

The Pennsylvania State University

The Graduate School

College of the Liberal Arts

**DIRECTIVES USAGE BY ITAS:  
AN APPLIED LEARNER CORPUS ANALYSIS**

A Thesis in

Applied Linguistics

by

Jonathon S. Reinhardt

Submitted in Partial Fulfillment  
of the Requirements  
for the Degree of

Doctor of Philosophy

August 2007

The thesis of Jonathon S. Reinhardt was reviewed and approved\* by the following:

Steven L. Thorne  
Assistant Professor of Applied Linguistics  
Thesis Adviser  
Chair of Committee

Karen E. Johnson  
Professor of Applied Linguistics

Paula Golombek  
Senior Lecturer in Applied Linguistics

Celeste Kinginger  
Associate Professor of French and Applied Linguistics

Michael McCarthy  
Adjunct Professor of Applied Linguistics

Joan Kelly Hall  
Professor of Applied Linguistics and Education  
Head of the Department of Applied Linguistics

\*Signatures are on file in the Graduate School.

## ABSTRACT

Many large American universities have developed courses for the preparation of international teaching assistants (ITAs) to teach undergraduates, which combine attention to the development of teaching skills such as lecturing and conducting office hours with the language features particular to those genres. For example, directive language is frequently found in the spoken academic genre of office hours consultations, but proves challenging for ITAs, even as its appropriate use is crucial to their success as future academic professionals.

This dissertation explores the nature of directive usage by learner and expert speakers in office hours contexts for the purpose of informing pedagogy. Within a social-functional framework, a series of analyses are conducted on learner data from a corpus of directive language produced by ITAs in preparation courses participating in office hour role plays and expert data from a genre-comparable subcorpus of office hours in MICASE, the Michigan Corpus of Academic Spoken English (Simpson et al., 2002). The five interrelated studies are comprised of 1) a grounded analysis to identify directive constructions, 2) a corpus-based register analysis to compare learner and expert use, 3) a genre analysis of the contextual features and text-based moves of office hours, 4) an analysis of individual usage focusing on three learners, and 5) an experimental study of the efficacy of a directives instructional unit. Findings show that compared to expert speakers, learners use fewer inclusion and independence appeals towards students, and use preferred constructions more frequently and from a smaller repertoire, often relying on multi-functional constructions like 'you can'. Findings from the analysis of individual usage show a disconnect among how ITAs have been socialized by their schooling, what

they are taught in ITA preparation courses, what they do and experience in their departments, classrooms, and offices, and the kinds of academic teaching professionals they say they want to become. Results from the experimental intervention show that instruction using a corpus-informed language awareness approach has a positive impact on use. Overall, the project has implications for ITA preparation, usage-based materials design, corpus-based pragmatics and politeness research, and the nascent field of applied learner corpus analysis.

## TABLE OF CONTENTS

List of Tables.....	ix
List of Figures.....	xi
List of Data Excerpts.....	xi
Acknowledgements.....	xii
1 Introduction.....	1
1.1 Background: International teaching assistants.....	1
1.2 ITA training and preparation.....	3
1.3 Purpose of the dissertation.....	6
1.4 Research question.....	7
1.5 Chapter overview.....	9
1.5.1 Chapter 2: Literature review.....	9
1.5.2 Chapter 3: Methods and data.....	10
1.5.3 Chapter 4: Directive language as an object and unit of analysis.....	11
1.5.4 Chapter 5: A corpus analysis of directive construction usage in office hours.....	11
1.5.5 Chapter 6: A genre analytic view of office hours.....	12
1.5.6 Chapter 7: Individual variation in directive construction usage.....	13
1.5.7 Chapter 8: An application of corpus-informed language awareness instruction.....	14
1.5.8 Chapter 9: Conclusion.....	15
2 Literature review.....	16
2.1 Object of analysis.....	17
2.1.1 A pragmatic approach.....	17
2.1.2 A grammatical approach.....	20
2.2 Method of analysis.....	23
2.3 Application to pedagogy.....	29
3 Methods and data.....	32
3.1 Mixed methods approach.....	32
3.1.1 Chapter 4 methods.....	33
3.1.2 Chapter 5 methods.....	33
3.1.3 Chapter 6 methods.....	34
3.1.4 Chapter 7 methods.....	35
3.1.5 Chapter 8 methods.....	35
3.2 Data.....	36
3.2.1 ITAcorp.....	37
3.2.1.1 ITAcorp development.....	37
3.2.1.2 Data gathering and preparation.....	40
3.2.1.3 Transcription and standardization.....	41
3.2.1.4 Database structure.....	43

3.2.1.5	Corpus status.....	45
3.2.1.6	Subject corpus.....	46
3.2.2	MICASE.....	47
3.3	Data quality.....	48
4	Directive language as an object and unit of analysis.....	51
4.1	A grounded analysis of directive language.....	51
4.1.1	Method.....	52
4.1.2	Results.....	53
4.1.2.1	Core lexical devices: modals, directive vocabulary, & imperatives.....	53
4.1.2.1.1	Modals & periphrastic modals.....	53
4.1.2.1.2	Irrealis.....	55
4.1.2.1.3	Directive vocabulary.....	57
4.1.2.1.4	Imperatives.....	58
4.1.2.2	Peripheral devices: Adjuncts & pronouns/subjects.....	59
4.1.2.2.1	Adjuncts.....	59
4.1.2.2.2	Pronouns & Subjects.....	61
4.2	Discussion.....	62
4.2.1	Directive construction: A social-functional unit of analysis.....	63
4.3	Conclusion.....	68
5	A corpus analysis of directive construction usage in office hours.....	69
5.1	A corpus analysis of register consistency.....	70
5.1.1	Method.....	71
5.1.2	Results: Learner group profile.....	72
5.1.3	Results: Expert group profile.....	74
5.1.3.1	External comparison.....	75
5.1.3.2	Internal comparison.....	77
5.1.3.3	Register consistency of the EOH corpus.....	78
5.2	Learner corpus analysis: Background.....	79
5.2.1	Comparative learner corpus analysis.....	80
5.3	Comparative analysis of learner and expert profiles.....	83
5.3.1	Method.....	83
5.3.2	Results: Comparative analysis.....	84
5.3.2.1	Core lexical devices.....	84
5.3.2.1.1	Modal and p-modal constructions.....	84
5.3.2.1.2	Directive vocabulary constructions.....	87
5.3.2.1.3	Imperatives.....	88
5.3.2.2	Adjuncts.....	89
5.3.2.2.1	Overall frequency.....	90
5.3.2.2.2	Frequency by adjunct type.....	91
5.3.2.2.3	Adjunct diversity.....	94
5.4	Conclusion.....	96

6	A genre analytic view of office hours.....	98
6.1	Office hours as genre.....	99
6.2	Contextual analysis.....	101
6.2.1	Name: ‘Office Hours’.....	101
6.2.2	Social context.....	102
6.2.3	Communicative purpose.....	102
6.2.4	Participant roles.....	103
6.2.5	Cultural values.....	104
6.2.6	Text context.....	105
6.2.7	Summary of contextual analysis.....	105
6.3	Genre moves analysis.....	106
6.3.1	Data.....	107
6.3.2	Method.....	108
6.3.3	Moves analysis: Office hours.....	108
6.3.3.1	Openings and closings.....	110
6.3.3.2	Topic identification.....	111
6.3.3.3	Diagnosis.....	113
6.3.3.4	Directive.....	115
6.3.3.4.1	Directive framing with politeness appeals.....	116
6.3.3.4.2	Directive framing with divergent or convergent actions.....	117
6.3.3.4.3	Framing to elaborate a directive.....	118
6.3.3.4.4	Framing to imply a directive.....	119
6.3.3.4.5	Use of pronouns to involve and exclude.....	120
6.3.4	Discussion.....	121
6.4	Conclusion.....	123
7	Individual variation in directive construction usage.....	126
7.1	Corpus-informed usage profiles.....	127
7.2	Variation analysis.....	129
7.2.1	Method.....	129
7.2.2	Results.....	130
7.3	Learner profiles of directives usage.....	134
7.3.1	Bing C.....	135
7.3.2	Yong W.....	137
7.3.3	Phongsak T.....	141
7.4	Discussion.....	143
7.4.1	Learner constructions.....	144
7.4.2	Contextual factors.....	146
7.5	Conclusion.....	148
8	An application of corpus-informed language awareness instruction....	150
8.1	Background.....	151
8.1.1	Directives instruction in ESL and ITA preparation curricula.....	151
8.1.1.1	Grammar texts.....	152
8.1.1.2	ITA texts.....	154
8.1.2	A corpus-informed language awareness approach.....	157

8.1.2.1	Corpus-based pedagogy.....	159
8.1.2.1.1	Corpus-driven pedagogy.....	159
8.1.2.1.2	Corpus-informed pedagogy.....	161
8.1.2.2	Language awareness.....	163
8.1.2.3	Curricular design principles.....	165
8.2	A unit of instruction on directives usage.....	167
8.2.1	Experimental implementation.....	169
8.2.1.1	Overall construction usage by groups.....	170
8.2.1.2	Mitigator and intensifier usage by groups.....	172
8.2.1.3	Individual usage.....	174
8.3	An expanded curricular re-design project.....	175
8.4	Conclusion.....	177
9	Conclusion.....	178
9.1	Implications for ITA instruction.....	179
9.2	Recommendations for future research.....	181
9.2.1	Participant variables.....	183
9.2.2	Listener variables.....	184
9.2.3	Situation and task variables.....	185
9.3	Contributions to applied linguistics.....	187
9.3.1	Usage-based language pedagogy.....	187
9.3.2	Pragmatics and politeness theory.....	188
9.3.3	Corpus-based examination of pragmatic usage.....	189
9.3.4	Applied learner corpus analysis.....	190
9.4	Conclusion.....	192
	.	
	Appendix A: Office hour role play chat scenarios.....	193
	Appendix B: Biographical, language learning, and computer use experience survey.....	194
	Appendix C: Survey on attitudes towards authoritarian teaching.....	196
	Appendix D: Adjuncts found in most frequent directive constructions in learner and expert corpora.....	197
	Appendix E: Interview questions on attitudes towards teaching and conducting office hours.....	198
	Appendix F: Written assessment of directives knowledge.....	199
	Appendix G: Unit of instruction on directives usage.....	200
	Bibliography.....	207

## LIST OF TABLES

Table 3.1. Number of participants, sections, and tokens per cohort in ITAcorp....	37
Table 3.2. Description of ESL 118 FA05 and SP06 activities entered into ITAcorp.....	39
Table 3.3. ITAcorp transcription conventions.....	42
Table 3.4. Data variables in ITAcorp.....	45
Table 3.5. Number of tokens in each ITAcorp activity subcorpus by section.....	46
Table 3.6. Number of tokens, roles, and encounters in the Learner Office Hours corpus derived from ITAcorp by activity.....	47
Table 4.1. Modal and p-modal exemplars from LOH and EOH data.....	54
Table 4.2. Irrealis indexical exemplars from LOH and EOH data.....	56
Table 4.3. Directive vocabulary exemplars from LOH and EOH data.....	57
Table 4.4. Imperative indexical exemplars from LOH and EOH data.....	58
Table 4.5. Adjunct exemplars from LOH and EOH data.....	59
Table 4.6. Pronoun exemplars from LOH and EOH data.....	61
Table 5.1. Total tokens in the entire LOH corpus and in the F05 and S06 cohort LOH corpus halves.....	73
Table 5.2. Number and frequency per 10K of modal and p-modal constructions in LOH corpus.....	74
Table 5.3. Total tokens in the MICASE ADV corpus, the entire EOH corpus, and the TA and FAC EOH corpus halves.....	75
Table 5.4. Number and frequency per 10K of modal and p-modal constructions in MICASE OFC (EOH) and MICASE ADV corpora.....	76
Table 5.5. Number and frequency per 10K of modal and p-modal constructions in EOH TA and FAC corpus halves.....	77
Table 5.6. Comparison of internal and external measures of EOH corpus register consistency.....	78
Table 5.7. Total tokens in LOH and EOH corpora.....	84
Table 5.8. Number and frequency per 10K of modal and p-modal constructions in LOH and EOH corpora in comparison.....	85
Table 5.9. Number and frequency per 10K of directive vocabulary constructions in LOH and EOH corpora in comparison.....	87
Table 5.10. Number and frequency per 10K of imperative indexicals in LOH and EOH corpora in comparison.....	88
Table 5.11. Number and frequency of MFCs and MFCs with adjuncts in EOH and LOH corpora in comparison.....	90
Table 5.12. Frequencies of individual MFCs with and without adjuncts in EOH and LOH corpora.....	91
Table 5.13. Frequencies of MFCs categorized by adjunct type in EOH and LOH corpora.....	93
Table 5.14. Number of tokens, types, and most frequent adjuncts for MFCs in EOH and LOH corpora.....	95
Table 6.1. Approach, object and unit of analysis in three approaches to analysis of directives usage in office hours.....	99

Table 6.2. EOH T participant information and office hours location and description.....	107
Table 6.3. Overview of common moves in office hours texts, including purposes and frequent constructions.....	109
Table 7.1. Use of most frequent constructions by all eight experts and eight chosen learners (freq. per 10K words).....	131
Table 7.2. Sources for biographical, attitudinal and assessment data used in learner usage profiles.....	135
Table 8.1. Number of students and corpus tokens in SP06 experimental and control groups.....	170
Table 8.2. Number and frequency per 10K of directive constructions used by SP06 EXP and CON groups.....	171
Table 8.3. Frequency of individual TCs and MFCs with and without MIT & INT in EXP and CON groups.....	173
Table 8.4 Mean uses and types of MFCs and TCs by partial EXP and CON groups	174

## LIST OF FIGURES

Figure 4.1. A social-functional model of directives usage.....	67
--	----

## LIST OF DATA EXCERPTS

Data excerpt 6.1. Informal nature of openings (T 280).....	110
Data excerpt 6.2. Informal nature of openings (T 320).....	110
Data excerpt 6.3. Informal nature of openings (T 575).....	110
Data excerpt 6.4. Informal nature of session transition (T 270).....	111
Data excerpt 6.5. Informal nature of session transition and immediacy of opening-topic ID transition (T 575).....	111
Data excerpt 6.6. Immediacy of opening-topic ID transition (T 280).....	112
Data excerpt 6.7. T request for topic ID (T 195).....	112
Data excerpt 6.8. S-initiated topic ID as statement (T 280).....	112
Data excerpt 6.9. S-initiated topic ID as declaration of intent (T 195).....	112
Data excerpt 6.10. T-initiated topic ID (T 300).....	113
Data excerpt 6.11. T framing of diagnosis (T 115).....	114
Data excerpt 6.12. T use of series of questions to diagnose and to frame critique (T 300).....	114
Data excerpt 6.13. Framing used with directive (T 115).....	116
Data excerpt 6.14. Framing and example of participant alignment (T 115).....	118
Data excerpt 6.15. Framing with elaborated directive (T 320).....	119
Data excerpt 6.16. Framing to imply directive (T 300).....	120
Data excerpt 6.17. Involvement appeals through use of pronouns (T 300).....	121
Data excerpt 7.1. Bing C's use of 'I hope' , 'I guess', and 'you can' (A2).....	144
Data excerpt 7.2. Youngkyu C's use of 'you can'.....	145

## ACKNOWLEDGMENTS

I would like first and foremost to acknowledge my advisor, Steve Thorne, for his support and mentorship throughout my graduate career and for his enthusiasm in supporting this project. From when a mutual friend first mentioned him as someone to talk to when considering graduate work in technology and applied linguistics, to taking his courses, to working with him on a variety of projects, he has always been right there in my ZPD, ready to not just talk about the academic discourse community, but to actually apprentice me into it as a colleague. For that I am very, very grateful.

I would also like to acknowledge the intellectual and practical guidance of Paula Golombek and Karen Johnson, who provided me with the opportunities and resources to actualize myself as an applied linguist and educator of educators. Their wisdom and insight regarding teacher development and language learning have inspired me greatly. I am also grateful to Julie Belz for her mentorship and guidance, and for the opportunity to develop and work on Telekorp, the German-English telecollaboration corpus, which inspired the learner corpus I developed for this project, ITAcorp. I thank all of my professors at Penn State for helping me build the intellectual foundation to write this dissertation and enter the field successfully.

Regarding the practicalities of the project, I am again very grateful to Paula Golombek for letting me disrupt an already highly effective ESL 118 curriculum, and working with me to integrate the new instruction smoothly. I am especially grateful to fellow graduate students Davi Reis and Emily Rine for implementing the experimental instruction in their ESL 118 sections, and accommodating the 118 curricular changes.

I would also like to acknowledge the Center for Language Acquisition and Gil Watz for supporting me during Fall 2005 when I initiated the corpus project. In addition, I am grateful that the Center adopted ITAcorp, and supported graduate assistants Gretchen Nauman and Davi Reis and me when we transcribed a portion of the data I use in this project. I am very excited to see the corpus expand and am looking forward to seeing other researchers make use of it. I am also very grateful to the designers and developers of MICASE for creating a well-designed and freely accessible resource. It is in their spirit of open access to the resources for scholarship that ITAcorp will also be freely accessible one day.

Finally I would like to thank my friends and family for their support during my graduate experience. Especially to my parents, thank you for believing in me.

## CHAPTER 1

### INTRODUCTION

#### 1.1 Background: International teaching assistants

In the United States in 2005, the North Dakota state legislature debated a bill that would have required the removal from teaching duties of any international teaching assistant (ITA) whose teaching caused more than 10% of that ITA's students to complain (North Dakota legislature bill HB1364, 2005). While this proposal was defeated, it reflects a tension over the dual mission of American academia<sup>1</sup> to serve society—by educating citizens on one hand, and by producing knowledge and capital through research on the other (Smith, Byrd, Nelson, Barrett, & Constantinides, 1992). In pursuit of research, many universities have become dependent on international graduate student enrollment over the last several decades, especially in engineering, mathematics, and many of the sciences. In many research-focused universities, these students are also employed, often secondarily, as teaching assistants. In this capacity, they are required to instruct, tutor and assess American undergraduates, many of whom may have only encountered non-native speakers of English in service interactions or filtered through Hollywood cinematic productions or news media. Faced with exposure to unfamiliar kinds of English, as well as to culturally diverse approaches to education, these undergraduates express their frustrations to parents, newspaper editors, university administrations, and elected representatives, some of whom respond with legislation that simplifies a highly complex issue and misdirects blame.

---

<sup>1</sup> 'American' here refers to universities worldwide organized according to a model originating in the USA in the late 19<sup>th</sup> and early 20<sup>th</sup> century, where undergraduate education, often focusing on liberal arts and citizen education from an Anglo-American cultural tradition, is combined with graduate and professional educational structures originating in German traditions (Smith et al., 1992).

Overall international graduate student enrollment in the USA has increased an average of 38% every 5 years for the last 50 years, from just over 12,000 in 1955 to nearly 260,000 today (Open Doors, 2006). In 1978, there were only 50 students from China (PRC) in the USA, (Burn, 1980, in Smith et al., 1992), but by academic year 2005-06, there were 62,582 (Open Doors, 2006), a 10% average increase every year over the past decade. In 2004-05, 50% of graduate students in engineering and 41% percent in math and physical sciences were foreign born (NY Times, June 24, 2005). There are no signs that the trend will stop, even with a brief downward trend caused by 9/11. This phenomenon has been attributed to a variety of causes, including the failure of the US educational system to produce students prepared for the rigors of graduate study in the sciences, the increasing dependence of public universities on private sector funding, and the internationalization of the American academy and professoriate. Whatever the cause, like many other American industries, academia has grown dependent on inexpensive, efficient research and teaching labor, much of which, as the above statistics indicate, are imported from outside of the US. The benefits to the US economy of this phenomenon have been great, as international students and their families brought 13.5 billion dollars to the economy in 2005-06 (Open Doors, 2006), and many graduates stay on in the USA for several years after graduation, contributing further to the economy. In recognition of this, government leadership has endorsed bipartisan legislation to support the development of a “strategic marketing plan to encourage foreign students to attend American schools” (Coleman-Bingaman, 2005).

While corporate and government interests focus on research and economic benefit, in the minds of much of the American public, the core mission of publicly funded universities is to educate American citizens. In large, often impersonal lectures, ‘sage-on-the-stage’ professors tell undergraduate students to save their questions for discussion and lab sections, and to seek help in

office hours if necessary. Once there, student comprehension of the course content and structure may be mediated by an ITA, whose unfamiliar language may then come to represent not a mediator for, but a barrier to, understanding. Unaware of the forces that put the ITAs in front of them, students then complain broadly about their ITAs' command of the English language (Bailey, 1982; Smith et al., 1993; Fitch & Morgan, 2003).

Unfortunately, undergraduate complaints have been championed by the 'Official English' movement (e.g., U.S. English, 2007) and jingoists emboldened by 9/11/01. For example, the current Wikipedia entry for 'foreign-born engineering and scientists' uses many of the above facts to fabricate a foreign conspiracy to overtake the 'upper echelons' of American academia and technology industries through "incentivized replacement of native talent" (Wikipedia, April 26, 2007)<sup>2</sup>. Xenophobic wing-nuts aside, lawmakers around the USA have responded with legislation in response to the concerns of their constituents, even as it seems at odds with the business-friendly legislation discussed above. While the supporters of the aforementioned North Dakota legislation were unsuccessful, as of this writing 20 US states require that the English abilities of international non-native speaking graduate students be assessed, and if necessary remediated, before they may assume teaching duties. Responsibility is thus passed to universities and the ITAs themselves.

## **1.2 ITA training and preparation**

Universities have responded with a wide variety of tests and programs for ITAs addressing both language and teaching skills (Smith et al., 1992). At many universities, all incoming international graduate students take an oral proficiency placement test and participate in ITA

---

<sup>2</sup> There is of course no real evidence for such a conspiracy, as US and naturalized citizen enrollment in science and engineering graduate programs fell only 1% from 1993 to 2003 (Oliver, 2005). It is simply that international enrollment has increased greatly over the same period.

preparation workshops or courses based on the results of these assessments. Since ‘too much of an accent’ is a major complaint, an evaluation with which many of the ITAs themselves self-consciously agree (Yook & Albert, 1999), much time and effort in many curricula has been given to issues of pronunciation. Many universities have also turned curricula towards development of the teaching skills necessary for ITAs to carry out their duties, including lecturing, discussion leading, tutoring, and conducting office hours, and the texts and activities traditionally used are based on the experiences and intuitions of the teachers, curriculum designers, and textbook authors.

Despite these persistent efforts and some successes, complaints about ITAs and ITA language use persist. As studies have shown that perception of ‘accent’ is not objectively measurable (e.g. Nisbett & Wilson, 1977), research has turned to examine the teaching aspect, focusing on classroom interaction and intercultural pragmatics (see chapter 2), and challenging the effectiveness of curricula based on designers’ intuitions and experiences (see chapter 8). Effective teaching includes the transactional skills to clearly explain and present information, as well as the interactional skills to negotiate intersubjectivity and evaluate learner needs. In terms of language, this involves making intentional linguistic choices at ideational, textual, and interpersonal levels. Seen from the perspective of genre (Swales, 1991), the parameters for these choices are afforded and constrained by the context of situation (Halliday & Hasan, 1985), and more broadly, culture. Because one’s conceptualization of teaching and education is mostly based on one’s own culture-embedded educational experience, ITAs’ understandings of teacher and learner roles may be quite different from the understandings and experiences of their American undergraduate students. For example, in East Asian educational contexts influenced by Confucian traditions, teachers are expected to be more explicitly authoritarian in their

interactions with students. Considering that in 2005-06 over half of all international graduate students were from educational cultures with Confucian influences<sup>3</sup> (Open Doors, 2006), the potential for misunderstanding based on cultural expectations is great (see also Bresnahan & Kim, 1993 and Jenkins, 2000 on this issue). This is compounded by the fact that ITAs are often the first, and sometimes only representative of a department that interacts at any length with undergraduates, especially in introductory-level courses.

While such misunderstanding can often occur in written communication (e.g. see Kruger, Epley, Parker, & Ng, 2005 for e-mail), it is more probable in spoken interaction, since these interactions tend to be taking place in real-time and to involve extemporaneous, unplanned, and at times unpredictable discourse. To succeed as academic professionals in the future, both US born and international graduate students must become proficient in many spoken academic genres<sup>4</sup> associated with research and teaching, including giving lectures and lecturettes, leading discussions, labs, and seminars, participating in colloquia, and advising, counseling, and conducting office hours. Even before standing in front of a class, ITAs are most likely to find themselves conducting office hours, as they tutor students on homework problems, prepare them for tests, and answer minor procedural questions on behalf of the supervising professor. In other words, they are asked to wield symbolic power<sup>5</sup> (Bourdieu, 1991) and perform precise professional and institutional roles, often with minimal preparation beyond pronunciation training.

---

<sup>3</sup> Students from China, Taiwan, Hong Kong, S. Korea, Japan, and Thailand totaled 58% that year (Open Doors, 2006).

<sup>4</sup> This is not to mention many written genres, like essay, paper, article, literature review, bibliography, review, proposal, and dissertation.

<sup>5</sup> 'Symbolic power' is defined here in Bourdieu's sense as the expression of social capital through linguistic means to reify existing power hierarchies and social structures.

In the exercise of this authority in consultation contexts like office hours, teachers often have to make suggestions, give advice, and direct students, i.e. tell them what to do. The office hours genre affords and even demands this particular kind of language. Since these directive speech acts inherently involve the imposition of the speaker-director's will, the speaker necessarily expresses his/her stance towards authority and the hierarchy of the teacher-student relationship in his/her choice of language, which may result in unintended face threats and misunderstandings. Examining and comparing how practicing teachers (expert speakers of English) and ITAs in preparation courses (advanced and near-expert learners of English<sup>6</sup>) use directive language in the particular genre of office hours consultations might thus provide insight into these issues, as a microcosmic representation of one-on-one teacher-student educational discourse in general, for both learners and expert speakers.

### **1.3 Purpose of the dissertation**

In view of the above, the purpose of this dissertation is to explore the nature of directives usage by experts and learners in office hours contexts, with the intent to inform pedagogy for ITAs in preparation. Research on non-native speaker discourse has looked at deontic speech acts like requests and directives, but not specifically in office hours contexts, and not specifically by learners preparing to be ITAs. In addition, the previous research has not used corpus analytic methods to examine this discourse (see chapter 3). Finally, this dissertation sets itself apart by having as its ultimate goal the development, implementation, and efficacy analysis of a unit of instruction informed by its findings. The goals and methods of analysis of other studies have been similar (see chapter 2 and §5.2.1), but to my knowledge none have gone beyond offering

---

<sup>6</sup> The terms 'learner' and 'expert' here refer to language proficiency, not teaching proficiency, which is a different matter—see §9.3. Arguments for the terms 'learner' and 'expert' are presented in the discussion of comparative learner corpus analysis in §5.2.1.

implications to the step of applying the findings to pedagogy and to examine the effect of that pedagogy. Practitioners often remark that applied linguistics research does not impact them because it remains theoretical and, ironically, inapplicable to their actual teaching situations. In response, the current project employs linguistic and applied linguistic theory to analyze empirical learner data, applies the findings to the design of classroom instruction, and then examines the effect of that instruction. The intention is thus to inform curriculum design and to inspire practitioners who work with and prepare ITAs, thus positively impacting the current and future academic lives of the students who are ITAs, and thereby the lives of those they will teach.

#### **1.4 Research question**

For the reasons explained above, it is hypothesized that misunderstandings occur at least partially due to cultural differences in teaching styles vis-à-vis the exercise of academic authority. Since a teacher's stance towards their own authority is expressed linguistically in directive language usage<sup>7</sup>, it is hypothesized that this stance would be expressed in office hours contexts where they are interacting with students. In view of this, the main research question is:

What is the nature of directive language usage in office hours contexts in relation to the exercise of academic authority? How can directive language use be effectively taught to L2 English learners preparing to be academic professionals?

---

<sup>7</sup> In this dissertation, the use of the term 'usage' is not meant to be the same as Widdowson's (1978) definitions of 'usage' as the ability to produce correct sentences, in contrast to 'use' as the ability to use the language effectively. All language use by the participants is 'use' in Widdowson's terms, and the use of 'usage' is simply a nominalization of the word 'use' that I use to avoid overuse (overusage?) of the word 'use'. See also §4.2.1 on 'usage-based'.

For a thorough answer, I examine directive language produced by both learners and experts in academic contexts from a variety of methodological perspectives, in a series of interrelated analyses that build on one another in an intentional order. I start with corpus analysis, as it provides a means to profile a given register, and to compare usage profiles from different users to find points of systematic variation. Specifically, I use contrastive corpus analysis (e.g. Granger, 1998, 2002) to compare a corpus of learner language produced by would-be ITAs in role play office hour contexts, with a corpus of language produced by expert speakers of English in actual office hours contexts<sup>8</sup>. While there are issues regarding the comparison of expert and learner usage, the purpose of these analyses are meant not to imply deficient usage by the learners, but to inform the development of ITA preparation materials that present directives usage as a matter of intentional choice and exercise of power, as opposed to a prescriptive set of usage rules.

To this end, I begin in chapter 4 by operationalizing directive language use in a way that forges relations between lexico-grammatical units and social function. In chapter 5, I use this unit of analysis to profile and compare directive usage by the experts and learners in the office hours register, defined here as a general kind of language associated with the office hours domain of use that systematically varies in regards to lexico-grammatical features (see also Biber, 2006). To relate lexically-defined functional usage to the social aspects of office hours, in chapter 6 I then conduct a genre analysis in two parts, a situational or contextual analysis (e.g. Tribble, 2001), and a moves or stage analysis (e.g. Swales, 1990). Genre is the ‘social’ side of a register, “a class of communicative events, the members of which share a set of communicative

---

<sup>8</sup> The term ‘expert’ refers to native speakers and near-native speakers in faculty and teaching assistant roles, as determined by the developers of MICASE (see also Simpson-Vlach & Leicher, 2006, p. 55). The term ‘learner’ means enrollment in a ITA advanced language and teaching preparation course, as determined by a standardized university English oral proficiency test (see §3.2.1.1). See §5.2.1 for more on these terms.

purposes” (Swales, p. 58)<sup>9</sup>. To balance the aggregate nature of the register and genre analyses, in chapter 7 I then examine the individual use of directives by 16 experts and learners in comparison, and by three learners in detail, using supplemental ethnographic data. Finally, in chapter 8 I present the unit of instruction on usage of directives for ITAs, and discuss the effect of its experimental implementation on learner directives usage.

## **1.5 Chapter overview**

The main research question can be broken down into five specific sub-questions, each of which provides a research objective for one of five interrelated studies that each correspond to a chapter. Each study is comprised of a series of analyses that address the objective. This series of studies, the main body of the dissertation, is first situated in a literature review (chapter 2) and a description of the methods and data used for each study (chapter 3).

### **1.5.1 Chapter 2: Literature review**

Since the first student complaint in the late 1970s (Smith et al., 1992), researchers have sought insight into the ITA issue, starting with Kathleen Bailey’s 1982 dissertation, where she correlated high levels of ITA ‘helpfulness’ and ‘interactiveness’ with positive evaluation of effective teaching by undergraduates. This spurred a formidable wave of research on ITAs in the late 1980s and early 1990s in applied linguistics and TESOL (Briggs, Clark, Madden, Beal, Hyon, Aldridge, & Swales, 1997), which influenced a strand of research on educational institutional discourse starting with the work of Bardovi-Harlig and Hartford (1990; 1993). The current project is situated in and contributes to these lines of research. The literature review will examine this research more closely in terms of object of analysis, method or approach to

---

<sup>9</sup> See §6.1 for more on the definitions and differences between ‘genre’ and ‘register’.

analysis, and application to pedagogy. The purpose of the review is to emplace the current project in the literature and to inform its design and identify approaches appropriate for analysis of directive language use by experts and learners in office hours contexts.

### **1.5.2 Chapter 3: Methods and data**

While chapter 3 provides a general overview of the variety of methods used throughout the dissertation, a more complete and focused methodology discussion is found in all the main chapters, each of which corresponds to an analytically independent research study. In chapter 3, the general methodology overview is followed by an extensive discussion of the source for learner data I designed and began developing for this project, called ITAcorp. ITAcorp is a corpus of advanced spoken learner English now housed permanently at Penn State that continues to grow in volume and expand in breadth, with in-kind support for continued development provided by Penn State's Center for Language Acquisition. I also discuss the source of the expert data, MICASE, the Michigan Corpus of Academic Spoken English (Simpson, Briggs, Ovens, & Swales, 2002), a resource generously made publicly accessible by its developers. The chapter concludes with a discussion of the quality of role play data, which have been criticized for use in interlanguage pragmatics research due to questions about authenticity (Bardovi-Harlig & Hartford, 2005). Briefly, I argue that the pedagogical situatedness and high stakes relevance of the role play tasks result in highly authentic data quality.

### **1.5.3 Chapter 4: Directive language as an object and unit of analysis**

Directives are traditionally considered to be speech acts reflecting speaker stance, which makes them difficult to identify and quantify for corpus analytic purposes. In addition, corpus analysis often results in decontextualization of specific lexical choice, and context of utterance is

necessary to relate those choices to context of situation and culture. A well-defined unit of analysis may alleviate this issue. Thus a guiding question for this study is:

What is ‘directive language’ and how can it be operationalized for both quantitative and qualitative analytic purposes?

To answer this, the concept of directive language is explored to the end of defining the object of inquiry, in this case, discrete strings of language used with directive intent, and operationalizing a unit of analysis for use in subsequent studies. For corpus analytic purposes, this lexico-grammatical unit must be searchable and quantifiable, while for genre analytic purposes, it must remain bound to its pragmatic function. With this in mind, I conduct a grounded analysis of transcripts of the expert and learner data. This analysis results in a list of units that are then grouped according to grounded theory into categories. The proposed units and categories are then discussed and supported with primary source linguistic and pragmatic theory. Finally, a usage-based ‘directive construction’ unit of analysis is proposed, grounded in a social-functional model of usage.

#### **1.5.4 Chapter 5: A corpus analysis of directive construction usage in office hours**

Once the unit of analysis is determined, corpus analyses are carried out to systematically describe directives usage in office hours. This corpus-based register analysis (e.g. Biber, Conrad, & Reppen, 1998) can then provide a means to profile and compare expert and learner usage of directives. The research question for this study therefore asks:

What are the systematic qualities of directives usage in office hours? To what extent does directive language usage by experts differ from usage by advanced ESL learners in this register?

To address this question, I explore the systematic qualities of directives usage in office hours through the individual and comparative corpus analysis of both the learner corpus from ITAcorp and the expert corpus from MICASE. Using the proposed ‘directive construction’ unit of analysis, I first examine the register consistency, i.e. the usage distribution of directive constructions, of both corpora separately, by comparing each corpora both externally and internally (see chapter 3 and 5 for more on the method of consistency analysis). I then compare directives usage in the two corpora using contrastive interlanguage analysis (Granger, 1998; 2002). I also compare adjunct usage in the two corpora, including mitigators and intensifiers, and examine the findings of the analyses using the social-functional theoretical model I developed in chapter 4 as an interpretive heuristic.

### **1.5.5 Chapter 6: A genre analytic view of office hours**

A thorough examination of language use as a social phenomenon examines it both as a register, i.e. a “functional variety of language—the patterns of instantiation of the overall system associated with a given type of context (situation type)” (Halliday and Matthiessen, 2004, p. 27), and as a genre, or the same text seen from the context of culture, with social purpose as its defining characteristic (see Swales, 1991). While a corpus-based register analysis can consider directive constructions in a text as a usage phenomenon, it does not relate that usage to the broader social purpose of that text. To this end, a genre analysis can complement a corpus-based

register analysis by exploring the contextual and textual features of office hours as a socially-defined genre. With this in mind, the research question asks:

What are the systematic contextual features of office hours as a genre? What systematic moves or stages are found in expert office hours texts and what is the role of directives in these moves?

To answer this, I first conduct a contextual analysis (e.g. Tribble, 2001) of the ‘office hours’ context through examination of a variety of features specific to the social participation of ITAs in these contexts, including its communicative purposes, participant roles, and cultural values. To complement this contextual analysis, I then conduct an empirically driven genre moves analysis (e.g. Swales, 1991) on the transcripts from the expert corpus. A series of obligatory and optional moves are proposed, including a ‘directive move’.

### **1.5.6 Chapter 7: Individual variation in directive construction usage**

While the empirical and heuristic strengths of corpus-informed register and genre analyses are useful, they still consider usage to be an aggregate phenomenon, and so are limited as to what they can explicate about how an individual’s usage may relate to his/her stance towards the exercise of academic authority. Corpus-informed studies of individual usage (Belz, 2004; Vyatkina & Belz, 2005) have provided insight into a variety of usage phenomena that would be untenable from analysis of aggregate data alone. The research question for this study therefore asks:

In office hours contexts, how does directives usage vary by individual? More specifically, how do individual learners use directive language in relation to the exercise of academic authority?

To address this question, I first present corpus-informed variation analyses of directives usage by the 8 experts and 8 of the learners in the subject corpora, examining and comparing frequency and variation. I then conduct in-depth exploration of directive language usage by 3 learner-participants in the learner corpus, using biographical, attitudinal, and assessment data<sup>10</sup> to inform the analysis. These portraits are presented not as case studies, since they are neither longitudinal nor inclusive of other language production types, but as qualitative usage profiles, since they are corpus-informed snapshots<sup>11</sup> of usage of a particular language type, directives. In alignment with the other studies, the social-functional framework developed in chapter 4 is used to interpret the findings.

### **1.5.7 Chapter 8: An application of corpus-informed language awareness instruction**

As mentioned above, the ultimate goal of this dissertation project is not simply to describe thoroughly the use of directive language by learners and experts in office hours contexts, but to apply the findings of that description to pedagogy. The question for this study thus asks:

How can ITA ‘office hours’ materials and curricula reflect findings on directive language usage, and how does that differ from how ‘office hours’ are traditionally taught? What is the influence of such an empirically-informed instruction on learner usage?

---

<sup>10</sup> Such data, beyond basic biographical data, was unfortunately unavailable for the MICASE participants.

<sup>11</sup> The influence of a pedagogical treatment on longitudinal usage by some of these learners is discussed in chapter 8.

To address this question, I first discuss background elements to situate an instructional unit, including a review of the current state of directives instruction in representative ESL grammar textbooks and ITA training materials, and a discussion of the corpus-informed language awareness approach (Carter, 1998; McCarthy, 1998; Gavioli & Aston, 2001; Flowerdew, 2005) on which the design of the unit is based. After the background discussion, I present the original instructional unit, which reflects preliminary comparative corpus findings, and was implemented in an experimental study in Spring 2006 in an advanced ITA language and teaching skills preparation course at a large, semi-public American research university. This is followed by a description of the study, controlled for teacher influence, conducted on two experimental sections that implemented the unit and one control section that did not. Finally, I discuss the curricular impact of the unit and describe a second version and a redesign of ESL 118 curriculum that are extensions of this project.

### **1.5.8 Chapter 9: Conclusion**

In the final chapter of this dissertation I present the summarized findings of the five component studies and points of discussion drawn from those findings and from the individual chapters. I then present implications for future research, especially for ITA preparation, usage-based language education, corpus-based research, and pragmatics and politeness research.

## CHAPTER 2

### LITERATURE REVIEW

As I briefly described in the introduction, since its advent in the late 1970s, researchers have sought solutions to the ‘ITA problem’, starting with Kathleen Bailey’s 1982 dissertation, where she correlated high levels of ITA ‘helpfulness’ and ‘interactiveness’ with positive evaluation of effective teaching by undergraduates. This spurred a formidable wave of research on ITAs in the late 1980s and early 1990s in applied linguistics and TESOL, summarized by Briggs, Clark, Madden, Beal, Hyon, Aldridge, and Swales (1997) in a highly comprehensive and thorough annotated bibliography of ITA research. One series of studies focused on correlating ITA attitude towards teaching and undergraduate evaluation of effective teaching (Bailey, 1982; Briggs & Hofer, 1991; Davis, 1991; Nelson, 1991; Brown, 1992; Rubin, 1992; Bauer & Tanner, 1993; Hendel et al., 1993; Inlis, 1993; Davies & Tyler, 1994; Smyrniou, 1995; all annotated in Briggs et al., 1997). Another group of studies focused on TA discourse, with more focus on language and genre, or ‘discourse domain’. For example, Byrd and Constantinides (1992, annotated in Briggs et al., 1997) focused on the language of math instruction, concluding that institutional discourse may be highly specific to field. Douglas and Selinker (1986; 1984; annotated in Briggs et al., 1997) looked at transferability of discourse strategies between domains and genres. There were also a series of studies looking at questioning in particular (Robinson, 1993; Myers, 1994; Rounds, 1994; annotated in Briggs et al., 1997), as well as textual features from an intercultural discourse analytic perspective (Tyler & Davies, 1990; Tyler, 1992; Williams, 1992; Tyler, 1994; annotated in Briggs et al., 1997). In sum, Briggs et al.

show a field trending towards genre-specific analysis of ITA discourse focusing on functional discourse features.

This trend influenced a strand of research on educational institutional discourse begun by Bardovi-Harlig and Hartford (1990; 1993) with which the current project is aligned. The review below will examine this strand more closely in terms of object of analysis, method or approach to analysis, and application to pedagogy. The purpose of this review is to situate the current project in the literature, to inform its design, and to identify approaches appropriate for analysis of directive language use by experts and learners in office hours contexts. Studies were selected by searching current journals, single-authored texts, and edited volumes, and were limited to those whose objects of analysis were directive or directive-like language use.

## **2.1 Object of analysis**

There have been many studies of the use of directive-like language by native (NS) and non-native (NNS) speakers in educational institutional discourse, including advising sessions, writing center interactions, and classroom instruction. This language is not always termed ‘directive’, and is most often understood in terms of pragmatic speech acts including suggestions, requests, and rejections (Bardovi-Harlig & Hartford, 1990; 1993; Thonus, 1999; Williams, 2005). Some studies look more generally at ‘modality’ (He, 1993; Iedema, 1996), while others specifically label their object of analysis ‘directives’ (He, 2000; Hyland, 2002; Yates, 2005; Dalton-Puffer & Nikula, 2006).

### **2.1.1 A pragmatic approach**

The most commonly used approach to directive language is to examine speech acts, including suggestions, requests, and rejections. Bardovi-Harlig and Hartford (1990) examine

how NS and NNS graduate advisees negotiate an interactional maxim of status congruence when making suggestions to and requesting advice from NS advisors in advising sessions. This conversational maxim states that a speaker will avoid or mitigate a speech act like a suggestion if it is incongruent with their speaker status in a particular context with a particular interlocutor. A few years later (1993), the researchers again analyze suggestions made by 6 NS and 10 NNS advisees in the same contexts, but also looked at rejection strategy usage and development over the course of several advising sessions. They show how NNS more often violate the congruence maxim when making suggestions and rejections by avoiding or violating status-preserving strategies, which include: a) appear congruent; b) mark your contribution linguistically (use mitigators); c) timing (do not begin with a noncongruent contribution); d) frequency (avoid frequent noncongruent turns); e) be brief; and f) use appropriate content.

