrin's values are based in oral samples.

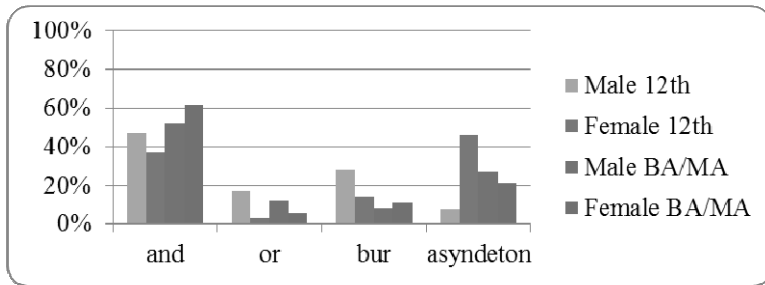
In relation to the external variables considered, i.e., gender and education, Graph 1-1 demonstrates the importance of education of the informants rather than the gender (cf. values for men ($\chi^2(1) = 11,650$, $p = .003$)) in the distribution of addition, alternative and contrast values. Thus, it is visible the contrast in addition and contrast values in Bachelor/Master informants and 12th grade informants. The lower values for alternative and contrast nexus in the informants with a Bachelor/Master degree may be related to the possibility of expressing these semantic relations through subordinate structures, namely concessive clauses, aspect that will not be explored in this paper.

Graph 1-1 Semantic nexus according to gender and education



In terms of the distribution of the type of conjunction, the disparity of the attested values (either by gender or education), as seen in Graph 1-2, raises the hypothesis that the use of these conjunctions in adult informants may be independent of the gender and education level, but rely on the individual style of each writer.

Graph 1-2 Distribution of the type of conjunction according to gender and education



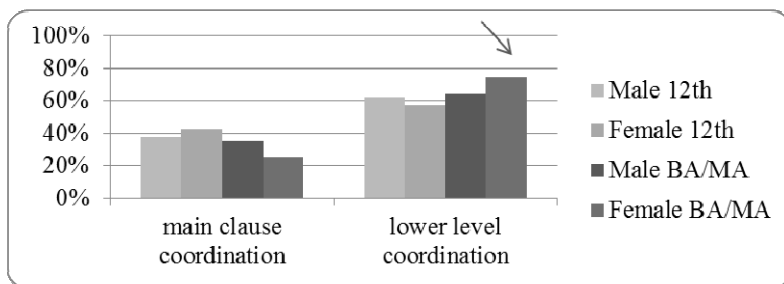
Our results for the distribution of the conjunctions according to gender contrast with the conclusions presented by Macaulay (2005: 91). According to this study, women tend to use *and* more often than men in oral samples. Also in Koppel et al. (2002) *and* is considered to be preferred by women either in fiction or non-fiction texts. Nevertheless, as stated by some authors (namely Cosme, 2008; Moder & Martinovic-Zic, 2004; Hasselgard et al., 2002; among others), the organization of the discourse, even in terms of frequency of clause structures (Chuquet & Paillard, 1987; Doherty, 2005; Fabricius-Hansen et al., 2005), is specific of each language, which can also explain the differences observed in our *corpus* and in the conclusions of Macaulay (2005) and Koppel et al. (2002), as well as the disparity of values for the ratio of conjunctions when comparing Portuguese, English and German (cf. Table 1-9 and 1-10).

Regarding the level of coordination (cf. Graph 1-3), it is visible a preponderance of lower level coordination in the texts of female informants with 12th grade, situation that contrasts with lower level coordination values in graduate women.

This may reveal that education seems to have a stronger influence on female informants as it increases the complexity of the coordinate constructions. Looking at male informants values, education is not an important factor, since the values are very similar. Similar results were documented by Scherre (1998: 241-243) in a work on nominal agreement in Brazilian Portuguese. According to her, when observing the data from male and female informants according to their education two distinct patterns arise: men's results are similar notwithstanding their level of

education⁸ and the percentage of agreement in women's data increases according to their level of education.

Graph 1-3 Distribution of the level of coordination according to gender and education



7. Conclusions

Taking into account the canonical coordinate structures in European Portuguese in argumentative written texts, we analyzed: (i) the percentage of occurrence of the canonical coordinate constructions; (ii) the frequency of occurrence of addition, contrast and alternative semantic nexus; and (iii) the level of coordination. Also, these results were crossed with extralinguistic factors such as education and gender of the informants.

The results indicate that despite significant differences in the male vs. female results, these differences are due to the level of education, rather than gender. Also, syntactic factors may interfere with the distribution of the semantic nexus as lower percentage of occurrence of alternative and contrast nexus in the informants with a Bachelor/Master degree may be explained by possibility of expressing these relations using subordinate structures, rather than coordinate ones.

Finally, the different values attested for conjunction types may be related to the individual style and the text typology as the cross-linguistic evidences illustrated.

In the future, it is our intention to: (i) widen the study to other levels of education and age intervals, and (ii) observe the distribution of other mechanisms of clausal connection according to the semantic nexus conveyed.

⁸ A further analysis indicates that men's results are influenced by the factor occupation (Scherre, 1998: 244-245).