Thonus (1999) examines suggestions and evaluations in 34 NS-NS and NS-NNS tutor-tutee writing center interactions. Like Bardovi-Harlig and Hartford, 'directive' is understood as a phase of an institutional interaction (Agar, 1985; see also chapter 6), where speech acts like suggestions and evaluations are most frequent. Thonus frames her analysis of the speech acts in terms of mitigation, theoretically aligned with Brown and Levinson's (1987) politeness theory, where a face-threatening act, like making an overt suggestion to a supposedly equal tutee in a writing center, can be avoided, performed off-record, redressed with mitigation, or performed on-record (p. 269). She determines that NS tutors more often violate these maxims when tutoring NNS than NS tutees, perhaps for the sake of clarity and communication, or because the NNS tutees do not accept the supposed equal status of their NS tutor as readily as NS tutees.

In her study of 8 NS and 8 NNS in-service Australian secondary school teachers' use of directives, Yates (2005) also combines pragmatic analysis of mitigation with politeness theory.

She categorizes mitigators according to Blum-Kulka's taxonomy (1989, in Yates 2005) of strategy of the directive head (level of directness), syntactic, lexical, and propositional modification, and relates each to positive, negative, mixed, or contextually dependent distancing (face) effect. For example, the directive head 'you sit down (plain stem)' is seen as distance maintaining, while 'I want (you to)' is less distancing. Politeness markers like 'please' are considered more distancing, while use of vernacular is less distancing. Approval moves like general praise are considered "warm but not solidary" (p. 74), while personal/cultural references are considered devices that are less distancing and rapport-building.

Williams (2005) examines suggestions and rejections in 10 writing center sessions, 4 between expert tutors and NS tutees, and 6 between expert tutors and NNS tutees. Like Bardovi-Harlig and Hartford and Thonus, her analysis focuses on the interactional, co-constructed nature of directive speech acts. She finds the structure of interaction to generally follow Agar's diagnosis-directive-report pattern, except for a dominance of the diagnosis and related problem identification phase, especially with the NNS tutee interactions. To analyze interaction, Williams goes a step further than mitigators and politeness, analyzing the frequency and nature of a variety of markers and conversational moves typically associated with expertise/dominance, including dominance markers like turn length, interruptions, and leading moves, dominant speech acts like speaker-initiated suggestions and rejections, and lexical indicators of dominance like lack of mitigators and presence of aggravators. In this sense, Williams' speech act analysis is embedded not in maxim violation or strategy avoidance and use, but in linguistic and quantitative evidence of the exercise of power.

### 2.1.2 A grammatical approach

While a pragmatics-based perspective sees directives as speech acts subject to mitigation, a more grammatical approach is to examine directive language as a form of modality. He (1993) examines how 8 advisors and 21 advisees, all native English speakers, use lexical, syntactic, and discourse-level modality, both deontic and epistemic, to construct their roles and the context itself. She defines modality as “the projected stance of participants towards claims regarding facts or reality ranging between the poles of affirmation and negation,” noting that it “concerns the interpersonal and interactional factors that affect how knowledge and knowing is displayed” (p. 504). She uses Halliday’s (1985, in He, 1993) description of grammatical modality as involving the relationship between a proposition and the context of a speech event by reference to the judgment of a speaker, as it is realized in the Finite of a Mood element. Because she feels his model does not specify the origins of the definitional meanings of modality and the causes of their variations, be they social/institutional, situational, or interactional, she expands Halliday’s model into a ‘socio-semantic’ analytic framework. This considers the interaction of modality with truth value of an utterance and discourse status of a speaker, each of which can be high or low value, so that high value modality (e.g. ‘must’ as opposed to ‘could’) combined with high value discourse status (e.g. a teacher speaking as opposed to a sibling) results in the co-construction of high truth value to the utterance. Modal devices include high and low value ‘operators’, i.e. the main modal structure, and high and low value ‘adjuncts’, i.e. mitigators and intensifiers. In this way the researcher incorporates what might be termed mitigation or intensification devices from a pragmatic-speech act perspective into a broader functional grammar-based definition.

Like He, Iedema (1996) uses a Hallidayan systemic-functional framework for critical analysis of teacher use of regulative discourse as a means to induct native English speaking primary and secondary school age children into “bureaucratic-pastoral self-discipline and self-determination” (p. 82). Iedema critically examines the nature of first-order regulative discourse, as opposed to second-order instructional discourse, as it involves ‘Command’, i.e. the demanding of goods and services in Hallidayan terms, and may be realized not only as imperatives but also metaphorically as declarative and interrogative metaphors of modulation and mood. Metaphors of modulation may be subjective, where the deontic force or ‘should-ness’ in Iedema’s terms, is realized in the Finite (e.g. ‘you should’ or ‘I want you to’), or objective, where the force is realized in the Predicator (e.g. ‘you are required to’ or ‘it’s necessary for you to’). Subjective modulation projects modal responsibility to the subject, while objective modulation distances it from the subject. The structures are also either implicit, where the requested action is part of the same verbal group as the modulation (e.g. ‘you should go’ or ‘you are required to go’), or explicit, where it is part of a projected clause (e.g. ‘I want you to go’ or ‘it’s necessary for you to go’). Thus, only explicit modulation is truly metaphorical. In complement to modulation, metaphors of mood involve the indirect realizations of performative speech acts, so that Commands are realized as declaratives (e.g. ‘I wish you would do this’ instead of ‘do this’) or interrogatives (e.g. ‘I asked where he was’ instead of ‘where is he?’).

Using this framework, Iedema shows how regulative discourse to young children tends to be more subjective implicit and less metaphorized than that addressed to older children, becoming ‘demodalized’ and ‘ideationalized’, as it moves from subjectification to factualization, passing through passivization, objectification, and nominalization phases. While these processes are intricate and beyond this literature review (though see chapter 4), the point is that the original

source of the command is backgrounded and eventually internalized by the student. For example ‘would you like to begin?’ is a metaphorized command in interrogative form that encourages interiorization of the source of compliance through subjective modulation. In this way, “sophisticated manifestations of social authority involve the construction of power as impersonal and the compliant individual as personal” (p. 97). Iedema is thus able to directly relate grammatical structure to the exercise of power in institutional contexts.

Several studies apply functional taxonomies of text type to analyze directive usage, noting differences of usage according to not just context and speaker, but the purpose or register of the directive. Dalton-Puffer and Nikula (2006) conduct a qualitative and quantitative analysis of directives, including requests, produced by NNS teachers and students in an EIL content-based secondary school environment. Going beyond the traditional influences of power, distance, and imposition, the authors show how the realization of a directive (e.g. authorship and degree of mitigation) may be influenced by the interplay of classroom register (regulative or instructional) and object of directive (to request information or direct action). He, in a later work (2000), presents directive use not critically as a means of social control, but positively as a means of socialization for heritage language learners. She analyzes the Chinese directives used by two teachers in weekend heritage school instruction, categorizing the directives according to function. Instructional directives include initiating directives, which are preceded by discourse markers and used to frame the class, socializing anticipating directives, which are simple imperatives sometimes embedded in a test question sequence, and modalized preference/permission statements, which involve the use of modals that while presenting the illusion of choice are in fact directives. Disciplinary directives include three phases, which may

be presented linearly or circular, of orientation, where the teacher orients a student's errant behavior, evaluation, where she presents the moral consequences, and the directive itself.

In sum, the studies above operationalize directive language from three major, interrelated perspectives: 1) pragmatically as speech acts with directive illocutionary force subject to mitigation and intensification, 2) grammatically as directive modality expressed through linguistic structure, and 3) generically as a register of institutional interaction involving the exercise of power. There are several notable common threads in all of these studies that should inform any similar study: 1) directives are pragmatic speech acts that express directive illocutionary force; 2) in institutional discourse<sup>1</sup>, directives may be a means of socialization, persuasion, and the exercise of academic power; 3) the grammatical form of a directive may vary widely, although it is usually centered on a verbal element that often contains a modal operator; and 4) a speaker may mitigate or strengthen the force of a directive by various lexico-grammatical means in the exercise of power. Relative participant status and directive purpose are determinative factors in this process. In consideration of these observations and a grounded analysis of the data, the object of analysis and an analytical framework for this project are laid out in chapter 4.

## **2.2 Method of analysis**

While the reviewed studies operationalize directive language differently, they use a variety of analytical methods and approaches. This has included discourse analysis of speech acts (Bardovi-Harlig & Hartford, 1990; 1993; Thonus, 1999; Farr, 2003; Dalton-Puffer & Nikula, 2006), sometimes with a major focus on frequency analysis (He, 1993; Williams, 2005; Yates, 2005). In addition, some have taken ethnographic approaches (Bardovi-Harlig, 1996; Jenkins,

---

<sup>1</sup> and certainly other macrogenres that are not the object of investigation here, for example political, medical, and parenting discourse

2000), experimental approaches (Garcia, 2004), or made extensive use of corpus analysis (Hyland, 2002; McEnery & Kifle, 2002).

A prevalent approach to the analysis of speech acts is the use of qualitative discourse analysis in combination with grounded categorization techniques and basic frequency analyses. Setting the precedent, Bardovi-Harlig and Hartford (1990, 1993, both above) conducted qualitative discourse analysis of 32 entire advising sessions, looking at maxim congruence vis-à-vis status (1990) and suggestion and rejection use in speech acts (1993). In both studies, the authors focused on group trends in strategy usage, illustrating these trends with transcript excerpts and supporting their findings with only basic frequency analyses. Using similar approaches, Thonus (1999, above) analyzed 34 writing center interactions between 11 NS tutors and 34 NS and NNS tutees, and Dalton-Puffer and Nikula (2006, above) looked at directive usage by two groups of NNS teachers and students. Farr (2003) used a similar approach to examine the strategic use of engaged listenership devices in NS student-tutor consultations<sup>2</sup>, including minimal (mhm, yeah, yes, etc.) and non-minimal response tokens (right, exactly, absolutely, etc.), interruptions, and overlaps. Her ‘mini-corpus’ is comprised of 10 interactions between 2 tutors and 9 tutees, totaling just under 70,000 words. In her analysis, she qualitatively analyzes the socio-pragmatic use of each device, using contextualized examples from the corpus for each.

More recently studies have continued to primarily use discourse analytic techniques, but with more attention to frequency analysis of the data. Whether or not the data is termed ‘corpus’ seems to be a function of whether the study is published after 2000 or so, regardless of whether corpus analytic techniques are actually used. He (1993, above) used a comprehensive frequency

---

<sup>2</sup> In an Irish academic setting, ‘tutoring sessions’ are more equivalent to American office hours than to American tutoring sessions.

analysis to look at use of high and low value modals by NS advisors and students, showing that counselors used low value modal elements as frequently, and high value elements twice as frequently, as students, reflecting the advisor's academic role as mediator between academic dictates and student wishes. In addition, she complements the quantitative with a close, qualitative 'socio-semantic' discourse analysis of exemplars. Williams (2005, above) uses discourse analysis to analyze the genre moves of NS-NS and NS-NNS writing center interactions and present exemplars, but focuses primarily on frequency when analyzing markers and conversational moves typically associated with expertise/dominance. In a sort of corpus-informed individual profile approach, Yates (2005, above, also §7.1) used discourse and frequency analytic techniques to look at the mitigation style of 8 NS and 8 NNS public school teachers. Yates stands out because her approach focuses on individual rather than group usage.

Other researchers have taken qualitative, ethnographic approaches to the study of institutional discourse. Bardovi-Harlig and Hartford (1996) interviewed participants from their earlier studies of academic advising sessions to determine the nature of pragmatic input available to NNS in institutional encounters. For example, negative feedback on inappropriate pragmatic usage is generally not available to learners, but positive evidence may be available from peers, and certain linguistic features of advisor talk itself may serve as input, like the use of modals and semi-modals. Jenkins (2000) investigates miscommunication between Chinese ITAs and professors in a mathematics department, using a qualitative approach based on interviews of 14 professors, and interviews, reflective journals, and classroom observations of 7 student-ITAs. She uses attribution theory and Gumperz' conversational inference theory to interpret the findings, which point towards professor misinterpretation of silence that resulted from student use of Chinese norms of interaction.

While experimental designs in intercultural pragmatics studies are rare (see Bardovi-Harlig & Hartford, 2005, for reasons why), Garcia (2004) used such an approach to examine differences between NS, high proficiency learner, and low proficiency learner awareness of pragmatic meaning in indirect speech acts. Fifty-six participants completed controlled listening tasks, which involved listening to spoken academic encounters and identifying and interpreting the involved speech act, including requests, suggestions, offers, and corrections. She found that the NS and the high proficiency learners significantly outperformed the low proficiency learners. Linguistic analysis of the problematic items showed them to exhibit unclear or indirect agents, false starts and hesitations, and unfamiliar uses of modals.

More prevalent are corpus-based research on academic language, although there have been no corpus-informed treatments of *spoken* academic learner language to my knowledge, a gap which the current project seeks to fill. Most corpus-informed studies of academic language are of native/expert speaker use of written language, for example textbook and academic journal language (e.g. Hyland, 2002; Biber, 2006—see §5.2), or of student writing (e.g. McEnery & Kifle, 2002; Granger, 2003). While corpus analysis and learner corpus analysis will be more thoroughly reviewed in chapter 5, for this review of methods, Hyland (2002) and McEnery and Kifle (2002) are especially relevant because they use a corpus approach to deal with directives and learner use of modality respectively, although in written genres.

Hyland (2002) examines the influence of written genre type and academic discipline on the use of directives in a 2.5 million word written corpus comprised of texts of research articles, academic textbooks, and L2 student senior theses. Hyland defines directive as an utterance that expresses “an obligation on the reader either to do or not do something” (p. 216). He excludes requests, invitations, and offers from his definition of directive because they “seem qualitatively

different forms of social action from utterances which strongly urge the reader to act in a certain way”. Directive force could be realized as (1) an imperative, (2) a modal of obligation addressed to the reader, or (3) a predicate adjective indexing directive force and controlling a complement to-clause. (p. 216). He recognizes that directives are used to mark and construct status differences in interaction, but that various directives index different weights of threat and imposition on the reader, depending on whether they direct the reader toward textual, physical, or cognitive action. Using basic frequency counts, Hyland determines that academic textbooks contain the highest percentage of directives, followed by research articles, then student theses. Imperatives were most frequent, followed by modals, especially ‘should’ and ‘must’, and then predicate adjectives. He found the proportion of directives guiding cognitive, textual, and physical action to be varied among genre types. Other linguistic features also account for differences, for example, textbooks do not use hedging, but do make use of inclusive pronouns, creating an authoritative yet inclusive tone. Hyland also finds that in general, writing in ‘hard’ disciplines like engineering and sciences, make more use of directives than ‘soft’ disciplines like sociology and business, although there are notable exceptions. He speculates this could be because directives may be one of the only means of reader engagement for science writing.

McEnery and Kifle (2002) used comparative learner corpus analysis to investigate the use of epistemic modality devices in the argumentative essays of NS (British) and NNS (Eritrean) students. They built two corpora, each of 22,000 words, and compared the number of modality devices of high, medial, and low probability. The NS writers used more devices of the first two types than the NNS writers, from a larger variety of parts of speech. Of all the devices, the modal ‘would’ was used more frequently by the NS at the highest significant difference. In contrast to previous studies (which were mostly of L1 Chinese writers), the authors found that the learners

used more low probability devices than the experts, giving the essays a more tentative, less certain feel. The authors attribute this not to culture but to the explanations of epistemic modality in the most common Eritrean EFL textbook, which focuses primarily on weaker modal verbs and advocates hedging.

In sum, the above methodological approaches each have advantages and disadvantages to their employment. Corpus analysis (Hyland, 2002; McEnery & Kifle, 2002; see also chapter 5) affords unexpected insights that would be difficult to ascertain based on discourse analysis alone, yet can miss the benefits of qualitative, ethnographic methods that afford consideration of individual variation and actual, as opposed to idealized, usage. Consideration of individual usage (Yates, 2005; see also Adolphs & Durow, 2004, and Belz 2004; 2006; see ch. 7) can counterbalance the distortion brought on by frequency analysis of a small corpus with a low number of tokens or contributors. Discourse analytic approaches (Bardovi-Harlig & Hartford, 1990; 1993; Thonus, 1999; Farr, 2003; Dalton-Puffer & Nikula, 2006) benefit from the naturalistic, contextualized quality of the data, and thus retain strong ecological validity. These approaches also allow for grounded techniques as a means to categorize data. On the other hand, approaches whose analytical frameworks are grounded in pragmatic politeness theory (*ibid.*) or functional linguistic theory (He, 1993, 2000; Iedema, 1996) have ontological and heuristic strength, especially if that theory has direct relation to the object of analysis. In view of these observations, an ideal methodological design for the current project uses both corpus analysis and discourse analytic techniques to look at both groups and individuals, using both quantitative and qualitative data, with a theoretical framework grounded in linguistic and pragmatic theory. Such a mixed methods approach is outlined in chapter 3, and fully described and applied in the main chapters.

### 2.3 Application to pedagogy

One final aspect to review in regards to studies of institutional discourse is the application to pedagogy that the study advocates or describes. Most of the studies above are research-oriented, and do not go further than description of the phenomenon to apply findings to the L2 classroom<sup>3</sup>. If pedagogical implications are mentioned at all, they are usually relegated to conclusions. While studies like Bardovi-Harlig and Hartford (1993), Iedema (1996), and Hyland (2002) do not explicitly offer pedagogical implications, they may be drawn through inference. In contrast, McEnery and Kifle (2002), Farr (2003), and Harwood (2005) offer more explicit implications, although they are not necessarily the express end goal of their studies.

In their study of NNS student advising interactions, Bardovi-Harlig and Hartford (1993) call for more research and instruction, noting “without explicit teaching and without necessary input, it is little wonder that NNSs learn to make suggestions before they learn the appropriate linguistic form for those suggestions” (p. 301). They note that because of this lack of input, NNSs are slow to use mitigators and avoid use of aggravators. As their focus is more on research, they do not offer any specific pedagogical implications. Similarly, Iedema (1996) does not offer specific implications, he concludes that for successful participation in school and work communities, individuals “need to be able to recognize how linguistic constructions realize different degrees of status and institutional familiarity, and they need to be able to respond ‘appropriately’” (p. 99). This involves “being aware of the various linguistic realizations (whether indirect or incongruent) through which social positioning occurs” (p. 99), and being aware of how those realizations background “or even suppress the identity of the source of control” and thus “enhance the authority implicit in the teacher’s pastoral role” (p. 100). While

---

<sup>3</sup> There have been several pedagogical studies of interlanguage pragmatics and development using corpus approaches that are reviewed in chapters 7 and 8, specifically Belz and Vyatkina, 2005 and Belz, 2006. These studies are not reviewed here because they are not specific to institutional discourse but rather L2 development.

this may be seen as a general endorsement of language awareness education, Iedema does not specify any particular approach, nor whether or not this awareness would involve use of analytic tools like those he used. This stance on awareness is similar to that of Hyland (2002), who concludes in his corpus analysis of the use of textbooks in academic writing that writers use directives in ways “intimately related to their assessments of appropriate reader-relationships in different generic and disciplinary contexts” (p. 236). While he does not make any pedagogical suggestions, his stance implies that awareness of this assessment process would be a goal in the instruction of directives usage, written or spoken.

Several authors offer explicit implications for pedagogy based on their findings. McEnery and Kifle (2002) concluded that their results on the underuse of epistemic devices by learners could be accounted for by the nature of the textbooks the learners had used. As a pedagogical implication, they mention instructor use of learner corpus analysis to customize instruction for immediate learner needs, including the use of comparative analysis. They note that treatment of epistemic devices in textbooks is usually syntactic in nature and not functional, and that modal forms are favored over lexical, adjectival, and noun forms. Similarly, Farr (2003) concludes with discussion of pedagogical implications in her exploration of the use of listenership devices in NS tutor-tutee interactions. She suggests that listenership be explicitly taught in academic (EAP) speaking courses as part of a productive listening skills unit. She advocates familiarizing students “with many spoken academic genres as each has its own characteristics and demands” (p. 82), as well as with production by speakers from a variety of backgrounds<sup>4</sup>, and from contexts that are locally relevant for learners. She also advocates corpus-based instruction, as it can allow for “the critical investigation of almost any aspect of language related issues” (p. 83), thereby improving

---

<sup>4</sup> MICASE, the source of the Expert Office Hours subcorpus used for the current project, seems to have been designed with this diversity in mind—see §3.3.2 and §6.3.1.

language awareness. Also quite explicitly, to the point of even suggesting specific activities, Harwood (2005) provides explicit pedagogical implications in his study of the use of inclusive and exclusive ‘we’ in academic writing. He promotes a corpus-based approach, and advises that teachers and classes create their own corpora and corpus-based activities. These activities include student examination of their own and others’ texts, consideration of collocations and phraseological units, and discussion of reader-writer relationships and differences in lexical paradigms<sup>5</sup>.

---

<sup>5</sup> Research on corpus-informed approaches is discussed in chapter 8.

## CHAPTER 3

### METHODS AND DATA

This chapter presents a brief overview of the variety of methods used in the five studies that comprise this dissertation. While the studies are presented separately, they are interrelated and cumulative in nature, and thus presented in an intentional order. The overview is followed by an extensive discussion of the source for learner data I designed and began developing, ITAcorp, a corpus of advanced spoken learner English still under development at Penn State as of this writing. I also discuss the source of the expert data, MICASE, the Michigan Corpus of Academic Spoken English (Simpson, Briggs, Ovens, & Swales, 2002). The chapter concludes with an argument for the use of role play data, which have been criticized for use in interlanguage pragmatics research because its effects are inconsequential to the role player (Bardovi-Harlig & Hartford, 2005). My argument is that, because the consequences of the role play tasks are high stakes and relevant for the ITAcorp participants, very authentic performances can be expected.

#### **3.1 Mixed methods approach**

Considering the breadth of the main research question and the variety of sub-questions in chapter 1, along with the research reviewed in chapter 2, a mixed methods approach is used for this project. Instead of one large and confusing explication of methods, followed by an equally unwieldy series of analyses, each of the descriptions below is only an overview of the methods used for the analyses in each study, for the sake of reader orientation. More in-depth discussions of the methods used are in that corresponding chapter.

### **3.1.1 Chapter 4 methods: Directive language as an object and unit of analysis**

In the first study, I explore the concept of directive language to the end of defining an object and operationalizing a pragmatic unit of analysis for use in subsequent studies. To identify these units, I conduct a semi-grounded analysis of the data, specifically 20 transcripts from the learner office hour (LOH) data, and 2 transcripts from the expert office hour (EOH) data. I term the analysis ‘semi-grounded’ because I understood to a degree that directives would exhibit directive illocutionary force (Searle, 1983), and would most likely be indexed by modals like ‘must’, ‘should’, and ‘have to’. Still, by examining the data, the concept of ‘directive’ emerges in unexpected forms. These forms are then grouped into categories of units, which are then discussed and supported with primary source linguistic and pragmatic theory. Finally, a ‘directive construction’ unit of analysis is proposed, grounded in a model of social-functional linguistic theory.

### **3.1.2 Chapter 5 methods: A corpus analysis of directive construction usage in office hours**

In the second study, I explore the systematic qualities of the ‘office hours’ text type as register through the use of corpus analytic techniques to examine and compare the frequencies of the directive constructions proposed in chapter 4 in both the LOH corpus from ITAcorp and the EOH corpus from the MICASE (Simpson et al., 2002). Corpus analysis involves the application of computer technology to count, collate, and examine the distinctive distributional and clustering patterns that typify a linguistic register (Biber, Conrad & Reppen, 1998), focusing on frequency and association patterns, “the systematic ways in which linguistic features are used in association with other linguistic and non-linguistic features” (p. 5). Using corpus analytic techniques, I first examine the register consistency of the two subject corpora to determine their

validity by analyzing usage distribution of the proposed directive constructions in the corpora. To this end, I conduct an internal comparison on the LOH corpus by splitting it in half and comparing the frequencies of the constructions in both halves. If the corpus is consistent and the constructions are more-or-less evenly distributed, there should be little or no significant difference, as measured by a log-likelihood analysis (Rayson & Garside, 2000). I conduct the same procedures on the EOH, and also compare it to a genre comparable external corpus, the MICASE advising (ADV) corpus. If there is less significant difference between the two halves of the EOH corpus than between the EOH and the ADV corpus, the EOH corpus can be considered register consistent. These analyses also create profiles of directives usage in the EOH and LOH corpora, which are then compared using contrastive interlanguage analysis (Granger, 1998; 2002), to determine the constructions with significantly different frequencies. These differences are then discussed using the social-functional theoretical model proposed in chapter 4 as an interpretive heuristic. Finally, adjunct use in the two corpora, including mitigators and intensifiers, are compared using contrastive procedures, and the findings are again interpreted using the social-functional model as a heuristic.

### **3.1.3 Chapter 6 methods: A genre analytic view of office hours**

In complement to chapter 5, where I examine the systematic qualities of directives usage in the office hours register using corpus analytic techniques, in chapter 6 I explore office hours and directives usage using genre analytic techniques that focus on the socially-defined qualities text and context. I first conduct a contextual analysis of the ‘office hours’ context through examination of a variety of features specific to the social participation of ITAs in these contexts, specifically the name of the genre, its social context, its communicative purposes, participant

roles, cultural values, and text context. To complement this situational approach, I then conduct an empirically driven genre moves analysis (Swales, 1990) to identify the moves in an expert office hours session, including diagnostic and directive moves (Agar, 1985).

### **3.1.4 Chapter 7 methods: Directives usage profiles of individual experts and learners**

To balance the register and genre-based analyses of group usage in chapters 5 and 6, in chapter 7 I offer usage profiles of individual expert and learner participants, for the purpose of exploring variation in usage and examining the relation of individual use to the exercise of academic authority. To analyze variation and distribution, I examine the usage profiles of the most frequent constructions produced by the 8 expert MICASE participants and 8 of the learner participants. The 8 learners were chosen from the consenting participants to assure an evenly-distributed representation of high to low frequency users. In addition, the directives usage of three of the 8 learners is profiled in depth, using biographical, attitudinal, and assessment data, including a post-semester interview and assessment data as well as a biographical survey.

### **3.1.5 Chapter 8 methods: An application of corpus-informed language awareness instruction**

The final study presents the theoretical background to, design of, implementation of, and efficacy analysis of experimental instruction on directives usage. I first discuss background elements to situate the instructional unit, including a review of the current state of directives instruction in representative ESL grammar textbooks and ITA training materials, a summary of recommendations from previous studies, and a discussion of the corpus-informed language awareness approach on which the design of the unit is based. After the background discussion

and presentation of the original unit on directives instruction, I analyze the efficacy of the instruction by presenting an experiment conducted in three sections of an ITA preparation course in Spring 2006 (see §3.2.1 below). Two sections were experimental (n=10+10), and one section was control (n=13); to control for teacher effect, the same teacher taught one of the experimental sections and the control section. A simple contrastive corpus analysis is then conducted on the post-implementation usage of directives by the sections in three office hours role play activities, to determine the extent and nature of instructional effect.

### **3.2 Data**

The following section contains a description of ITAcorp, the source of the Learner Office Hours (LOH) corpus used in the analyses, and includes an explanation of its development and composition, which is based on the spoken and synchronous computer-mediated activities completed in ESL 118, American Oral English for ITAs III, the advanced level ITA training course at Penn State University. An explanation of the corpus data gathering, preparation, transcription, and standardization procedures, including transcription conventions, is followed by an explanation of the structure of the database that houses the corpus and of the corpus variables that can be used as metadata search parameters. A description of the current status of the corpus is followed by a brief description of the subject corpus derived from ITAcorp, the LOH corpus. A description of the baseline corpus used for the development of the Expert Office Hours (EOH) corpus, the Michigan Corpus of Academic Spoken English (MICASE) (Simpson et al., 2002) is also presented.

### 3.2.1 ITAcorp

The data for this project are located in ITAcorp, a corpus of advanced spoken learner English under development at Penn State. After gaining experience developing Telecorp, a telecollaborative learner corpus of English and German (Belz, Reinhardt, & Rine, under review), I initiated the design and development of the corpus in Summer 2005 as a means for data collection for the current project, with the assistance and cooperation of the curriculum coordinator and instructors of ESL 118, American Oral English for ITAs III. Data collection began in Fall 2005, and the project received Penn State IRB approval as part of project #21429. Informed consent was requested from all ESL 118 course participants, but two in the Fall 2005 cohort and one in the Spring 2006 cohort declined. In Fall 2006, the corpus was transferred to the auspices of the Penn State Center for Language Acquisition, where it will be used for a variety of research purposes and eventually made publicly available. Plans are to continue development of the corpus for several years and expand data types.

Table 3.1. Number of participants, sections, and tokens per cohort in ITAcorp

cohort	no. of partic.	no. of sections	no. of tokens
Fall 2005 (F05)	40	4	43295
Spring 2006 (S06)	32	3	59511

#### 3.2.1.1 ITAcorp development

The pedagogical framework of ESL 118 informed the design of the corpus, as it determined the activity types that provided the data to populate the corpus. The course is an oral skills development course for international graduate students who either advanced from another course or were placed there by the Penn State American English Oral Communication Proficiency Test (AEOCPT), which all non-American graduate students at Penn State must pass

in order to become teaching assistants, per legislative mandate<sup>1</sup>. The goal of ESL 118 is to develop spoken academic discourse competence in the course participants, especially in the genres of advising and lecturing. At the end of the course, in lieu of the AEOCPT, the participant takes a face-to-face post-evaluation exam, conducted by two external examiners, comprised of an open-ended question about teaching, an office hours role play with an examiner playing the student role, and a lecturette presentation which the participant prepares beforehand.

To meet course goals and prepare participants for the post-evaluation exam, the course activities included a face-to-face office hours role play and two lecturette presentations in front of the class, all three of which were recorded onto mini DVD-R discs for the participants. In addition, I initiated and facilitated three computer-mediated activities, each preceding a corresponding face-to-face activity, to provide practice for the in-class version, and to allow for the production of computer-mediated data. Finally, for the Spring 2006 cohort, the three face-to-face post-evaluation activities were also added to the corpus, after experimental implementation of a unit of instruction for 2 of the 3 sections (see chapter 8). A description of each activity is provided in Table 2.

Table 3.2. Descriptions of ESL 118 FA05 and SP06 activities entered into ITAcorp

<p><b>A1: office hours role play (chat)</b> In this online synchronous chat activity, participants alternated playing the role of TA and student for 10-minute sessions, each session with a different partner. Each TA was assigned an ‘office’ or separate chat room, where they would advise a ‘student’ who had been provided with a role play scenarios (see appendix A), reflecting problems common to American undergraduates. I developed the scenarios in conjunction with the course instructors.</p>
---

---

<sup>1</sup> More information on the Penn State ITA Program and the AEOCPT are available on the Penn State Department of Applied Linguistics website.

**A2: office hours role play (in class)**

In this face-to-face in-class activity, participants took turns playing the role of TA while another participant played the role of 'student' using a scenario that s/he had invented. Each 'office hours session' took between 5 and 8 minutes and was recorded onto a DVD-R disc to provide the 'teacher' with a means of self-evaluation. For simple feedback purposes, the course instructor and other participants also evaluated the 'teacher' using an evaluation rubric.

**A3: concept presentation (chat)**

In this online synchronous chat activity, participants alternated playing the role of TA explaining a concept to a 'class' of 2 or 3 other participants playing the roles of students in a 'presentation' lasting 15-20 minutes. The TA was instructed to use interactive discussion involvement strategies learned in class, and 'students' were instructed to answer and ask questions.

**A4: lecturette presentation (in class)**

In this face-to-face in-class activity, participants each presented an 8-12 minute lecturette on a concept from their field of their choice, using presentation strategies discussed in class. For self-evaluation purposes, the presentation was recorded onto a DVD-R, while the course instructor and other participants evaluated the presenter for feedback purposes.

**A5: TA conversation (Fall 2005) or article discussion (Spring 2006) (chat)**

This online synchronous chat activity was different for the two semesters. In Fall 2005, participants were grouped into 'teachers lounge discussion' groups, where they were instructed to discuss their perceived strengths and weaknesses as TAs. In Spring 2006, participants lead 10-minute discussions as 'teachers' with 2-3 'students' on one of 4 one-page articles on current topics that all participants had been provided with earlier.

**A6: interactive lecturette presentation (in class)**

This face-to-face activity similar to A4, participants again each presented an 8-12 minute lecturette on a concept from their field of their choice, focusing on involving the 'student' audience. For self-evaluation purposes, the presentation was again recorded onto a DVD-R, while the course instructor and other participants again evaluated the presenter for feedback purposes.

**A7: open-ended question (post-evaluation)**

This post-evaluation exam activity was conducted face-to-face by examiners, who presented the participant with an open-ended question about the participant's attitude towards teaching and learning. Responses ranged from 3 to 10 minutes.

**A8: office hours role play (post-evaluation)**

This post-evaluation exam activity was also conducted face-to-face by examiners, who presented the participant with a role play scenario in which a 'student' played by one of the examiners would come to the participant playing a TA. The scenarios, developed by the ESL 118 course coordinator, reflected problems typical of those faced by American undergraduates. The duration of each role play was between 5 and 12 minutes.

**A9: lecturette presentation (post-evaluation)**

In this face-to-face post-evaluation exam activity, participants each presented a 10-15 minute lecturette on a concept from their field of their choice to the examiners, who asked questions as if they were the participant's students.

**3.2.1.2 Data gathering and preparation**

For the face-to-face activities (A2, A4, A6, A7, A8, and A9), data gathering involved collecting the DVDs, extracting the .wav audio files from each using freeware, assigning file IDs, and archiving the files digitally. The files were then transcribed, converted to MS Excel files, and imported into FileMakerPro. For the chats (A1, A3, and A5), data gathering involved copying and pasting the chat transcripts into MS Excel files and cleaning the data so that each row included the metadata associated with the data, including chat ID, scenario ID, participant, role, time of turn, and turn number (see table 3 below).

Each instance of an activity completed by a participant (in the teacher role) was considered one file, and a coding scheme was developed so that file had its own identification code or file ID. Each semester was assigned a letter (F=Fall, S=Spring, M=Summer) followed by a two-digit year (e.g. 05, 06) and a number identifying section (e.g. 1, 2, 3, 4, etc.). Each student was assigned a number, unique for the entire semester (e.g. in S06, students were numbered 01-32). All activities (A) were assigned a number (1-9), as in table 2. For example, according to this system, file S06314A4 is Spring 2006, section 3, student 14, activity 4. For the chats, because

participants role played a teacher multiple times, the student number is not used in the file ID but rather a random number, assigned in order the chat was entered into the corpus.

Because the current and possible future projects involve qualitative and socio-cognitive approaches towards data analysis, including longitudinal studies of individual development, all data is associated with meta-data on the participant who produced it (Belz, 2002). As per IRB requirements, individual identifiers were removed and pseudonyms were assigned, using a scheme adapted from that used for Telekorp (Belz, Reinhardt, & Rine, under review). This procedure was different depending on the participant's national origin. For Chinese names, the second syllable of the given name was used with the first initial of the family name, so that Mao Zedong would have been Dong M. In the case of all other names and single syllable Chinese given names, a pseudonym beginning with the same letter was randomly chosen, so that Cesar Chavez might have been Carlos C.

### **3.2.1.3 Transcription and standardization**

A team of Center for Language Acquisition research assistants transcribed all the corpus activities; I transcribed most A2 and all A8 activities, since these two activities and A1, which was chat-based and thus needed no transcribing, were the source of project data. Transcription involved listening to the files and transcribing the spoken data according to conventions adapted from MICASE conventions (table 3.3)<sup>2</sup>. Theoretically, each data unit can be associated or tagged with infinite meta-data, so that each sound, word, phrase, utterance, turn, or larger discourse unit can be associated with speech quality information like tone or pitch and information like speaker and task. Practically speaking, however, the logistics of time and expertise prohibit close

---

<sup>2</sup> While the transcription conventions were adhered to as best as possible, individual transcriber discretion was necessary on many features, for example turn boundaries, repeated utterances, overlapping turns, intonation contours, and length of pauses. Because this is an issue with any project involving multiple transcribers, I transcribed most (80%) of the data used for the current project.

transcription of the amount of data required for corpus analytic purposes. For the purposes of the current project, a small corpus of at least 100,000 tokens to match the expert corpus was desirable, the close transcription of which would have taken more time and resources than were available.

Transcription conventions (table 3.3) were adapted from MICASE conventions, to allow for maximum ease of data comparability between the two corpora. Transcriptions were saved in text files and then converted into Excel files so that each row corresponded to one turn and included the metadata associated with the data, including file ID, participant, role, and turn number. Data population involved exporting from Excel files and importing into a FileMaker Pro database.

Table 3.3. ITAcorp transcription conventions

- Regular spelling is used, not phonetic, except for ‘gonna’ and ‘wanna’, if there is noticeably elision.
- Every word and utterance is transcribed, even repetitions.
- Each participant is marked consistently as: T, S1, S2, CLASS, etc.
- Each turn starts with (speaker):\_\*  
T: yadda yadda
- The first word of sentences is not capitalize. Only ‘I’ and proper nouns are capitalized.
- Numbers are spelled out with dashes.
- Except for numbers, dashes are not used; a space is used between words that usually have dashes.
- Turn boundaries are determined by floor taking and holding.
- Inter-turn feedback is marked with brackets: [(speaker):\_xxx xxx]  
T: blah blah blah [S: yadda?] blah blah
- Overlapped turns are marked with **bold**.  
T: yadda yadda **yadda** [S: **blah**] yadda
- Laughter or other sounds are marked with a verbalization of the sound in capital letters, in carrot brackets: <SOUND>  
[CLASS: <LAUGH>] S: <COUGH>
- Falling sentence final-like intonation is marked with a period: .\_  
T: so. tell me about your day.
- Rising intonation (not necessarily all questions) is marked with ?\_  
S: do you like it? or not.
- A one-two second pause is marked with a comma: ,\_

S: um, I don't know

- A three-four second pause is marked with an ellipsis: ...\_

S: uh... I'm not sure

- A five second or longer pause is marked with a capital letter P followed by a colon and the number of seconds, in carrot brackets: <P: ##>\_

T: let me see <P: 07> well, I I think

- Common feedback or smallwords: mhm, uh huh, yeah, uh, um, mmm

\* underscore represents a single space

Because the corpus contains both computer-mediated and spoken data, turns were chosen as the basic data unit for the corpus, so that each record would correspond with one turn.

Computer-mediated discourse is forced into linear sequential turns, and since it is the actual medium of interaction for those activities, it is desirable to keep it in that form for analysis.

Spoken data, however, poses more of a challenge, and could theoretically be stored as words, phrases, or sentences, etc. Because the spoken data is mostly conversational and needed to use the same unit as the computer-mediated data, turns were seen as an acceptable basic unit, except that the corpus would then be unable to adequately represent overlapping or subtle distinctions like when an interjected turn was actual floor-taking or simply interjected comment or feedback (although this is often the subjective decision of the transcriber). An adequate compromise was reached by retaining this information in an unadulterated transcript file associated with the parsed turn-based data units. This data can be saved for later export into more sophisticated parsing and analytical tools than were available for the current project.

#### **3.2.1.4 Database structure**

The corpus data are archived in a FileMakerPro database, because of the software's versatility and customizable design, search, and import/export features. The database is relational and contains three tables: a main data table containing production data from all the activities, a

transcripts table containing the full unparsed transcripts of the face-to-face activities and a participant table containing data on each participant, comprised of their answers to a course biographical survey (appendix B)<sup>3</sup>. The main data table contains the production data from all activities so that each database record corresponds to a single turn. Each record/turn contains the main data, participant role (T=teacher, S=student, E=evaluator), file ID, turn number, and time of turn, which was only available for computer-mediated activities. The transcripts table contains the full unparsed transcripts of the activities (so that the main data field corresponds to an entire transcript), participant names corresponding to participant roles, the file ID, and the role play scenario ID if part of that activity. The participant table contains data on all participants, including their pseudonym, participant ID, name, age, gender, degree, and country of origin. It also includes data that participants completed on a student information survey on language learning and computer use experience (see appendix B), which as of the S06 cohort included a section on attitudes towards teaching (see appendix C). Table 4 contains a list of all data and meta-data fields, which correspond to corpus variables and search parameters. Researchers can use the database by first setting the search parameters by specifying any of the data fields. A search then finds all records containing the desired item. Any or all of the found data can then be exported as a subcorpus into comma or tab-delimited text, which can then be imported into Excel, WordSmith, or other corpus software for further analysis.

---

<sup>3</sup> This design is based on my work re-designing Telekorp, the telecollaborative corpus of German and English (Belz, Reinhardt, & Rine, under review).

Table 3.4. Data variables in ITAcorp

main data table	transcripts table	participant table
data file ID participant role (T or S) turn number time (chat only)	file ID participant 1 (T role) participant 2 (S role) scenario ID full transcript	pseudonym participant ID given name family name age gender degree dept country (30 variables regarding language learning history and computer experience— appendix B, and attitudes towards teaching— appendix C)

### 3.2.1.5 Corpus status

Table 5 shows the number of tokens in each activity subcorpus of the subject corpus, the unevenness reflecting the problems and difficulties of data gathering. One problem was that the Fall 05 cohort recordings suffered from poor sound quality, as microphones were not used. This was rectified for subsequent cohorts, except that occasionally a microphone would not function or stop functioning halfway through a recording. The post-evaluations for Fall 05 are not yet available because they were recorded onto audiotape and must be digitized and segmented. Lessons learned from these issues will be applied to future semester cohorts as the corpus is augmented.

Table 3.5 Number of tokens in each ITAcorp activity subcorpus by section

section <sup>4</sup>	A1	A2	A8	total
F051	4399	5196	(3)	9595
F052	4648	(1)		4648
F054	6726	8756		15482
F055	5485	8085		13570
S061	10023	7822	3256	21101
S063	(2)	8277	3620	11897
S064	6434	13616	6463	26513

(1) Data is not available due to poor sound quality of recordings.

(2) Sections 1 and 3 were combined for the computer-mediated activities (A1), and are recorded in the S061 row.

(3) Post-evaluation data was not gathered for the Fall 05 cohort.

### 3.2.1.6 Subject corpus

Data from the F05 and S06 cohort production in A1, A2, and A8 are considered the learner office hours (LOH) corpus, which will be subject to analysis in the current project. Separately, these three subcorpora will be compared to each other, for analysis of the influence of pedagogical treatment (A1+A2 vs. A8; see ch. 8). Together, they will be compared with the EOH. Table 6 shows the total number of tokens in the subject corpus according to activity, including the total number of tokens spoken by participants when playing faculty or TA (T) roles and when playing student (S) roles. Number of encounters indicates the total number of enacted role plays, or role-played encounters between one T and one S<sup>5</sup>.

<sup>4</sup> Ordinal inconsistencies with the section numbers originate from university registrar and scheduling; they are maintained here for simplicity's sake.

<sup>5</sup> It should be noted that in A1 all participants played each role several times, that in A2 each participant played the T role once, and that A8 only contains the post-evaluation role plays of most, but not all, of the S06 ESL 118 participants.

Table 3.6. Number of tokens, roles, and encounters in the Learner Office Hours corpus derived from ITAcorp by activity

	A1	A2	A8	LOH total
total tokens	37715(2)	51752(3)	13339	102806(2,3)
faculty or TA (T role)	18417	28361	7810	54588
students (S role)	17442	23305	5529(1)	46276
number of encounters (files)	112	73	27	212

(1) All A8 S roles were played by external evaluators, not ESL 118 student participants

(2) Total includes 1856 tokens of off-task interaction in A1

(3) Total includes 86 tokens produced by course instructors in A2

### 3.2.2 MICASE

For the current project, the Michigan Corpus of Academic Spoken English (MICASE) (Simpson et al., 2002; Simpson-Vlach & Leicher, 2006) was utilized as baseline data for comparison with the learner data<sup>6</sup>. MICASE is comprised of the transcriptions of 152 academic speech events from a variety of text types<sup>7</sup>, e.g. advising, colloquia, discussion sections, lectures, office hours, and tutorials, from a balance of university academic disciplines. It contains approximately 1,695,540 tokens spoken by 1571 individuals. Corpus development began in 1997 and was completed in 2002. The purpose of the corpus is fourfold: one, to fill a gap in available corpus resources, since no spoken academic corpus has existed previously; two, to develop descriptions of language features that may diverge from current explanations of grammar and vocabulary, since those explanations may have been based on intuition and written language; three, to analyze development of academic speech patterns over time; and four, to allow for the development of more appropriate, corpus-informed ESL and EAP materials than are currently available. Participants include 160 faculty, 257 graduate students, and 782 undergraduates, and are each identified with the metadata categories of rank (senior faculty, graduate student,

<sup>6</sup> The database is free and publicly accessible on the WWW. The developers were informed of the current project and will be provided with a copy of this dissertation upon completion.

<sup>7</sup> These could be considered 'genres' or 'registers'. See chapters 5 and 6 for relevant discussion.

undergraduate student, etc.), age-group, gender, and native-speaker status (native, near-native, and non-native). The corpus was accessed through its web interface, which allows searching for words or phrases and co-text with any combination of metadata parameters.

MICASE was chosen for the current project because of the genre compatibility of its office hours (OFC) subcorpus (120,629 tokens) with ITAcorp. For the purpose of the project, this subcorpus would serve as the baseline Expert Office Hours (EOH) corpus for the register analysis of directives usage (chapter 5), as well as the source of the transcripts for the semi-grounded analysis (chapter 4), genre analysis (chapter 6), and pedagogical materials in the intervention (chapter 8). It should be noted that while restrictions on data because of lack of consent and genre compatibility issues limited the use of some transcripts for these individual usage profiles (chapter 7), all of the data in the subcorpus produced by those in a T role (faculty or TA) was used for the register analysis (chapter 5). MICASE is organized so that each transcript file is one faculty or TA's office hour session, comprised of multiple encounters with individual students stopping in. Because of the structure of the MICASE web interface, it was prohibitively difficult to determine the number of tokens produced by those in T roles. In the statistics made available by the MICASE designers, graduate students, even if in a TA role, were counted as students and lumped together with undergraduates, so those statistics are not used here. Because of the limited number of transcript files, however, it was possible to determine the number of participants assuming T and S roles, as well as the total number of encounters. The individual EOH participants are discussed in more detail in chapter 6.

### **3.3 Data quality**

I maintain that the LOH role play data exhibits qualities of educational institutional discourse and is valid as a data type for this kind of research. Bardovi-Harlig and Hartford

(2005) discuss the merits of institutional talk as a source of data for interlanguage pragmatics research. They note three qualities of institutional talk, defined as interaction between an institutional representative, e.g. a teacher, interviewer, or doctor, and a client, e.g. a student, a job applicant, or a patient, that contribute to the comparability of multiple sources. First, it involves restricted goal, task, or identity orientation; second, it may involve mutually understood constraints on allowable contributions; and third, it may involve contextually dependent inferential frameworks. Institutional interactions “often include expected norms of interaction such as turn-taking, constant social relations/roles, and asymmetrical power relationships” (p. 9). Regarding role play data, Kasper and Dahl (1991) note that of all the sources of interlanguage pragmatics data, including discourse completion tasks, role plays, and observation of authentic speech, role plays offer the replicability of discourse completion as well as the open-ended authenticity of spoken interactions. Role plays, if designed to be open-ended, allow “examination of a speech act behavior in its full discourse context”, including naturalistic turn-taking, impromptu planning decisions, and negotiation of meaning (p. 228).

Bardovi-Harlig and Hartford (2005) also discuss the relative merits and drawbacks of three widely used data sources, conversation, discourse completion tasks, and role plays, which all suffer some limitation that authentic institutional interaction does not. Conversation is interactive and consequential yet difficult to compare, discourse completion tasks and questionnaires are highly comparable but neither interactive nor consequential for the participants, and role plays and simulated tasks are interactive and comparable, but not so consequential. Role plays are conducted between speakers who have no social connection, and thus have little consequence beyond the task. In line with this argument, a criticism of the current project might be that the LOH role plays would be inconsequential to the participants and thus

any observed generic qualities invalid, even though the EOH participants in the current project, as actual faculty, grad students, and undergraduates, both native and near-native speakers, were engaged in unquestionably authentic discourse.

In reply to this criticism, it is maintained that for this project, the pedagogical embeddedness of the role play data actually enhances, rather than detracts from the goal of the research to effect pedagogical change and result in positive assessment outcomes for students. Whether or not they are producing ‘authentic data’, students are assessed by their performance inside class, not outside class. Granger (2002) explains that nearly all learner production in a classroom “usually involves some degree of ‘artificiality’ and that learner data is therefore rarely **fully** natural” (p. 8; scare quotes and bold in original). There are different degrees of authenticity in classroom learner production, but nearly all are bounded by task variables to varying degrees. If classroom role play data with authentic consequence must be excluded because it is artificial to a degree, then the authenticity of all classroom data and validity of performance assessment must be called into question. In any case, for the LOH corpus participants enrolled in ESL 118, their performances in the role plays have real consequences, i.e. whether or not they are allowed to assume TAs duties and continue their academic careers, even if the role plays contain a degree of inauthenticity. In addition, the role play simulates a task they will most likely be participating in similar office hours situations as TAs and faculty in the future. Thus, because the role play task content is relevant and the consequences are high stakes for the ITAcorp participants, I maintain that very authentic performances can be expected.

## CHAPTER 4

### **DIRECTIVE LANGUAGE AS AN OBJECT AND UNIT OF ANALYSIS**

In this chapter, the concept of directive language is explored to the end of defining an object and operationalizing a pragmatic unit of analysis for use in subsequent studies. For corpus analytic purposes, this unit must be quantifiable into discrete units, while for purposes of qualitative analysis, it must be relatable to interpersonal function. To identify these pragmatic units, I conduct a semi-grounded analysis of transcripts of the subject data. This analysis results in a list of units that are then grouped according to grounded analysis into conceptual categories. The proposed units and categories are then discussed and supported with primary source linguistic and pragmatic theory (Searle, 1976; Brown & Levinson, 1987; Blum-Kulka, House, & Kasper, 1989; Scollon & Scollon, 1995; Palmer, 2001; Halliday & Matthiessen, 2004). Finally, a usage-based (Langacker, 2000; Tomasello, 2003) ‘directive construction’ unit of analysis is proposed, grounded in a social-functional linguistic model of usage.

#### **4.1 A grounded analysis of directive language**

The relationship of pragmatic force to the lexicon is not discrete and clearly one-to-one. The studies reviewed in chapter 2 came to certain conclusions, reviewed here for reader convenience, that 1) directives are pragmatic speech acts that express directive illocutionary force; 2) in institutional discourse<sup>1</sup>, directives may be a means of socialization, persuasion, and the exercise of academic power; 3) the grammatical form of a directive may vary widely, although it is usually centered on a verbal element that often contains a modal operator; and 4) a

---

<sup>1</sup> Directives have this quality in other macrogenres that are not the object of investigation here, for example political, medical, and parenting discourse.

speaker may mitigate or strengthen the force of a directive by various lexico-grammatical means in the exercise of power. In consideration of these findings and the differences between the subject genre of office hours and those in the reviewed studies, which used units of analysis defined with *a priori* frameworks (see §2.1), a semi-grounded approach was chosen to identify a unit of analysis for ‘directive’. Semi-grounded refers to the fact I do not deny that the reviewed research and a certain amount of previous knowledge influenced the analysis. I maintain that a unique conceptualization of ‘directive’ and proposed unit of analysis emerged from the data informed by this knowledge.

#### **4.1.1 Method**

An arbitrary sample of twenty transcripts of Fall 2005 class office hours role plays from the LOH corpus (all of F054A2) and a sample of two transcripts from the EOH corpus (115 & 300) were selected for analysis. Because large amounts of data are desirable for corpus analysis, and it is physically impossible to read an entire corpus<sup>2</sup>, these transcripts total only approximately one-eighth of the data. The transcripts were read, and words or phrases with directive force were underlined, the concept of ‘directive force’ informed by the literature and understood generally as “where we try to get others to do things” (Searle, 1983). Requests, invitations, and offers were not noted, as they are considered qualitatively different in the urgency of speaker stance to get the listener to act (see also Hyland, 2002, p. 216). The units resulting from the reading were grouped into meaningful categories that correspond to devices on three levels: core lexical (usually verbal) devices, syntactic-level devices including adjuncts and subject (usually pronominal) devices, and discourse-level devices, specifically the use of propositional irrealis. To the end of defining a unit of analysis, after presentation of exemplars of

---

<sup>2</sup> Reading an entire corpus would in fact defeat the purpose of corpus analysis.

each proposed category, primary source grammatical and pragmatic theory is presented that supports considering these devices to be ‘directive language’.

#### **4.1.2 Results**

The unit categories are presented below in list format, with up to four randomly selected exemplars for each category, chosen simply to represent the variety of usage. Unless there are less than four, the exemplars listed are not all the directive language found in the transcripts, which again represent only about 1/8<sup>th</sup> of all the data. The number four was chosen simply because of space constraints. Unless a bullet point precedes a given exemplar, a single line is equivalent to one exemplar.

##### **4.1.2.1 Core lexical devices: Modals, directive vocabulary, & imperatives**

The reading identified several categories of lexical-level items that directly index directive force, including modals and periphrastic modals, irrealis devices, imperative forms, and directive vocabulary including verbs, nouns, adjectives, and adverbs. It is proposed that these devices be the core part of the unit of analysis.

###### **4.1.2.1.1 Modals & periphrastic modals**

Modals and semi- or periphrastic modals (p-modals) were the most obviously common directives in both corpora. Besides traditional modals of obligation (must, should, have to, have got to/gotta, had better), there were modals of necessity (need to), suggestion (might, could, can), and permission (can). ‘Gonna’ as a commissive was rare in the samples, and one p-modal-like unit with definite directive force that stood out, especially in the EOH sample, was ‘want to’. The subject of the modals was usually ‘you’, and was sometimes followed or preceded by an

adjunct like ‘probably’ or ‘I think’. The apparent differences in frequency of usage between the learner and expert samples suggest a productive comparative corpus analysis.

Table 4.1. Modal and p-modal exemplars from LOH and EOH data

modal or p-modal	exemplars from LOH	exemplars from EOH
have got to ('ve gotta)	(none noted)	you gotta be a little more specific you gotta start writing right away you've gotta figure out you've gotta include sentences
can	you can bring your proposal you can come back and ask me you can go to the library you can't borrow them	you can just cut that you can make the point about you can sort of look back over you can use them in your analysis
could	you could finish visual documents	you could also ask them you could have a paragraph you could just use what's here you could tie this need
be going to ('re gonna)	(none noted)	you're gonna have to do some you're gonna need to keep this you're gonna wanna give some
had better	you'd better do that you had better reading other	(none noted)
have to	you have to arrange it earlier you have to derive the equation right you have to include the procedure you have to overcome impressionism	you almost have to read aloud you have to repeat that you have to tell me you'll have to change that
might	(none noted)	that you might do to help yourself you might even rewrite these you might even tell me you might wanna think about
must	you must be very careful you must read about the questions	(none noted)
need to	you don't need to draw you don't need to finish you need to make a copy you need to study very hard	you just need to emphasize you need to catch these sentences you need to show us you need to think about
should	you should get another time you should give a very good excuse you should go over the textbook you should go to the library	you should absolutely feel free to you should do it or whatever you should probably be able to you should think of punctuation
want to (wanna)	(none noted)	I want you to go home and think you might wanna like highlight you wanna maybe look at you would want to look at

According to a traditional semantics-based grammatical framework (Palmer, 2001), English represents modality grammatically using modal systems rather than mood. Directives are considered to be a kind of event modality, which refers to events that are not yet actualized, in contrast to propositional modality, which refers to a speaker's attitude to the truth-value of a proposition. Within event modality, directives are a kind of deontic modality, where "the conditioning factors are external to the relevant individual" (p. 9), which is in contrast to dynamic modality concerned with ability and willingness, or internal factors. Directives involve obligatives, permissives, and commissives, although the latter are sometimes considered separate. Palmer notes that 'must' is the prototypical obligative marker and 'may/can' the permissive marker, and that the source of obligation or permission may emanate from within (subjective) or from outside (objective) the speaker. With periphrastic equivalents like 'have to', the source tends towards the objective. Finally, obligation equates with deontic necessity, which accounts for 'need to', while permission equates with deontic possibility, which accounts for 'might' and 'could'.

#### **4.1.2.1.2 Irrealis**

In the transcripts, irrealis appeared to co-occur with directive language, indexed by the phrase like 'if you'. There were also many uses of 'I would' in the EOH sample that had indirect directive illocutionary force.

Table 4.2. Irrealis indexical exemplars from LOH and EOH data

indexical	exemplars from LOH sample	exemplars from EOH sample
if you	<ul style="list-style-type: none"> <li>• <i>If you</i> have an emergency you can finish the homework before the deadline</li> <li>• <i>If you</i> don't understand about the question yeah, you can come back and ask me</li> <li>• <i>If you</i> have any problem with these questions just come to see me.</li> </ul>	<ul style="list-style-type: none"> <li>• <i>if you</i> wanna talk about, the mall being successful as bringing people there to socialize you should talk about it where you're talking about that</li> <li>• figure out <i>if you</i> can what its history is you may be able to even just look on the web actually</li> <li>• <i>if you</i> do a colon, to say alright now what follows is an explanation, this has to be lower case</li> </ul>
I would	(none noted)	<ul style="list-style-type: none"> <li>• that's where <i>I would</i> put the emphasis</li> <li>• then <i>I would</i> put the quotation</li> <li>• maybe <i>I would</i> take</li> <li>• if I were to rephrase this <i>I would</i></li> </ul>

While 'I would' might be categorized under modal, it is not because conditional is traditionally considered another mood altogether, and the 'I' subject sets it apart. Irrealis in English is usually marked with hypothetical statements and the use of conditional 'would'. It is not surprising that irrealis is used with directives, as directing someone to do something may require posing alternate realities and suggesting or obliging the listener to future action. In some languages, irrealis is marked by subjunctive mood, as both involve non-assertion, in contrast to indicative mood. In many languages the subjunctive is used to express directive force, including polite commands and weak obligatives (Palmer, 2001), besides marking irrealis. In spite of claims that the English subjunctive is dead, it can still be used for formulaic ('heaven help us'), volitional ('if I were you'), or mandative ('I suggest that he leave') purposes (Berk, 1999), the latter which overlaps functionally with directives. No such subjunctive uses were found in the current data.

#### 4.1.2.1.3 Directive vocabulary

In spite of the comprehensiveness of Palmer’s taxonomy, the grounded analysis found that directives are also made without use of the modal system. The analysis shows that particular lexical items, including verbs, nouns, and adjectives or participles, can index directive force, often involving subject pronouns other than ‘you’.

Table 4.3. Directive vocabulary exemplars from LOH and EOH data

root	exemplars from LOH sample	exemplars from EOH sample
suggest	I suggest that you try this my suggestion is that	what I’ve been suggesting
recommend	(none noted)	what I recommend doing
necessary	(none noted)	words aren’t really necessary
encourage	(none noted)	I don’t want to encourage you to

A systemic-functional framework encompasses directive vocabulary beyond traditional modal systems. In SFL, modality is part of the MOOD system and refers to the intermediate degrees between positive and negative poles, construing “the region of uncertainty that lies between ‘yes’ and ‘no’” (Halliday & Matthiessen, 2004, p. 147). This scale of modulation is marked by degrees of obligation in a command, while it is marked by degrees of inclination in an offer. Besides modulation, obligative modality is subject to a system of VALUE (see He, 1993, and ch. 2), where high value is expressed with ‘must’, median value with ‘may’, and low value with ‘can’. Finally, obligation can be expressed by a finite modal operator, e.g. ‘you *should* know that’, or by an expansion of the Predicator, often in the passive, e.g. ‘you’re *supposed to* know that’ (p. 147). The latter occurs through a system of ORIENTATION, which plays a part in metaphorization (see Iedema, 1996, and ch. 2), as the source of the obligation is objectified (be supposed to) and implicit (you), though it could also be subjectified and explicit, e.g. ‘I expect that you know that’.

#### 4.1.2.1.4 Imperatives

While use of imperatives seems relatively common in the transcripts, bare imperatives pose a problem for corpus analysis because the English form is the plain root stem and is therefore indistinguishable from the infinitive and most simple present conjugations. For corpus analytic purposes, several indexicals that seemed to co-occur with imperatives, specifically ‘please’, ‘don’t’, and ‘just’ were noted to allow for analysis of at least a representative sample.

Table 4.4. Imperative indexical exemplars from LOH and EOH data

indexical	exemplars from LOH sample	exemplars from EOH sample
please	please come in please come to me please let me know please sit down	(none noted)
don't	don't worry about that	don't do it don't proofread don't get caught don't go looking
just	just come to see me just feel free to stop by just give me the paper just write it down	just give the quotation just make the observations just move through them just read sentence by sentence

Traditionally, imperative is a grammatical mood, like indicative or subjunctive. Palmer (2001, p. 80) considers English imperative closely associated with directive modality but not part of the modal system, noting that it can express both obligation and permission, and that it is performative in nature. It is closely related to the jussive, which can be understood as the first or third-person imperative, including ‘let’ phrases. In some languages, jussive and imperative overlap functionally with subjunctive in expressing commands and obligation (p. 138). From a universal grammar perspective, Han (2000) maintains that imperatives canonically express directive illocutionary force, and that an imperative contains a directive and an irrealis feature, which is expressed in some languages as subjunctive or infinitive forms. From a systemic-

functional perspective, the imperative is part of the MOOD system network. An imperative is a realization of a command involving the demanding of goods-and-services in an interpersonal exchange, in contrast to offers (giving goods-and-services), statements (giving information), and questions (demanding information). An obligative statement (i.e. in the indicative) is actually a realization of a modulated imperative, and could thus be considered an interpersonal metaphor of mood.

#### 4.1.2.2 Peripheral devices: Adjuncts & pronouns/subjects

Devices peripheral to the core directive elements, specifically adjuncts and pronouns/subjects were also found. Although they are discrete from the core element, pragmatic theory (see below) maintains that these adjuncts index interpersonal tenor, reflecting speaker intent and influencing listener interpretation. In light of this understanding, these adjuncts are considered objects of analysis, and so part of the main unit of analysis, for this project.

##### 4.1.2.2.1 Adjuncts

Adjuncts or ‘small words’ (Hasselgreen, 2002; see also Schiffrin, 1988) that intensified or strengthened the force of a particular directive, as well as language that mitigated or weakened it, were found in the data, especially in the EOH sample.

Table 4.5. Adjunct exemplars from LOH and EOH data

category	exemplars from LOH sample	exemplars from EOH sample
intensifiers	(none noted)	you should <i>absolutely</i> feel free
mitigators	<i>maybe</i> you should get another time	<i>I think</i> you should expand this I also <i>just</i> want you to be making <i>I mean</i> you need to <i>you know</i> fiddle you <i>might</i> wanna <i>like</i> make this you <i>almost</i> have to read aloud you can just drop this, <i>right</i> you can <i>sort of</i> look back over <i>maybe</i> you <i>would</i> want to

Brown and Levinson's politeness theory (1987) provides a framework for understanding mitigation and aggravation/intensification devices. Based on Goffman's notions of face, politeness theory maintains that human interaction is a constant interplay of redressing (potential) affronts to face, or face-threatening acts (see also Scollon & Scollon 1995). A directive would be considered an act that threatened the receiver's negative face, i.e. "the basic claim to territories, personal preserves, rights to non-distraction, i.e. to freedom of action and freedom from imposition" (Brown and Levinson, 1987 p. 61). Mitigation devices then are redresses to negative face threats. While directives aren't necessarily threatening to positive face, an office hours consultation could be, as it involves criticism and disapproval. Scollon & Scollon (1995) explain that hierarchical systems, where power is unequal and distance is maximal or minimal, is where miscommunication can easily occur, especially in intercultural communication events, since "calculation of the appropriate face system is always inextricably tied to the expression of the hierarchical system of relationship between or among the participants" (p. 49).

From a pragmatic speech act perspective, Blum-Kulka, House, & Kasper (1989), offer a taxonomy, used by several of the studies reviewed in chapter 2, that considers mitigation as part of a scale of indirectness strategies. The most direct head act strategy is the use of imperatives, followed by increasingly more indirect strategies: performatives like 'I suggest' or 'you are going to', obligatives, 'want statements', suggestory formula, indirect preparatory queries, strong hints, and mild hints. Any head act can be made more indirect through the additional use of mitigation strategies.

While traditional grammatical approaches are silent on connection between structure and use, an SFL perspective, concerned with language as a social semiotic, addresses issues of status, politeness, and formality through the concept of interpersonal metaphors of mood. In this system,

a bald command ('go!') can be metaphorically realized as a modulated indicative ('*you should go to school*'), or through the use of projection nexus clauses ('*I want you to go to school*'). This process of syntagmatic extension serves to distance speaker and addressee, as "metaphorical variants create a greater semiotic distance between meaning and wording" (Halliday & Matthiessen, 2004, p. 631).

#### 4.1.2.2.2 Pronouns & Subjects

In the analysis, it was noted that while the majority of directives seemed to be used with 'you', many were made using other subjects, usually pronouns, and very occasionally directive vocabulary itself.

Table 4.6. Pronoun exemplars from LOH and EOH data

pronoun	exemplars from LOH sample	exemplars from EOH sample
(null)	just write it down	don't go looking for more
2 <sup>nd</sup> person	<i>you</i> have to include the procedure	<i>you</i> wanna make a distinction
1 <sup>st</sup> person	<i>I</i> suggest that you try this let's get back to the question <i>we</i> want you to derive the answer <i>we</i> 'll consider about that	<i>I</i> don't want to encourage you <i>I</i> 've suggested places let's look at this one <i>we</i> have to bring whatever
3 <sup>rd</sup> person	<i>my suggestion</i> is that	<i>it</i> 's better to look to the future

Pronoun choice may also have a relation to positive and negative face appeals. Harwood (2005) surveyed use of 'I', exclusive 'we', and inclusive 'we' usage in a 300,000-token corpus of academic journals. He notes the relationship between the use of 'I' and establishment of authorial presence and the use of exclusive/inclusive 'we' to persuade readers. For negative politeness purposes, he observes that 'I' can be used as a mitigation device and exclusive 'we' can exclude the reader, while use of inclusive 'we' or 'let's' is a positive appeal towards involvement. In traditional grammar, first person is generally subjective and third person

objective, while ‘you’ can be second person as well as impersonal, referring to ‘(any) one’, the use of which could be considered an appeal towards negative face or independence.

## **4.2 Discussion**

The results of the grounded analysis shows that for the needs of the current project, no one singular theoretical perspective may be adequate for the operationalization of a unit of analysis that must remain based in pragmatic use. Traditional grammatical taxonomies, for example Palmer’s modality framework, offer insight into the semantics of the core devices, but are divorced from usage and do not include adjuncts and other devices that arguably influence directive force. In consideration of the reviewed literature, a pragmatics-based approach towards defining a unit of analysis might be most appropriate.

In a traditional pragmatics framework, Searle (1976) defines a directive illocutionary speech act as “attempts (of varying degrees...) by the speaker to get the hearer to do something. They may be very modest ‘attempts’, as when I invite you to do it or suggest that you do it, or they may be very fierce attempts as when I insist that you do it” (p. 355). Being directive, the speech act reflects the speaker’s wish or desire, and being illocutionary it commits the hearer to future action. A speech act may be indirect, in which case it has primary (what is said) and secondary (what is meant) illocutionary force, which is often the case with requests. It is successful depending on whether or not its illocutionary force is taken up by the hearer, and so may be co-constructed over a series of turns. Finally, a speech act may be marked by an illocutionary force indicator or device, for example word order or mood.

For the purposes of this project, the devices identified in the above analysis are considered indicators of directive illocutionary force. However, it is not clear that the objects of analysis can simply be equated with ‘speech acts’ for several reasons related to the corpus nature

of the project. Indicators like gesture and intonation, as well as evidence of listener uptake determinative of a speech act's success, are for all practical purposes impossible to consider with corpus data of this nature. In addition, the secondary force of any indirect speech acts in the data may be impossible to know without checking the intent of each speaker. For the purpose of this study, the act of contextually appropriate usage in performance is considered evidence of directive force, as its utterance reflects directive intent on the part of the speaker.

Another candidate for unit of analysis is the SFL concept of 'mood element' (Halliday & Matthiessen, 2004), which contains the subject, finite, and mood adjunct of a given clause, and as such "carries the burden of the clause as an interactive element" (p. 120). As SFL considers language first and foremost a social semiotic, it is not separated from pragmatic usage as other grammatical frameworks are. However, as mentioned above, directive force is often projected out of the finite through metaphorization into the mood residual, which would mean some items would be excluded from consideration, if analyses were limited to discretely defined mood elements. Instead of adopting the exact definition, the concept of mood element, specifically its inclusion of the subject and mood adjunct, might be used to inform a construct more useful for the needs of the current project that includes devices whose directive force is a result of clausal projection.

#### **4.2.1 Directive construction: A social-functional unit of analysis**

For the purpose of this project, directive language use is operationalized as 'directive construction', which is defined as a social-functional device comprised of one or more separate, lexico-grammatical units that index directive illocutionary force. As social-functional devices, they may be considered social instantiations of speech events and may correspond to pragmatic speech acts. A directive construction contains at the minimum a core lexical device, which may

be a modal or p-modal, directive vocabulary, or imperative. Peripherally, a directive construction may include the subject, usually a pronoun, and often an adverbial adjunct, as they influence the illocutionary force or politeness of the directive.

The term ‘construction’ is preferred as it corresponds to a usage-based approach to language cognition and development (Langacker, 2000; Tomasello, 2003), where the term refers to patterns of paired form and semantic or discourse function learned through frequency of exposure (Goldberg, 2003). The term is also appropriate because of the structural diversity of the findings, as a ‘construction’ is not defined in terms of underlying syntactic structure but rather a ‘what you see is what you get’ approach (p. 219), which matches the empirically grounded nature of the project. Langacker explains that in a usage-based model, “substantial importance is given to the actual use of the linguistic system and a speaker’s knowledge of this use” (2000, p. 1). Tomasello (1998) explains that usage-based, or cognitive-functional linguists, “language structures are symbolic instruments that convey meaning, from the smallest morphemes to the most complex constructions” (p. xi), such that meaning and use are inextricably intertwined. Kemmer & Barlow (2000, introduction) explain that usage-based approaches share certain common conceptions about language and learning, including 1) the relation between structures and instances of use or ‘usage events’, 2) the importance of frequency, 3) a focus on learning and experience in language acquisition, 4) the importance of usage data, 5) the relation between usage, variation, and change, 6) the interconnectedness between the linguistic system and other cognitive systems, and 7) the crucial role of context. Because these conceptualizations underlie this project theoretically, the term ‘construction’ is preferred to other terminology.

Seen from a corpus or frequency perspective, directive constructions may correspond to formulaic sequences (Schmitt, 2004) or lexical bundles (Biber, Conrad, & Cortes, 2004), except

that they are based on function rather than frequency. Biber, Conrad & Cortes (2004) ascribe the variety of terms for multi-word units to differences in research goals adopted, identification criteria, formal characteristics of the units, corpus size, and recognition of register (p. 372). These criteria certainly hold true for the current study, and the intention of this study is not to promote the concept of ‘directive construction’, but rather to operationalize a relevant unit of analysis for localized purposes.

While a ‘construction’ is a usage-based unit of analysis for quantitative purposes, for qualitative purposes the force of a directive in use is fluid, dynamic, and recursive, with less distinct boundaries than the term ‘construction’ may imply. As an interpersonal, conversational phenomenon, it can be negotiated and manifest at an extra-sentential discourse level. One directive construction, containing core and peripheral elements, can itself be core in relation to other peripheral directives, and its force may be influenced by other co-textual interpersonal elements. The force of several directive constructions can thus overlap and compound one another. Chapter 6 uses genre analysis of text and context and chapter 7 uses discourse analysis of individual usage to explore this quality in more depth.

The social aspect of ‘social-functional’ can be understood in consideration of the social aspects of the functional linguistic theory discussed above as well as genre theory (see chapter 6), which affords analysis of the socio-cultural aspects of the office hours texts and contexts where directive constructions are used. The use of a directive construction indexes a power relationship between the ‘director-speaker’ and the ‘directed-listener’, which in an office hours context reflects the role hierarchy inherent to institutional discourse, where the more powerful T (teacher, in this case professor or TA) is the director-speaker, and the less powerful S (student) the directed-listener.

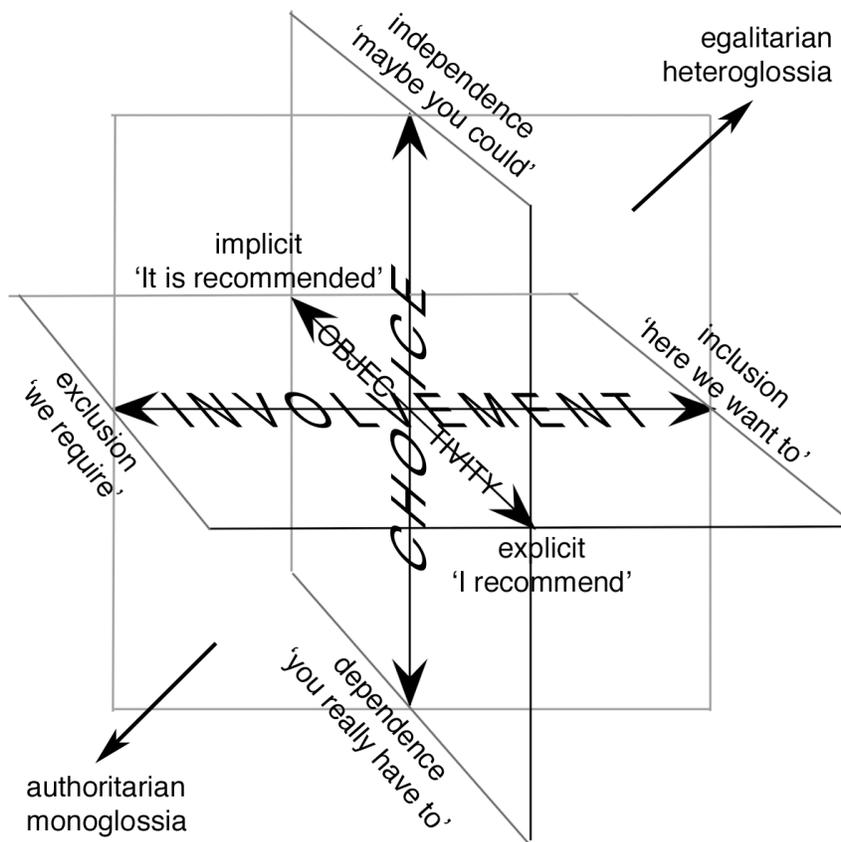
As a result of the above semi-grounded analysis, I propose a social-functional model of directive usage (figure 4.1), also informed by the above discussed linguistic and pragmatic theory. In this model, the force of a directive construction reflects the stance of the director-speaker vis-à-vis power, and can be seen to operate on three interdependent planes: choice (dependence-independence), involvement (exclusion-inclusion), and objectivity (implicit-explicit)<sup>3</sup>. When conceptualized as a space transected by three axes, X, Y, and Z, the independence-inclusion-explicit corner would represent heteroglossia and equal, reciprocal distribution of symbolic power, while the dependence-exclusion-implicit corner would represent monoglossia and unequal, one-way distribution. On the plane of choice, the director-speaker's choice of strong or weak directives and use of politeness strategies like mitigating or intensifying adjuncts reflect the degree of choice, of independence of action, that the director-speaker wishes the directed-listener to apprehend, in the sense of acknowledge, accept, and act upon. For example, 'you have to' implies less choice than 'you could' on a cline of obligation, and 'you really could' less than 'maybe you could'. On the plane of involvement, the director-speaker's choice of pronoun reflects the degree of inclusion that s/he wishes the directed-listener to apprehend. For example, 'we require' with an exclusive 'we' referring to a department or institution, is more exclusionary than 'here we want to' with an inclusive 'we' referring to both the director-speaker and the directed-listener. Finally, on the plane of objectivity, the use of more objectivizing, implicit forms<sup>4</sup> serves to conceal the source of authority or power, while the use of explicit, unprojected forms serves to make it available for scrutiny. For example, 'the recommendation is that you' is more implicit than 'I recommend that you'.

---

<sup>3</sup> It should be noted that the terms 'independence' and 'involvement' are from Scollon & Scollon (1995), and that several other terms have been suggested for these phenomena, like 'respect' and 'rapport'. See also chapter 9.

<sup>4</sup> While periphrastic modals might serve to objectify the source of power when compared to their true modal counterparts, for example, 'you have to' in comparison to 'you must', historical shifts towards regularization in usage patterns towards may complicate this trend (Collins, 2005).

Figure 4.1. A social-functional model of directives usage



In the space represented in the model (fig. 4.1), a director-speaker may intentionally or unintentionally evoke authority, pulling the force in one direction, which may evoke dependence on, exclusion from, or obfuscation of the source of, power on the part of the directed-listener; in the model, this would be represented by movement down, to the left, and away from the viewer respectively. Alternatively, the director-speaker may intentionally or unintentionally downplay authority, pushing the force in the other direction, which may foster independence, inclusion, and awareness; in the model (fig. 4.1), this would be represented by movement up, to the right, and towards the viewer respectively. In this manner the director-speaker initiates directive force with

a construction at one point but then may adjust the directive force in negotiation with the directed-listener. The directive is construed on all three planes simultaneously, but one or two may be more prominent, depending on director-speaker intention. Power is thus constructed and negotiated through the use of directive constructions.

### **4.3 Conclusion**

As Biber and his colleagues note, “the overall importance of multi-word units in discourse can be fully understood only by undertaking empirical research studies from different perspectives” (2004, p. 372). In response, this chapter has defined the object of study of this dissertation, and operationalized a unit of analysis for use in the quantitative corpus and qualitative discourse analyses of subsequent chapters. The proposed unit of directive construction was developed through grounded analysis of the subject corpus transcripts, and categories were supported with primary source linguistic and pragmatic theory. Finally, the unit of analysis is grounded in a heuristic model of social-functional linguistic theory that considers the multi-level and dynamic nature of power. This heuristic will now be used for analysis of the data.

## CHAPTER 5

### **A CORPUS ANALYSIS OF DIRECTIVE CONSTRUCTION USAGE IN OFFICE HOURS**

In this chapter, I explore the systematic qualities of directives usage in office hours through the individual and comparative corpus analysis of both the Learner Office Hours (LOH) corpus from ITAcorp and the Expert Office Hours (EOH) corpus from MICASE (Simpson et al., 2002). I first examine the genre consistency, or usage distribution, of the subject corpora, by comparing the corpora both externally with other corpora and internally with themselves. I then discuss theory and research on learner corpus analysis, followed by a comparison of directives usage in the two corpora using contrastive interlanguage analysis (Granger, 1998; 2002). I also compare adjunct use in the two corpora, including mitigators and intensifiers, and examine the findings of the analyses using the social-functional theoretical model proposed in chapter 4 as an interpretive heuristic.

As discussed in chapter 2, a variety of methodological approaches have been used to analyze language use in office hours-like institutional discourse. In her review of the growing number of corpus analyses of academic language, L. Flowerdew (2001) explains that while general corpora have gone towards ever larger, multi-million word corpora, academic corpora have gone the route of smaller, register-based corpora, but that there is “a dearth of spoken corpora for academic purposes” (p. 110; see O’Keeffe, McCarthy, & Carter, 2007, ch. 10 for a more recent review). Flowerdew mentions MICASE and a more recent, but not publicly available, corpus, the TOEFL 2000 Spoken and Written Academic Language Corpus (T2K-SWAL), compiled by Biber and colleagues. In his 2006 work, Biber uses corpus-based register

analysis to profile spoken and written university registers which comprise T2K-SWAL, including textbooks, lectures, and office hours. The office hours subcorpus is comprised of 11 texts and 50,400 tokens<sup>1</sup>. Biber's analysis determines that similar to other spoken registers, there are a predominance of verbs and pronouns in office hours. He notes that 'so' is the most common discourse marker (p. 66), as well as 'okay'. Biber also notes the use of modals for expressions of directive stance, and that of all the spoken academic registers in the corpus, conditional voice is most frequent in the office hours corpus; almost half of all conditional clauses in the corpus begin with 'if you'. Biber reasons that conditional clauses "cushion the force of directives by providing possible options and anticipating alternatives" (p. 77).

### **5.1 A corpus analysis of register consistency**

Similar to Biber's investigation of office hours, a general register analysis would investigate keywords, frequencies, and collocations. For the current project, however, the object of corpus investigation is the use of directive language in office hours, not office hours in general, although that is an important aspect of the project. More important to the current project is to first establish consistency of register, to provide evidence that directives in the corpora are more or less evenly distributed throughout each corpus, before any comparisons are made. When dealing with small corpora, the issue of distribution is crucial, since high word count cannot be relied on to even out individual variation. One relatively straightforward method to provide evidence for consistency is to halve the corpora and then compare the two halves. The fewer statistically significant differences there are between parallel items in the two halves, the more likely the items are evenly distributed, and thus any further analysis using the corpora is valid.

---

<sup>1</sup> This is less than half the size of either MICASE's or ITAcorp's office hours subcorpus.

### 5.1.1 Method

The LOH and EOH corpora were searched for instances of directive constructions produced by those participants in the TA or professor (T) role. In the case of the LOH corpus, each construction was searched for in the ITAcorp FMPro database; the resulting data was exported and imported into MS Excel spreadsheets, which allows for more customizable sorting and more transparent counting than either WordSmith or FMPro. In the case of the EOH corpus, MICASE's web interfaces allows for custom searches, which were then copied and pasted into Excel spreadsheets. For each corpus, each construction was given its own spreadsheet, and each instance was identified with immediate context (5 words to the left and 5 to the right), file ID number, and speaker. Frequency counts were then established for each construction, and rates per 10,000 words were relativized by dividing the count of a construction by the total number of tokens in its corpus.

To provide evidence for the register consistency of the LOH corpus, it was split into its two cohorts (F05 and S06) so that frequency and log-likelihood analyses could be run between the two halves. If the two halves of each corpus were not statistically different, then, it can be argued that the whole corpus was internally register-consistent. For each construction with frequency over 1 instance per 10K, a log-likelihood statistic comparing its frequency in the two halves was then calculated (Rayson & Garside, 2000)<sup>2</sup>. Since a log-likelihood statistic over 3.84 indicates significant difference equal to a p-value less than .05 between two halves of a whole corpus, a statistic lower than 3.84 on parallel items is considered evidence here for statistical similarity between the two halves, and thus register consistency of the whole LOH corpus.

---

<sup>2</sup> Calculations were done using Paul Rayson's online Log-Likelihood Calculator at: <http://ucrel.lancs.ac.uk/llwizard.html>

It should be noted that the office hours subcorpus of MICASE (i.e. the EOH corpus) was categorized as such by the creators of MICASE based not on register analysis but on genre analysis, in that all the transcripts shared parameters of what is termed ‘office hours’ (see chapter 6, table 6.2 for a breakdown of the corpus). By comparing the profile of this subcorpus with another, genre comparable subcorpus (see §5.2.2 below on genre comparability), namely the advising subcorpus (ADV), the register consistency of the EOH can be argued<sup>3</sup>. To this end, the EOH corpus is first compared to an external baseline provided by the advising subcorpus. The EOH is then split into two halves, so that each half is comprised of four transcripts (see table 6.2). The first half is comprised of office hour sessions conducted by 4 TAs, 3 females and 1 male, all of which were a mix of one-to-one (T-S) and one-to-many (T-Ss) sessions. The second half is comprised of office hour sessions conducted by 4 faculty, again 3 females and 1 male, all of which were again a mix of one-to-one (T-S) and one-to-many (T-Ss) sessions. This analysis provides an internal profile that can be compared to the external baseline analysis. If the internal comparison showed less significant difference than the external comparison, it is then argued that the EOH corpus shows a consistent distribution of directive constructions.

### **5.1.2 Results: Learner group profile**

The LOH corpus was searched for modals and p-modals as operationalized in chapter 4. Of the other core lexical devices, directive vocabulary were omitted because of low frequency numbers, and imperative indexicals were omitted because they were only partially representative of all imperatives. To provide evidence for the register consistency of the corpus, the two cohort (F05 and S06) halves were searched separately, and the results subjected to a log-likelihood analysis to provide negative evidence of similarity. The two cohorts were a natural split, as each

---

<sup>3</sup> No such corpus for external comparison was available for the LOH; the other ITAcorp subcorpora, lecture and discussion, are not similar enough to warrant comparison.

corresponded to a semester of ESL 118 instruction. Table 5.1 presents the total tokens in each half.

Table 5.1. Total tokens in the LOH corpus and in F05 and S06 cohort corpus halves

all LOH	F05 cohort half	S06 cohort half
102806	43295	59511

The two cohort F05 and S06 LOH halves were first searched for modal and p-modal (periphrastic or semi-modal) constructions, primarily with ‘you’ as a subject. The irrealis construction ‘I would’ was also included in this search. Each set of results was then vetted for and rid of non-directive uses of the constructions, for example epistemic uses like ‘you must have been happy’. One exception to this was ‘you can’, where it proved very problematic to distinguish between cases of dynamic ability, deontic permission, and directive force. For comparative purposes, however, as long as the uses of can in the comparison corpus are also not differentiated, the comparison will remain valid, if not generalizable. In any case, for the sake of the corpus profile, totals with and without ‘can’ were calculated. The negative forms of the constructions were treated separately, as these constructions have different, not necessarily opposite<sup>4</sup>, directive illocutionary effect than their affirmative forms.

---

<sup>4</sup> For example, ‘you have to’ is obligative, but ‘you don’t have to’ is not, while both ‘you should’ and ‘you shouldn’t’ are obligative.