References

- Argamon, S., M. Koppel, J. Fine & A.R. Shimoni (2003). Gender, genre, and writing style in formal written Texts. *Text* 23: 321-346.
- Aries, E.J. & Johnson, F.L. (1983). Close friendship in adulthood: Conversational content between same-sex friends. *Sex Roles* 9: 1183-1196.
- Bernstein, B. (1971). *Class, Codes and Control*, vol. 1. London: Routledge & Kegan Paul.
- Biber, D. (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber, D. (1995). *Dimensions of Register Variation: A Cross-linguistic Comparison*. Cambridge: Cambridge University Press.
- Biber, D., S. Conrad & R. Reppen (1998). *Corpus Linguistics Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Biber, D. & E. Finegan (eds.) (1994). *Sociolinguistic Perspectives on Register*. New York: Oxford University Press.
- Camacho, J. (2003). *The Structure of Coordination: Conjunction and Agreement Phenomena in Spanish and Other Languages*. Dordrecht: Kluwer.
- Chambers, J.K. (2003). *Sociolinguistics Theory*, 2nd edn. Malden, MA: Blackwell.
- Chaves, R. (2007). Coordinate Structures: Constraint-based Syntax-Semantics Processing. PhD dissertation, Lisbon University.
- Cheshire, J. (2002). Sex and gender in variationist research. In J.K. Chambers, P. Trudgill & N. Schilling-Estes (eds.), *The Handbook of Language Variation and Change*. Blackwell Publishers, pp. 423-43.
- Cheshire, J. (2005). Syntactic variation and beyond: gender and social class variation in the use of discourse-new markers. *Journal of Sociolinguistics* 9: 479-507.
- Cheshire, J. (2009). Syntactic variation and beyond. In N. Coupland & A. Jaworski (eds.), *The New Sociolinguistics Reader*. Basingstoke: Palgrave Macmillan, pp. 119-135.
- Cheshire, J., P. Kerswill & A. Williams (2005). On the non-convergence of phonology, grammar and discourse. In P. Auer, F. Hinskens & P. Kerswill (eds.), *Dialect Change: Convergence and Divergence in European Languages*. Cambridge: Cambridge University Press, pp. 135-167.
- Cheshire, J. & A. Williams (2002). Information structure in male and female adolescent talk. *Journal of English Linguistics* 30: 217-238.

- Chuquet, H. & M. Paillard (1987). *Approche linguistique des problèmes de traduction anglais – français*. Paris: Ophrys.
- Coates, J. (2004). *Women, Men and Language*. 3rd edn. Harlow: Longman.
- Colaço, M. (2005). Configurações de Coordenação Aditiva: Tipologia, Concordância, e Extracção. PhD dissertation, Universidade de Lisboa.
- Cosme, C. (2008). A *corpus-based* perspective on clause linking patterns in English, French and Dutch. In C. Fabricius-Hansen & W. Ramm (eds.), 'Subordination' vs. 'Coordination' in *Sentence and Text from a Cross-Linguistic Perspective*. Amsterdam: Benjamins.
- Doherty, M. (2005). Language-specific conditions for discourse linking and appropriateness. *Linguistics* 43: 1-16.
- Eckert, P. (1989). The whole woman: Sex and gender differences in variation. *Language Variation and Change* 1: 245-268.
- Erman, B. (1992). Female and male usage of pragmatic expressions in same-sex and mixed-sex interaction. *Language Variation and Change* 4: 217-234.
- Fabricius-Hansen, C., W. Ramm, K. Solfeld & B. Behrens (2005). Coordination, discourse relations, and information packaging – cross-linguistic differences. In M. Aurnague, M. Bras, A. Le Draoulec & L. Vieu (eds.), *Proceedings of the Symposium on the Exploration and Modelling of Meaning (SEM-05)*, Biarritz, France, 14-15 November 2005, pp. 85-93.
- Johannessen, J. (1998). *Coordination*. New York: Oxford University Press.
- Halliday, M.A.K. (1987). Spoken and written modes of meaning. In R. Horowitz & S.J. Samuels (eds.), *Comprehending Oral and Written Language*. San Diego, CA: Academic Press. pp. 55-87.
- Hasselgård, J., B. Behrens & C. Fabricius-Hansen (2002). *Information Structure in a Cross-Linguistic Perspective*. Amsterdam: Rodopi.
- Henry, A. (2002). Variation and syntactic theory. In J.K. Chambers, P. Trudgill & N. Schilling-Estes (eds.), *The Handbook of Language Variation and Change*. Blackwell Publishers, pp. 267-282.
- Herring, S. & J. Paolillo (2006). Gender and genre variation in weblogs. *Journal of Sociolinguistics* 10: 439-459.
- Holmes, J. (1995). *Women, Men and Language*. Harlow: Longman.
- Holmes, J. (1998). Women talk too much. In L. Bauer & P. Trudgill (eds.), *Language Myths*. New York: Penguin Books, pp. 41-49.
- Holmes, J. & M. Meyerhoff (eds.) (2004). *The Handbook of Language and Gender*. Oxford: Blackwell Publishing.