Table 5.2. Number and frequency per 10K of modal and p-modal constructions in LOH corpus

modal or p-modal construction	all LOH		LOH F05 half		LOH S06 half		log-l F05vS06	
	total	freq/10K	total	freq/10K	total	freq/10K		
you can (undiff.)	536	52.137	236	54.510	300	50.411	0.80	
you should	103	10.019	54	12.473	49	8.234	4.43*	
you have to	55	5.350	23	5.312	32	5.377	0.00	
you need to	41	3.988	15	3.465	26	4.369	0.52	
you had better	29	2.918	9	2.079	20	3.361	1.51	
you don't need to	17	1.654	6	1.386	11	1.848	0.33	
you want to (wanna)	12	1.167	1	0.231	11	1.848	6.87*	
you could	9	0.875	1	0.231	8	1.344	n/a	
I want you to	9	0.875	4	0.924	5	0.840		
you must	6	0.584	4	0.924	2	0.336		
it is/would be better if you	4	0.389	0	0.000	4	0.672		
you don't have to	3	0.292	0	0.000	3	0.504		
you should not	2	0.195	1	0.231	1	0.168		
you may	2	0.195	0	0.000	2	0.336		
I would	1	.097	0	0.00	1	0.168		
TOTAL w/can	828	80.540	354	81.765	474	79.649		0.14
TOTAL w/o can	292	28.403	118	27.255	174	29.238		0.35

\* A log-likelihood greater than 3.84 indicates a p-value less than .05, and thus, a significant difference between F05 and S06 cohort production of this construction. Statistical difference was not calculated for constructions whose overall LOH frequency were less than 1 per 10K.

The results are presented in table 5.2, in descending order of overall frequency in the whole LOH corpus. Most notably, there is no evidence for significant difference in modal and p-modal use between the two halves (log-l = 0.14 w/ 'can', 0.35 w/o 'can'), which may be interpreted as evidence for register consistency. The notable exceptions are 'you should', which is more frequent in the F05 half, and 'you want to' and 'you could', which are more frequent in the S06 half. These differences may be a result of the Spring 2006 pedagogical intervention, which is the topic of chapter 8.

### 5.1.3 Results: Expert group profile

As with the LOH corpus, the EOH corpus was searched for modals and p-modals as operationalized in chapter 4. To provide evidence for the register consistency of the corpus, as discussed in the methods section above, the corpus was first compared with another MICASE

subcorpus, the advising (ADV) corpus. The corpora were searched separately, and the results subjected to a log-likelihood analysis to provide negative evidence of similarity. The EOH corpus was then split in half into TA and faculty (FAC) corpus halves; total tokens in the ADV corpus, total EOH and TA and FAC halves are in table 5.3. The two halves were a natural split, as explained in the methods section above and chapter 3. As with the previous analyses, the two halves were searched separately, and the results subjected to a log-likelihood analysis to provide negative evidence of similarity. Finally, the log-l results from both analysis are compared, as evidence of the genre consistency of the EOH corpus.

Table 5.3. Total tokens in the MICASE ADV corpus, the entire EOH corpus, and the TA and FAC EOH corpus halves

ADV	EOH	EOH TA half	EOH FAC half
61383	125343	79195	46148

The first analysis compares the EOH corpus with the ADV corpus, both of which are subcorpora of MICASE. Because office hours and advising are functionally similar, comparison of their profiles should show some, but not great differences, in any case larger than the differences found in the internal comparison.

### 5.1.3.1 External comparison

As with the LOH corpus, the EOH and ADV corpora were searched for modal and p-modal constructions, primarily with ‘you’ as a subject. Again, the irrealis construction ‘I would’ was also included. Each set of results was then vetted for and rid of non-directive uses of the constructions. As with the LOH, ‘can’ was left undifferentiated, and the negative forms of the constructions were treated separately.

Table 5.4. Number and frequency per 10K of modal and p-modal constructions in MICASE OFC (EOH) and MICASE ADV corpora

modal or p-modal construction	EOH (MICASE OFC)		MICASE ADV		log-l OFCvADV	
	total	freq/10K	total	freq/10K		
you can	340	27.126	150	24.437	1.15	
you want to (wanna)	173	13.802	50	8.146	11.81*	
you need to	76	6.063	33	5.376	0.34	
you have to	87	6.941	34	5.539	1.28	
you could	59	4.707	64	10.426	19.12*	
you should	46	3.670	37	6.028	4.91*	
I would	38	3.032	26	4.236	1.68	
you (not) need to	16	1.276	7	1.140	0.06	
you have got to ('ve gotta)	12	0.957	7	1.140	n/a	
I want you to	6	0.479	3	0.489		
I don't want you to	4	0.319	0	0.000		
you are better off	4	0.319	0	0.000		
it is/would be better if	1	0.080	2	0.326		
you may not	1	0.080	0	0.000		
you must	0	0.000	2	0.326		
you may	0	0.000	2	0.326		
you had better	0	0.000	1	0.163		
TOTAL w/can	863	68.851	418	68.097		0.03
TOTAL w/o can	523	41.726	268	43.660		0.36

\* A log-likelihood greater than 3.84 indicates a p-value less than .05, and thus, a significant difference between OFC and ADV production of this construction. Statistical difference was not calculated for constructions whose overall OFC frequency were less than 1 per 10K.

The modal and p-modal frequencies in the EOH and ADV corpora are presented in table 5.4 above, in descending order of overall frequency in the EOH (OFC) corpus. The two corpora showed significant differences in 3 out of the 8 most frequent constructions, but no overall significant difference. The use of 'want to' is more frequent in the EOH corpus, while 'you could' and 'you should' are significantly more frequent in the ADV corpus. A cursory explanation may be that 'you want to' is more frequent in the EOH corpus because the T authority in office hours tends towards alignment and persuasion. In advising sessions (ADV), 'you could' may be very frequent because options are being offered to the advisee, while 'you should' is less frequent in office hours because it subjectivizes and thus emphasizes the authority of the T director, which may be counterproductive to persuasion as it explicitly restricts the directed S sense of independence (i.e. 'avoid sounding too preachy').

### 5.1.3.2 Internal comparison

The differences between the EOH and the ADV corpora are only sometimes significant and are explored in §5.1.3.3 below. Before this, however, it is useful to also conduct an internal consistency analysis, as with the LOH analysis above. As described in the methods section above, the EOH was split into TA and FAC halves. Each half was then searched for modal and p-modal constructions, primarily with ‘you’ as a subject. Each set of results was then vetted for and rid of non-directive uses of the constructions. As with the other analyses, ‘can’ was left undifferentiated, and the negative forms of the constructions were treated separately.

Table 5.5. Number and frequency per 10K of modal and p-modal constructions in EOH TA and FAC corpus halves

modal or p-modal construction	EOH TA half		EOH FAC half		log-l TA <sub>v</sub> FAC	
	total	freq/10K	total	freq/10K		
you can (undifferentiated)	214	27.022	126	27.303	0.01	
you want to (wanna)	116	14.647	57	12.352	1.13	
you need to	48	6.061	28	6.067	0.00	
you have to	57	7.197	30	6.501	0.21	
you could	29	3.662	30	6.501	4.81*	
you should	26	3.283	20	4.334	0.86	
I would	17	2.147	21	4.551	5.32*	
you (not) need to	15	1.894	1	0.217	8.29*	
you have got to (‘ve gotta)	10	1.263	2	0.433	n/a	
I want you to	4	0.505	2	0.433		
I don't want you to	0	0.000	4	0.867		
you are better off	0	0.000	4	0.867		
it is/would be better if	0	0.000	1	0.217		
you may not	0	0.000	1	0.217		
you must	0	0.000	0	0.000		
you may	0	0.000	0	0.000		
you had better	0	0.000	0	0.000		
you might	0	0.000	0	0.000		
TOTAL w/can	536	67.681	327	70.859		0.43
TOTAL w/o can	322	40.659	201	43.556		0.58

\* A log-likelihood greater than 3.84 indicates a p-value less than .05, and thus, a significant difference between TA and FAC production of this construction. Statistical difference was not calculated for constructions whose overall EOH frequency were less than 1 per 10K.

The modal and p-modal frequencies in the EOH TA and FAC corpus halves are presented in table 5.5 above, in descending order of overall frequency in the TA corpus. By measure of the

four most frequent constructions and the totals, there is no evidence for significant difference, and thus evidence for consistency of register. Of the remaining constructions, ‘you could’ and ‘I would’ are used more significantly in the FAC half. It should be noted also that in the TA half, a single individual made 10 of the 15 instances of ‘you don’t need to’ while another single individual made all 10 instances of ‘you’ve got to’. Individual differences are explored more in chapter 7.

### 5.1.3.3 Register consistency of the EOH corpus

As shown in table 5.6 below, there is more significant difference between the EOH corpus and the external, yet similar ADV corpus, than there is between the two halves of the EOH corpus itself, the TA and FAC subcorpora (i.e., the values in the 3<sup>rd</sup> column are almost all higher than in the 4<sup>th</sup>). This holds true for the all but 2 (‘I would’ and ‘you don’t need to’) of the top 8 most frequent modals and p-modals. The EOH and ADV corpora are nonetheless very similar, as five of the eight measures show no statistical difference.

Table 5.6. Comparison of internal and external measures of EOH corpus register consistency

construction or measure	EOH freq.	log-l EOHvADV	log-l TAvFAC
you can	27.126	1.15	0.01
you want to (wanna)	13.802	11.81*	1.13
you have to	6.941	1.28	0.21
you need to	6.063	0.34	0.00
you could	4.707	19.12*	4.81*
you should	3.670	4.91*	0.86
I would	3.032	1.68	5.32*
you don’t need to	1.276	0.06	8.29*

While it may seem self-evident that the two register profiles of language use from the same genre are more similar than two register profiles from different genres, it is nevertheless important to establish consistency of distribution so that any further analyses are valid. The analyses in table 5.6 above show that both the LOH and EOH corpora are consistently

representative of directives usage and thus valid for the purposes of comparison, in spite of their relatively small sizes. Still, since low frequency constructions may be more affected by individual variation than higher frequency constructions, a systematic picture of usage may be unbalanced without analysis of individual usage, which is the subject of chapter 7. Before that, however, a contrastive learner corpus analysis may be useful for determining exactly where to examine more in-depth.

## **5.2 Learner corpus analysis: Background**

In what might be considered ‘corpus linguistics applied’, researchers have in the last decade or so (see Nesselhauf, 2004, for a survey) begun to apply principles and techniques from corpus linguistics to the study of second/foreign language (L2) learning and teaching, as these techniques allow for the computerized analysis of learner language and development using far larger amounts of data than previously possible. A pioneer in this field, Granger (1998; 2002) claims that before learner corpus analysis, there had been avoidance of natural language use data in studies of L2 learning because some language features are very infrequent, variables are numerous and uncontrollable, and learners may avoid troublesome features. She defines a learner corpus as an “electronic collection of authentic FL/SL textual data assembled according to explicit design criteria for a particular SLA/FLT purpose. They are encoded in a standardised and homogenous way and documented as to their origin and providence” (2002, p. 7).

Mark (1998, in Granger, 2002, pp. 6-7) notes that many studies of L2 learning inform instruction through description of target language and characterization of learners, but not by description of learner language. Many traditional studies rely on experimental set-ups, which produce results that may be replicable, but are ecologically invalid because of the artificiality of the set-up. In contrast, learner corpus analysis offers the promise of learner instruction based on

analysis of empirical data that may be considered valid because it is collected under naturalistic classroom conditions; in addition, high ecological validity is possible when the results inform the curriculum used by the very students that produced the data (e.g. Belz, 2004; Belz & Vyatkina, 2005). This idea of applied learner corpus analysis, where results of a particular learner corpus analysis are fed back into curriculum design for particular learners, is discussed further in chapters 8 and 9.

### **5.2.1 Comparative learner corpus analysis**

Researchers believe they have found in learner corpus analysis a relatively theoretically neutral methodology, and have increasingly used it in combination with a variety of frameworks, including genre theory (Upton and Connor, 2001), cognitive linguistics (Waara, 2004), Relevance Theory (Hasselgreen, 2002), developmental sequence theory (Housen, 2002), systemic-functional APPRAISAL theory (Flowerdew, 2003), and sociocultural theory (Belz, 2004). Fundamental to most approaches to learner corpus analysis is the comparison of learner data to expert or other learner data, i.e. comparative learner corpus analysis (CLCA), using techniques pioneered by Granger (1998; 2002)<sup>5</sup>. Granger's approach is something of a hybrid between error analysis and contrastive analysis, focusing on aggregate synchronic learner language data. Granger explains that her technique retains the analytic power of error analysis without the drawbacks. Early error analysis involved corpora that were small and not systematic, and served more as "depositories of errors, only to be discarded after the relevant errors had been extracted" (1998, p. 6). It focused on decontextualized errors and produced irreproducible results, where new CLCA approaches allows consideration of not only what learners get wrong

---

<sup>5</sup> Granger's school uses the term 'contrastive interlanguage analysis', using Selinker's term 'interlanguage', which is dispreferred for the current project because of its association with a specific psycholinguistic approach to SLA, as 'inter' can be construed to imply incompleteness or deficiency.

but what they get right (Leech, 1998, p. xvii), since “both the context of use and the linguistic context (co-text) is permanently available to the analyst” (Granger, 1998, p. 14).

CLCA has been criticized for comparing production of non-native speakers with that of native speakers. Critics argue that because the term ‘learner’ does not construe ‘nativeness’ as an untenable teleologic goal (see Block, 2003), it is preferred to ‘non-native’, although in some situations, when the corpus participants do not consider themselves learners, it too may be inappropriate. Similarly, since ‘expertise’ with a language is an achievable state, the term ‘expert’ is preferred to ‘native’. More controversial is the use of expert production as the baseline corpus against which learner production is compared. In response to this criticism, I argue the purpose of applied learner corpus analysis is not to imply deficient usage by learners, but to describe learner language and apprise studies of L2 learning and development. The ultimate goal is to inform a choice-based pedagogy where learners are provided the resources to see language use as a matter of choice-making, rather than a prescriptive pedagogy where learners are given a list of rules to follow (or break). Being comprised of classroom language use data is thus a strength of a learner corpus rather than a weakness, as it is more ecologically valid to base instruction on what actually occurs in the target genre and L2 classroom, than on what others believe occurs in the genre, and should occur in the classroom.

Several studies using CLCA methodology are particularly relevant to the current project. Ringbom (1998) found overuse of certain forms in the essays of learners from a variety of L1s in the International Corpus of Learner English (ICLE), a collection of over 3600 essays written by EFL learners from 11 different L1s, totaling over 2.5 million words (Granger, 2003). For all backgrounds, learners tended to overuse verbs like ‘think’ and ‘get’, as well as first and second person pronouns, which resulted in vagueness. For some learners, he sees evidence of covert

transfer, where L1 procedures are used “in the absence of appropriate L2 procedures being available” (p. 49), in avoidance (underuse). In confirmation of claims that advanced learner language makes inappropriate use of vague language, he finds overuse of vague elements like quantifiers (e.g. ‘more’ and ‘very’) and general nouns (e.g. ‘people’ and ‘things’).

Adding the strengths of contrastive analysis to CLCA, Altenberg (2002) conducted a very detailed analysis of the ICLE production of Swedish L1 writers. He attributed overuse of a variety of constructions with ‘make’ to intra- and interlingual transfer. Besides conducting an analysis of an NS and IL corpus, he conducted a contrastive analysis on a bilingual corpus, a corpus containing two versions of a text, which allowed him to analyze how translators had represented English target items using ‘make’ in their Swedish L1. Using the three corpora together allowed him to come to a more complete and nuanced understanding of the complexities of transfer.

Using CLCA, Petch-Tyson (1998) discusses the various degrees of interpersonal involvement, or writer/reader visibility, that can be found in written discourse. She presents a list of corpus-retrievable items that characterize this visibility, which includes presence of first and second-person pronouns, emphatic particles, vagueness, imperatives, and direct questions (p. 111). She analyzed ICLE writing and found that use of these features was more frequent in non-native writing from a variety of L1s than in native English writing. Upon horizontal analysis of co-text<sup>6</sup>, she found qualitative differences in the uses of these features based on L1. Cobb (2003) replicates Petch-Tyson’s study by focusing on the use of first and second-person pronouns in his Quebecois ESL learner corpus. His results show learner language that is even more divergent from NS use than in Petch-Tyson’s.

---

<sup>6</sup> This refers to the analysis (horizontal reading) of the linguistic context of a particular term that has first been identified in a vertical concordance.

Aijmer (2002) illustrates how CLCA methodology can be used with different theoretical frameworks, in this case modality theory, although she does not identify it as such. By looking at the ICLE production of L1 Swedish, German, and French writers, she analyzed the use of a variety of modals and modal devices, and concluded that learners overuse speech-like forms in writing. She found considerable overuse of some forms, and attributed differences between learner groups on L1 transfer of form and register/genre. She also compared epistemic and deontic uses of modals and modal phrases, and found considerable differences, attributing the differences to student difficulty in expressing degrees of commitment and vagueness.

### **5.3 Comparative analysis of learner and expert profiles**

In the tradition of the above literature, this section presents a comparative analysis of learner and expert directives usage in office hours contexts, to the end of informing pedagogical treatment (chapter 8). This comparative analysis will inform further analyses, especially of genre and discourse patterns (chapter 6) and individual variation (chapter 7), by pointing towards areas where expert and learner use differs significantly. The analysis goes beyond the above consistency analyses and focuses not only modals and p-modals but on all core lexical devices and peripheral devices, including adjuncts, as operationalized in chapter 4.

#### **5.3.1 Method**

The comparative analytic techniques used here involve comparing frequency rates of a set of items in genre-comparable corpora, one produced by learners (LOH) and the others by experts (EOH), using the comparative learner corpus analytic methods described in §5.2.1 above. To determine the overuse or underuse ratio of construction use by the learners, a log likelihood

analysis was run to compare usage frequencies with that of experts. Total tokens in each corpus are shown in table 5.7 below.

Table 5.7. Total tokens in LOH and EOH corpora

all LOH	all EOH
102806	125343

## **5.3.2 Results: Comparative Analysis**

The results of the comparative analysis are presented below in sections, on the core lexical devices of modals and p-modals (§5.3.2.1.1), directive vocabulary (§5.3.2.1.2), imperatives (§5.3.2.1.3), and on the peripheral adjuncts (§5.4.2.2). Each table of results is followed by interpretation informed by the social-functional model of directive use presented in chapter 4.

### **5.3.2.1 Core lexical devices**

#### **5.3.2.1.1 Modal and p-modal constructions**

The results in table 5.8 are presented in two sections. The top section lists the most frequent constructions (MFCs), i.e. those constructions with frequencies over 1 per 10,000 tokens when the two corpora are combined, ordered by log-likelihood from most overused by learners to most underused by learners (i.e. overused by experts). The bottom section lists those constructions with combined frequencies not great enough to warrant a log-l calculation, ordered by descending combined frequency.

Table 5.8. Number and frequency per 10K of modal and p-modal constructions in LOH and EOH corpora in comparison

modal or p-modal construction	all LOH		all EOH		log-l LOHvEOH*
	total	freq/10K	total	freq/10K	
you can (undiff.)	536	52.137	340	27.126	+91.67*
you had better	29	2.918	0	0.000	+46.24*
you should	103	10.019	46	3.670	+35.13*
I want you to	9	0.875	6	0.479	+1.35
you don't need to	17	1.654	16	1.276	+0.14
you have to	55	5.350	87	6.941	-2.32
you need to	41	3.988	76	6.063	-4.84*
you could	9	0.875	59	4.707	-31.87*
I would	1	0.095	38	3.032	-45.52*
you want to (wanna)	12	1.167	173	13.802	-137.51*
you have got to ('ve gotta)	0	0.000	12	0.957	
you must	6	0.584	0	0.000	
I don't want you to	0	0.000	4	0.319	
it is/would be better if you	4	0.389	1	0.080	
you are better off	0	0.000	4	0.319	
you don't have to	3	0.292	0	0.000	
you should not	2	0.195	0	0.000	
you may	2	0.195	0	0.000	
you may not	0	0.000	1	0.080	
TOTAL w/can	828	80.540	863	68.851	+10.37*
TOTAL w/o can	292	28.403	523	41.726	-28.58*

\* A plus sign indicates LOH overuse, while a minus sign indicates LOH underuse, in comparison to EOH. A log-likelihood greater than 3.84 indicates a p-value less than .05, and thus, a significant difference between LOH and EOH production of this construction. Statistical difference was not calculated for constructions whose combined EOH+LOH frequency were less than 1.0 per 10K.

The results show that in aggregate, the learners generally underuse modal and p-modal constructions in comparison to experts (log-l = -28.58), unless 'you can' is included, in which case there is actually slight overuse (log-l = +10.37). In other words, in office hours contexts, learners seem to prefer 'you can', 'you had better', and 'you should' where experts prefer 'you could', 'you want to', and the irrealis construction 'I would'. Most obviously, this warrants further investigation of learner use of 'you can'. It could be that the learners are using 'can' in an appeal towards S independence, since 'can' in its dynamic 'to be able to' sense infers choice and possibility. However, when spoken by someone in a position of authority relative to the listener, 'you can' may be interpreted as permission rather than possibility, which may be more clearly

expressed with ‘you could’, which is perhaps not coincidentally underused by the learners.

Chapter 7 explores the nature and consequences of learner overuse of ‘you can’ and ‘you had better’ in more depth by examining individual cases. The use of ‘had better’ by the learners, where the experts didn’t use it at all, also serves to imply authority and promote dependence and lack of choice, as it implies undesirable consequences with non-compliance.

Of the other constructions, there seems to be learner underuse of p-modals, including ‘you need to’, ‘you have to’, and ‘you’ve got to’, while there is relative overuse of ‘you should’. The use of ‘you should’ generally implies that the speaker agrees morally with the source of the obligation, if he is not that source himself. As explained in chapter 4, in some grammars p-modals have the effect of objectivizing or distancing the speaker from the source of the directive (Palmer, 2001), and may be considered a grammatical metaphor (Iedema, 1996). Using p-modals instead of traditional modals may thus be a strategy to indirectly build solidarity and involvement, as the director/speaker can oblige the directed/listener, but distance himself from the source and avoid appearing to make a moral judgment. In office hours situations, the effect of the director/speaker (T) preferring ‘you should’ to p-modal equivalents, then, could be that directed/listener (S) perceives the T as more authoritarian and judgmental than the T intends to be (see chapter 7 on the disconnect between ITA intention and use).

In a similar vein, the experts use ‘you want to’ very frequently<sup>7</sup>, while the learners use it very infrequently, perhaps because assuming to know another’s desires is considered inappropriate in many cultures. Experts may use ‘you want to’ to distance themselves from the source of the obligation while indirectly building solidarity and involvement, while projecting a

---

<sup>7</sup> Biber (2006) must not consider ‘you want to’ a directive, as he reported no such constructions in his T2K-SWAL office hours corpus, although he mentions the presence of ‘we want to’ and ‘I want you to’ (p. 124). Similarly, he mentions ‘I would’, but only where ‘would’ acts as a mitigator with the lexical verb holding the directive force (suggest, recommend, etc.).

measure of subjectivity onto the directed/listener. This ‘persuasive want to’ may be highly prevalent in office hours because of this projection quality. ‘I would’ is also very frequent in the EOH while not at all in the LOH. As an unrealis device, it may function to build solidarity and involvement through the implication of reciprocity (Brown & Levinson, 1987), since it indirectly infers that the directed-listener should do as the director-speaker would do, which in an unequal institutional power relationship would be very risky to decline<sup>8</sup>. These persuasive involvement strategies are explored and discussed in the genre moves analysis in chapter 6.

### 5.3.2.1.2 Directive vocabulary constructions

Table 5.9. Number and frequency per 10K of directive vocabulary constructions in LOH and EOH corpora in comparison

construction	all LOH		all EOH		log-1 LOHvEOH
	total	freq/10K	total	freq/10K	
I suggest	19	1.848	3	0.239	+16.36*
I recommend	17	1.654	1	0.08	+20.58*
(s.o.) they suggest	0	0.000	3	0.239	
my suggestion	9	0.875	0	0.000	
(a/some/two) suggestion	7	0.681	0	0.000	
(I/we) allow	5	0.486	0	0.000	
(sthg) is required	4	0.389	0	0.000	
you are not allowed	2	0.195	0	0.000	
(student/this) is not allowed	2	0.195	0	0.000	
I encourage	2	0.195	0	0.000	
it is not necessary	2	0.195	0	0.000	
you are supposed to	2	0.195	2	0.16	
I do not recommend	1	0.097	0	0.000	
you are not supposed to	1	0.097	1	0.08	
TOTAL	73	7.101	10	0.798	+67.29*

Comparison of the use of directive vocabulary constructions by learners and experts in table 5.9 above shows that learners use these constructions much more frequently than experts (log-1 = +67.29). Blum-Kulka, House, and Kasper (1989) note that performatives like ‘I suggest’ or ‘I recommend’ are nearly as direct as bald imperatives. From an SFL perspective, directive

<sup>8</sup> It would, in fact, require the directed-listener to say something like ‘well, you’re not me’ and claim independence, which would have the effect of rejecting involvement overtures.

vocabulary used by an institutional authority with a third person subject would distance the listener from the power source of the directive (e.g. ‘it is required’ or ‘the administration suggests’), as a result of metaphorization. Such directive vocabulary uses thus influence the objectivity level, while pronoun choice influences the directive on the level of involvement.

### 5.3.2.1.3 Imperatives

Table 5.10 Number and frequency per 10K of imperative indexicals in LOH and EOH corpora in comparison

indexical	all LOH		all EOH		log-l LOHvEOH
	total	freq/10K	total	freq/10K	
please...	60	5.836	0	0.000	+95.66*
don't worry	35	3.404	0	0.000	+55.80*
let me...(1)	42	4.085	20	1.596	+12.95
let's/let me see	38	3.696	24	1.915	+6.57*
just...	42	4.085	46	3.670	+0.25
don't...(2)	17	1.684	28	2.234	-0.98
let's...(1)	24	2.334	78	6.223	-20.40*
TOTAL (all)	258	25.096	196	15.637	+25.23*
TOTAL (w/o 'please' and 'don't worry')	163	15.884	196	15.637	0.02

(1) not including ‘see’; (2) not including ‘worry’

Table 5.10 above shows that of the imperative indexicals listed here, overall there was no significant difference in usage by the learners and the experts, if ‘please’ and ‘don’t worry’ are omitted. It is striking that the learners used ‘please’ so frequently when the experts used none. Yet MICASE experts do use ‘please’, just not in office hours, but rather in lectures<sup>9</sup>. Use of ‘please do (something)’ to an individual in office hours may construe negative politeness and thus independence, but it may serve to distance speaker and listener (Yates, 2005), thus inferring exclusion. In light of the lack of any expert uses, the overuse of ‘don’t worry’ by learners may be an artifact of the task, or it may indicate a comforting, yet parental attitude towards the listener, which would again serve to construe authority. The indexicals ‘just’ and ‘don’t’ represent

<sup>9</sup> There were 33 uses of ‘please (do something)’ in the MICASE lecture subcorpus.

imperatives in general, and are not significantly over or underused by the learners compared to the experts. However, it is notable that ‘let me’ is overused by the learners, while ‘let’s’ is underused, since ‘let me’ could be interpreted as negatively polite and exclusive, while ‘let’s’ is inclusive. Thus, learners may construe independence but not involvement.

In summary, the overall, aggregate picture of learner core directive usage, in comparison to expert usage, is that learners, whether intentionally or unintentionally, may evoke exclusion, but not necessarily dependence, on the part of the directed/listener, when playing the T role in office hours contexts. There are many caveats, however, when each construction and individual is considered separately. This rough picture can now be brought into focus by an investigation of peripheral constructions like adjuncts (below), complemented by genre analysis (chapter 6), and enhanced by qualitative analysis of individual expert and learner usage (chapter 7), before discussion of the pedagogical treatment (chapter 8).

### **5.3.2.2 Adjuncts**

As discussed in chapter 4, adjuncts to a core lexical element (usually verbal) are understood to play a role in the directive force of a construction, and thus can be considered part of a construction. To determine this role, and the differences between usage in the two corpora, a comparative corpus analysis of adjunct use was conducted on the most frequent constructions (MFCs), except for the ‘can’ constructions. As explained above, these MFCs included any modal, p-modal, and directive vocabulary construction (but not imperative indexical) with a frequency of over 1 per 10K in the combined EOH and LOH corpora. Since searching every

directive construction would have been prohibitively time consuming, these MFCs are considered representative of overall adjunct use<sup>10</sup>.

The MFCs were searched for the presence of adjuncts or ‘small words’ (Hasselgreen, 2002; Schiffrin, 1988) that modified the directive force of the construction in any way, within 5 places to the left or right of the core element of the construction. The adjuncts are discussed below, and a full list is presented in appendix D. The adjuncts in the MFCs were tallied, and frequencies and percentages were calculated. To present a full picture of adjunct use, adjunct type and diversity were also measured and compared.

### 5.3.2.2.1 Overall frequency

Table 5.11. Number and frequency of MFCs and MFCs with adjuncts in EOH and LOH corpora in comparison

	EOH		LOH		log-l LOHvEOH
	count	freq per 10K	count	freq per 10K	
all MFC uses	461	36.78	300	29.18	9.87*
all MFC uses w/ adjuncts	280 (61%)	22.34	166 (55%)	16.15	11.24*

As presented in table 5.11 above, overall the learners used adjuncts with the MFCs less frequently than did the experts (LOH = 16.15, EOH = 22.34), which is a significant difference (log-l = -9.87). However, the learners included adjuncts 55% of the time that they used these constructions, while the experts included adjuncts 61% of the time, which is not significantly different (log-l = -0.91). The two measures are important to note separately, as frequency measures in relation to all production, while percentage measures production in relation to only the production of MFCs. This means that the learners, with several exceptions, aren’t producing as many constructions with adjuncts, but when they do use those constructions, they apply adjuncts nearly as frequently as the experts. In other words, it may not necessarily be that

<sup>10</sup> Imperatives were also omitted from this analysis because they were already representative of overall use. One of the indexicals, ‘just’, would itself be considered a mitigating adjunct.

learners do not know how to use adjuncts, but rather that they possibly do not understand the use of directives, or that they are simply choosing to use them at this frequency, which is lower than that of experts. For a closer analysis, it is useful to look at the most frequent constructions separately, and the types of adjuncts used. Table 5.12 below presents the differences in the frequency adjuncts were found in the co-text of each of the most frequent constructions, as well as the percentages of constructions found with adjuncts.

Table 5.12. Frequencies of individual MFCs with and without adjuncts in EOH and LOH corpora

construction	EOH freq per 10K			LOH freq per 10K		
	w/ ADJ	w/o ADJ	% with	w/ ADJ	w/o ADJ	% with
you should	1.76	1.91	48%	5.35	4.67	53%
you have to	4.15	2.79	60%	2.82	2.53	53%
you need to	3.27	2.79	54%	2.33	1.65	59%
you don't need to	0.88	0.4	69%	1.07	0.39	73%
you want to	9.49	4.31	69%	0.88	0.29	75%
you could	2.63	2.07	56%	0.68	0.19	78%
you had better	0	0	0	1.36	1.46	48%
I recommend	0	0.08	0	0.97	0.68	59%
I suggest	0.16	0.08	67%	0.68	1.17	37%
TOTAL	22.34	14.44	61%	16.15	13.05	55%

Of the MFCs, 'you should' is notable because not only do the learners use it more frequently than the experts, they use it with adjuncts more of the time. In other words, while learners may be overusing 'you should' in comparison to experts, they are also modifying it more frequently. It is also interesting to note that although the learners use 'you could' and 'you want to' far less frequently than the experts, the learners who do use them make use of adjuncts. On their part, the experts use adjuncts with 'you want to' and 'you don't need to' more than two-thirds of the time. The matter of which adjuncts are used in all of these cases requires closer multidimensional investigation of different constructions and adjunct type.

### 5.3.2.2.2 Frequency by adjunct type

To determine and compare the type of adjuncts used, the adjuncts were categorized into five categories (see appendix D for complete list): mitigators, intensifiers, hesitation markers, feedback-oriented discourse markers, and conjunctive discourse markers. Each MFC in the above analysis was assigned only one adjunct; for constructions with multiple adjuncts, the overall effect based on the adjunct closest to the core device was designated as the adjunct for that construction. Mitigators (MIT), which include words and phrases like ‘maybe’, ‘probably’, ‘I think’, and ‘kind of’, as in constructions like ‘maybe you should’, are used for purposes of negative politeness, and serve to appeal towards independence and lessen imposition. In contrast, intensifiers (INT), which include words and phrases like ‘absolutely’, ‘do’, ‘really’, and ‘still’, as in constructions like ‘you really need to’, are also used for negative politeness, but serve to strengthen the speaker’s commitment to the proposition, and in a directive context, curb choice available to the directed/listener<sup>11</sup>. The other three categorized were separated out from mitigators and intensifiers, since many of them appeared to serve multiple purposes, and primary speaker intent was impossible to tell. Hesitation markers (HES), like ‘um’, ‘uh’, and ‘mhm’, as in constructions like ‘um you don’t need to’, can act as mitigators, but they can also be simple disfluencies. Feedback-oriented discourse markers (FOD) like ‘okay’, ‘right’, ‘yeah’, and ‘you know’, as in constructions like ‘yeah, you want to’, serve as discourse markers as well as simple positive feedback that appeal to involvement or solidarity, i.e. positive politeness, while conjunctive discourse markers (COD) like ‘so’, ‘and’, ‘but’, and ‘or’, as in constructions like ‘but you have to’, cohere the construction to the surrounding co-text, and may act as negative politeness markers depending on that co-text.

---

<sup>11</sup> It should be noted that use of intensifiers with ‘don’t need to’ serves to mitigate and offer more choice to the listener, rather than to intensify and restrict choice.

Table 5.13. Frequencies of MFCs categorized by adjunct type in EOH and LOH corpora

construction	freq. MIT		freq. INT		freq. HES		freq. FOD		freq. COD	
	EOH	LOH								
you should	0.56	3.02	0.32	0.19	0.16	0.58	0.32	0.1	0.4	1.46
you have to	1.44	0.78	0.56	0	0.16	0.49	0.4	0	1.6	1.56
you need to	1.6	0.97	0.16	0	0.56	0.39	0.24	0.1	0.72	0.88
you don't need to	0.8	0.58	0	0	0.08	0	0	0.1	0	0.39
you want to	5.58	0.88	0.48	0	0.4	0	0.96	0	2.07	0
you could	1.04	0.39	.24	0.1	0.4	0.1	0	0.1	0.96	0
you had better	0	1.07	0	0	0	0.1	0	0	0	0.19
I recommend	0	0	0	0.68	0	0	0	0.1	0	0.39
I suggest	0.08	0.19	0	0	0.08	0	0	0.1	0	0.39
TOTAL	11.09	7.88	1.76	0.97	1.83	1.65	1.91	0.58	5.74	5.06

Table 5.13 shows the frequencies of all the various types of adjuncts used with the most frequent constructions. Mitigators were used with these constructions less frequently by the learners (freq MIT = 7.88) than by experts (freq MIT = 11.09). As with their choice of directive (see §5.3.2.1), this could mean, then, that the learners, when in T roles, are perceived as more authoritative and offering less choice by their ‘students’. Regarding the constructions, it is interesting that the learners use mitigators with ‘you should’ much more frequently (freq MIT = 3.02) than the experts (freq MIT = 0.56), which again speaks towards the learners’ productive use of that construction. Also, the experts mitigate ‘you want to’ very frequently (61% of its adjuncts are mitigators)—nearly half of those with ‘might’ (‘you might want to’)<sup>12</sup>. It is also noteworthy that overall, the learners use intensifiers less frequently (freq INT = 0.97) and with fewer constructions than the experts (freq INT = 1.76). This may indicate intentional avoidance of intensifiers in a reverse attempt to show respect, although it could be argued that the directed/listener might then interpret the avoidance not as true mitigation but rather detachment. Regarding hesitations, it is interesting that the learners use these markers less frequently (freq

<sup>12</sup> While the high frequency of ‘you might want to’ may not be intuitively surprising, what is more surprising is that this construction is not taught as a means to give a directive or suggestion, maybe because it is not traditionally considered a p-modal.

HES = 1.65) than the experts (freq HES = 1.83), albeit not significantly. While a hesitation might indicate general disfluency (which might be better gauged by an overall count), hesitation as an adjunct to a directive might also be an intentional mitigation strategy, in which case the learners once again would be mitigating less than the experts. The learners also use feedback-oriented discourse markers like ‘okay’ and ‘yeah’ less frequently (freq FOD = 0.58) than the experts (freq FOD = 1.91) in the co-text of directive constructions. Because these markers are generally used to show listenership and involvement (Farr, 2003), as with hesitators, an overall count might be more telling. Still, the lower frequency of FOD use by the learners is congruent with the other findings that the learners make less overall use of politeness strategies than the experts. Finally, regarding the use of conjunctive discourse markers, the learners use them slightly less frequently (freq COD = 5.06) as the experts (freq COD = 5.74). To interpret these results, the function of these markers would need to be examined on the level of broader co-text, since sometimes, for example, the use of ‘but’ in ‘but you need to’ could be interpreted as an intensifier, and sometimes as simply a cohesive device between phrases. This is a subject for future investigation.

#### **5.3.2.2.3 Adjunct diversity**

In another measure, the diversity of adjuncts used was tallied. All adjuncts were counted separately, so that the token counts are greater than the number of constructions above, since many constructions had multiple adjuncts. The types of adjuncts were also counted, as a direct measure of diversity. Finally, the five most frequent adjuncts were noted for each MFC. Table 5.14 below presents the results.

Table 5.14. Number of tokens, types, and most frequent adjuncts for MFCs in EOH and LOH corpora

construction	EOH adjuncts		LOH adjuncts	
	no. tokens (no. types)	most frequent adjuncts by type (no. tokens)	no. tokens (no. types)	most frequent adjuncts
you should	27(17)	so (6); uh/um/mm (3); probably, still, yeah (2 ea.)	57(12)	I think (15); but, maybe (11 ea.); uh/um/mm (7); so (4)
you have to	73(26)	so (17); just (6); but, okay, right, yeah (5 ea.); then, well (4 ea.)	29(7)	so (10); but, I think, uh/um/mm (5 ea.)
you need to	52(20)	uh/um/mm (11); so (6); you know (5); would (4); but, I think, just (3 ea.)	27(9)	but (8); I think, uh/um/mm (4 ea.); so, yeah (3 ea.)
you don't need to	11(6)	really (5); you know (2)	12(7)	so (4); I think (3)
you want to	200(29)	so (33); might (32); uh/um/mm (23); right (16); but (12)	9(3)	might (4); may (3); maybe (2)
you could	37(17)	I mean, uh/um/mm (6 ea.); but, just (3 ea.)	7(4)	maybe (4)
you had better	0	n/a	19(8)	I think (7); maybe (3)
I recommend	0	n/a	11(6)	will (4); so, strongly (2 ea.)
I suggest	2 (2)	might wanna (1); um (1)	7(4)	but, just, so (2 ea.)
TOTAL	402(46)	so (64); uh/um/mm (46); might (33); but (24); just(19)	178(25)	I think (34); so (27); maybe (21); uh/um/mm (19); yeah (7)

The results show that the experts used 46 different types of adjuncts 402 times in the 280 MFCs with adjuncts they produced, an average of 1.44 adjunct per MFC. In comparison, the learners used 25 different types of adjuncts 178 times in the 166 MFCs with adjuncts they produced, an average of 1.07 adjunct per MFC. In other words, the learners used fewer adjuncts per construction, and from a smaller repertoire. For both groups, the most frequent conjunctive discourse marker was 'so'. The most frequent mitigating adjuncts produced by the experts were 'might' and 'just', while the learners produced 'I think' and 'maybe' most frequently. An exemplar of the most common learner construction with a mitigating adjunct would be 'I think you should', while for the experts it would be 'you might want to'. It is notable that nine learners

but only one expert used ‘I think you should’, while a single learner produced all four learner uses of ‘you might want to’<sup>13</sup>.

In summary, learners make less frequent use of fewer types of adjuncts, most importantly mitigating and intensifying adjuncts, than do the experts, at least by the measure of adjunct use with the constructions they use most frequently. While the above conclusions on overall directive use were that learners did not evoke inclusion and solidarity to the degree that the experts did, this analysis adds the observation that the learners did not evoke independence and choice through mitigating adjunct use as frequently as experts did. This accords with the findings of the reviewed corpus research above and research on English language learner production in institutional discourse (chapter 2).

#### **5.4 Conclusion**

This chapter used corpus-based register analysis to present a picture of the systematic usage of directive constructions by learners and experts in office hours contexts, both internally to establish consistency, and externally for comparative purposes. Using contrastive interlanguage analysis (Granger, 1998; 2002), I compared directive construction usage in the Learner Office Hours (LOH) corpus from ITAcorp with the Expert Office Hours (EOH) corpus from the Michigan Corpus of Academic Spoken English (MICASE) (Simpson et al., 2002). Seen through the social-functional theoretical framework developed in chapter 4, an aggregate picture of learner directive construction usage emerged, which showed that learners, in comparison to experts, whether intentionally or unintentionally, do not as frequently invoke inclusion on the part of the directed/listener through their choice of core directive device, when playing the T role in office hours contexts. Regarding adjuncts, learners did not evoke independence and choice

---

<sup>13</sup> One expert used 27 of the 32 expert uses of ‘you might want to’, while 3 others produced the other 5 uses, out of a total of 173 uses of ‘you want to’. The phenomenon of preferred forms is discussed in chapter 7.

through mitigator use as frequently as experts did. In addition, they also made less use of intensifying adjuncts, which could be interpreted as detachment rather than politeness<sup>14</sup>.

While this chapter demonstrated the heuristic strengths of aggregate corpus analysis of register, it approached the data from the functional entry point of the proposed social-functional framework. To complement the above register analysis and complete the picture of systematic usage, a genre analysis that approaches the data from a social entry point will be presented next in chapter 6. By considering directives usage at the level of text and context, this analysis will explicate more thoroughly the relationship between directives usage and the exercise of symbolic authority and power.

---

<sup>14</sup> Whether a teacher/speaker not evoking independence and inclusion is the same as evoking dependence and exclusion is an empirical question to ask of students/listeners. It is also arguable whether there is such thing as 'politeness-neutral' discourse when there is unequal distribution of power among interlocutors (Bourdieu, 1991).

## CHAPTER 6

### A GENRE ANALYTIC VIEW OF OFFICE HOURS

In complement to chapter 5, where I examined the systematic qualities of directives usage in the office hours register using corpus analytic techniques, in this chapter I explore office hours and directives usage using genre analytic techniques that focus on both text and context. As based on the research question posed in chapter 1, the goal of this chapter is to explore the text type of ‘office hours’ as a genre through both contextual analysis and a genre moves analysis (Swales, 1990). I first conduct a contextual analysis of the ‘office hours’ context through examination of a variety of features specific to the social participation of ITAs in these contexts. To complement this situational approach, I then conduct an empirically driven genre moves analysis on several transcripts from the EOH corpus from the MICASE. The two approaches to genre analysis have complementary purposes. The purpose of the contextual analysis is to critically examine the office hours contexts in which ITAs find themselves to illuminate the interrelationships and points of tension. The purpose of the moves analysis is to explore the relationship between directive construction usage and textual moves, thereby relating functional language use realized through register and social activity realized through genre. The overarching purpose of both analyses is to identify possible concepts for inclusion in an ITA preparation curriculum.

In chapter 4, I developed a social-functional framework for analysis of directive usage, and the concept of directive construction was proposed as a multi-unit analytical unit that would allow interpretation of the corpus analysis findings to consider social aspects of usage. This

social-functional framework is now extended to consideration of genre, for analysis of larger chunks of text as moves and the role of speech act constructions within those moves. The relationship among the three approaches is presented in Table 6.1 below, with the broadest analysis, contextual, at the top, and the most narrow, register, at the bottom, mediated by text-based genre analysis.

Table 6.1. Approach, object and unit of analysis in three approaches to analysis of directives usage in office hours

<b>Analytic approach</b>	<b>Object of analysis</b>	<b>Unit of analysis</b>
genre (contextual)	context	situation (office hours)
genre (linguistic)	text	moves or stages (greetings, diagnosis, directive, closings, etc.)
corpus-based register	discourse (lexico-grammatical units)	directive constructions (core and peripheral)

## 6.1 Office hours as genre

Biber (2006) explains that for him, the difference between genre and register is that genre is the ‘context-plane’ of register, while register is the ‘expression-plane’ of genre. Genre can be considered a culturally-recognized ‘message type’ that is concerned with ideology and social power, while register is a general kind of language associated with a domain of use that systematically varies in regards to lexico-grammatical features. From a systemic-functional tradition, register is a text type seen from the ‘system pole of the cline of instantiation’, or a “functional variety of language—the patterns of instantiation of the overall system associated with a given type of context (situation type)” (Halliday and Matthiessen, 2004, p. 27). Genre is the same text type seen from the context of culture, with social purpose as its defining characteristic. Martin (1985, in Swales, 1990 p. 40) explains that genres are realized through registers, which are realized through language. Thus, looking at language use through the lens of

genre allows us to apprehend the social, while corpus-based register analysis of the same language use allows consideration of the functional at the lexico-grammatical construction level.

Swales (1990) discusses the definition of genre from four different traditions: folklore studies, where it has been understood as a classificatory category, an ‘ideal type’, and an evolving form; literary studies, where it has been understood as an unstable, clarificatory form; rhetoric, where it has been understood as an historical, heteroglossic form defined by action and purpose; and linguistics, where it has been understood as a type of communicative event defined by a particular speech community. He notes that ‘genre’ can be characterized by several features, most notably a shared set of communicative purposes, prototypicality relationships, constraints on membership (not always defined), and a community label. Swales defines ‘genre’ most succinctly as comprising “a class of communicative events, the members of which share a set of communicative purposes” (p. 58), as well as structure, style, content, and intended audience. He notes that a genre is most clearly recognized by expert members of the community that engenders it.

In a more recent edited volume, Johns (2002) explains that applied linguistic study of genre falls into several schools or approaches. Paraphrasing Flowerdew (2002, in Johns, 2002), she explains that linguistic approaches tend to “foreground language and the texts themselves” (Johns, 2002, p. 4), and use functional grammar theory to analyze genres primarily from a lexico-grammatical perspective. These approaches include the Sydney School, led by Halliday and Martin, and studies of English for Specific Purposes, which include Swales’ genre moves analysis framework. Johns explains that in contrast to these linguistic approaches, researchers adopting contextual approaches like New Rhetoric see genre from a more critical perspective, and argue an inextricable relationship between context and power (Johns, 2002, p. 9). Genres are

thus “textual tools, exploited for social, and sometimes hegemonic, purposes within communities by knowledgeable experts” (p. 9). Johns introduces a pair of studies she considers to be transitional between the linguistic and the contextual, one that advocates using both approaches in the classroom, and another that uses both in the analysis of the texts and context of a university graduate program.

## **6.2 Contextual analysis**

In view of the strengths of both approaches to genre analysis, a dual approach similar to those labeled by Johns as transitional is taken here. For the first, contextual analysis, Tribble (2001) provides a useful framework in his analysis of a scientific journal article. Before conducting a linguistic register analysis of the article, he examines the contextual, generic features of the article focusing on a variety of features: the name of the genre, its social context, its communicative purposes, participant roles, cultural values, text context, and formal text features. This framework is applied here for examination of office hours contexts as they relate to expert and novice (in this case learner) participants.

### **6.2.1 Name: ‘Office Hours’**

In American academia<sup>1</sup>, the name of the genre under study, ‘office hours’ is institutionally defined as an officially recognized time period when an instructor can receive a student in his/her office. The name is not particularly descriptive, considering the variety of activities that occur in these sessions, which include advising, tutoring, counseling, chatting, apprenticing, reviewing, and evaluating. Swales points out that “the fact that a communicative event is labeled by the institution as being an event of such-and-such a kind does not necessarily

---

<sup>1</sup> As noted in chapter 1, ‘American’ here refers to universities structured according to a model originating in the USA in the late 19<sup>th</sup> and early 20<sup>th</sup> century, prevalent in the US but increasingly influential of higher education structures world-wide.

mean that it will be so” (p. 55). In Swales’ own MICASE corpus, in fact, the office hours texts vary in terms of topic, and include discussions of research papers, homework problem solving sessions, and preparations for exams. What the sessions have in common is that their purposes are both evident and constrained (Swales, 1990, p. 55) to the expert participants, who know what content is appropriate and what is inappropriate for such events.

### **6.2.2 Social context**

The social context of the office hours sessions is one where there is a single T participant, usually an instructor or teaching assistant, and a single S participant (though sometimes more), usually a student, in a semi-public, conversational setting, usually the T’s office, but sometimes in public ‘third space’ environs friendly to open conversation, like a local coffee shop or bar. Expert participants would most likely find a private setting like an instructor’s home, or a quiet setting like a library inappropriate for office hours. In North American academic settings, the ‘open door’ policy represents the semi-public, conversational nature of the office hours event. The setting thus constrains certain features like session length and appropriateness of purpose and topic.

### **6.2.3 Communicative purpose**

The communicative purpose of the office hours session is typically to provide an opportunity for the S to meet with the T in a one-on-one environment to discuss academic matters, outside of the classroom setting and its constraints. Although office hours may include academic advising<sup>2</sup>, they are usually directly connected to a particular course. While they may involve a litany of purposes as mentioned above, like advising, tutoring, counseling, etc., they

---

<sup>2</sup> Although office hours sometimes includes advising, students in American academic contexts are often assigned a separate advisor with whom they do not take classes, and thus not in a T-S relationship with them.

are also defined by what purposes are inappropriate, for example providing answers to homework, being intimate, or somehow ‘getting off track’. For the S, the purpose may be to ask questions s/he could not ask for whatever reason during class, while for the T, the purpose may be to provide personalized feedback and evaluate student learning. For both, there is the shared goal of the resolution of specific and predictable kinds of academic problems and issues. The knowledge of what is appropriate and what is inappropriate is a matter of expertise, as these purposes may be only partially recognized by less experienced participants (see also Bardovi-Harlig & Hartford, 1993), which would include new students, and in the case of the current project, new international teaching assistants.

#### **6.2.4 Participant roles**

The office hours text type shares many features with other forms of educational institutional discourse (Agar, 1985; Bardovi-Harlig & Hartford, 1993), in that the institutional roles of the participants afford the purposes with which they enter the event. Unlike in other institutional discourse genres like writing center interactions (Thonus, 1999; Williams, 2005) or even advising sessions<sup>3</sup>, the participant relationship in office hours is unequal and hierarchical, since the T retains the power of evaluation and assessment in the associated course, even as the relationship is extended outside the classroom. This power structure affords the T to assume the role of director, while the S assumes that of the directed, even though the S often initiates the interaction. As Bardovi-Harlig & Hartford (1990; 1993) observed in advising sessions, this interaction becomes even more complex as it involves intricate sociopragmatic awareness of when to observe maxims and when to negotiate. On the other hand, the office hours context is for some a less formal environment, where the T and S can build a more personal relationship

---

<sup>3</sup> In undergraduate and lower graduate advising sessions, the advisor does not necessarily have assessment power over a student in the same way that student’s teacher does.

than may be possible in the classroom. Individuals, both T and S, with more experience in the American tradition develop office hours genre expertise through years of apprenticeship, while individuals coming from other traditions may not understand how office hours both reflects and helps define T-S relationships.

Another important issue for consideration is the role of teaching assistants in contrast to the role of teachers or professors who retain evaluative power. Teaching assistants may discuss homework, tutor students on course content, preview or review exams, explain course requirements, and grade exams, but may deflect some questions and comments they feel unauthorized to answer. In this sense, TAs act as mediators and in some ways spokesperson, as they may be the first, or even only representative of a department with which a student may have lengthy interaction. International students may be unfamiliar with, and are often untrained in, this mediating role where they are expected to conduct triage and diagnosis, as it were, for their course and department.

### **6.2.5 Cultural values**

Cultural values are at the core of office hours involving international participants, as these values influence their understanding of roles, role expectations, and purposes. ‘Office hours’ may not exist outside American or American-influenced academia, or they may be very different in form and purpose<sup>4</sup>. For example, in comparison to American traditions, academic T-S relationships have been reported in Asian/Confucian traditions to be more hierarchical and authoritarian, while relationships in Latino traditions may be more formal yet more implicit. While these are broad generalizations and it is not the intention of this project to explore these

---

<sup>4</sup> My own experience in Japan was that they did not exist there, even in a college whose curriculum was heavily influence by American liberal arts traditions. I would expect that in the United Kingdom as well, office hours as would be supplanted by tutoring sessions.

specific cultural influences in-depth, recognition of these influences is important in a contextual analysis. The concept of office hours as an officially recognized S right to T access, or as a means for personal relationship building, reflects certain egalitarian and consumerist interpersonal values that are specific to American culture. In chapter 7, cultural values vis-à-vis institutional roles will be discussed in conjunction with the individual case studies of learner usage of directives.

### **6.2.6 Text context**

In Tribble's (2001) framework, text context refers to knowledge of other texts that may be required of participants. In the case of office hours, this would refer to spoken and written text types like institutional regulations, class participation protocols, syllabi, exams, discussion, presentations, homework, and research papers. As with office hours, context of situation and culture are constitutive of each of these. To participate expertly in office hours, familiarity with each of these may be required to a greater or lesser degree.

### **6.2.7 Summary of contextual analysis**

In sum, office hours can be seen as an academic discourse genre engendered by the affordances and constraints of an academic ecology or context, that includes "expected norms of interaction such as turn-taking, constant social relations/roles, and asymmetrical power relationships" (Bardovi-Harlig & Hartford, 2005, p. 9). The expertise and knowledge of the rules and practices of office hours, understood by both T and S core participants, is seldom made explicit to newcomers, who may come with very divergent conceptualizations of T-S roles and relationships. Office hours play an important part in American academia, as they are, as much as the classroom, a context for the reproduction and reification of academic cultural structures and

values. The linguistic means for this reproduction in office hours may be more fully understood by examining office hours as texts, and the relationship between moves and speech act constructions. As the following analysis shows, directives are a primary means for this function.

### **6.3 Genre moves analysis**

In complement to a contextual, situational approach, a more text-based approach to genre analysis is to consider the moves or stages that typically comprise the genre. Labov (1972, in Swales, 1990) first conducted such an analysis on inner-city spoken personal narratives, and identified a series of moves including abstract, orientation, complicating action, resolution, and coda. Swales (1990) conducted moves analysis on research genres like academic journal articles, and broke each ‘move’ into ‘steps’. Unlike Swales’ approach where moves and steps occur in a sequential order, systemic-functionalists analyze ‘stages’ of a given genre that may not occur in any particular order, like recount, report, explanation, procedure, and narrative.

As a text type, an office hours text has certain predictable phases like openings and closings, and as a type of institutional discourse similar to writing center interactions and advising sessions, it may have phases like diagnosis, report, and directive (Agar, 1985; Williams, 2005). Agar describes ‘diagnosis’ as “the process through which the institutional representative fits the client frame to the institutional frame” (p. 149) by asking questions and making meta-commentary to both elicit information and control the interaction. In the ‘report’ phase, Agar explains that the institutional representative fits the client frame into the institutional frame to document or archive the visit. In office hours and other educational institutional discourse text types, this phase may be less prominent, it may occur after the visit, or it may be a matter of informal evaluation or formative assessment on the part of the T participant. Finally, the

‘directive’ phase, which Agar explains is not always part of an institutional interaction, involves the representative directing the client into responsive action.

### 6.3.1 Data

The data for the moves analysis is from the EOH corpus in MICASE as described in chapter 3. As discussed in that chapter, the EOH transcripts consist of a variety of office hours purposes, including tutoring, advising, instruction, and counseling. The transcript numbers, names, T information, location, and descriptions (purposes) are presented in Table 6.2 below.

Table 6.2. EOH T participant information and office hours location and description

Trans no.	Name (subject)	T status	T gender	T age	Location	Description (purpose)
175	Intro Biology Exam Review	JG	F	17-23	classroom	T review exam Qs with 10 Ss in a group, very mixed interaction
195	Heat and Mass Transfer Office Hours	JG	M (NNS)	24-30	classroom	T reviews exam with 20 Ss individually and in small groups
115	Anthropology of American Cities Office Hours	SG	F	24-30	coffee shop	T conferences with 4 Ss on their course research papers individually
270	Computer Science Office Hours	SG	F	24-30	office	T reviews various homework problems with 11 Ss individually
575	Statistics Office Hours	JF	F	31-50	office	T reviews homework with 10 Ss in a group
320	Art History Office Hours	JF	F	31-50	office	T helps 4 Ss choose a topic for their final project individually
280	Economics Office Hours	JF	F	31-50	office	T helps 9 Ss with various topics, individually and in small groups
300	Intro to Poetry Office Hours	SF	M	51-over	office	T reviews course work with 2 failing Ss, reviews exam and work with 1

In testament to the thoughtful design of MICASE, the representativeness of the EOH corpus is broad in field and level (sciences and humanities, introductory and advanced courses), T status (junior and senior graduate TAs, junior and senior faculty), T age, and location. Six of the eight T participants are female, which is irrelevant at this point as gender differences are not an object of this moves analysis, although gender differences may play a role in the interpretation of the learners profiled in chapter 7. Most apparent in the above table is the diversity of purpose

among the various office hours sessions, including exam review, paper conferences, advising, homework review, project topic selection, and even failure intervention. Representing the diverse, interactional modes of office hours, four Ts (115, 270, 320, 300) held one-on-one, successive sessions, two Ts (195, 280) held one-on-several mixed with one-on-one successive sessions, and two Ts (175, 575) held a one-to-many session.

### **6.3.2 Method**

To explore the moves present in office hours texts, as well as any order in which they may occur, transcripts from the EOH were analyzed for common moves using a grounded approach, informed by the definitions of Agar's (1985) institutional discourse 'phases'. The data excerpts were chosen from any of the 8 transcripts because of their suitability to illustrate the points of discussion. For directive moves, discourse analytic techniques were used to code directive constructions and the features identified in chapter 4 to be core or periphery to those constructions, including politeness appeals to choice and involvement. Content analysis was also used to analyze and code the propositions in the co-text of the constructions according to whether or not the actions described diverged from or converged with the will of the T (see §6.3.3.4.2 below). As mentioned above, the purpose of the analysis is to bridge context and register, identify a baseline for comparison with the learner data (chapter 7), and implicate a corpus-informed language awareness pedagogy that considers aspects of genre (chapter 8).

### **6.3.3 Moves analysis: Office hours**

In spite of the differences among participants and purposes, there is evidence for the existence of several common move types among the transcripts, including opening, closing, topic identification, diagnosis, and directive moves. These moves roughly correspond to those outlined

by previous research (Agar, 1985; Williams, 2005; Bardovi-Harlig & Hartford, 2005). However, considering the breadth of purpose in office hours and the descriptive nature of this analysis, it is not implied that these are the only types of moves in office hours sessions; for example, there may be a more general ‘instructional’ meta-move. Moreover, these moves include many other speech act construction types other than directives, like information provision, explanation, question, response, praise, and critique. Directive constructions tend to be most prominent in directive moves, just as question constructions might be most prominent in diagnosis moves, and thanking constructions in closing moves (see table 6.3).

Table 6.3. Overview of common moves in office hours texts, including purposes and frequent constructions

<b>Move</b>	<b>Purpose</b>	<b>Common constructions</b>
opening	to begin interaction	greeting
closing	to end interaction	thanking, leave-taking
topic identification	to identify reason for S visit	question, responses, clarification
diagnosis	to clarify S reason or understanding, to align S frame with T frame	question, explanation, instruction
directive	to direct, regulate, and instruct S	directive, praise, critique, explanation, instruction

Except for the obvious placements of openings and closings, and an initial topic identification move, the moves in the ‘body’ of an office hours session do not necessarily occur in a set order. Directive moves are usually intertwined with diagnosis moves, and both are negotiated throughout the body, as the T fine tunes the macro-level ‘take-home directive’ message by diagnosing, instructing, and directing at a micro-level. Reflecting the unequal power relationship of academic interactions, the T controls the content of most moves, although the S can initiate a move by asking or responding to a question.

### 6.3.3.1 Openings and closings

Unfortunately because of the method of data collection, many of the initial openings were not recorded in MICASE. Still, a variety of openings are available for analysis because each T participant had multiple, usually successive sessions. The openings in the EOH were usually very informal and quick, often no more than a simple one word 'hi' (data ex. 6.1), sometimes followed by a mild directive by the T to the S to have a seat or pull up a chair (data ex. 6.2: 'if you wanna', data ex. 6.3: 'you can').

#### Data excerpt 6.1. Informal nature of openings (T 280)

T: mhm. <P :08> hi were you waiting to see me for office hours?

S1: um, yeah i'm a G-S-I.

T: okay. [S1: i just just to make sure ] no you're who are you?

S1: Andrew

#### Data excerpt 6.2. Informal nature of openings (T 320)

T: Emily if you wanna hop, over, <S1 LAUGH> while Janet does some sh- photographs maybe i can get you started,

#### Data excerpt 6.3. Informal nature of openings (T 575)

S1: thirty thirty observations is the rule of thumb the book uses and so that's the rule of thumb we'll use too. you guys can pull up a table top, you can sit on the floor over here, maybe we can move this case possibly... is that alright? just [R1: yes ] so it's out of the way [R1: yes. ] we can get more people here.

The informality of some of the encounters is apparent in analysis of transitions between two students, as a T closes a session with one S and another immediately begins another (data ex. 6.4, 6.5).

Data excerpt 6.4. Informal nature of session transition (T 270)

T: okay

S1: thank you.

T: sure.

S2: hi.

Data excerpt 6.5. Informal nature of session transition and immediacy of opening-topic ID transition (T 575)

T: great.

S1: thank you.

T: okay. next.

T: do you have a question?

S2: i'll, i'll go after them

### **6.3.3.2 Topic identification**

A topic identification move is apparent in all of the EOH sessions in a variety of forms, serving the purpose of making explicit to both S and T the reason for the session. The topic ID may occur after small talk or a lengthier opening/greeting, or it may be immediately after the opening (data ex. 6.5, 6.6), depending on time constraints and the nature of the S-T relationship.

Data excerpt 6.6. Immediacy of opening-topic ID transition (T 280)

S: hi Miss Gaston, [T: hi ] i'm in your uh section for class my name's Simon

T: hi Simon

S: um i have a question about um, my econ major i'm gonna, study abroad at uh, University of New South Wales [T: okay ] and i kinda wanna know which uh courses are acceptable.

The T can initiate the topic ID move by simply asking the S his/her reason for attending (data ex. 6.5, 6.7).

Data excerpt 6.7. T request for topic ID (T 195)

T: so, who came here first? i don't know, um, you wanna start, do you have a question?  
specific question?

With or without this T initiation, the S can identify the topic move, often as a statement (data ex. 6.6), an indirect speech act (data ex. 6.8), or a declaration of intent (data ex. 6.9). The declarative forms of these statements are a form of indirect politeness, as they avoid direct question forms that may be seen as disrespectful face threats. They may be received as permission to ask a question as evidenced by the T response (data ex. 6.9)

Data excerpt 6.8. S-initiated topic ID as statement (T 280)

S: um, Friday, [T: yeah ] i'm not going to be able to be at lecture, (xx)

Data excerpt 6.9. S-initiated topic ID as declaration of intent (T 195)

S: yeah i have one other question.

T: sure, mhm

Sometimes the T has asked the student to attend the session, in which case the topic ID is usually initiated by the T (data ex. 6.10).

Data excerpt 6.10. T-initiated topic ID (T 300)

S2: ...<sup>5</sup>okay. uh, the reason i asked you to come in is that, i- i'm looking at the grades and i'm looking at at this paper and, you're at the point where i don't want you to, fall off the edge. uh and and get a grade that's not gonna be, supportive...

### 6.3.3.3 Diagnosis

An initial diagnosis move often follow the topic ID, as the T often asks a series of questions to clarify the reason for the S visit. Agar notes this diagnosis phase is where the institutional representative fits the client frame into the institutional frame. In the case of office hours, this would involve the use of framing techniques by the T to delineate the topic, reason, or problem in terms of T and community practices. This is often accomplished through questioning and re-framing of S answers to those questions. In data excerpt 6.11, for example, the T has begun the session by asking the S for specific questions regarding the essay the S has written. The T frames the diagnosis in terms of a dichotomous choice, i.e. what the S feels is weak and/or strong about the composition, leaving the direction of the session open to the S in an appeal towards S choice. The S, however, relegates control over session topic to the T.

---

<sup>5</sup> Ellipsis in the data excerpts does not represent a 2-3 minute pause as in the transcription conventions, but rather unrelated text that would distract from the illustrated point in this particular analysis.

Data excerpt 6.11. T framing of diagnosis (T 115)

T: right good, alright so what do you- do you have specific questions about this for me like are there parts that you think are, weak and parts you think are particularly strong?

S: i dunno yeah i was just kinda hoping you'd read over and say this, has to be changed or <T LAUGH> you know whatever

T: okay no specific questions for me?

S: nope nope.

T: alright. let me take a look at this um

The T may conduct something that truly resembles a diagnosis by asking a series of questions in order to assess the S. In the case of data excerpt 6.12, the questions are used also to frame a general critique.

Data excerpt 6.12. T use of series of questions to diagnose and to frame critique (T 300)

T: okay good. [S: yeah ] alright... okay first, tell me wh- wh- what your experience has been in writing classes in the past?

S: okay. um you mean from like high school?

T: no, m- mostly here

S: or just in college well i've tak-...

T: well where did you go to high school?

S: i went to uh\_ ...

T: who was the instructor?

S: the instruc- it was...

...

T: and how are you doing in that other course? [S: uh ] have you done much writing?

S: it's pretty...

T: okay and who's teaching that course?

S: uh it's uh, ...

...

T: okay. alright and in this class, things aren't going quite that well. <LAUGH>

While the above diagnosis moves occurred at the beginning of a session, diagnoses are used throughout a session in conjunction and intertwined with directives in negotiation with the S, allowing the T to adjust his/her directive. The T may initiate the diagnosis with a question, or a student question, for example, can lead the T to adjust the instruction and/or directive, and so serves as a sort of diagnostic. In this sense, a diagnosis move may involve formative assessment, while its directive counterpart involves instruction.

#### **6.3.3.4 Directive**

A directive move involves the T directing the S, to both regulate S actions and instruct him/her. It may be centered on one directive construction, or several overlapping constructions, and can be best described as a chunk of discourse. As discussed in chapter 4, a directive construction is comprised of a core, usually verbal element, and may also include peripheral elements like a subject pronoun and immediate adjuncts, e.g. mitigators, intensifiers, and feedback-oriented discourse markers. A genre moves analysis, however, in order to ascertain the textual function of a move and relate lexico-grammar to context, requires consideration of elements beyond the level of core and immediate periphery. The boundaries of a directive move, when seen from a genre perspective, are not particularly distinct. The core directive construction

may be framed (Goffman, 1974) in advance, or elaborated after the fact, often with politeness appeals, praise, critique, and future divergent or convergent action<sup>6</sup>.

#### 6.3.3.4.1 Directive framing with politeness appeals

Discourse analysis of directive moves shows the use of choice appeals like mitigators (code: MIT) and involvement appeals (code: involve), to frame the core directive construction and mitigate its face threat potential. Data excerpt 6.13 shows how a T uses these devices to frame a directive, by false-starting the directive, pre-emptively employing a variety of appeals, and then following through with the core construction. The sum effect of this is to make the directive more acceptable, as it is framed in terms of listener choice (independence) and involvement (inclusion).

Data excerpt 6.13. Framing used with directive (T 115)

T: great (praise). now, um, you m- **you might** (begin directive) though i don't know i mean (MIT) and this is not so important for the paper (MIT) but, i think this introduction sounds really good. um, you\_ it's written really clearly and stuff (convergence, praise). um, now i worry (involve), whether you are capturing by interviewing your grandmother, um, that you're capturing what people do in Howell (convergence), or whether you're capturing an older, form of socializing, that still remains in Howell especially among older people (divergence, critique). um, so you might (MIT) i- it doesn't change your argument necessarily (MIT), but **you might** (MIT) **wanna qualify** (directive) it in that kinda way [S: okay ] you see what i mean? (involve)

---

<sup>6</sup> While a directive move in some ways can be understood as a frame, framing is neither a construction nor a move, but rather a technique used for directive or other purposes. Some cognitive linguists and communicative theorists would argue that framing is elemental to all human cognition and communication (see §6.3.4).

#### 6.3.3.4.2 Directive framing with divergent or convergent actions

Analysis also shows a framing technique where the T construes past or potential future S actions that either diverge from (code: **divergence**) or converge with (code: **convergence**) the will of the T and the academic discourse community the T symbolically represents. The divergent action is something the S has done in the past or *could* do in the future, while the convergent action is implied to be what the S *should* do in the future. A convergent action done by the S in the past is praised (code: **praise**), while a divergent action is critiqued (code: **critique**). Because use of irrealis is necessary to construe possible actions, conjunctions like ‘if’, ‘whether’, and ‘unless’, and hypothetical modals like ‘would’ and ‘could’ are found often in these frames. In many languages, this would be the realm of the subjunctive mood (see §4.1.2.1.2).

Data excerpt 6.14 presents an example of alignment, where the S states the convergent action and aligns with the directive. The T presents the divergent action with a critique after a series of choice and involvement appeals. The S then deduces the convergent action, which the T then repeats as the directive. In this way, the T avoids the directness of stating the directive, while moving the S towards self-direction.

Data excerpt 6.14. Framing and example of participant alignment (T 115)

T: ...alright now this is right (praise). this is totally right (praise). right. okay so here's\_ right... i think (MIT) you can totally make this point (convergence, praise) but i don't think (MIT) the um Upscaling Downtown helps you make it (divergence, critique).

S: so i just throw that out (alignment)

T: yeah i'd just (MIT)[S: okay ] throw that out (directive) [S: okay ] um because uh... and i don't know how you wanna organize this (MIT), but

#### 6.3.3.4.3 Framing to elaborate a directive

Data excerpt 6.15 presents an example of the use of an elaborated directive and the use of convergent and divergent future possibilities to frame the directive in a post hoc fashion. This shows that framing techniques and directive constructions do not necessarily occur in a predictable sequence. In the data excerpt, the T makes frequent use of hypothetical language to explain what the student may find in her research that would converge with T expectations as an explanation of her initial but vague directive ('you wanna keep it in perspective').

#### Data excerpt 6.15. Framing with elaborated directive (T 320)

T: i think (MIT) **you wanna, wanna keep** it in perspective here (directive). [S: okay ] i mean (MIT) **noting** that they vary in in number (directive), [S: and f- ] would be, could (MIT) be of of import if if you find that **they** (pro) say it's auspicious to have a variable number of threads (convergence) or you find that **someone** (pro) starts talking about, stripes, and the number of stripes meaning something (convergence), [S: mokay ] that you could tell it was for a woman that had um, one child or two children depending on how many you know mid stripes (convergence), i don't know (MIT) if you find something of that in your research than you'd say oh, i wish i'd counted (divergence)[S: mhm ] so sometimes **doing** a little more specific n- no- taking notes on a piece (directive), you may or may not (MIT) use it all, in the paper you may\_ or not find that it's it's worthwhile (convergence). i'm not gonna wanna read, <SS LAUGH> minute details (divergence) unless they have [S: importance, okay ] a lot of im- yeah importance for the rest of the paper (convergence).

#### 6.3.3.4.4 Framing to imply a directive

Data excerpt 6.15 also illustrates the use of framing with convergent and divergent possibilities to imply a directive, in the last few phrases where the T states she doesn't want to read details unless they're important. In other words, the S must infer that s/he should include only important details to please the T audience, as the T does not say so explicitly. This also shows the overlapping, recursive nature of directive force, as the options could be considered an elaboration of the preceding directive to take specific notes. Data excerpt 6.16 illustrates this more clearly, as a seemingly unrelated directive "if you look...and if you just read" follows a series of divergent critiques and convergent possibilities. This frame of 'more than I need' and

‘disguising the meaning’ versus ‘bringing the meaning out’ implies its own directive for the S to follow the convergent option. In this way, the convergent possibility becomes the directive<sup>7</sup>.

Data excerpt 6.16. Framing to imply directive (T 300)

T:...and there are some cases where you're where you're looking, or where you seem to be saying something um, and i think i know what i know what you want to say, but because you've sort of (MIT), you've given me more than than i need (divergence, critique) you're really disguising the meaning (divergence, critique)[S: McKay ] rather than bringing the meaning out (convergence). so that, if y- if you look at this sentence and if you just (MIT) r- read that sentence aloud (directive).

#### 6.3.3.4.5 Use of pronouns to involve and exclude

Data excerpt 6.15 also shows how the T uses deictic pronouns (code: *pro*) to refer to the academic community when presenting a convergent option, the effect of which is an involvement appeal to the S. In the elaborated directive in data excerpt 6.15, the T presents ‘they’ and ‘someone’ as the experts on tribal weaving, whose authority should be taken into consideration when the S analyzes the textile for her paper. In contrast, the T in data excerpt 6.16 uses ‘I’ and ‘you’ to exclude and critique the S. The T construes himself as being the object of the divergent action.

In data excerpt 6.17, which is the continuation of 6.16, the T directive to ‘look back and say...’ coincides with the divergent S past action. In other words, the S is to look back and realize that his use of the word ‘elegy’ was wrong. The T follows the critiques with the

---

<sup>7</sup> In fact, the directive ‘doing more note taking’ in data excerpt 6.15 could be coded as convergence rather than a directive.

convergent alternative. He uses 'we' ambiguously, as it could exclude the S, since the S did not use the term 'elegy' correctly, or it could include the S, as the convergent option for the S to take in the future. In this way, the T appeals to S involvement and the convergent possibility becomes the directive.

Data excerpt 6.17. Involvement appeals through use of pronouns (T 300)

T: ... so that, **if y- if you look** at this sentence and **if you** just (MIT) r- **read** that sentence aloud (directive).

S: <READING> poetry is the best representation of the elegy... </READING>

T: okay. now, part of part of what **you need to do** is to **look back** at the beginning of this and **say** (directive), w- how are you using the word elegy? (divergence, critique) and how can poetry be a [S: (representation) ] representation of the elegy? (divergence, critique) it, **we** (pro) can talk about, the feelings that are associated with an elegy or **we** (pro) can talk about mourning or **we** (pro) can talk about a lament for for the dead (convergence). but, by definition an elegy is [S: is a poetry, form of poetry. yeah ] a poetry is a is a p- a form of poetry right...

#### 6.3.4 Discussion

It is clear looking at the above data excerpts that the boundaries of a directive construction are porous, and that a directive move can be framed through a variety of techniques. In office hours contexts, the T may frame the directive with divergent or convergent S actions, whether past actions that are critiqued or praised, or future possibilities that are construed as undesirable or desirable. In addition, framing may be used to elaborate on or imply a directive. T

use of basic choice appeals like mitigators and involvement appeals using pronouns can make a directive less of a threat to the S. Framing in this sense is not a move or a construction, but rather a technique (see again footnote 6).

Including the sense discussed by Agar (1985; §6.3.3.3 above), framing has been studied widely from a variety of related perspectives. Goffman (1974) explains framing as the schematic means for interpreting and assigning meaning to experience. Framing-as-schematization may also be a prime means of genre acquisition and socialization (Swales, 1990). Bernstein (1996) considers framing a more abstract principle, concerned with the internal logic of pedagogic practice and the nature of control over selection, sequencing, pacing, criteria, and the social base of communication. In cognitive linguistics (e.g., Lakoff & Johnson, 1980), framing is considered a powerful rhetorical means of persuasion, as one party frames an argument using terms and metaphors that the other must use and operate within, and is therefore constrained by. Building on these definitions, the current project defines directive framing as a rhetorical, or genre-based, means of construing the past and hypothetical future actions of the directed according to the will of the director. In educational contexts, directive framing allows pedagogic discourse to be simultaneously instructional and regulative, as a T can use content to frame past and hypothetical S action.

In institutional discourse where power is hierarchical and unequally distributed, the director T frames future S action as choice, and the directed S must choose to converge and be involved, or diverge from the will of the T and the community s/he represents. Students may accept negotiated direction as part of the process of academic apprenticeship, especially when they are offered true options. For others, the only truly independent action may be to drop out. Seen from this perspective, the T use of directive constructions and moves, in conjunction with

politeness appeals and framing, can thus be understood as a means for the exercise of symbolic authority and power (Bourdieu, 1991). Why a T should couch their directives in politeness appeals at all to an inferior may be difficult to understand for those unfamiliar with the American cultural mythos of egalitarianism and informality, one that privileges involvement and downplays power hierarchies in academic relationships.

#### **6.4 Conclusion**

In this study, two approaches to genre analysis were used to examine the social activity of office hours. A contextual analysis examined the name of the genre, its social context, its communicative purposes, participant roles, cultural values, and text context. The conclusion drawn from this analysis is that office hours is a complex educational institutional discourse type involving spoken interaction, as it fills an important role by providing opportunities for learning, interaction, and relationship building not found in the classroom. It differs from other discourse types like service encounters, writing center sessions, and advising sessions in that the participant power relationship is unequal and hierarchical, and it extends beyond the temporal boundaries of the session. Because participant expectations of purpose and participant roles vary by cultural and individual historical background, office hours have the potential to be the locus of misunderstanding and miscommunication, especially when those backgrounds vary. Knowledge of genre expectations and practices is the privilege of core T and S participants (Lave & Wenger, 1991).

In complement to this contextual analysis, a language-focused moves analysis examined the different moves or stages that comprise an office hours session, using data excerpts from the EOH corpus from MICASE. Analysis showed openings, closings, topic identification, diagnosis, and directive moves to occur in a semi-fixed, dynamic and partially negotiable order. Closer

discourse analysis of directive moves identified the use of framing as a means for politeness, to elaborate or imply a directive, and to present a directive in terms of divergent or convergent action. In addition, pronouns were used strategically in framing to exclude or involve the addressee. The rhetorical effect of directive framing in office hours is persuasion and the exercise of symbolic power.

There is perhaps nothing surprising about these findings, since guidance and direction are major purposes of education. In addition, directive framing is certainly not exclusive to office hours or educational discourse. Nevertheless, for the purpose of teacher education and ITA preparation especially, the above explications are necessary and telling. Implications for curriculum design are that ITAs need to develop awareness of the varying expectations regarding purpose and T-S roles, as well as awareness of the moves in an office hours session. When making a directive move, for example, critiques might be better framed in involvement appeals, while directives might be framed with multiple convergent options if possible, to provide a student with a measure of independence. In the preparation of curricula, this information should be embedded in discussions of the contextual features discussed above, especially role expectations and cultural issues, since learners may come from academic cultures where gestures towards egalitarianism in the form of politeness appeals to inferiors are unexpected and perhaps even inappropriate.

This completes the corpus and genre-based examination of the systematic use of directives in office hours from a social-functional perspective. A major caveat to these findings is that averages and prototypical structures are ideal, and no one individual exists who perfectly represents that ideal. For a complete picture, analysis of individual variation can provide balance to analysis of systematic usage by groups. In chapter 7, individual usage is examined by

comparing individual expert and learner profiles, and by exploring directives usage by 3 learners in detail, complemented by survey and interview data.

## CHAPTER 7

### INDIVIDUAL VARIATION IN DIRECTIVE CONSTRUCTION USAGE

To complement the aggregate nature of the corpus-based register analysis (chapter 5) and genre analyses (chapter 6), in this chapter I analyze the variable nature of directives usage by individuals in office hours contexts. The strengths of corpus analysis are that with the aid of computer technology, usage patterns of frequency and co-occurrence can be identified that even the closest discourse analysis would miss. Moreover, corpus analysis can be used as a heuristic for further research, pointing the analyst in a fruitful direction. A criticism of corpus analysis, however, especially of those using small corpora, is that mean frequency data captures usage patterns of an ideal, but not real, individual. Moreover, because individual variation is, of course, a feature of all language usage, it may be that all, or a large portion, of the production of a token or construction in a corpus was produced by only a few individuals, which would negatively affect the generalizability of the findings. While examination of register consistency as was done in section 5.2 may address this criticism to an extent, only consideration of individual usage in combination with corpus and genre analysis can provide a more complete descriptive picture of usage. The purpose of this chapter, then, as introduced in the research question in chapter 3, is to examine the variable qualities of directives usage by experts and learners, to balance the exploration of systematic qualities in previous chapters.

To this end, I first present corpus-informed variation analyses of directives usage by the 8 experts and 8 of the learners in the subject corpora. I then conduct in-depth exploration of directive language usage by 3 learner-participants in the LOH corpus, using biographical,

attitudinal, and assessment data<sup>1</sup>. These portraits are presented not as case studies, since they are neither longitudinal nor inclusive of other language production types, but as qualitative usage profiles, since they are corpus-informed snapshots<sup>2</sup> of usage of a particular language type, directives. In alignment with previous chapters, a social-functional framework is used to interpret the findings.

## 7.1 Corpus-informed usage profiles

Recently, researchers have begun to employ what might be termed corpus-informed usage profiles for a variety of purposes, including the analysis of directives use (Yates, 2005) and the documentation of learner language development (e.g. Belz, 2004; 2006; Belz & Vyatkina, 2005). In these studies corpus data is used to profile language use by one or more individuals, sometimes in conjunction with ethnographic data, and often contextualized within analysis of aggregate data. These studies demonstrate the limitations and possible pitfalls of small corpus studies that rely solely on quantitative analysis.

Focusing on individual variation, Yates (2005; see also chapter 2) compares the use of mitigation in directives by 9 NS and 9 NNS (L1 Chinese) secondary teachers in Australia and its relation to construction of social identity. She finds that nearly the same number of directives were used by the NS (1137) as by the NNS (1126), that a greater percentage of the NS directives were indirect or mitigated (69%) than the NNS directives (57%), and that in both groups the female teachers mitigated slightly more than the males. To balance this aggregate interpretation, she explains that two individual NNS mitigated more than seven of the NS, although the NNS used fewer involvement approaches, instead relying on formal positive politeness markers like

---

<sup>1</sup> Such data, beyond basic biographical data, was unfortunately unavailable for the EOH participants.

<sup>2</sup> The influence of a pedagogical treatment on usage by some of these learners is discussed in chapter 8.

'please'. She presents profiles of each teacher's usage with data examples to demonstrate how mitigating style influences the construction of the teacher's social identity.

Building on her work on pragmatic development and the use of T/V address systems, Belz (2004; see also Kinginger and Belz, 2005) demonstrates the use of a microgenetic approach, grounded in sociocultural theory, to investigate development of the usage of German da-compounds<sup>3</sup> by a class of German learners in a telecollaboration course, collected in a learner corpus. She analyzes every use of da-compounds in the corpus, finding a general increase in the number of compounds used with increasing proficiency, but with considerable variation by users over time. She makes use of ethnographic information to inform her interpretation of learner uses and misuses of these structures, implicating a variety of sociocultural and interpersonal influences on development, especially interaction with expert speakers. In a related later work (2006), Belz traces the da-compound usage of a single German learner over the span of several years and notes the influence of a timely pedagogical intervention on the learner's usage. In addition, she attributes part of the learner's development on sociocultural factors like learner socioeconomic background. Belz and Vyatkina (2005) examine the influence of pedagogical intervention and telecollaboration on German learner's use of modal particles, small interactional words speakers use to reflect their stance towards propositions. Using corpus-based data and coursework to develop their interpretations, the authors create and analyze profiles for eight German-American pairs of students, tracing learner development over the course of the semester.

## **7.2 Variation analysis**

Yates' work shows that aggregate data analysis of small corpora may be misleading, while Belz' work demonstrates the heuristic strength of individual usage profiles to examine

---

<sup>3</sup> These are equivalent to English there-compounds like thereby and therewith, but much more productive and common.

learner language. In other words, consideration of usage by individuals can provide balance and clarity into the insights offered by corpus analysis of group usage. With this in mind for the current project, I first follow Yates in conducting a corpus-based analysis of variation in directives usage by both experts and learners. Then, following Belz' example, I create and examine corpus-informed profiles of 3 of the LOH participants.

### **7.2.1 Method**

To analyze variation, I examined the frequencies of the most frequent constructions (MFC) produced by the 8 EOH participants and 8 of the LOH participants. As explained in §5.4.2.2, MFCs included any modal, p-modal, and directive vocabulary construction with a frequency of over 1 per 10K in the combined EOH and LOH corpora, except for 'you can' and imperatives. Since searching every directive construction would have been prohibitively time consuming, these MFCs are considered representative of overall directives use. Eight learner participants were chosen to balance the eight experts. The 8 learners were chosen by first considering all 32 consenting S06 participants (the semester of the extended project, see §7.3 below). Participants who were absent for any of the 3 data collection sessions (A1, A2, and A8) were omitted. The remaining 24 were sorted according to frequency of MFCs, and an evenly distributed array of most-to-least frequent producers were chosen. Participants 6, 7, 10, 13, and 19 were chosen because they participated in the extended portion, and the remainder, participants 1, 4, and 16, were chosen to evenly fill in the distribution.

### **7.2.2 Results**

The frequency of MFC use per 10,000 words for each individual was calculated using the overall word count, not simply an individual's use in the T role, to allow for comparability

between the two groups, since T token counts were unavailable for the EOH. Results are presented in Table 7.1.

Table 7.1. Use of most frequent constructions by all eight experts and eight chosen learners (freq. per 10K words)

individual expert or learner	total all words (T + S)	freq. per 10K	total MFCs	you should	you have to	you need to	you don't need to	you want to	you could	you had better	I recommend	I suggest	I want you to	I would
320	9099	57.15	52	1.10	5.50	2.20		30.77°	10.99				2.20	4.40
115	30328	51.77	157	4.62	1.32	8.90	1.65	26.71°	4.29		0.33	0.66	1.32	1.98
300	12126	47.01	57	6.60	1.65	11.55°		4.95	11.55°				1.65	9.07
575	11221	46.34	52	8.02	10.69	7.13		13.37°	1.78			0.89	0.89	3.56
270	19742	40.52	80	2.53	18.24°	3.04		9.12	6.08					1.52
195	20299	37.44	76	2.96	7.39	9.85°	4.93	9.36	0.99					1.97
280	13702	26.27	36	1.46	8.03°	5.84	0.73	5.84	2.92					1.46
175	8826	21.53	19	1.13	3.40	5.67°		3.40	4.53					3.40
Meng Z*	1824	82.24	15	21.93					16.45	38.38°		5.48		
Xin Z	1966	55.95	11	40.69°						5.09	5.09	5.09		
Yuan D*	2066	48.40	10	29.04°	4.84					14.52				
Bing C* <sup>1</sup>	2584	30.96	8	3.87	11.61°	3.87	3.87	3.87					3.87	
Taehoon J	1915	26.11	5			5.22	5.22				10.44°	5.22		
Phongsak T* <sup>1</sup>	1967	25.42	5			5.08		20.34°						
Youngkyu C*	1650	18.18	3			6.06				6.06	6.06			
Yong W <sup>1</sup>	2475	16.16	4	4.04			4.04					8.08°		

\* participated in experimental ESL118 pedagogical treatment; ° preferred MFC for this individual; <sup>1</sup> profiled below in section 7.3.

A comparison between the two groups shows that the overall frequency of expert use of MFCs ranges from 21.53/10K to 57.15/10K, while for the learners, it ranges more widely, from 16.16/10K to 82.24/10K. A developmental explanation for this might be that learners are unfamiliar with the interpersonal and genre expectations of office hours, and so are overusing and underusing directive constructions as they develop an individual usage style or preference. Another explanation is that there were insufficient data for each learner for an accurate measure. It is also notable that the experts use a greater variety of constructions than the learners, which would reflect a smaller repertoire on the part of the learners, although this could again be due to insufficient amounts of learner data.

There is much variability in individual expert use. While participant (P) 175 produced 'you want to' at 3.4/10K, P115 used it at 26.71/10K, over 8 times as frequently. P270 used 'you have to' at 18.24/10K, nearly 14 times more frequently than P115. P575's use of 'you should' is 7 times more frequent than P320 or P175's use. Each individual expert has his/her own preferred MFCs, so that P300 prefers 'you need to', 'you could', and 'I would', while P270 prefers 'you have to', 'you want to', and 'you could'. One explanation for the variety is the differences in disciplines, for example the four most frequent users of 'you have to' are computer science, economics, statistics, and heat and mass transfer. It may be that in the hard sciences, procedures and rules are more rigid than in the social sciences and humanities. Another might be similarity in session purpose. The two experts in the two exam reviews, for biology (175) and heat and mass transfer (195), are similar in that 'you need to' is the most preferred construction; in spite of the similarities, however, the former next prefers 'you could' and 'I would', while the

latter ‘you want to’ and ‘you have to’. This could reflect differing styles in presenting a sense of independence or choice to students. Most likely, individual use is influenced by both topic constraints and individual preferences.

It is difficult to assess the extent to which interplay between knowledge of genre conventions and individual style or preference influences learner usage. On one hand, corpus analysis of learner usage shows less use in general of language that evokes independence or involvement; on the other hand, the results of all learner participants on the survey of attitude towards teacher authority (3.23 out of 5, 1 being more authoritarian and 5 being less authoritarian; see appendix D), as well as the responses of the majority of interview participants, shows that the learners as a whole wish to be less authoritative and more involved as TAs and instructors.

Learners show variety in use of directive constructions, but less so than experts. The eight representative learners here used on average 3.875 different constructions, while the experts used on average 7.25. While this may again be due to insufficient opportunity, the reliance of some of the learners on some of the constructions (what Hasselgreen, 1994 and others call ‘lexical teddy bears’) is apparent. For example, Meng Z used 7 instances (38.38/10K) of ‘you had better’, while his next MFC was 4 instances of ‘you should’ (21.93/10K). Xin Z used 8 instances (40.69/10K) of ‘you should’, and only 1 each of three others. Another student not presented here used 10 instances of ‘you have to’, and only 1 instance of one other MFC while playing T in the three office hour role plays. Only when these frequencies are compared to other MFC usage frequencies, i.e. to other learners, to experts, and to their own usage of other MFCs, does any reliance stand out. Meng Z’s usage of ‘you had better’ is 38 per 10K and Xin Z’s usage of ‘you

should' is 41 per 10K, both of which are more frequent than the uses of any of the experts, who had 4 to 15 times as much opportunity to repeat construction types.

In sum, variation analysis as presented in Table 7.2 shows that learners seem to rely on certain directive constructions when playing the T role in office hours sessions, avoiding some constructions entirely (e.g. 'I would') and others partially (e.g. 'you want to'), while using other constructions that the experts avoided (e.g. 'you had better'). As discussed in chapter 5, the effect of this may be that an ITA's intentions and use of politeness strategies are misinterpreted.

### **7.3 Learner profiles of directives usage**

Participants from the Spring 2006 ESL 118 cohort were recruited to participate in an extended portion of the ITAcorp project that involved an interview and assessment of directives knowledge after the course had finished<sup>4</sup>. Seven individuals participated, three of whom, Bing C, Yong W, and Phongsak T, were chosen for in-depth learner profiles that would illustrate the nature of individual variation in directives usage<sup>5</sup>. To complete an individual profile, social-functional interpretation of that individual's corpus-based usage was supported by biographical, attitudinal, and assessment data, including the abovementioned interview and assessment data as well as the biographical survey and survey on teaching completed as part of regular ESL 118 coursework. The instrument, design and medium of implementation for each source are presented in Table 7.2 below, and the instruments themselves are provided in the appendices listed.

---

<sup>4</sup> These participants were compensated for their participation. This extended portion is covered by the original Penn State IRB approval, project #21429.

<sup>5</sup> These 3 were chosen from the 7 because their usage profiles best illustrate the points of discussion.

Table 7.2. Sources for biographical, attitudinal and assessment data used in learner usage profiles

instrument	design	medium
Biographical survey (appendix B)	13 items, including biographical information, language learning experiences, overseas experiences, and computer use experience (FA05 only)	online, part of ESL 118 course work
Survey on attitudes towards teacher authority (appendix C)	9 Likert-scale items	online, part of ESL 118 course work
Interview on attitudes towards teaching and conducting office hours (appendix E)	5 questions on teaching	oral, post-semester
Assessment of directives knowledge (appendix F)	3 items asking participants to list directives, list usage contexts, and contrast functions of different forms	paper, following post-semester interview

### 7.3.1 Bing C

Bing C is the pseudonym of a 25-year-old female Chinese Ph.D. biology graduate student who was enrolled in, and passed, ESL 118, and was in the experimental section for directives instruction (see chapter 8). In the three course role play interactions in which she played the T role, she used a wide variety of directive constructions, including ‘you can’ (9)<sup>6</sup>, ‘you have to’ (3), ‘you should’, ‘you want to’, ‘I want you to’, ‘you need to’, ‘you don’t need to’, and ‘don’t (+ imperative)’ (1 each). Discourse analysis of her transcripts shows several directive constructions not captured in the corpus analysis, specifically 3 more imperative constructions and a ‘what is needed is for you to’ cleft construction. In addition, she used 7 ‘if you’ constructions as part of her directives.

Corpus analysis of Bing’s usage shows she used the most frequent directive constructions and ‘you can’ more frequently (182 per 10K) than either the learners (79

---

<sup>6</sup> For the learner profiles presented here, uses of ‘you can’ were disambiguated to include uses that indexed possibility and/or permission, and to exclude uses that clearly indexed ability with little or no directive force. Because of the work this disambiguation would have been for the entire corpus, however, as discussed in chapter 5, ‘you can’ is not analyzed in the comparative analysis, and is therefore not considered an MFC.

per 10K) or the experts (67 per 10K). While this may demonstrate a directive-oriented stance towards interaction with students in office hours, it does not necessarily demonstrate an authoritarian stance. A closer look shows that of her nine uses of MFCs, she used adjuncts with six of them (66%), which is a higher percentage than either learners (55%) or experts (61%). Four of the six adjuncts are mitigators, including ‘I think’, ‘I guess’, ‘just’, and ‘may’ (1 each). In addition, she used adjuncts with all nine uses of ‘you can’, and eight of those adjuncts are mitigators, and include ‘I hope’ (3), ‘I think’ (1), and ‘just’ (4).

The corpus findings above show that Bing knows a variety of directive constructions and has a good understanding of the interpersonal need for appeals to independence. In the post-instruction assessment, she offered, “I think it’s better to” and “I think it would be good to” as examples of directive constructions, and commented on the difference between ‘I would’ and ‘you should’ with: “(I would) is more suggestive. While suggest this is a good choice, the TA let the students to choose. However, for (you should), the TA just ask the students to ‘put the emphasis here” (post-instruction assessment, 5/4/06). This demonstrates that although she did not use ‘it’s better’, ‘it would be good’, or ‘I would’, she understands their usage, and furthermore believes that ‘I would’ allows more choice than ‘you should’<sup>7</sup>.

This assessment is congruent with Bing’s score on the survey on attitudes towards teacher authority, where she scored 3.625 on a 5 pt. scale, 0.39 above the average. While this measure was not formally standardized, her interview responses are also congruent with her directive but polite approach. She explains “in China, teacher and TA have to

---

<sup>7</sup> In other words, she understands that ‘I would’ has politeness potential, but believes it to be negative rather than positive. This sort of partial understanding indexes development.

(be given) authority, so students never challenge, however here you have to be prepared for any question” (interview, 5/4/06). She thinks a professor should be “good at guidance, he never blame you, says you are the most hardworking, always encourage you, praise you, if you do something wrong, he just comfort you, never say no” (interview, 5/4/06). She prefers a less authoritarian style of teaching, but feels it’s important that students like the professor and like the class, but that the TA has to be firm, and “let them know what they should do, you cannot be too kind” (interview, 5/4/06). It’s difficult for her because “if I was in China, I know the level, in what extent I should do, however here...I don’t know where I should be” (interview, 5/4/06).

In sum, Bing’s use and demonstrated knowledge of a variety of directive constructions, combined with her use of mitigators and other adjuncts, presents a very advanced picture of directive usage, although she is still a learner by her own account. In consideration of her interview and survey responses, her language use indicates an awareness of the social and interpersonal aspects of T-S relationships, including the challenges posed by the intercultural nature of the T-S interactions in which she participates. Certain elements, however, especially her choice of the construction ‘you can’ and the mitigators ‘I guess’ and ‘I hope’, may result in miscommunication. As learner constructions (Waara, 2004; below), they are comprehensible to a listener, but may not reflect the meaning she intends to convey.

### 7.3.2 Yong W

Yong W is the pseudonym of a 25-year-old male Chinese PhD student in economics, who was enrolled in, but did not pass<sup>8</sup>, ESL 118, and was not in the experimental section for directives instruction (chapter 8). In the three course role play interactions in which he played the T role, he used ‘you can’ (6), ‘I suggest’ (2), ‘you should’ (1), and ‘you don’t need to’ (1). This puts him at the low frequency end of directives usage in comparison with the other learners<sup>9</sup>. A closer analysis of his transcripts, however, shows several directive-like constructions not captured in the corpus analysis, including a ‘why not’, a ‘it will be helpful for you’, a ‘students are required to’, and a ‘would you please’, all of which he listed as examples of directive language in the post-semester assessment. In addition he uses 9 ‘if you’ constructions, although not all were clearly associated with directives.

Of his four uses of MFCs, Yong used an adjunct with two of them, saying “I think you don’t need to worry” and “I mean you should you can think about this question”. In the latter example, the mixing of forms may index development. Besides these, his use of ‘it will be helpful’ is preceded by ‘I think’, but none of his six uses of ‘you can’ are clearly mitigated. In total he mitigates 50% of his MFCs, which is lower than both experts and learners, and when directive uses of ‘you can’ are considered, the percentage is even lower. Still, in the examples he provides on the post-assessment, he includes ‘please’, ‘strongly’, and ‘I think it may be’ in the example of directive language, which

---

<sup>8</sup> I do not know the reasons why Yong did not pass, and they may not have anything to do with his directives usage. As per the IRB agreement, none of the analysis here was available to the instructor, and was in fact conducted 10 months after the course ended.

<sup>9</sup> By this measure Yong W is 8<sup>th</sup> out of the 8 chosen for the variation analysis, but 19<sup>th</sup> out of the 24 overall.

demonstrates at least declarative knowledge of these forms. In addition, analysis of his entire discourse shows that he does make use of some mitigators.

In his post-semester written assessment, Yong does not list modal or p-modal forms as examples of directives; instead, he lists the phrases noted above and “I (strongly) suggest that”, “you are not allowed to”, and “this is prohibited”. When asked to provide a situation where he might use directives, he offers, “some students want to postpone homework. I may say ‘sorry, it is not allowed to do it except for valid excuses’” (post-instruction assessment, 5/5/06). When asked the difference between ‘you could’ and ‘you can’, he first wrote that ‘could’ “may be more suitable because it is not so strong”, but then crossed it out. In its place he wrote that ‘can’ “is more suitable, because it is just an explanation of extra option, not some comments on student’s activity” (post-instruction assessment, 5/5/06). This may be interpreted to mean that he believes a TA should be neutral, and that ‘could’ implies judgment while ‘can’ does not. In other words, it may be that Yong believes a TA should not give directives, but rather simply act as a mediator, not assessor. This interpretation would explain his relative lack of directives usage in the role play activities, and his reporting of constructions that index permissive rather than obligative force.

Moves analysis of Yong’s three interactions shows that he is familiar with the opening-topic ID-diagnosis-directive pattern. His choices for topic identification phrases are interesting, and show the irregular variation typical of pragmalinguistic development. In the first role play (A1), he uses ‘So what can I help you’, which is comprehensible but ungrammatical. In the second (A2) he uses ‘What’s your problem’, which is sociopragmatically inappropriate, and in the last (A8), he uses ‘What can I do for you’,

which is grammatical but perhaps overly formal<sup>10</sup>. He also seems to refer often to the authority of the department or professor. In the chat role play (A1), he answers questions regarding the syllabus of the course he is teaching for an interested student, explaining that ‘students are required’ to do something rather than using ‘I require’ or ‘you are required’. In the class role play (A2), he goes over a test score with a student, and explains that the student cannot have credit for an incorrect calculation because of ‘our policy’ on test correction. In the post-evaluation role play (A8), he offers advice on how the student (the evaluator) can get extra help, but explains that he must first check with departmental policies. In other words, he seems very aware of his mediating role as a TA, and reflects this awareness with language choices that implicitize the departmental source of authority (Iedema, 1996; see also §4.2.1).

While other learners in this project also referred to departmental authority, they did not do it so frequently and in combination with low usage of directives. In Yong’s case, it may be that he was operating with an idea of what happens in office hours based on his own experience. It turns out that Yong was actually a TA in his department when he was taking ESL 118, although he only held office hours and graded tests, and did not teach<sup>11</sup>. He also explained in his post-semester interview that he had been a TA in China, where TAs were more like student supervisors, as compared to the US where TAs “are more equal to students” (interview, 5/5/06). As an example, he explained how in the US, students can come to your office “and argue all kinds of things they want to argue; in China, not quite common” (interview, 5/5/06), and that in China most students just accept

---

<sup>10</sup> Also in A8, Yong allowed the ‘student’ to explain her entire issue without interruption, which, along with the formality of the help offer and his deference to department authority, may indicate that her true status as evaluator influenced his choice of language more than the parameters of the role play.

<sup>11</sup> It is not clear why Yong had assumed TA duties when this was technically illegal; his department may have decided separately that he was fit for non-teaching assistantship duties.

what the TA says. He said that he didn't have trouble with this difference, since he had learned in English class about cultural differences, so he was trying to get used to it. When asked if there were advantages to an authoritarian style of teacher-student relationships, he explained that in a more American style, students may have too much power and argue about their courses, when they do not know what is appropriate for them to study, but that if they have no power at all, the department won't know what they really need, and so there should be balance between both approaches. When asked which style he preferred, he said a less authoritarian, American style, consistent with his score of 3.75 out of 5 (0.52 above the average) on the survey on attitudes towards teacher authority.

In sum, it may be that Yong W felt that he could be less authoritative, or even build 'we're in the same boat' solidarity with his students, by avoiding the use of directives himself and implicitly invoking departmental authority instead of his own. His language choices in the role plays reflected his own experience as an actual TA, where the risky use of face-threatening directives can be avoided by invoking departmental authority and taking the safe, but disempowered, role of mediator. While there were many other factors besides use of directives, e.g. pronunciation, that determined his final grade in ESL 118, it is ironic that he failed the course, considering that he was simply interpreting the TA role his department had constructed for him, one where a TA maybe does not really need to use directives.

### **7.3.3 Phongsak T**

Phongsak T is the pseudonym of a 27-year-old Thai student getting his Ph.D. in Computer Science and Engineering, who had been in the US for 2 1/2 years getting his

Master's before enrolling in ESL 118<sup>12</sup>. He participated in the second experimental section, which may explain his accurate answers on the post-instruction assessment of directives knowledge. In the role plays, he used 'you might want to' four times<sup>13</sup>, once in A1 (chat role-play), twice in A2 (in-class role play), and once in A8 (post-evaluation role play), besides 'you need to' once in A1 and 5 directive uses of 'you can'. In addition to constructions caught by the corpus analysis, a close reading of his transcripts show the use of several imperatives ('turn them in', 'please finish', 'don't hesitate'), several 'if you' phrases (e.g. "if you want more reference maybe I can give you..."), and a 'it's better if' construction. Phongsak also demonstrates near-expert use of mitigators, as not only are all of his uses of 'you want to' hedged, but he also mitigates appropriately throughout the transcripts, for example, in A2 he says "because it's not too long from the due date maybe you can...". Although a complete analysis of all adjunct usage by all learners was not available for comparison, Phongsak hedged with 'maybe' 8 times in all 3 activities combined, which seems high.

While his use of the measured directive constructions is not as varied as, for example, that of Bing C, his answers on the post-instruction assessment show that he was paying attention during the experimental instruction (see chapter 8). He offers all of the directives covered in the instruction as examples, and explains, "must/need/have to are strong words. In many situations, TAs might want to avoid using them. 'Should' is somewhat strong. When using strong words, TAs or professors usually use a pronoun 'we' for direction. 'Might want to' is a mild word for most suggestions we make" (post-instruction assessment, 5/5/06). He explains the difference between 'you want to use' and

---

<sup>12</sup> This is in contrast to both Bing C and Yong W, who had both been in the US for one semester before enrolling in ESL 118.

<sup>13</sup> Of the 8 individuals who used 'you want to' 12 times, Phongsak's 4 times was the most.

‘do you want to use’, in that the former implies that the suggestion “is the best choice in TA’s opinion and the TA may not want to discuss about other choices” (post-instruction assessment, 5/5/06)<sup>14</sup>. It is interesting that he makes use of ‘might want to’ even in this written assessment.

In his interview, Phongsak explained that although he has not taught, he will probably be a TA, depending on his department’s needs, and that upon completion of his degree he will return to his country and be a professor. He explained that Thai student-professor relationships are much more formal than those in the U.S., and that students there are more respectful. When asked the advantages and disadvantages of egalitarian and authoritarian teaching styles, he answered that egalitarian might be appropriate for courses that require a lot of discussion, or for grad-level courses, but that authoritarian styles might be appropriate for larger courses where the subject matter was less negotiable, for example mathematics. He said that in the future, his style will depend on the kind of course he teaches, though he thinks his personality fits better with discussion. In view of his usage and assessment, this final comment can be seen to summarize Phongsak’s directive usage, in that he seems to understand that contextual variables influence his choice of language.

---

<sup>14</sup> Bing is also one of the 8 ITAcop users of directive ‘you want to’, and one of the 5 who employed it after receiving explicit instruction in its usage (see chapter 8). She explains in her post-instruction assessment her understanding of the difference between ‘you want to use this equation’ and ‘do you want to use this equation’, in that the former is a suggestion, while the latter lets “the student think first, rather than told them to do so directly” (post-instruction assessment, 5/4/06).

## 7.4 Discussion

The individual profiles above offer insight into learner usage phenomena of interest to the development of ITA preparation curriculum, specifically the use of learner constructions and the influence of socio-cultural factors.

### 7.4.1 Learner constructions

A phenomenon of advanced proficiency, learner constructions are utterances considered comprehensible and syntactically correct but pragmatically inappropriate (Waara, 2004). A closer look at directives usage by Bing C and others shows evidence of learner constructions that may cause misunderstanding in listeners. In data excerpt 7.1, Bing attempts to use ‘I guess’ as a mitigating appeal towards independence, but it has the effect of lack of commitment to the force of her directive (e.g. ‘I guess you’ll be prepared...’). In this utterance, ‘I think’ would have been more appropriate as a mitigator.

Data excerpt 7.1. Bing C’s use of ‘I hope’, ‘I guess’, and ‘you can’ (A2)

T: I have assigned the page numbers for you **I guess**

S: ok that’s good

T: and you are\_ **I hope you can** be more careful on class so **if you** take my notes and under- understand what I (taught) in the classroom [S: yeah] **I guess you’ll be** you’ll be, prepared to work this uh homeworks quickly

Data excerpt 7.1 also shows an example of Bing’s overuse of ‘you can’, which she mitigates with ‘I hope’. In this example, because of the ‘I hope’, a listener might interpret the ‘you can’ as a simple wish or plea rather than a directive, most likely less forceful than Bing intends. In the post-instruction assessment, she lists ‘I hope you can’

as a directive construction, and comments on the difference between ‘you could’ and ‘you can’ that “(you could) is more suggestive. Whether or not to do this is determined by students. For (you can), it seems the student ‘should’ do so” (post-instruction assessment, 5/4/06).

There are dozens of examples in the LOH transcripts of this ‘you can’ learner construction, where experts might choose ‘you could’ or ‘you should’. As noted above, Yong W, in the post-evaluation role play (A8), said “I mean you should you can think about this question”, which may be a mixing of forms that index development, or an attempt to mitigate the force of ‘you should’. Another learner, Yue Z, also in his post-evaluation role play, said “also I think, mm, you can you need to you may need to read a book” when giving advice, which may again index development, or be an attempt to strengthen the directive.

More revealing is the intonation of one learner, Youngkyu C, in his post-evaluation. In data excerpt 7.2, he is telling the student, played by the evaluator (E), that she can ask him for help anytime. He first says ‘you can email me’ with little stress on ‘can’, but then repeats himself, saying ‘if you want to know something, you can email me’ putting clear, intentional stress on ‘can’, as if he intends it to be more forceful than his other uses of ‘can’. It is also interesting that he combines this usage with an ‘if you’ phrase.

Data excerpt 7.2. Youngkyu C’s use of ‘you can’

T: oh okay, you can email me, and then, you just, if you, want to know something [E: mhm] if you, you <u>can</u> email me [E: mhm] yeah I can, reply your email so, you don’t need to hesitate email me, also you can, stop by my office, at office hour [E: mhm] so I can help you [E: okay] yeah
--

The usage examples above, in combination with the corpus results in chapter 5 showing significant overuse, are evidence that learners use ‘you can’ in pragmatically inappropriate ways, and so may be misinterpreted if their intentions are truly directive. In the dynamic sense, the use of ‘you can’ could be an appeal to independence as it seems to offer the directed-listener more choice. However, Bing’s response in the assessment data above, and especially Youngkyu’s use of heavy stress, is counter-evidence to this interpretation. Alternatively, the usage may be attributed to the L1 influence of Chinese, which has several different equivalents of ‘you can’ (Zhu, 1996), none of which overlap exactly with the permission-possibility-ability conceptual field for English ‘can’. This may be a factor, but it is worth noting that Youngkyu’s L1 is Korean. It may be that ‘can’ is the prototypical modal that learners fall back on because of its diversity of function. As the ‘swiss army knife of modals’, as it were, ‘can’ is very appealing to users who may need to express both deontic and dynamic meaning without necessarily committing to either.

#### **7.4.2 Contextual factors**

Besides the existence of learner constructions in the language of these advanced learners, the above profiles show the influence of a variety of contextual or socio-cultural factors that influence learner history, including frequency of ambient input, length of sojourn, vocation, instruction, gender, and the development of identity as an academic professional. Phongsak’s abilities may be due to the fact he had been in the US much longer (2 years) than either Bing or Yong (1 semester each), and so perhaps the instruction resonated with growing pragmatic awareness of the language ambient to the

environments he found himself. Considering the frequency of ‘might wanna’ in expert usage, it could be that Phongsak had been noticing it all along, and just needed it to be confirmed. While some research shows that attitudinal orientation towards the host culture has a greater effect on the development of expert-like usage than sojourn length (Adolphs & Durow, 2004), it may be that sojourn length itself influences that orientation. In addition, that orientation is probably influenced by vocation, which in Yong’s case, may have been sending a message contrary to that of his ESL course. As a graduate assistant who was already conducting office hours, Yong was gaining actual experience that was probably more meaningful than role plays. In contrast, Phongsak had a scholarship from his government and had never actually taught or conducted office hours, and so had no experience with which to compare.

Gender issues may also be a consideration, as many studies have found that women tend to mitigate more than men (Yates, 2005). While this was not a focus of the analysis, it is notable that Yong mitigated less than average and Bing did more than average, although so did Phongsak. Regarding influence of instruction, Phongsak and Bing demonstrate positive effects from the implementation of the directives unit, which is the topic of the next chapter. Yong’s situation emphasizes the need for a true needs analysis and coordination with other departments on ITA preparation curriculum development. Finally, it is interesting to consider how the learners’ usage profiles provide insight into the development of their identities as academic professionals. If teachers teach only based on their own apprenticeships-of-observation (Johnson, 1999), ITAs would teach according to how they are taught, which in the case of many would be rather authoritarian. The findings from the corpus analysis are evidence that they may in fact

tend to, at least in comparison to experts in office hours contexts. At the same time, all three profiled learners said that they wanted to interact with students in a more less authoritarian, American style in the future. Furthermore, some learners like Yong may find a less directive, mediating style more useful as a TA, a possibility not addressed in most TA preparation courses. It is this disconnect among several factors that is perhaps the most important finding of the project: how ITAs have been socialized by their schooling, what they are taught in ITA preparation courses, what they actually do and experience in their departments, classrooms, and offices, and the kinds of academic teaching professionals they say they want to become.

## **7.5 Conclusion**

The above analyses of individual usage add texture and detail to the framework provided by the corpus and genre analyses of group usage, which are heuristically powerful but only part of the picture. The variation analysis shows that learners and experts alike have preferred forms that may depend on both individual affect and genre conditions, although some learners over-rely on these preferred forms. Meng and Xin seem to rely very much on their ‘lexical teddy bears’. While Bing used a larger variety of directive forms, Phongsak’s post-instruction assessment showed he had at least declarative knowledge of as many as Bing did. One issue with corpus analysis of learner language this underscores is that when a learner does not use a construction, it is impossible to know whether s/he did so as a matter of incomplete understanding or ability, simple avoidance, or intentional choice.

The individual profiles offer insight into learner usage phenomenon like learner constructions, in this case ‘you can’, which may be overused not because it is a ‘teddy

bear' but rather a multi-functional 'swiss army knife'. In terms of the social functional framework described in chapter 4, Yong's case shows the role of the third, as yet little discussed, plane of objectivity, as he often (but not always) projects agency to his department, which has the effect of implicitizing the source of authority from the learner. While Yong is simply fulfilling the role expected of him, he becomes another gatekeeper to a frustrated student, a disempowered doorman for his department as it were. Bing's usage profile shows understanding on the plane of choice and the interactional need for appeals to independence, while Phongsak's use of 'you want to' and explanation of 'we' shows, albeit less clearly, and understanding of the need for appeals to involvement. Optimistically, their understanding is evidence for the effectiveness of the pedagogical intervention, which is the subject of the next chapter.

## CHAPTER 8

### **AN APPLICATION OF CORPUS-INFORMED LANGUAGE AWARENESS INSTRUCTION**

The purpose of this dissertation, as an endeavor in applied linguistics, is not only to analyze the usage of directive constructions by experts and learners in office hours contexts, but also to apply the findings of this analysis to instruction. As discussed in chapter 2, many studies similar to this project describe the pragmatics of educational institutional discourse but relegate pedagogical implications to their conclusions. In contrast, this project has as its primary goal the application of the findings of previous chapters to the development of a unit of directives instruction for ITA pedagogy, based on a corpus-informed language awareness approach that is theoretically commensurate with the social-functional framework proposed and applied in the previous chapters. To this end, in this chapter I present a directives instructional unit for ITA preparation curricula. I first situate the instructional unit by reviewing directives instruction in popular ESOL grammar textbooks and ITA preparation texts materials, and by discussing the corpus-informed language awareness approach on which the unit is based. After this background discussion, I present the original instructional unit, based on preliminary research and implemented in an experimental study in ESL 118 in Spring 2006. This is followed by a description of the study conducted on two experimental sections that implemented the unit and one control section that did not, taught by the instructor of one of the experimental sections. Finally, I discuss the transformative impact that the instruction had on the overall ITA preparation curriculum.

## **8.1 Background**

To ground the design and development of the instructional unit, I first review the state of directives instruction in advanced ESL and ITA preparation curricula by examining treatment in a representative selection of popular texts on grammar and ITA preparation. Next, I present the pedagogical approach in terms of its constituent elements, corpus-informed pedagogy and language awareness pedagogy, followed by guiding principles for curriculum development distilled from this discussion.

### **8.1.1 Directives instruction in ESL and ITA preparation curricula**

A few of the studies reviewed in chapter 2 found negative influences by textbooks on learner language use. In his study of the use of inclusive and exclusive pronouns in academic writing, Harwood (2005) examined how pronoun usage was taught in several EAP instructional texts. He determined that instruction on the subtle differences among different fields and genres was usually bypassed in favor of prescriptive guidelines to avoid use of pronouns in general. Similarly, McEnery and Kifle (2005) found their participant usage had been heavily influenced by one nationalized textbook that simplified instruction on the usage of epistemic devices, which subsequently had a negative influence on learner understanding. Regarding the current project, directives too are subject to oversimplification, decontextualization, and prescriptivism. In this section, I critically examine the pedagogical approach towards directives and office hours instruction reflected in one TESOL pedagogical grammar text, two ESL grammar textbooks, and two ITA training texts. The purpose of this examination is to present the need for the instruction as well as to inform its design. In a more thorough examination, I would survey ITA preparation curricula in a number of universities to determine which texts and approaches are actually used; unfortunately, time and resources for a survey of adoption rates were unavailable.

### 8.1.1.1 Grammar texts

Most ESOL grammar textbooks that cover advanced grammar have a section on modals that include directive uses. In his popular *English Grammar in Use*, Murphy (2004) explains the usage of several directive modals. He explains the semantic difference between ‘have to’ for obligation and ‘must’ for opinion, as well as the British difference between ‘didn’t need to’ and ‘needn’t have’. He explains that ‘have got to’ can be used for ‘have to’, while ‘ought to’ may be used for ‘should’. He gives very thorough explanations for situations where ‘must’ or ‘should’ are used. His explanation of ‘had better’ is clear and comprehensive, that “with *had better*, there is always a danger or a problem if you don’t follow the advice. *Should* only means ‘it is a good thing to do’” (p. 70). In her well-known blue-covered *Understanding and Using English Grammar*, Azar (1989) presents the modals ‘must’, ‘have to’, and ‘have got to’ for expressing necessity, that ‘have to’ is most common, that ‘must’ is a stronger ‘have to’, and that ‘have got to’ is more informal ‘have to’. Azar explains that advisability is expressed with ‘should’, ‘ought to’, and ‘had better’, and that ‘should’ and ‘ought to’ have the same meaning, ranging from a suggestion to a “statement about responsibility or duty” (p. 78). She explains that ‘had better’ is usually stronger, and that it often “implies a warning or a threat of possible bad consequences” (p. 78). In the activities, students complete fill-in-the-blanks, for example, and discuss the difference in strength between ‘should’ and ‘had better’, or ‘should’ and ‘must’. She presents ‘to be to’ and ‘to be supposed to’ for expectations, and ‘let’s’, ‘why don’t’, ‘shall I/we’, and ‘could’ for making suggestions. The meaning of ‘let’s’ is presented as “I have a suggestion for us” (p. 86), and ‘why don’t’ for making friendly suggestions. ‘Could’ is contrasted with ‘should’, as the former “offers suggestions or possibilities” (p. 87), while the latter “gives definite advice” (p.

87). Activities practice this distinction by contrasting parallel statements made with ‘could’ or ‘should’.

While thorough and comprehensive, the coverage of modals in both Murphy’s and Azar’s texts are organized according to a mix of structure and function, so that all the modals are clustered in one section of each book, but also so that each page in those sections covers a different function. So, for example, ‘must’ appears several times, in treatments of its epistemic and directive meanings. This ordering may confuse a student going through the text linearly, although the texts do not require that instructors do so. More problematic is the fact that the rules and sample sentences are not based on actual empirical usage, but rather on Murphy’s and Azar’s intuitions. This sanitation is misleading to students for several reasons. First, it decontextualizes usage by mixing examples from a huge variety of genres, fields, and speakers. Second, it portrays spoken language as sentences rather than utterances, and notes very few distinctions between spoken and written usage. Third, it simplifies qualities and tendencies into reified rules; for example, the use of ‘had better’ implies that the speaker may believe the addressee is already doing something they shouldn’t, which is impossible to discern without understanding of the discourse-level context of the utterance. Simply saying it is a ‘strong warning’ is inaccurate. Finally, these treatments give rare forms like ‘ought to’ and genre-specific forms like ‘to be to’ as much treatment as common forms like ‘have to’, which may mislead students.

One can be optimistic that instructors are aware of these issues, which may be the case if they are trained in pedagogical grammar. One popular text in such courses is Celce-Murcia and Larsen-Freeman’s *The Grammar Book* (1983). In their treatment of deontic modality, the authors discuss not only true modals but also periphrastic (have to, have got to) and modal-like (e.g. had better) forms. They present deontic modals as ‘social interactional’ modals, and explain that “a

person using them properly take into account the characteristics of the social situation” (p. 83), and that knowledge of the formality of a particular context is necessary for their use. They explain that for giving advice, modals should be presented as a system ordered “according to the speaker’s degree of authority and/or conviction, or the urgency of the advice” (p. 85), and advocate use of a gradient like ‘you might/you could, you should, you had better, you must, you will’, with the caveat that the distance between the modals is not equal regarding authority or urgency. They explain that periphrastic modals are usually less formal than their equivalents.

Celce-Murcia and Larsen-Freeman’s approach is sensitive to genre choices and constraints, and their explanation of modality as a gradient system provides teachers with a conceptual basis for their understanding. However, by their own admission their explanation is not comprehensive, and while they mention the importance of the characteristics of the social situation, speaker authority, and formality of context in modal choice, they do not explain how or that the modality system interacts with other systems also dependent on these issues, like politeness (mitigators) or deictic (pronoun) systems. In addition, like the student grammar texts, they fail to mention some common modal-like forms (esp. ‘need to’ and ‘want to’), as well as nominal and adjectival forms of modality.

#### **8.1.1.2 ITA texts**

In the popular ITA training text *Teaching Matters: Skills and Strategies for International Teaching Assistants* (1990), authors Teresa Pica and G. Barnes devote an entire 14-page chapter entitled ‘Working with Students on a One-to-One Basis’ to discussion of office hours. The stated purposes of the chapter are to help ITAs understand “the reasons students must be encouraged to come to office hours, the TA’s role in the office-hour meeting, strategies for using time effectively, (and) ways to conduct a tutorial” (pp. 54-55). Areas of discussion include

‘establishing office hours’, ‘relationships outside the classroom’, ‘guiding the student through office-hour meetings’, and ‘individualized instruction’, which covers tutorials. The ‘relationships’ section discusses ‘office-hour protocol’, and includes dozens of points of what to do and what not to do, for example establishing authority, knowing student names, not keeping students waiting, showing interest, letting them ask questions, seeming eager to see students, minimizing interruptions, and avoiding the roles of counselor, romantic interest, and close friend. The ‘guiding’ section presents what could be seen as genre moves: openers, getting to the point, returning to the point, requesting clarification, and bringing closure. Each one of these sections includes cultural information and 7 to 9 suggested phrases, like ‘Hello. Would you like to see me?’, ‘I suppose we’d better get down to business’, ‘Let’s go back now’, ‘I didn’t quite follow that’, ‘I can see your point, but’, and ‘Do you think we’ve covered everything?’. None of these phrases were found in the MICASE EOH corpus. Moreover, directives are not treated explicitly. The chapter concludes with 10 multiple choice comprehension questions and 11 discussion questions, including several role play scenarios.

In its treatment of office hours, *Teaching Matters* covers most aspects discussed in the contextual analysis (chapter 6), albeit not particularly critically. The prose is clear and didactic, and unambiguously aimed at an audience unfamiliar with American academic culture. ITAs are told what office hours are about, and how to conduct them in a polite and professional manner. In the ITA development classroom, an instructor would assign the reading, students would answer the multiple choice questions on their own, and the class would do the discussion questions together. In spite of its comprehensive nature, however, the text is not based on empirical language use, but rather expressions the authors “have heard teachers say to their students that seem to work well for them” (p. 60). Of the 47 phrases offered, only two showed up

in the MICASE EOH corpus, ‘come on in’ twice, and ‘have a seat’ once. In addition, although it may be important for students to be aware of them, neither corrected topic digressions nor major clarification requests were noted in the genre analyses conducted for chapter 6.

In *Teaching American Students: A Guide for International Faculty and Teaching Assistants in Colleges and Universities* (1997), author Ellen Sarkisian offers a 5-page appendix (pgs. 67-71) on office hours based on the reflections of a math and a social science professor. The first professor discusses reasons why students come to office hours, including picking up tests, difficulties with course work, and getting to know students, while the second discusses the parts of an office hour session, namely opening, setting a climate, listening to the student, and closing. For each of these reasons and stages, each professor offers a series of questions s/he might use like ‘How was that quiz for you? Any surprises so far?’, ‘OK, well, keep trying. If you run into problems like this again, drop by’, or ‘Do you want to pull a chair over to the table?’. In comparison to *Teaching Matters*, Sarkisian’s text is less didactic and more descriptive, focusing on the content of teaching issues and not on language. It is based on intuitions, although the authors are not linguists, and so the text includes more idiolectic and idiomatic examples.

In sum, the above representative textbooks on the grammar of directives and ITA preparation may be inadequate in their coverage for several reasons. Most importantly, they do not base instruction on actual empirical findings nor are the language examples actual usage. In addition, directives are not explicitly related to office hours genres at all. The ESOL grammar texts, even if partially functional, isolate forms and systems from authentic usage contexts, so that functional relationships are indiscernible. While the pedagogical grammar text mentions the role of social contexts, like the ESOL texts it is constrained by its own table-of-contents, which forces a structural-grammatical organization onto the contents. The ITA texts are very

comprehensive, but present academic structures as benign and unproblematic, and avoid discussion of issues regarding authority and power. In addition, they do not show the relationship between academic language use and academic genres and contexts, which might empower ITAs as learners and future professionals by developing language awareness. As a caveat, it should be noted that this small review does not imply that past or current ITA preparation curricula are ineffective, since there are many other factors that influence effectiveness<sup>1</sup>. Still, ITA preparation, and all teacher preparation for that matter, is about teaching individuals to wield symbolic power. The danger of allowing uncritical access to power is that it may be wielded uncritically.

### **8.1.2 A corpus-informed language awareness approach**

As the above review of textbooks shows, common approaches to instruction on office hours and directives for ITAs are transmission-based and relatively uncritical with regards to discussion of the exercise of academic authority and power. While the texts certainly include lots of useful information, and individual instructors may very well use a variety of methods with these texts, it is difficult to see how the material on office hours and directives as presented in the texts would be commensurable with the implications of the studies reviewed in chapter 2.

According to these studies, effective instruction of directives and office hours should involve:

- explicit teaching of pragmatic forms and usage, including mitigators (Bardovi-Harlig & Hartford, 1993)
- learner recognition of role of status in institutions (Iedema, 1996)

---

<sup>1</sup> In my opinion, most of the reviewed texts reflect the prevailing pedagogy at the time, and their popularity is testament to their clear and comprehensive design.

- learner awareness of how authority and social positioning are realized in linguistic expression (Iedema, 1996)
- learner awareness of the role of generic and disciplinary contexts on user-addressee relationships and thus the use of directives (Hyland, 2002)
- instructor use of learner corpus analysis to customize instruction for immediate learner needs, including the use of comparative analysis (McEnery & Kifle, 2002)
- treatment of non-verb forms of modality (McEnery & Kifle, 2002)
- learner familiarization with multiple genres (Farr, 2003)
- explicit teaching of pronoun forms (Harwood, 2005)
- learner examination of their own and others' texts, and consideration of collocations and phraseological units (Harwood, 2005)

Informed by these findings and by research on corpus-informed and language awareness approaches, I developed a unit of directives instruction for ITAs in response to this need, to be implemented in ESL 118 in Spring 2006. In these materials, learners are offered corpus-informed tools that allowed them to examine contextualized directive language usage, as used by experts and learners like themselves. Use and discussion of these resources promote learner awareness of genre features and language as discourse, leading to a critical understanding of their roles as TAs and future academic professionals. The instruction is corpus-informed, meaning that the materials are corpus-based and contextualized for specific purposes, and also based on principles to develop language awareness, including awareness of language-as-discourse and of genre features. These concepts are presented below, followed by a summary of curricular design principles that informed the development of the instructional unit.

### **8.1.2.1 Corpus-based pedagogy**

As commercial lexicographers in the field of L2 education have begun using genuine language derived from corpora for their work, materials and curriculum designers have also turned to corpus studies for ideas and inspiration. Some claim that corpus-based pedagogy is something of an educational panacea for L2 teaching (Johns, 1991; 1994; Bernardini, 2004). Regarding corpus research, Sinclair (2004) and Tognini-Bonelli (1996) make the distinction between ‘corpus-based’ and ‘corpus-driven’, the latter referring to a more open, inductive approach to analysis of untagged corpora. Regarding pedagogy, ‘corpus-driven’ is reserved for approaches that explicitly teach corpus or concordance reading skills, while ‘corpus-aided’ (Bernardini, 2004) or ‘corpus-assisted’ is used as an umbrella term, much as ‘computer-aided/assisted’ is for the highly diverse field of CALL. McCarthy (1998) adds ‘corpus-informed’, to refer to “what we do with the insights (from corpus research) in pedagogy, since insights alone are no guarantee of good teaching, and must be mediated in some way to create models that are meaningful and useful to language learners” (p. 23).

#### **8.1.2.1.1 Corpus-driven pedagogy**

A pioneer in corpus-driven language pedagogy, Johns (1991) describes a pedagogical application of concordance analysis for ‘data-driven learning’ (DDL), where the language learner is “essentially a research worker whose learning needs to be driven by access to linguistic data” (p. 2). In a DDL approach, learners examine concordances of data from expert speakers to discover language facts and rules inductively, developing metalinguistic awareness based on evidence from authentic language use (p. 3). Bernardini (2004) enthusiastically touts a ‘discovery learning’ form of DDL, in which “learners are guided to browse large and varied text collections in an open-ended, exploratory way” and encouraged to pursue corpus-aided language

discovery projects of their own trajectories. She sees in this approach a panacea of “knowledge restructuring, critical autonomy, researcher skills, language awareness, opportunities for communicative interaction and so on” (p. 27).

Although corpus-driven activities are usually seen as an advanced activity, St. John (2001) concluded that there is a place for corpus-based instruction in beginning language instruction after examining how a beginning German learner, albeit a linguistics major, successfully used a parallel corpus and a concordancer. Kennedy and Miceli (2001) reported mixed results with a corpus-based course for intermediate Italian learners, finding that some learners successfully made use of corpus tools while others did not. They concluded extensive training is necessary for learners to benefit from corpus-driven approaches. Using DDL as a theoretical framework, Cheng, Warren, and Xu (2003) combined graduate-level IT and discourse analysis courses to teach corpus analysis and have students complete corpus investigations on EFL. Chambers (2005) taught a short course for undergraduate language majors on corpus consultation, grounded in DDL and principles of learner autonomy. Student experiences were generally positive, although most felt that corpus consultation was at best a supplement to dictionary and grammar text use.

In a corpus-driven approach, learners may also compare corpus-based data from expert speakers with that from learners, including their own production (what Seidlhofer, 2002 terms ‘learner-driven data’). Meunier (2002; see also Granger & Tribble 1998) notes the benefits of exposing learners to both expert and learner corpora and corpora-based materials, including the opportunity for learners to access to appropriate, expert production, to notice the differences between their own and expert production, and to negotiate and interact with other students, teachers, and experts during the learning process. Gabel (2001) advocates providing guidance to

learner exploration by presenting guidelines, first through use of statistical procedures to determine true overuse and underuse, followed by qualitative contextual analyses for interpretation.

Building on the learner corpus analysis work of Belz (e.g. 2004; 2006), Belz and Vyatkina (2005) presented learners involved in a telecollaborative partnership with corpus-informed activities developed from the learners' own immediate production, in just-in-time fashion, so that learners were able to develop awareness of the pragmatic consequences of their own usage, in this case, of German modal particles. While the authors call this approach corpus-driven, it is highly relevant for the learners since it is their own production, unlike other corpus-driven approaches where the subject corpora are drawn from users and genres far removed from the learners.

Lee and Swales (2006) report on a corpus-based EAP course that in effect trained NNS doctoral students to use corpus analysis on corpora of general expert usage, expert academic speaking, expert academic writing in their field, and small corpora of their own writings. The effect was 'technology-enhanced rhetorical consciousness raising' (p. 72) through decentering, as learners discovered language usage by exploring texts by multiple authors on their own, rather than by relying on the authority of grammar textbooks or a single instructor. The authors admit that course success was contingent on a number of convergent factors, including motivated doctoral students who were highly genre-acclimated, computer proficient, and familiar with quantitative data (p. 72).

#### **8.1.2.1.2 Corpus-informed pedagogy**

Aston (1995) explains that corpus analysis and language pedagogy may find common ground in a non-componential schematic view of language based on use, as "corpus linguistics

sees such schemata primarily as social facts of language, evidenced in patterns of use in the community as a whole” (p. 262). He advocates the use of smaller, genre-specific corpora, since they allow for easier pragmatic and semantic interpretation and the data remains contextualized within a discourse, and suggests that corpus-based tasks be embedded in communicative tasks to remain relevant to learning. He explains that “corpora should be seen as providing opportunities to engage in discourse as well as to analyse it” (p. 270).

Braun (2005) argues that ‘pedagogic mediation’ is required for teachers and learners to overcome the shortcomings of corpora as they are used for research and lexicography, specifically in terms of size, design, annotation scheme, and data format. She contends that many corpus-driven activities, especially those involving concordancers, de-contextualize text from discourse, which must be re-constructed or authenticized (Widdowson 2003, in Braun 2005 p. 53) for language learning to occur. In design and in application, a corpus-based approach should be combined with a discourse-based approach “which focuses on the analysis of linguistic means of expression in relation to their communicative (situational) and cultural embedding” (p. 54). This may be achieved by designing smaller corpora that are focused on the genres and registers relevant to learners, annotated for pedagogical use, and supplemented with audiovisual versions if possible. Discourse-focused enrichment activities like exploratory tasks and explanations should be integrated with the corpus, and learners should read texts in their entirety to contextualize corpus findings.

L. Flowerdew (2005) presents a survey of studies that successfully combine corpus-based approaches with genre-based pedagogy. She argues that small corpora are useful for pedagogic purposes, explaining that “the more the corpus draws on features from the students’ own socio-cultural environment, the easier it should be for the teacher to act as a kind of mediating

specialist informant of the raw corpus data, thereby authenticating the data for classroom use to fit the students' reality" (p. 329).

Gavioli and Aston (2001) also advocate instructor involvement, and argue that corpus tools can build learner autonomy and language awareness if the corpus and activity are carefully structured and graded. They maintain that discourse authentication can occur when a corpus-based activity forces the learner into the objective role of text observer, alternating with the role of discourse participant (p. 241). Activities like writing an essay, engaging in a discussion, and collaborating with others serve to mediate interaction between the learner and the corpora and provoke a 'pragmatic reaction' as the learner alternates between observer and participant. Language awareness is effectively raised, for instance, through discussion of corpus data, enabling "learners to develop their own descriptive frameworks, and to question and critique those of teachers, textbooks, and reference materials" (p. 242).

### **8.1.2.2 Language Awareness**

Because of their focus on raising consciousness and making explicit knowledge about language, language awareness principles are very compatible with corpus-informed approaches. Carter (1998) explains that language awareness is about "understanding of tendencies, variable rules, and choices according to context and interpersonal relations" (p. 52). Language awareness has its roots in the U.K. in the 1970's, when there was a movement to introduce knowledge about language across the curriculum in response to a fall in literacy rates. Foreign language education and awareness of linguistic principles were seen as a means to both address literacy and prejudice (Hawkins, 1998). Concomitant with this movement was the growing popularity of Chomsky-inspired mentalist models and Krashen's Monitor Model, both of which downplay any role of higher cognitive functions in acquisition, which, along with misinterpretations of

communicative language teaching principles that grammar did not need to be taught, may account for why language awareness did not gain popularity in the U.S.

One aspect of language awareness is the understanding that spoken language is highly divergent from written language, a fact that is often ignored in language pedagogy, even though discourse competence is often a stated pedagogical goal. Corpus research has found that spoken language is full of vague language, interpersonal exchange, repetition, and formulaic, fixed phrases (Carter, 1998). Interactional spoken language, because it is extra-referential and clause- rather than sentence-based, and highly context-, participant-, and culture-dependent, can be better understood as discourse (McCarthy 1998). Interactional features of discourse include turn-taking and interrupting, discourse markers, information staging, generic features, and contextual constraints like ellipsis and deixis. Crucial to discourse awareness is understanding of spoken genres, which have core yet negotiated structures, are comprised of stages, and are highly contingent on participant goals and relationships (McCarthy, 1998, p.47).

A view of language as discourse “focuses, where appropriate, on complete spoken and written texts and on the social and cultural contexts in which such language operates” (McCarthy & Carter, 1994, p. 1). Translated into a pedagogical approach, building learner discourse awareness means raising learner consciousness of the role of choice in language use, as “different choices from within the grammatical system realize different meanings”, both textual and ideological (p. 29). To achieve this, learners explore the relationship between features of spoken discourse and the social contexts in which they function, i.e. genre (p. 38). McCarthy and Carter (1994) offer several key curricular principles for application to a curriculum that develops awareness of language-as-discourse. The contrastive principle involves “contrasting treatments of the same or related content enable a focus on language difference and can do much to promote

language awareness” (p. 166). The continuum principle involves exposure to a variety of texts in the same genre by different authors. The inferencing principle involves teaching strategies for cultural and literary understanding, or interpretative skills.

### 8.1.2.3 Curricular design principles

To facilitate the development of the unit of instruction, the above literature is distilled into several interrelated curricular design principles.

- A corpus-informed language awareness (CILA) approach uses corpus data selected to illustrate specific learning objectives, and so is **objective-driven**. Unlike data-driven, inductive approaches, CILA instruction does not present students with unorganized corpus data with the expectation for emergent understanding. The corpus data, either statistical findings or specific texts, are chosen to illustrate a specific learning objective.
- The primary objective of CILA instruction is to develop **metalinguistic awareness** of ideational, interpersonal and textual qualities of language in use. This is accomplished through the use of corpus-based tools like concordances and frequency analyses, as well as text and discourse analytic tools and activities, to compare and contrast a variety of texts. This reflects the aforementioned contrastive principle (McCarthy & Carter, 1994).
- A CILA approach uses corpus data from a **variety** of users, modes, media, and related genres. This variety may include individuals from different backgrounds, L1s, ages, and genders, using spoken and written forms, in a variety of spoken, written, visual, aural, and computer-mediated media, and from genres of a variety of types similar or different

to the genre of study. This reflects the aforementioned continuum principle (McCarthy & Carter, 1994).

- Variety serves the objective of developing learner awareness of language **choice**. An inherent danger to presenting expert corpus data to learners is the implication that expert usage is prescription. Instead, presentation of a variety of usage examples implies that meaning making is a matter of choice, and that individual choice and style are co-determinant factors in usage patterns.
- A CILA approach uses corpus data from genres comparable to those in which learners will need to participate, produced by learners like themselves wherever possible. This establishes **relevance** for learners, as they see immediate application in the texts they are learning, and can identify with participants.
- Relevance serves the objective of developing learner awareness of genre **expectations** and practices, as learners have a vested interest in understanding these expectations. Learning genre practices and parameters in tandem with the concept of choice allows learners to see language as dynamic, negotiated, and creative.
- Unlike corpus pedagogy approaches that de-contextualize usage excerpts, a CILA approach treats excerpts as texts embedded in co-text and socio-cultural contexts. This **contextualization** allows learners to see language as discourse and meaning as situated

and context-dependent. This reflects the aforementioned inference principle (McCarthy & Carter, 1994).

- Contextualization serves the objective of promoting critical reflection and learner **transformation**. Learners can undergo transformation by experiencing usage as both observer and participant, by noticing similarities and differences between usage events, by considering the tension between language as choice and expectation, and by reflecting critically on the interdependence of context, participants, and usage.

## **8.2 A unit of instruction on directives usage**

Guided by these principles and encouraged by the findings of previous studies, I designed a unit of instruction on directives usage (appendix G) in early Spring 2006, for implementation in the three sections of ESL118 by the regular instructors. The unit was designed to be part of the course's office hours unit, and to take approximately two 70-minute class periods. It was preceded by other activities and readings related to office hours, and was followed by the office hours chat activity (A1, see §3.2.1.1), which was followed by the face-to-face office hours role play (A2). Importantly, the instructors were asked not to explicitly tell the student to use the directives they had just learned in the activities. An instructor's version of the unit framed each activity with instructions and suggestions for implementation. So that the instructors would remain invested in the activities, they were encouraged to adapt the activities as they saw fit to their own teaching styles.

The unit is split into four activities. The objective of the first activity (A) is to introduce students to directives, and to present them with an example of how unintended directive usage can lead to misunderstanding, thus establishing relevance for the students. To this end, students

are presented with an excerpt from an ITAcorp role play between two Fall 2005 students where the T reacts somewhat rudely and abruptly to an S request for credit on a test answer. The instructor is provided with questions that ask students to identify contextual features, to identify several bolded phrases (which are directive constructions) and discuss their function, and to discuss how professors and TAs usually make suggestions.

In the second activity (B), the objective is to introduce learners to the factors that are involved in choosing a directive. To this end, students are presented with a chart (appendix G, activity B) representing the considerations involved in choosing a directive. This representation acts as a materialization that mediates the learners' understanding of directive functions. In the diagram, a list of core directive verbal elements ('must', 'have to', '(have) got to/ gotta', 'had better/best', 'need to', 'want to', 'should', 'might want to', 'ought to', 'why not/don't'), are presented in the middle of the page with the 'strongest' at the top and 'weakest' at the bottom. Arranged around this list are sidebars presenting a variety of features for consideration ('context', 'participants', 'purpose', 'medium', 'meaning', 'source', 'strength', 'hedging and intensifying'), each with questions that the user theoretically considers when choosing a directive. The instructor is directed to go through each construction and features with the students, calling attention to the various qualities and importance of each. In a follow-up activity, students return to the first activity and discuss how the use of different constructions would affect the interpersonal outcome.

The objective of the third activity (C) is to provide students with practice identifying directives and discussing related features. For this activity, students are presented with 7 excerpts of MICASE usage for analysis, as a class and then individually or in pairs. Students are asked to identify context, directives, and hedges, to discuss how the related features influence speaker

choice, and to discuss how different choices would affect meaning and effect. Excerpts were chosen from MICASE for their diversity of topic and directives usage, to illustrate differences in the features and how they affect choice.

The fourth activity's objective (D) is to provide students with data on overall trends in directive usage patterns in American English, and for them to identify differences in frequency, meaning and mode among various modals. Students are presented with data on modal usage trends, summarized from an article on necessity and obligative modal usage trends (Collins, 2005). The data is in chart form, accompanied by questions on differences in the modals. Comparative corpus activities comparing ITAcorp and MICASE were not included, since the data had not yet been collected, although this approach was built into the second version of the unit (Thorne, Reinhardt, & Golombek, forthcoming; see also §8.3).

### **8.2.1 Experimental implementation**

The first version of the directives unit as described above was implemented in Spring 2006. The purpose of the experiment was to determine the effect, if any, of such an instructional unit on learner usage of directives in the role play activities that followed (A1, A2, and A8—see table 3.2 for descriptions). As mentioned above, an experiment was set up so that two different instructors implemented the unit in two of the three ESL118 sections (1 and 3)<sup>2</sup>, one of whom also taught an office hours unit to the third section (4) using the regular materials, without the directives unit, as a control. The instructors were videotaped and interviewed afterwards.

---

<sup>2</sup> As mentioned in chapter 3, there was no section 2, an anomaly of the university scheduling system.

### 8.2.1.1 Overall construction usage by groups

The total number of students in, and tokens produced by, the experimental and control groups is presented in table 8.1 below. Although the control group had only 67% the participants as the combined experimental group, it produced 80% as much language. Informal observers of the control group noted that they seemed more cohesive and dynamic as a group<sup>3</sup>. It should also be noted that not all 33 students participated in all three activities (see §8.2.1.3 below).

Table 8.1. Number of students and corpus tokens in SP06 experimental and control groups

group(s)	no. of students	no. of tokens (A1+A2+A8)
Experimental (sections 1 + 3)	20 (10+10)	32998
Control (section 4)	13	26513

Table 8.2 shows the number and frequency of a selection of directive construction used by the two groups in comparison. The selection was comprised of those constructions given treatment on the advance organizer<sup>4</sup> activity of the instructional unit (TC=treated construction), the most frequent constructions (MFC) as identified in chapter 5<sup>5</sup>, and ‘you can’. The constructions are ordered by log-likelihood from most overused by the experimental group to most overused by the control group (i.e. underused by the experimental group).

---

<sup>3</sup> I don’t think this had an influence on their directives usage; I just thought I’d mention it.

<sup>4</sup> Four of the constructions listed on the advance organizer are not in this chart. The construction ‘might want to’ was listed separately from ‘want to’, while ‘why not/don’t’, ‘(have) got to’ and ‘ought to’ were also listed but not used by anyone in either group.

<sup>5</sup> except ‘I would’, which was not used by any learners. It is considered a MFC because its usage frequency by experts alone was over 1 per 10K.

Table 8.2. Number and frequency per 10K of directive constructions used by SP06 EXP and CON groups

directive construction	Experimental (EXP)		Control (CON)		log-l EXPvsCON
	total	freq/10K	total	freq/10K	
you had better ①②	19	5.758	1	0.377	+16.09*
you have to ①②	28	8.485	4	1.509	+15.38*
you could ②	8	2.424	0	0.000	+9.44*
I want you to ②	5	1.515	0	0.000	+5.90*
you want to ①②	9	2.727	2	0.754	+3.42
you must ①	1	0.303	1	0.377	
you need to ①②	14	4.243	12	4.526	-0.03
you should ①②	22	6.667	27	10.184	-2.19
I recommend ②	3	0.909	10	3.772	-5.66*
you don't need to ②	2	0.606	9	3.395	-6.48*
you can	138	41.821	162	61.102	-10.76*
I suggest ②	1	0.303	13	4.903	-15.00*
all TC (①) uses	93	28.19	47	17.73	+7.01*
all MFC (②) uses	111	33.65	78	29.42	0.83

\* A plus sign indicates EXP overuse, while a minus sign indicates CON underuse, in comparison to EXP. A log-likelihood greater than 3.84 indicates a p-value less than .05, and thus, a significant difference between EXP and CON production of this construction. Statistical difference was not calculated for 'you must', since its combined EXP+CON frequency was less than 1.0 per 10K.

① TC: indicates this construction was treated in the materials (i.e. listed on the advance organizer)

② MFC: indicates this construction was a most frequent construction (MFC) as measured by overall use in the combined EOH and LOH corpora. See chapter 5.

The experimental groups used the TCs significantly more frequently (28.19 per 10K) than the control group (17.73 per 10K)<sup>6</sup>, but not the MFCs, for which there was no significant difference (log-l=0.83). Specifically, in comparison to the control group, the experimental groups made significantly more use of the TCs 'you had better', 'you have to' and, if 'you want to' and 'I want you to'<sup>7</sup> are combined, 'want to', which is actually how it was presented on the organizer. The experimental group made significantly less use of several of the constructions that were not treated in the instructional unit, namely 'I suggest', 'I recommend', and 'you can', all of which, coincidentally, the experts also used less frequently. An interpretation of these findings is

<sup>6</sup> When only the 24 participants who completed all 3 activities are considered, the results are similar. Thirteen experimental participants produced the TCs at a frequency of 49.11 per 10K, while eleven control participants produced them at 28.58 per 10K.

<sup>7</sup> When the two constructions are combined as 'want to', the difference between the groups becomes significant at +7.69. Seven different individuals from the experimental groups used 'you want to', 'I want you to', or both—see §8.2.1.4

that the experimental groups were choosing from the repertoire they had just been presented, while the control group was making more use of constructions with which they were perhaps already familiar. In other words, the instructional unit seems to have had an effect on the experimental groups in regards to the kinds and amount of directive constructions they used in subsequent role play activities<sup>8</sup>. Again, the experimental participants were not explicitly told to use the directives in those activities.

#### **8.1.1.2 Mitigator and intensifier usage by groups**

Since only mitigators (MIT) and intensifiers (INT) and not the other adjunct types were part of the instructional unit, only the use of these adjunct categories by the two groups is presented here. The methods described in §5.4.2.2 were again applied to the experimental and control groups use of MFCs and TCs<sup>9</sup>. The usage frequency of each construction is listed in table 8.3 below, as well as the frequency and percentage of the construction used with mitigators and intensifiers. The constructions are presented in descending order of most frequent in the combined EXP and CON corpora, i.e. the entire SP06 corpus.

---

<sup>8</sup> An important caveat is that the control group may have been using other directive constructions, e.g. imperatives, more frequently. Since these other construction types were not part of the instruction, they are not analyzed here.

<sup>9</sup> As with the chapter 5 analysis, 'you can' was omitted from this analysis.

Table 8.3. Frequency of individual TCs and MFCs with and without MIT and INT in EXP and CON groups

construction	EXP freq per 10K (%)			CON freq per 10K (%)		
	all	w/ MIT	w/ INT	all	w/ MIT	w/ INT
you should ①②	6.67	3.64 (55)	0.61 (9)	10.18	1.89 (19)	0
you have to ①②	8.49	1.21 (14)	0	1.51	0.38 (25)	0
you need to ①②	4.24	0.91 (21)	0	4.53	1.82 (42)	0
you had better ①②	5.76	2.12 (37)	0	0.38	0	0
I suggest ②	0.3	0	0	4.9	0	0
I recommend ②	0.91	0	0.61 (67)	3.77	0.38 (10)	1.13 (30)
you don't need to ②	0.61	0	0	3.4	1.51 (44)	0.75 (22)
you want to ①	2.73	2.73 (100)	0	0.75	0	0
you could ②	2.42	1.21 (50)	0.30 (13)	0	0	0
I want you to ②	1.52	0.30 (20)	0	0	0	0
you must ①	0.3	0	0	0.38	0	0
all TC (①) uses	28.19	10.61 (38)	0.61 (2)	17.73	4.09 (23)	0
all MFC (②) uses	33.65	12.12 (36)	1.52 (13)	29.42	5.98 (20)	1.88 (6)

Of the treated constructions, the experimental group used mitigators more frequently than the control group with three of the five constructions that both groups used ('you should', 'you had better', 'you want to'), while the control group used them more frequently with two ('you have to', 'you need to'). Considering the low frequency of the control group's use of 'you have to' (only 4 instances), however, this statistic may not be useful. More telling might be the statistics regarding all TC and MFC uses, which show that the experimental groups not only used mitigators and intensifiers more frequently (TC: EXP=28.19 per 10K, CON=17.73 per 10K; MFC: EXP=33.65 per 10K, CON=29.42 per 10K), but they also used them a greater percentage of the time with the constructions (MIT: EXP=38%, 36%, CON=23%, 20%; INT: EXP=2%, 13%, CON=0, 6%). This can be considered evidence for positive influence of the instructional intervention on the experimental groups' use of mitigators and intensifiers with these directive constructions, insofar as more, as opposed to fewer, appeals to listener independence is a desirable learning outcome with these learners. While more is not necessarily better, in this case it is more expert-like regarding mitigator use with MFCs (experts=11.09 per 10K, experimental=12.12 per 10K, control=5.98 per 10K).

### 8.1.1.3 Individual usage

The statistics above do not show how individual differences may have influenced the findings. As discussed in chapter 7, individual differences show a much more complex picture of usage than can be captured in frequency statistics. As shown in table 8.4, when the 24 participants who completed all 3 activities are considered separately from the 33 in the group analyses above<sup>10</sup>, analysis shows that the experimental participants used a larger variety of construction types more times on average than the control participants. As with the larger group analyses, there is a greater difference between the two groups with regards to the TCs; the difference in the MFCs may again not be significant.

Table 8.4 Mean uses and types of MFCs and TCs by partial EXP and CON groups

partial group	treated constructions (TCs)		most frequent constructions (MFCs)	
	uses	types	uses	types
Experimental (n=13)	5.2	2.2	6.4	2.9
Control (n=11)	3.3	1.5	5.6	2.8

Of the eight learners in table 7.2 (see p. 131) chosen from the 24 participants, the five denoted by asterisks in that table participated in the experimental groups. All five of these individuals and the three control participants each had different preferred MFCs, denoted by superscript circles, except for two individuals who shared the same preferred MFC. Of the three individuals who are profiled in chapter 7, Bing C and Phongsak T were in the experimental groups, while Yong W was in the control groups. As discussed in those profiles, Bing used 5 TCs 8 times, and Phongsak used 2 TCs, one of them the most uses of ‘you want to’ of any participant at 4 uses. In comparison, Yong W used only 2 TCs one time each. In sum, participants from both groups used a variety of directive constructions in the role play activities, but experimental participants used the constructions given treatment in the materials more

---

<sup>10</sup> With such small total token numbers per individual, this is important when doing comparisons of individuals. See also footnote 5.

frequently. Insofar as usage, as opposed to non-usage, indexes awareness, this is further evidence for the effectiveness of the instruction in raising learner awareness of the usage of directive constructions.

### **8.3 An expanded curricular re-design project**

My experience with the first version of the materials led to a second version of the materials, which is not fully discussed here, since it is an outgrowth, and not technically part, of this project. It is more fully described in Thorne, Reinhardt, and Golombek, forthcoming. Developed in late Spring 2006 and implemented Summer 2006, it is mentioned here to illustrate the cyclical process of curricular design and redesign in an applied learner corpus analytic approach. After the Spring 2006 implementation, I consulted with two faculty members and asked them if they were interested in collaborating on the continued development of the learner corpus and ITA preparation curriculum, the first faculty being a corpus analysis expert and the second being the curriculum coordinator for ESL 118. I was conducting initial versions of the analyses for this dissertation project, and was eager to incorporate the findings into future implementations, even though I had reached my data collection goals. Per the findings of the first implementation, where the impact of the instruction was seen as contingent on a broader approach across the whole semester, an entire revision of the ITA preparation curriculum was envisioned, with the directives unit I would refine as an initial piece.

The second version reflects the dissertation research I had continued after I had designed and implemented the first version above, which reflects only preliminary findings. The second version more directly reflects the social-functional model of directives usage (chapter 4), presenting politeness as a key feature of directives choice, including pronoun/subject choice as related to involvement appeals, and adjunct choice as related to independence appeals. It also

includes more statistical charts comparing ESL 118 learner usage to MICASE expert usage, and features conceptual ‘flow chart’ materializations based on sociocultural principles, specifically the use of explicit and conceptually framed objectifications of the units of language that expert speakers utilize to achieve particular social actions (e.g. Neguereula, 2003). These materializations also address Gal’perin’s emphasis on efficiency and establishing a systemic “orienting basis” as necessary elements for generating qualitative shifts to new forms of cognitive functioning. Informal evaluation of the second version’s Summer and Fall 2006 implementation was very positive, especially by the ESL 118 students, many of whom were especially interested in the contrastive corpus analysis, perhaps because of their backgrounds as graduate students in mathematics, engineering, and the hard sciences. Future plans are to evaluate the second version using methods developed in this project, and to evaluate development of genre awareness through pre- and post-semester assessment measures.

The expansion and re-development of the ITA preparation curriculum at Penn State continues as of this writing. The new curriculum builds on the above corpus-informed language awareness approach by framing the entire ESL 118 course curriculum in the development of learner awareness of spoken academic discourse genres, including office hours as well as lecturing and discussion-leading genres. Starting in Fall 2006, the curriculum coordinator began developing and implementing more corpus-informed language awareness activities for all the genres. The Fall 2006 implementation also fronted several broad genre awareness activities, intended to frame the new approach. In addition, a future teaching research project activity was introduced, where students would explore the discipline-specific written and spoken genres relevant to their specific needs.

Another change implemented after for the Fall 2006 implementation was the revision of the task descriptions for the office hours chat role play (A2) and post-semester evaluation office hours role play (A8), to better reflect what actually goes on in office hours, based on content analysis of the EOH corpus from MICASE. The content of many office hours sessions, and one of the common actual duties of a TA, is the tutoring and reviewing of specific, course-related materials. This poses a challenge to TA preparation courses like ESL 118, because the students are from a variety of fields in which the instructor and many of the other students do not participate. To meet this challenge, starting in Fall 2006, the A2 S role is being played by an expert speaker, who improvises the role of an undergraduate in need of tutoring on a specific topic.

#### **8.4 Conclusion**

In this chapter I presented an instructional unit on directives usage for ITAs. I first provided background information on how several popular textbooks present directives and office hours, and on literature explaining the corpus-informed language awareness approach used to develop the unit. I then described the unit of instruction and its experimental implementation, and how the instruction was found to have an impact on directives usage. I then discussed and described the transformative effect of this instruction on the ITA preparation curriculum.

In conclusion, the end goal of this dissertation project, to inform classroom instruction, has evolved into an ongoing, cyclical journey of research and application to practice. The purpose is to provide current and future academic professionals from international backgrounds the resources to be aware of and use English academic language like directives to their benefit and the benefit of their students and communities. In chapter 9, the project conclusion, I will discuss remaining questions regarding these issues.

## CHAPTER 9

### CONCLUSION

In this dissertation I examined directive language usage by learners and experts in office hours contexts from a variety of analytic perspectives, in a series of separate but interrelated studies. As per the guiding research question, the purpose was to describe the nature of directive language usage in relation to the exercise of academic authority, and to explore how directive language use could effectively be taught to L2 English learners preparing to be academic professionals. In seeking the answer to this question, the findings go beyond description of directive language usage as a register, to relate usage to the texts and contexts of the office hours genre, and to examine usage by both groups and individuals. As the end goal of the project was to inform language awareness instruction for ITAs, there are a number of implications for ITA instruction based on the project findings. These findings lead to recommendations for future research on advanced learner language development and interaction. In addition, the dissertation contributes broadly to several areas in applied linguistics. Most significantly, the dissertation stands as an applied example of usage-based language pedagogy. The dissertation also speaks to pragmatics and politeness theory, offering several important correctives to current conceptualizations. Methodologically, the dissertation offers a replicable approach to corpus analysis of pragmatic usage, and also offers a major contribution to the nascent field of applied learner corpus analysis. These contributions are presented after discussion of specific implications for ITA instruction and recommendations for future research.

## 9.1 Implications for ITA instruction

The findings of this dissertation have significant implications for ITA education. A major finding is that there are measurable group and individual differences in directives language usage among ITAs in preparation and practicing experts in office hours contexts, which suggests the need for investigation into possible reasons for these differences. My response in chapter 7 outlined core features of ITA directive language use, including a smaller repertoire of constructions, a wider range of usage frequencies of particular constructions, and reliance on fewer constructions (the ‘teddy bear effect’) (Hasselgreen, 1994), especially highly multi-functional ones (the ‘swiss army knife effect’). These phenomena should be taken into consideration when designing instruction for ITAs.

In chapter 5, I showed that learners use certain constructions, especially those that appeal to listener involvement, and to a lesser extent those that appeal to listener independence, less frequently than do experts. A related finding from the surveys and interviews discussed in chapter 7 is that for the most part, ITAs in preparation do not want to come across as authoritarian to their future students. If ITAs are not aware of the pragmatic effect of their language choices on listeners, when they are in teacher roles, their language choices may have undesired effects, or they may avoid taking pragmatic risks. In view of this finding, ITAs should be made aware of the pragmatics of giving directives. The efficacy analysis in chapter 8 showed that language awareness instruction could have a positive influence on pragmatic awareness.

Using a genre approach in chapter 6, I showed that contextual features play an important role in office hours, and that there are distinct textual moves in office hours involving directives. These findings should be incorporated into ITA preparation curricula. This focus would address the concerns of one of the instructors in this project that the genre aspect of office hours,

including discussion of openings, diagnosis, and closing moves, was not integrated with the unit of instruction. This focus would also serve to bridge lexical use to discourse and situation through discussion of features like social context, purpose, roles, and values. For example, discussion of the cultural differences, especially the issue of why those in American academic T roles even make choice and involvement appeals to students, or frame directives in terms of divergent or convergent options, can bring students to a ‘third place’ (Kramsch, 1993) where they can might see American and their own cultural practices in new light.

I showed in chapter 7 that individual differences play a role in directives language usage, especially in relation to contextual factors like educational background, gender, and sojourn length. The learner profiles showed that ITAs in preparation and acting ITAs may have very different understandings of the role of a TA based on contextual factors and learning histories, which influence their use of directives. The implications of this are that ITA instruction should take into consideration these contextual factors for individual learners, and that materials should be designed that raise learner awareness of these factors and allow for customized instruction when possible.

The version of the materials presented in chapter 8 was designed with corpus-informed language awareness principles in mind. Because of the logistics of project implementation, only very preliminary findings and initial research went into the design of the materials<sup>1</sup>. In addition, the efficacy analysis was restricted to treated constructions (TCs) rather than the most frequent constructions (MFCs) as eventually determined by analysis of the entire corpus. With this in mind, the unit of instruction of directives usage should be updated to make the TCs match MFCs as described in the analyses. Among the MFCs, several particular constructions warrant special

---

<sup>1</sup> The second version of the materials presented in Thorne, Reinhardt, and Golombek (under review) are a better reflection of the findings of the studies in this dissertation project, most notably the model in fig. 4.1.

treatment, specifically those that were significantly overused ('you can') or underused ('you want to', 'I would') by the learners. Based on the positive response of ITAs participating in the training course, contrastive corpus activities are suggested as an effective means for this focused treatment. A related recommendation is to incorporate the findings on modifier use in chapter 5. In the second version of the materials, subject/pronoun choice is associated with involvement appeals, and adjunct choice is associated with independence or choice appeals. The adjunct choice activities could be supplemented with the findings on diversity and frequency of adjunct type.

The most radical implication for ITA instruction is that ITA programs in American academia should consider re-evaluating their entire ITA preparation curricula in terms of a usage-based approach, and re-design portions of their curricula according to corpus-informed language awareness principles. Such a revamp at Penn State was inspired by the findings of this project, but met with difficulties because of the scale of the project and the necessity for instructor retraining. Implementing a curriculum based in a usage-based approach in many ways requires a radical re-thinking of the nature of language and learning that goes to the heart of personal philosophies of teaching and learning. Because this is a major contribution of this dissertation, I discuss it further in §9.3.1 below.

## **9.2 Recommendations for future research**

There are many possible directions for research implied by the findings of this dissertation. Two areas of research with perhaps the most heuristic potential involve learner language development and interaction. The current project is highly descriptive in nature, and makes only a few claims regarding development. In chapter 7, it was claimed that in comparison to experts, learners tended towards greater variation in preference of and reliance on certain

constructions. Development refers to change in communicative competence over time. Each learner's developmental path is unique and may involve phenomena like experimentation with forms and under- and overgeneralization, which would appear in a corpus as non-expert patterns of preference and reliance.

To investigate development, it would be necessary to develop a longitudinal corpus and trace learner development over time using the corpus data (for an example, see Belz, 2006)<sup>2</sup>. Ideally, the corpus would capture every utterance created by the learners, or as much as possible, to minimize extrapolating between two temporal points, which can miss certain variable features of development. One logical outgrowth of the current project would be to follow ITAs from the preparation courses into their actual teaching assignments to document their language use in a variety of contexts, with students, fellow graduate students, and professors (see Jenkins, 2000 for the latter), preferably over several semesters. This would provide the ultimate means to trace development and determine if any pedagogical intervention has lasting effects on usage. Directives or other pragmatic constructions could be the unit of analysis, and with enough data, one could investigate a variety of influences on usage, for example variables related to participant, listener, situation, and task. These variables are discussed below.

Another feature to investigate would be the usage of directives or other pragmatic constructions as a dynamic, negotiated process involving interaction. The force of a directive can transform in strength and direction during the course of an interaction depending on a variety of factors. The model of directives construction usage proposed in chapter 4 (fig. 4.1) illustrates the space where this negotiation occurs (see §9.3.2 below for discussion of the contribution this model makes). One possible method not explored in the current project that would be suitable to

---

<sup>2</sup> Longitudinal corpora are rare because of the logistics in following individuals over time. See Cutting (1999) for an example.

this examination of interaction might be conversation analysis (e.g., Benwell & Stokoe, 2002). Regardless of the methodology employed, it would be very interesting to examine strings of directives as they are given and negotiated in the course of actual office hours sessions. Interaction patterns would most likely be influenced by the same variables that impact development.

### **9.2.1 Participant variables**

Participant variables like gender, age, and identity would most likely influence either development or interaction. Some research has been done on these variables and ITAs (see Briggs et al., 1993), but none used corpus methods to my knowledge. One project in this area could look at a teacher's use of directives reflects the development of his/her teacher identity within a single interaction, across interactions with different learners, or over time. A related project could look at attitudes towards authoritarian teaching in depth, and separate attitudes towards formality from attitudes towards authority, which the current project did not do. To investigate formality and secondary school teacher use of directives, Yates (2005) offers a framework for analysis of adjunct usage that could be applied to ITAs and higher education.

A set of related variables is the influence of the relationship of the T and S participants on usage, especially the relative status of the T compared to the S participant. One limitation to the current project is that I sometimes conflated teaching assistants and professors into the same 'T-participant' category<sup>3</sup>. As shown by the case of Yong W, this may be inaccurate and misleading, as TAs in some departments act more as mediators than instructors. TAs have less academic authority than professors, and some TAs have more or less authority than other TAs based on

---

<sup>3</sup> The MICASE web interface allows for searching by role, but it does not distinguish between graduate students that are students, and graduate students that are TAs. For the corpus analyses, I separated the T-participants production from others.

their assignments as instructors, tutors, or graders. Although the register consistency analysis of the TA and faculty halves of the EOH corpus showed very little significant difference in usage for many constructions (§5.1.3.2, table 5.5), it would be interesting to see an analysis of more usage by these subgroups. Interlocutor familiarity is a related variable that deserves additional attention. It might be expected that the level of familiarity between the T and S interlocutors in an office hours interaction would influence formality levels and presumably the type of directive constructions used. Again, this could be investigated in a single or series of interactions, or as a matter of development over time, as relationships usually evolve over the course of an instruction period (for example, over a single office hours session or across a semester of interaction).

### **9.2.2 Listener variables**

Another area to investigate, especially for an ITA or ESL learner population, is the perception of the listener, i.e. the student in office hours interactions with ITAs, with regards to several more language-related variables. This is closer to the mainstream strand of ITA language research, and it is also related to assessment, since it is the perception of an evaluator that determines the fate of many learners who would be ITAs. For the ESL 118 students at Penn State in the current project, the AEOCPT test that they must pass presents several notions worth examining from an empirical usage-based perspective. The AEOCPT website<sup>4</sup> tells prospective test-takers that the exam will assesses “pronunciation (the articulation of specific sounds, and the stress and intonation of your speech); fluency (the rate and appropriateness of pauses in your speech); and comprehensibility (how extensive and appropriate your usage of vocabulary and grammar is)”. While pronunciation has been the object of much research on ITAs, it would be

---

<sup>4</sup> At the time of writing, this information was on [http://lals.la.psu.edu/ita\\_aeocpt.php](http://lals.la.psu.edu/ita_aeocpt.php)

very difficult to examine with corpus methods, while research on fluency and lexico-grammatical elements are better suited to this method.

Perceived fluency has been shown by some research to be related to speaker use of certain language forms like smallwords (Hasselgreen, 2002) or what I have termed adjuncts. If this is the case, there is even more reason for ITAs to become aware of the role of adjuncts, as appeals to independence, and as markers of fluency to listeners. From a corpus perspective, markers of fluency like length of utterance, word gap, and pause length are difficult to transcribe accurately, but could be assessed by measuring words per turn and turn length. With regards to interaction, one could assess listener response through examination of interactional elements like syntactic and lexico-grammatical complexity of response, to determine if a listener simplifies or alters responses in reaction to perceived low fluency. Using smallwords as an indexical feature, the development of a single learner's fluency could be examined by measuring change in frequency and diversity of use over time.

### **9.2.3 Situation and task variables**

A variety of situational variables also influence directive usage patterns. Most notably, office hours sessions vary vis-à-vis field of study, course level, and reason for visiting, all of which might affect the type of directives used by the T participant. As noted in §6.3.1, each of the MICASE sub-corpora were designed with a balance of these variables, so that analyses of a sub-corpus as a whole would produce even results. However, the differences among the discipline sub-genres within the office hours macro-genre might be telling, as Hyland (2002) and Harwood (2005) found in academic writing. Directives usage may also be affected by the location of an office hours session, whether the session is held in a private office, a common office, or a public location.

In the ESL or ITA preparation classroom, situation variables correspond to task variables, which could be examined separately for their influence on directives usage, again with regard to interaction or development. Task type and task medium, for example, may have had an effect on directives usage in the current project. A1 was computer-mediated practice, A2 was face-to-face class performance, and A8 was face-to-face post-evaluation. The role play scenarios were all slightly different, and although realistic, some parameters and topics could have induced more directive language on the part of the T than others. Informal observations were that there were far more uses of directive constructions in A1 and A2, which immediately followed instruction, than in A8, which and took place 6 weeks later. There could be a retention issue, as students may have forgotten what they had learned over time, which would be a point of investigation for a longitudinal learner corpus.

In addition, the stakes of the three activities varied greatly. A1 was low-stakes with the only audience being the interlocutor playing the student role, A2 was medium-stakes performed with visual cues and an audience of classroom peers, and A8 involved high-stakes testing, since the learner's performance was directly scored by two evaluators, one of whom played the S role. In this way, task authenticity might play a role after all in the high stakes post-evaluation role plays, since the student role playing the T was well aware that although the evaluator was role playing the S, s/he was in control of the student's immediate academic fate, i.e. whether or not the student would be allowed to assume TA duties. In that case, the pragmatics of the actual situation could trump those of the role play situation, and the student T could avoid directives so as not to imply disrespect to the evaluator. Once more of ITAcorp data is available for analysis, these topics will be more available for investigation.

### **9.3 Contributions to applied linguistics**

As described above, the findings of this dissertation implicate a rather full agenda for future research, especially in the areas of development and interaction. While these recommendations are specific and detailed in nature, the dissertation has broader, more important implications that can be considered contributions to applied linguistics and related fields. Most significantly, this project contributes an applied example of usage-based language pedagogy. In application to theory, this dissertation also makes a significant contribution to pragmatics and politeness theory in the model of directives usage presented in figure 4.1 (p. 67). The project contributes methodologically to corpus linguistics in that it offers a replicable approach for the analysis of pragmatic usage, and also to applied corpus linguistics as a model of learner corpus analysis that bridges research and classroom practice.

#### **9.3.1 Usage-based language pedagogy**

The first, and perhaps most significant contribution of this dissertation is as an example of a usage-based, social-functional conceptualization of language and learning. In contrast to some psycho-structural approaches to language (e.g. Pinker, 1994), from a usage-based perspective (e.g. Langacker, 2000; Tomasello, 2003), language structure does not underlie use, but rather emerges from use (Hopper, 1998). Grammar is an always developing, never completed or fully static resource, even for native/expert speakers. Forms exist in prefabricated formulaic constructions, but through social use in various contexts, deconstruct and reconstruct into new forms. Metaphorically speaking, language learning is better understood as participation rather than acquisition (see Thorne & Lantolf, 2007, Sfarid, 1998, for discussions).

In social-functional approaches to language learning and teaching, function and genre rather than grammatical categories are the curricular organizing principles. Learning occurs

through meaningful, contextualized use and social participation, and learners are empowered to see language as a matter of making meaningful linguistic choices rather than making correct or incorrect choices. In contrast to approaches that decontextualize language forms for analysis, and then attempt to reanimate them for use<sup>5</sup>, in a social-functional approach, the explicit relationships between the social and functional aspects of use remain ecologically intact throughout the learning experience, and remain relevant to learner needs. The challenge arises when teachers and learners are asked to re-conceptualize their understandings of language, teaching, and learning with these notions, as their beliefs are often subconsciously embedded and even self-contradictory. In the minds of many administrators and teachers, perhaps since they are constantly measuring it for testing purposes, language is better understood as a distributed object, rather than a shared resource. Changing this understanding would be the subject of another dissertation.

### **9.3.2 Pragmatics and politeness theory**

Another significant contribution of this dissertation is the social-functional model of directives usage I developed (fig. 4.1) that advances the complexity and comprehensiveness of pragmatics research in the area of speaker modality. The model is intended to reflect the dynamic stance of the director-speaker using directive constructions, operating on three interdependent planes: choice (dependence-independence), involvement (exclusion-inclusion), and objectivity (implicit-explicit). There are several innovative and corrective aspects to the model. Since it has little to do with politeness in the sense of ‘well-mannered’, the model avoids the terms ‘positive’ and ‘negative’ politeness by using more transparent terms suggested by the literature, especially ‘independence’ and ‘involvement’ (Scollon & Scollon, 1995). By having the choice and

---

<sup>5</sup> Also known as ‘voodoo language pedagogy’.

involvement planes perpendicular to each other, the model corrects the misleading assumption that ‘positive’ and ‘negative’ politeness are two ends of a linear continuum. In my formulation, I also add a third dimension of explicit and implicit objectivity based on systemic-functional notions (Iedema, 1996; Halliday & Mattheissen, 2004) of grammatical metaphor and the obfuscation of power. I argue that this interdependent planes model more accurately describes the intricacies of institutional discourse, and pedagogically, provides a more powerful heuristic for visualizing and objectifying linguistic resources used to wield symbolic power.

A director-speaker (i.e. the teacher) implies and negotiates symbolic power with the directed-listener (i.e. the student) through the use of directive constructions in this space. Because it is three-dimensional, the model allows for the illocutionary force of a single directive, in isolation or in negotiation, to be dynamically represented on all three planes simultaneously. While the model was developed for directive constructions usage, it may be applicable as a model representing the linguistic exercise or negotiation of institutional power in general. As such, it may be useful as a theoretical framework for understanding the dynamic exercise of power in other forms of institutional discourse such as advising, mentorship, or apprenticeship.

### **9.3.3 Corpus-based examination of pragmatic usage**

A major methodological contribution of this dissertation focuses on corpus analysis of pragmatic usage. Corpus linguists have generally avoided analysis of pragmatic usage because speech acts rarely correspond perfectly with grammatical or lexical classes, and so are difficult to identify with traditional corpus methods. In chapter 4, I outlined a method that may be replicated for corpus-based examination of pragmatic usage that avoids manual tagging and requires reading only a sample portion of the corpus. First I used a grounded approach to identify as many functional units as possible in the sample, and used those particular constructions to inform

definition of the unit of analysis. I then took great care in defining a unit of analysis, that would relate the social and the functional aspects of pragmatic usage, but that would also remain quantifiable and searchable using corpus analytic methods. This unit of analysis is grounded in a usage-based construction theory of language (Tomasello, 2003), and so is not bound to any pre-existing syntactic structure, but rather relies on attested units of language use and their frequency and social function.

For items like imperatives that were not easily identifiable, words like ‘don’t’ or ‘please’ were used as indexicals; these instances were considered representative rather than comprehensive of all imperative constructions. To allow for consideration of cultural and situational factors, I then complemented this register-based analysis with a genre-based perspective, analyzing the same data in fully contextualized chunks of discourse. The moves analysis showed that a corpus-based register analysis alone would have failed to capture some directives that were marked by framing rather than explicit construction use (§6.3.3.4.4). An analysis of individual usage showed its variable nature, especially in the case of learners. In consideration of these factors, it may be that corpus analysis of pragmatic usage is necessarily combined with other methods. In any case, the use of mixed, yet complementary methods adds to the heuristic strength of the approach.

#### **9.3.4 Applied learner corpus analysis**

To examine pragmatic usage from multiple perspectives (register, genre, group, and individual), the corpus-based mixed methods approach described above can be applied to anyone—experts or learners, in isolation or in comparison. When applied to learners for the purpose of informing pedagogy, the approach can be considered applied learner corpus analysis, a nascent field with much potential for applied linguistics research on language development,

assessment, and materials and curriculum design. Applied learner corpus analysis usually involves the use of small, localized, genre-specific learner corpora designed for a particular applied linguistics research or pedagogical goal. It has the benefit of being immediately applicable and relevant to the instruction for which the analysis is done, and as an ongoing process, embodies the cyclical nature of research informing practice. As an example of this approach, this dissertation provides a major methodological contribution to the emerging field.

As described above as a recommendation for future research, applied learner corpus analysis can be used to examine language development empirically using a longitudinal corpus, in a learner-specific manner sensitive to individual differences and microgenetic change over time (for an example, see Belz 2006). For assessment purposes, applied learner corpus analysis can be used to develop corpus-based profiles of learners at different stages of development, without necessarily using the high-pressure and time-bounded environment of traditional testing. These profiles could be used over the course of a semester or more to track an individual learner, not against an unrealistic ideal, but against her/himself. Similarly, in materials and curriculum design, applied learner corpus analysis can be used to inform instructional resources that are customized specifically for a learner or group of learners. In the case of the current project, analysis of the production of one semester's learners was used to inform the design of the next semester's materials. Because learners readily identify with learners from previous semesters like themselves, they can benefit greatly from such an approach. If students can be provided with their own production, the benefits can be immediate and substantial (see Belz & Vyatkina, 2005 for an example).

## 9.4 Conclusion

The end goal of this dissertation project was to inform language awareness instruction for ITAs. The findings led to a series of recommendations for future ITA instruction and research on advanced learner language development and interaction. In addition, the dissertation contributes to several areas in applied linguistics, specifically usage-based language pedagogy, pragmatics and politeness theory, corpus analysis of pragmatics usage, and applied learner corpus analysis. The project was based on the notion that when future teachers are made aware of how their language choices influence their teaching and student understanding, they are empowered to make those choices more intentionally and to create communicative conditions that are less prone to misinterpretation, thus positively impacting their professional lives and their students' educational experiences. This awareness can also promote conscious reflection on their role and purpose in education and the responsibility inherent in wielding symbolic power in educational contexts.

The goal of learner empowerment through language awareness may be noble, but it is challenged by formidable institutional forces and conceptual frames (Lakoff & Johnson, 1980). For example, when language is conceived of as a transmissible product, it becomes an object of consumption. This fits into the conceptual frame of understanding for many Americans, who have come to see education as a product, and students as consumers with rights to high quality products. The issue would be alleviated if language awareness were a goal for all learners, undergraduate and graduate, native and non-native speakers alike, since it provides a means to see how language is used to frame and simplify very complex issues like communication and education. We can work towards this goal by conducting sound research and designing effective

instruction for our own students based on these principles. It is my hope that this dissertation serves as an example that will inspire future research to this end.

## **Appendix A**

### Office hour role play chat scenarios (Spring 2006 activity A1)

1. You took a quiz in class and you got a bad grade. You think some of your answers were good and you that you deserve at least partial credit. Discuss this issue with your TA.
2. You were sick for two weeks, and during that time you missed a mid-term. Now you need to talk to your TA to discuss what course work you missed and to schedule a make-up test.
3. You are having trouble understanding the professor's lectures because you don't think that the TA prepares you enough in the discussion sections. Discuss the problem with your TA.
4. You have been having a problem with your roommate, and it's been very distracting for you and you haven't been able to complete your homework assignments for class. Ask your TA for some advice.
5. You need to take a course in order to graduate on time, but you have not met the pre-requisites for the course, although you feel you know the material. Ask the TA of the course if you can take the course.
6. You visit your TA at the beginning of the semester to ask him / her about the course syllabus and learn more about the requirements of the course, because you may want to drop the course for another.

## **Appendix B**

### Biographical, language learning, and computer use experience survey

#### A. ABOUT YOU

1. Name:
2. Class section and meeting time:
3. Instructor's name:
4. Sex:
5. Age:
6. Email address(es):
7. URL of personal website (if you have one):
8. Degree you are working towards:
9. College and department:
10. What is your country or culture of origin?

#### B. ABOUT YOU AND ENGLISH

1. What are your native language(s)/dialect(s)?
2. If they are different from your native language, what are the language(s)/dialect(s) of your parents, extended family, community, region, or country?
3. What languages have you learned besides English? To what extent? Please consider reading, writing, listening, speaking, grammar, etc., as well as functional ability.
4. How long have you been in the United States? If you have been elsewhere in the U.S. besides PSU for study, work, or residence, where was it and for what purpose?
5. What countries besides the U.S. have you visited? For what purpose?
6. Where did you learn English? Please list all the schools and the length of time you studied English there. Please also include any non-school experiences.
7. Please describe your study of English up to this point. Please include as many details as possible, e.g. your approach to learning the language, the role of the teacher, aspects you consider important for learning a language, etc.

C. ABOUT YOU AND TECHNOLOGY (Fall 2005 only)

1. Do you have your own personal computer?
2. How much time per day do you spend on a computer:
  - a. using email?
  - b. searching/browsing the Web?
  - c. instant messaging/chatting?
  - d. using other tools?
3. Where do you usually access the Internet?
4. For what purposes do you use a computer? the Internet?
5. How many years have you been using a computer? the Internet?
6. Did you use computers in elementary/middle/high school? If you did, for what purposes?
7. In what country was this schooling primarily?
8. Did you use computers at university for your undergraduate work? If you did, for what purposes?
9. In what city and country was your undergraduate university?
10. Rate your ability with the following application types (expert semi-expert average novice beginner):
  - a. word processors like MS Word
  - b. instant messengers like AOL IM
  - c. web browsers like Explorer or Firefox
  - d. presentation software like MS Powerpoint
  - e. webpage authoring software like Dreamweaver
  - f. course websites like Angel

## Appendix C

### Survey on attitudes towards authoritarian teaching

Please rate your agreement with these statements according to your opinions about the roles of teachers and students in a college situation:

1- strongly disagree   2- somewhat disagree   3- neutral   4- somewhat agree   5- strongly agree

1. Relationships between teachers and students in my home country are more formal than in the U.S.\*
2. It is normally good for students to ask the teacher questions during class.
3. It is acceptable to eat or drink during class.
4. Students can learn a lot by working with each other.
5. The teacher and text are not necessarily the source of all classroom knowledge.
6. It is wrong for a student to challenge a teacher during class.\*
7. It's okay for a teacher to change the syllabus based on student needs and desires.
8. If a student doesn't understand, it's okay for him/her to ask a fellow student for help during class.
9. If a student misses a class, the teacher should make sure s/he gets the missed assignment.

\* Question 1 was not counted, and question 6 was reversed, in calculating the average score to determine attitude towards authoritative teaching

## **Appendix D**

Adjuncts found in most frequent constructions in expert and learner corpora (within 5 tokens to the left or right of main element)

COD – Conjunctive discourse markers:

also, and, but, now, or, so, then

FOD – Feedback-oriented discourse markers:

mhm, mkay, okay, right, right?, yeah

HES – Hesitation markers:

uh, um

INT – Intensifiers

absolutely, actually, always, certainly, do, really, see, seriously, still, strongly, will

MIT – Mitigators

all, almost, even, I don't think, I guess, I mean, I suppose, I think, I'm afraid, in general, in my opinion, just, kinda, like, may, maybe, might, only, or whatever, probably, sort of, usually, well, would, you know

## **Appendix E**

Questions for interview on attitudes towards teaching and conducting office hours

1. What kind of teacher do you most admire now and why?
2. What kind of teacher do you see yourself being in the future?
3. What do you think will be (or is) the most challenging thing about advising students and conducting office hours with them?
4. What are the challenges of being a TA?
5. Teaching styles can range from authoritarian to egalitarian. What do you think are the advantages and disadvantages of each style? If you could categorize yourself what would you say?

## Appendix F

### Written assessment of directives knowledge

1. People use directive language when they want someone to do something. Directives are usually phrases like 'you should (do something)' or 'I suggest you (do something)'. In the space below, please provide as many directive phrases as you can think of.
2. Please describe one or two situations where you think you would use directives.
3. What are the differences, if any, between A and B?

Situation 1: A math TA is explaining to a student that the use of a certain equation in a problem is appropriate at the point of reference.

- A. You want to use this equation.
- B. Do you want to use this equation?

difference:

Situation 2: A writing TA is explaining to a student that the use of emphasis is appropriate at the point of reference.

- A. I would put the emphasis here.
- B. You should put the emphasis here.

difference:

Situation 3: A TA is explaining an option for extra credit to a student.

- A. You could propose an extra project.
- B. You can propose an extra project.

difference:

## Appendix G

Unit of instruction on directives usage (Instructor's version— student version included all text below except italicized portions)

### A. Introduction

*Time: 10 minutes. Pass out the conversation and have students read. Questions to ask students are as follows, but please feel free to ask others or present the material as you see fit.*

- 1. Where is this taking place? Who are the participants? What is the purpose of the meeting?*
- 2. Look at the bolded words/phrases. Who makes them? What is their function? (directives; to give advice, make suggestions, etc.)*
- 3. How does the teacher react to the student's requests? (the teacher becomes upset) Why do you think so? (maybe because the student uses inappropriate directives with the teacher)*
- 4. What words can a professor use to make a suggestion to a student? What words can a student use to make a suggestion to a professor? Why is it different?*

Look at the sample chat conversation between ESL 118 participants in an office hours role play.

1. T: how are you?
2. S: good. And you?
3. T: good.
4. S: I have a question about the quiz
5. T: go ahead.
6. S: I got a bad grade. But I think some of my answers were correct, and should at least have partial credit. Do not you think so?
7. T: i do not understand what you are saying.
8. T: my quiz is multiple choice. it means that computer made scores.
9. T: so, there is no mistake and no partial points.
10. S: You do not give partial credits? Although my answers were not totally right, some of them are good. **Why don't you give me** some partial credits?
11. S: I understand the process and methods very well. It is not reasonable
12. T: anyway, there is no partial credits. it is my rule to every student.
13. T: please don't argue any more.
14. T: if you don't like my policy, you can drop off this course because it is not mandatory course.
15. S: But some students do not know the methods any more. Some know the methods, but only did the wrong calculation, just like me. **You should differentiate** those two kinds
16. S: Why drop? I like the course so much.
17. T: the result is the most important factor.

18. T: **you should not do** wrong calculation.
19. S: I do not agree. I think the process is the most important part
20. S: Ok. I will try next time
21. T: if you think so, drop off my course.

## B. Chart: Choosing a Directive

*Time: 15 min. Explain to students that we consider many factors when choosing a directive word or phrase, and that as TAs and future professors this is relevant to them, since making suggestions and giving advice is a large part of their duties. Use the chart to examine the conversation in activity A, maybe questioning how it could have played out differently. Go through each factor and discuss its role in choice.*

*FYI:*

- *Directives refer to the explicit or implicit use of 'you' as a subject.*
- *There are non-directive uses of all these words, e.g. 'must' and 'should' are used for deduction.*
- *In an informal situation we are more likely to make casual use of directives, whereas in a formal situation we are more likely to be aware of social differences, so that we might hedge what we say, or avoid stronger directives unless we are the ones with the authority.*
- *There are differences in spoken and written modes vis-à-vis frequency. See the chart in activity D.*
- *A friend giving advice is different from a parent, teacher, or doctor giving advice.*
- *Use of 'you must' usually infers that the speaker is the subjective source of the directive, while 'you have to' does not infer this. Most other directives can be either subjective or objective.*
- *'External' and 'internal' is similar, but more generally refers to the source of authority. So a TA might say to a student: 'The professor says you must' but 'I think you might want to', if the TA is trying to be build solidarity/appeal towards equality.*
- *'Had better' implies a negative consequence, and 'have got to/gotta' is especially emphatic.*
- *'Want to' is peculiar to speech. For example, in helping with homework, a TA might say 'Here I think you want to write X instead of Y'.*
- *A strong directive can be softened with hedging, and a weak directive can be strengthened with it. Consider the social effect of this strategy, and the consequences of the opposite.*

## CHOOSING A DIRECTIVE

### CONTEXT

Where and when is the exchange occurring?  
How formal or informal is the situation?

### PARTICIPANTS

Who is taking part and what is their social relationship? How equal or unequal is it?

### PURPOSE

Why is the exchange occurring? Help?  
Advice? Suggestion?

### MEDIUM

Is the context written or spoken?

**must**

**have to**

**(have) got to/ gotta**

**had better/best**

**need to**

**want to**

**should**

**might want to**

**ought to**

**why not/don't**

### MEANING

What meaning is desired? Prohibition?  
Obligation? Necessity?

### SOURCE

What is the source of the directive? Is it subjective or objective?  
Is it external or internal?

### STRENGTH

How strong is the desired directive? Is it weak, neutral, or strong?

### HEDGING & INTENSIFYING

Does the directive need softening (e.g. like, kinda) or intensifying (e.g. really)?

### C. Identification in Context

*Time (Mon): 10 min. (Wed): 20-30 min. For this activity, the first example should probably be done together as a class, and the rest should be assigned as homework. On Wednesday, then, students could break into pairs and compare their answers, reporting back to the class the answers to the questions. As students present them, write the directives and hedges on the board. The purpose is for students to see the diversity of expert use of directives and how they relate to the factors on the chart.*

*Students could underline or highlight the directives and circle the hedges. Where a directive starts or ends, or what is a hedge and what is not, is often not clear, and can be a topic of discussion. You could also have students identify things like backchanneling, comprehension checks, or discourse markers, to understand how directives fit into an entire utterance—heck, have them parse the whole thing and provide you with a discourse analysis of it! Just kidding.*

Look at the following examples of actual expert speaker usage. Why do the speakers choose to use the directives that they do?

1. Identify the context, participants, and purpose.
2. Identify the directive words or phrases. Remember that not all uses of modals are directive, even primarily directive ones like must or should. The subject of a directive is usually 'you', the addressee.
3. Identify any hedges that the speakers use.
4. How do the factors in the chart influence the speakers' choices?
5. How could the choices be expressed differently, and how would that change the meaning and effect?

Example 1:

T: um, okay that's good, okay here you go... okay this is a very awkward sentence.  
<LAUGH>

S1: yeah it is <SS LAUGH>

T: right. okay, i um... okay i want you to... well maybe i should see how you, do this okay. um...it seems to me like in this paragraph there are lots of totally different things going on. um, and in general you should do this throughout the paper too. you need to go through and ask yourself what the point of each paragraph is, right? um and make sure it has a point, and make sure it says what its point is, okay? um, cuz here we've got, okay. the first sentence which is to me a really important point that you should talk more about, right?

Example 2:

T: but you still need to be able to say them correctly. rather\_ you can't just say what are the assumptions random sample normal, no that's not enough you have to tell me what's a random sample, what's normally distributed what are you talking about there so you need to be able to say it in a sentence. alright?

Example 3:

S1: so you just square 'em and add 'em?

T: so yep you wanna do the same thing for how you do a T-test here you have to get your S-P... you have to get your S-P, and then have one over N-one plus one over N-two. so you have to get S-P from your two Ss, for that formula for combining the two.

S2: the S-P would be the S-one minus S-two?

T: nope S-P has the formula and, it's a pretty complicated one but it's about, averaging or pooling, we wanna take a look at page...

S3: you have to take the square of that?

T: yes you have to square each of them you weight 'em by the degrees of freedom

Example 4:

S: it's so easy when you explain it. <S LAUGH>

T: that's the problem with this course i mean you lear- you don't learn it\_ that's what i mean about you have to do it until it's really, second nature to you cuz you can't learn it watching, me you've got to work through it and really, pick up a lot of the fine points..

Example 5:

T: okay well then i would take that one. to give you a sense of of how that goes, and you need to remember that you have all of this year and all of next year to explore. you don't have to be in a rush you can just kinda be like i'm gonna take creative writing i'm gonna take math, the French you gotta keep going with unless you wanna start another language. so you know but really just you should take things that you like, things that might be interesting to you, and if you\_ cuz people graduate from the university with degrees in history and sociology and anthropology and English and, they get jobs

Example 6:

T: you might wanna... pull that out a little bit. um... now do you think\_ again, do you think, cuz i think this is an older person thing. do you think this is an older person thing, or do you think kids do this too?

S: about, like church and volunteering at the hospital? [T: uhuh ] i would just assume it's an older

T: alright. so, let me know it's an older, person thing. [S: okay ] right? it doesn't undermine your argument for, that to be the case. [S: okay ] right? in a way, it's just that you're s- eh then you just need to emphasize more clearly at the outset that, the comparison is partly cuz Howell is a more um, has historically been a more blue-collar, middle-class, place.

Example 7:

T: so first thing you do draw a picture... and you don't know the prices so just make some, downward sloping budget line. [S: mkay. this is what i want ] and the only thing is on\_ we have to be a little careful cuz we know that, this is quasi-linear so the indifference curves actually, curve but hit the axes. so if you wanna\_ here let's a... this is one thing that it sounds like i'm being picky [S: mhm ] but you really have to be careful about this cuz this is subtle things that are gonna matter later. [S: okay ] so you wanna get more, like that.

S: more curved for the quasi-linear

T: yeah cuz the quasi-linear's shifted straight up so you wanna be sure that you, you're looking kind of like that.

### D. Chart: Facts about the usage of directives in everyday American English

*Time: 10 min. This final activity is mainly as a time filler if the previous activity is finished in time. It presents corpus-informed factoids about directive usage by expert/native speakers.*

The following facts about directive usage in American English were derived from a corpus of written and spoken English (1). Use the chart to answer the following questions:

1. What is the most frequently used directive?
2. Which directive is most common in writing?
3. Which directives are more or less popular in British English?
4. What special meanings do 'gotta' and 'had better' have?
5. Which is more frequently used in speech, 'must' or 'have to'? Which is declining in usage?
6. Which is more frequently used in speech, 'should' or 'need to'? Which is declining in usage?
7. If the source of obligation for 'must' tends to be subjective, meaning that the speaker holds the authority, what difference could there be between 'you must' and 'you have to'?

modal/semi-modal	freq. per million	written: spoken ratio	trend in usage	notes/source of obligation or necessity
must	244*	3.5:1	↓	more subjective
have to	1364	1:7	↑	
(have) got to/gotta	173*	1:7	↑	emphatic
need to	473**	1:3.6	↑	more objective
should	662	1:1.25	↓	
ought to/oughta	51	1:2.7	↓	more subjective
had better	41	1:1.1		implies negative consequence

↓ historical decline in usage

↑ historical increase in usage

\* notably less frequent than British usage

\*\* notably more frequent than British usage

(1)Collins, P. (2005). The Modals and Semi-Modals of Obligation and Necessity in Australian English and Other Englishes. *English World-Wide* 26 (3), 249-273.

## BIBLIOGRAPHY

- Adolphs, S., & Durow, V. (2004). Social-cultural integration and the development of formulaic sequences. In N. Schmitt (Ed.), *Formulaic Sequences* (pp. 107-126). Amsterdam: Benjamins.
- Agar, M. (1985). Institutional discourse. *Text*, 5, 147-168.
- Aijmer, K. (2002). Modality in advanced Swedish learners' written interlanguage. In S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 55-76). Amsterdam: John Benjamins.
- Altenberg, B. (2002). Using bilingual corpus evidence in learner corpus research. In S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 37-54). Amsterdam: John Benjamins.
- Aston, G. (1995). Corpora in language pedagogy: matching theory and practice. In G. Cook & B. Seidlhofer (Eds.), *Principle & Practice in Applied Linguistics* (pp. 257-270). Oxford, UK: Oxford.
- Azar, B. (1989). *Understanding and Using English Grammar*. Harlow, UK: Longman.
- Bailey, K. (1982). *Teaching in a Second Language: The Communicative Competence of Non-Native Speaking Teaching Assistants*. Unpublished Ph.D. Dissertation, University of California, Los Angeles.
- Bardovi-Harlig, K., & Hartford, B. (1990). Congruence in native and nonnative conversations: Status balance in the academic advising session. *Language Learning*, 40, 467-500.
- Bardovi-Harlig, K., & Hartford, B. (1993). Learning the rules of academic talk: A longitudinal study of pragmatic change. *Studies in Second Language Acquisition*, 15, 279-304.
- Bardovi-Harlig, K., & Hartford, B. (1996). Input in an institutional setting. *Studies in Second Language Acquisition*, 18, 171-188.
- Bardovi-Harlig, K., & Hartford, B. (2005). Institutional discourse and interlanguage pragmatics research. In K. Bardovi-Harlig & B. Hartford (Eds.), *Interlanguage Pragmatics*. Mahwah, NJ: Lawrence Erlbaum.
- Belz, J. (2004). Learner corpus analysis and the development of foreign language proficiency. *System*, 32(4), 577-591.

- Belz, J. (2006). At the intersection of telecollaboration, learner corpus research, and L2 pragmatics: Considerations for language program direction. In J. Belz & S. Thorne (Eds.), *Internet-Mediated Intercultural Foreign Language Education* (pp. 207-246). Boston, MA: Heinle & Heinle.
- Belz, J., Reinhardt, J., & Rine, E. (under review). Telekorp: The telecollaborative learner corpus of English and German.
- Belz, J., & Vyatkina, N. (2005). Learner corpus research and the development of L2 pragmatic competence in networked intercultural language study: The case of German modal particles. *Canadian Modern Language Review/Revue canadienne des langues vivantes*, 62(1), 17-48.
- Benwell, B. & Stokoe, E. (2002). Constructing discussion tasks in university tutorials: Shifting dynamics and identities. *Discourse Studies*, 4(4), 429-453.
- Berk, L. (1999). *English Syntax: From Word to Discourse*. Oxford: Oxford University Press.
- Bernardini, S. (2004). Corpora in the classroom: An overview and some reflections on future developments. In J. Sinclair (Ed.), *How to Use Corpora in Language Teaching* (pp. 15-38). Amsterdam: John Benjamins.
- Bernstein, B. (1996). *Pedagogy, Symbolic Control and Identity Theory*. London: Taylor and Francis.
- Biber, D. (2006). *University Language: A Corpus-based Study of Spoken and Written Registers*. Amsterdam: John Benjamins.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25(3), 371-405.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics: Investigating Structure and Use*. Cambridge, UK: Cambridge University Press.
- Block, D. (2003). *The Social Turn in Second Language Acquisition*. Washington, D.C.: Georgetown University Press
- Blum-Kulka, S. (1989). Playing it safe: The role of conventionality in indirectness. In S. Blum-Kulka, J. House & G. Kasper (Eds.), *Cross-cultural Pragmatics: Requests and Apologies*. Norwood, NJ: Ablex.
- Bourdieu, P. (1991). *Language and Symbolic Power*. Cambridge, MA: Harvard University Press.
- Braun, S. (2005). From pedagogically relevant corpora to authentic language learning contents. *ReCALL*, 17(1), 47-64.

- Bresnahan, M., & Kim, M. (1993). Factors of receptivity and resistance toward international teaching assistants. *Journal of Asian Pacific Communication*, 4, 1-12.
- Briggs, S., Clark, V., Madden, C., Beal, R., Hyon, S., & Aldridge, P. (1997). *The International Teaching Assistant: An Annotated Critical Bibliography*. Ann Arbor, MI: The English Language Institute, The University of Michigan.
- Brown, P., & Levinson. (1987). *Politeness: Some Universals in Language Usage*. Cambridge, UK: Cambridge University Press.
- Carter, R. (1998). Orders of reality: CANCODE, communication, and culture. *ELT Journal*, 52(1), 43-56.
- Celce-Murcia, M., & Larsen-Freeman, D. (1983). *The Grammar Book: An ESL/EFL Teacher's Course*: Newberry House.
- Cheng, W., Warren, M., & Xu, X. (2003). The language learner as language researcher: Putting corpus linguistics on the timetable. *System*, 31, 173-186.
- Cobb, T. (2003). Analyzing late interlanguage with learner corpora: Québec replications of three European studies. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 59(3), 393-423.
- Coleman-Bingaman American Competitiveness Through International Openness Now, US Senate Bill (2005).
- Collins, P. (2005). The modals and semi-modals of obligation and necessity in Australian English and other Englishes. *English World-Wide*, 26(3), 249-273.
- Cutting, J. (1999). The grammar of the in-group code. *Applied Linguistics*, 20(2), 179-202.
- Dalton-Puffer, C., & Nikula, T. (2006). Pragmatics of content-based instruction: Teacher and student directives in Finnish and Austrian classrooms. *Applied Linguistics*, 27(2), 241-267.
- Farr, F. (2003). Engaged listenership in spoken academic discourse: The case of student-tutor meetings. *Journal of English for Academic Purposes*, 2, 67-85.
- Finder, A. (2005, June 24). Unclear on American campus: What the foreign teacher said. *New York Times*, p. A1.
- Fitch, F., & Morgan, S. (2003). "Not a lick of English": Constructing the ITA identity through student narratives. *Journal of Communication Education*, 52(3 & 4), 297-310.

- Flowerdew, L. (2002). Corpus-based analyses in EAP. In J. Flowerdew (Ed.), *Academic Discourse* (pp. 95-114). Harlow, UK: Longman.
- Flowerdew, L. (2003). A combined corpus and systemic-functional analysis of the problem-solution pattern in a student and professional corpus of technical writing. *TESOL Quarterly*, 37(3), 489-511.
- Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: Countering criticisms against corpus-based methodologies. *English for Specific Purposes*, 24(2005), 321-332.
- Gabel, S. (2001). Over-indulgence and under-representation in interlanguage: Reflections on the utilization of concordancers in self-directed foreign language learning. *Computer Assisted Language Learning*, 14(3-4), 269-288.
- Garcia, P. (2004). Developmental differences in speech and recognition: A pragmatic awareness study. *Language Awareness*, 13(2), 96-114.
- Gavioli, L., & Aston, G. (2001). Enriching reality: Language corpora in language pedagogy. *ELT Journal*, 55(3), 238-246.
- Goffman, E. (1974). *Frame Analysis*. New York: Harper.
- Goldberg, A. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Science*, 7(5), 219-224.
- Granger, S. (1998). The computer learner corpus: A versatile new source of data for SLA research. In S. Granger (Ed.), *Learner English on Computer* (pp. 3-18). New York: Longman.
- Granger, S. (2002). A bird's-eye view of learner corpus research. In S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 2-33). Amsterdam: John Benjamins.
- Granger, S. (2003). The international corpus of learner English: A new resource for foreign language learning and teaching and second language acquisition research. *TESOL Quarterly*, 37(3), 538-546.
- Halliday, M., & Hasan, R. (1985). *Language, Context, and Text: Aspects of Language in a Social-semiotic Perspective*. Oxford, UK: Oxford University Press.
- Halliday, M., & Matthiessen, P. (2004). *An Introduction to Functional Grammar*. London: Arnold.
- Han, C.-H. (2000). *The Structure and Interpretation of Imperatives: Mood and Force in Universal Grammar*. London: Routledge.

- Harwood, N. (2005). 'We do not seem to have a theory...the theory I present here attempts to fill this gap': Inclusive and exclusive pronouns in academic writing. *Applied Linguistics*, 26(3), 343-375.
- Hasselgreen, A. (1994). Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with vocabulary. *Applied Linguistics*, 4(2), 237-260.
- Hasselgreen, A. (2002). Learner corpora and language testing: Smallwords as markers of learner fluency. In S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition, and Foreign Language Teaching* (pp. 143-174). Amsterdam: Benjamins.
- Hawkins, E. (1999). Foreign language study and language awareness. *Language Awareness*, 8(3&4), 124-142.
- He, A. (1993). Exploring modality in institutional interactions: Cases from academic counselling encounters. *Text*, 13, 503-528.
- He, A. (2000). Grammatical and sequential organization of teachers' directives. *Linguistics and Education*, 11(2), 119-140.
- Hopper, P. (1998). Emergent grammar. In M. Tomasello (Ed.), *The New Psychology of Language* (Vol. I, pp. 155-166). Mahwah, NJ: Lawrence Erlbaum.
- Housen, A. (2002). A corpus-based study of the L2-acquisition of the English verb system. In S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 77-116). Amsterdam: John Benjamins.
- Hyland, K. (2002). Directives: Argument and engagement in academic writing. *Applied Linguistics*, 23(2), 215-239.
- Iedema, R. (1996). 'Save the talk for after the listening': The realisation of regulative discourse in teacher talk. *Language and Education*, 10(2&3), 82-102.
- Jenkins, S. (2000). Cultural and linguistic miscues: A case study of international teaching assistant and academic faculty miscommunication. *Intercultural Education*, 24, 477-501.
- Johns, A. (2002). Introduction. In A. Johns (Ed.), *Genre in the Classroom* (pp. 3-13). Mahwah, NJ: Lawrence Erlbaum.
- Johns, T. (1991). Should you be persuaded - two examples of data-driven learning materials. In T. Johns & P. King (Eds.), *Classroom Concordancing. English Language Research Journal* (Vol. 4, pp. 1-16). Birmingham, UK: Birmingham University.

- Johns, T. (1994). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T. Odlin (Ed.), *Perspectives on pedagogical grammar* (pp. 193-312). New York: Cambridge University Press.
- Johnson, K. (1999). *Understanding Language Teaching*. Boston: Heinle and Heinle.
- Kasper, G., & Dahl, M. (1991). Research methods in interlanguage pragmatics. *Studies in Second Language Acquisition*, 13(215-247).
- Kemmer, S., & Barlow, M. (2000). Introduction: A usage-based conception of language. In S. Kemmer & M. Barlow (Eds.), *Usage-based models of language*. Palo Alto, CA: Leland Stanford Junior University.
- Kennedy, C., & Miceli, T. (2001). An evaluation of intermediate students' approaches to corpus investigation. *Language Learning and Technology*, 5(3), 77-90.
- Kinging, C., & Belz, J. (2005). Sociocultural perspectives on pragmatic development in foreign language learning: Microgenetic case studies from telecollaboration and residence abroad. *Intercultural Pragmatics*, 2(4), 369-422.
- Kramsch, C. (1993). *Context and Culture in Language Teaching*. Oxford: Oxford University Press.
- Kruger, J., Epley, N., Parker, J., & Ng, Z. (2005). Egocentrism over e-mail: Can we communicate as well as we think? *Journal of Personality and Social Psychology*, 89(6), 925-936.
- Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.
- Langacker, R. (2000). A dynamic usage-based model. In S. Kemmer & M. Barlow (Eds.), *Usage-based models of language* (pp. 1-64). Palo Alto, CA: Leland Stanford Junior University.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
- Lee, D., & Swales, J. (2006). A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes*, 25, 56-75.
- Leech, G. (1998). Preface. In S. Granger (Ed.), *Learner English On Computer* (pp. xiv-xx). New York: Longman.
- McCarthy, M. (1998). *Spoken Language & Applied Linguistics*. Cambridge, UK: Cambridge University Press.

- McCarthy, M., & Carter, R. (1994). *Language as Discourse: Perspectives for Language Teaching*. London: Longman.
- McEney, T., & Kifle, N. A. (2002). Epistemic modality in argumentative essays of second-language writers. In J. Flowerdew (Ed.), *Academic Discourse* (pp. 182-194). London: Longman.
- Meunier, F. (2002). The pedagogical value of native and learner corpora in EFL grammar teaching. In S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 119-142). Amsterdam: John Benjamins.
- Murphy, R. (2004). *English Grammar in Use*. Cambridge, UK: Cambridge University Press.
- Nesselhauf, N. (2004a). Learner corpora and their potential for language teaching. In J. Sinclair (Ed.), *How to Use Corpora in Language Teaching* (Vol. 12). Amsterdam: John Benjamins.
- Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgements. *Journal of Personality and Social Psychology*, 35(4), 250-256.
- O'Keeffe, A., McCarthy, M. J. & Carter, R. A. (2007). *From Corpus to Classroom*. Cambridge, UK: Cambridge University Press.
- Oliver, J. (2005). *Graduate Enrollment in Science and Engineering Programs Up in 2003*. Washington, DC: National Science Foundation.
- Open Doors. (2006). *Open Doors 2006*. Retrieved April 27, 2007, from <http://opendoors.iienetwork.org/>. New York: Institute of International Education.
- Palmer, F. (2001). *Mood and Modality*. Cambridge, UK: Cambridge University Press.
- Petch-Tyson, S. (1998). Writer/Reader visibility in EFL written discourse. In S. Granger (Ed.), *Learner English On Computer* (pp. 107-118). New York: Longman.
- Pica, T., & Barnes, G. (1990). *Teaching Matters: Skills and Strategies for International Teaching Assistants*: Heinle & Heinle.
- Rayson, P., & Garside, R. (2000). *Comparing corpora using frequency profiling*. Paper presented at the Association for Computational Linguistics, Hong Kong.
- St. John, E. (2001). A case for using a parallel corpus and concordancer for beginners of a foreign language. *Language Learning and Technology*, 5(3), 185-203.

- Sarkisian, E. (1997). *Teaching American Students: A Guide for International Faculty and Teaching Assistants in Colleges and Universities*. Cambridge, MA: Derek Bok Center for Teaching and Learning.
- Schiffrin, D. (1988). *Discourse Markers*. Cambridge, UK: Cambridge University Press.
- Schmitt, N. (Ed.). (2004). *Formulaic Sequences*. Amsterdam: John Benjamins.
- Scollon, R., & Scollon, S. W. (1995). *Intercultural Communication*. Oxford, UK: Blackwell.
- Searle, J. (1976). A taxonomy of illocutionary acts. In K. Gunderson (Ed.), *Minnesota Studies in the Philosophy of Science: Volume VII, Language, Mind, and Knowledge* (pp. 344-369). Minneapolis: University of Minnesota Press.
- Searle, J. (1983). *Intentionality*. Cambridge, UK: Cambridge University Press.
- Seidlhofer, B. (2002). Pedagogy and local learner corpora: Working with learning-driven data. In S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 213-233). Amsterdam: John Benjamins.
- Sfard, A. (1998). On two metaphors for learning and on the dangers of choosing just one. *Educational Researcher* 27, 4-13.
- Simpson, R. C., Briggs, S. L., Ovens, J., & Swales, J. M. (2002). *The Michigan Corpus of Academic Spoken English*. Ann Arbor, MI: The Regents of the University of Michigan.
- Simpson-Vlach, R., & Leicher, S. (2006). *The MICASE Handbook*. Ann Arbor, MI: University of Michigan Press.
- Sinclair, J. (2004). *Trust the Text: Language, Corpus and Discourse*. London: Routledge.
- Smith, R., Byrd, P., Nelson, G., Barrett, R., & Constantinides, J. (1992). *Crossing Pedagogical Oceans: International Teaching Assistants in U.S. Undergraduate Education*. Washington, DC: ERIC.
- Swales, J. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge, UK: Cambridge University Press.
- Thonus, T. (1999). How to communicate politely and be a tutor, too: NS-NNS interaction and writing center practice. *Text*, 19, 253-279.
- Thorne, S. & Lantolf, J. (2007). A linguistics of communicative activity. In S. Makoni & A. Pennycook (eds.), *Disinventing and Reconstituting Languages* (pp. 170-195). Clevedon, UK: Multilingual Matters.

- Thorne, S., Reinhardt, J., & Golombek, P. (forthcoming). Mediation as objectification in the development of professional discourse: The case of international teaching assistants. In J. Lantolf & M. Poehner (Eds.), *Sociocultural Theory and the Teaching of Second Languages*. London: Equinox.
- Tognini-Bonelli, E. (1996). *Corpus Theory and Practice*. Pescia, Italy: Tuscan Word Centre.
- Tomasello, M. (1998). Introduction: A cognitive-functional perspective on language structure. In M. Tomasello (Ed.), *The new psychology of language* (Vol. I, pp. vii-xxiii). Mahwah, NJ: Lawrence Erlbaum.
- Tomasello, M. (2003). *Constructing a Language: A Usage-based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
- Tribble, C. (2001). Corpora and corpus analysis: New windows on academic writing. In J. Flowerdew (Ed.), *Academic Discourse*. Harlow: Addison-Wesley Longman.
- Upton, T., & Connor, U. (2001). Using computerized corpus analysis to investigate the textlinguistic discourse moves of a genre. *English for Specific Purposes*, 20, 313-329.
- US English. (2007). *U.S. English*. Retrieved May 4, 2007, from <http://www.usenglish.org>
- Waara, R. (2004). Construal, Convention, and Constructions in L2 Speech. In M. Achard & S. Niemeier (Eds.), *Cognitive Linguistics, Second Language Acquisition, and Foreign Language Teaching*. Berlin: Walter De Gruyter.
- Widdowson, H. (1978). *Teaching Language as Communication*. Oxford, UK: Oxford University Press.
- Wikipedia. (2007). *Wikipedia: Foreign-born Engineers and Scientists*. Retrieved April 26, 2007, from <http://en.wikipedia.org>
- Williams, J. (2005). Writing Center Interaction: Institutional discourse and the role of peer tutors. In K. Bardovi-Harlig & B. Hartford (Eds.), *Interlanguage Pragmatics* (pp. 37-65). Mahwah, NJ: Lawrence Erlbaum.
- Yates, L. (2005). Negotiating an Institutional Identity: Individual Differences in NS and NNS Teacher Directives. In K. Bardovi-Harlig & B. Hartford (Eds.), *Interlanguage Pragmatics* (pp. 67-97). Mahwah, NJ: Lawrence Erlbaum.
- Yook, A., & Albert, R. (1999). Perceptions of international teaching assistants: The interrelatedness of intercultural training, cognition, and emotion. *Journal of Communication Education*, 48(1), 1-17.

Zhu, Y. (1996). Modality and Modulation in Chinese. In M. Berry, C. Butler & R. Fawcett (Eds.), *Meaning and Form: Systemic Functional Interpretations* (pp. 183-209). Westport, CT: Greenwood.

## Jonathon Sean Reinhardt

---

jonreinhardt@yahoo.com

### Education

---

2007	Ph.D. in Applied Linguistics, The Pennsylvania State University
1992	M.A. in Applied Linguistics/TESOL, University of Illinois at Chicago
1989	B.A. in German, University of Illinois at Urbana-Champaign

### Academic Positions

---

Fall 2007	Assistant Professor of Foreign Language Education Southern Connecticut State University, New Haven, CT
2006-present	Co-director of the Internet Technologies for Advanced Foreign Language Proficiencies project Center for Advanced Language Proficiency Education and Research, University Park, PA
2002-2007	Graduate Assistant The Pennsylvania State University, University Park, PA
2001-2002	Language Technology Consultant Colby College, ME
1998-2001	Fellow in Comparative Culture Miyazaki International College, Miyazaki, Japan
1995-1997	English-as-a-Second-Language Instructor Tutorium in Intensive English University of Illinois at Chicago
1992-1994	English-as-a-Foreign-Language Instructor M.I.L. The Language Center, Funabashi, Japan

### Selected Publications

---

- Reinhardt, J. (2007). "Negotiating meaningfulness: an enhanced perspective on interaction in computer-mediated foreign language learning environments". In S. Magnan (Ed.), *Mediated Discourse Online*. Amsterdam: John Benjamins.
- Thorne, S., Reinhardt, J., & Golombek, P. (2007). "Mediation as objectification in the development of professional discourse: The case of International Teaching Assistants". In J. Lantolf & M. Poehner (Eds.), *Sociocultural Theory and the Teaching of Second Languages*. London: Equinox.
- Belz, J. A. and Reinhardt, J. (2004). "Aspects of advanced foreign language proficiency: Internet-mediated German language play". *International Journal of Applied Linguistics* 14 (3), pp. 324-362.
- Reinhardt, J. and Nelson, B. K. (2004). "Instructor use of online CALL resources: A survey of socio-institutional and motivational factors". *ReCALL* 16 (2), pp. 292-307.
- Isbell, K. and Reinhardt, J. (2000). "Web integration: A model for task-based learning". In E. Hanson-Smith (Ed.), *Technology-Enhanced Language Environments: Case Studies in TESOL Practice*, pp. 45-55. Alexandria, VA: TESOL.