

北京外国语大学

博士学位论文

中文题目 基于语料库的英语指令类言语行为
间接性变异研究

外文题目 A corpus-based study of variation in
indirectness across English directive
speech acts

姓 名	周顾盈
学 号	202120300074
导 师	许家金 教授
研究方向	语料库语言学
专 业	外国语言学及应用语言学
系 别	中国外语与教育研究中心

2025 年 5 月 2 日

**A corpus-based study of variation in
indirectness across English directive speech acts**

ZHOU Guying

**A dissertation submitted in conformity with the requirements
For the degree of Doctor of Philosophy
In Linguistics and Applied Linguistics**

**Supervised by Professor XU Jiajin
National Research Center for Foreign Language Education
Beijing Foreign Studies University**

2 May 2025

摘要

间接表述广泛存在于日常交流中，体现了人类语言的复杂性和语境依赖性。在语言哲学和语用学领域，因字面主义与语境主义的对立，催生了对言语行为的间接性作为二元还是级差概念的争论，也引发了学界对间接性语境制约因素的探讨。

本研究聚焦指令类言语行为间接性的变异情况。以往间接言语行为研究多基于内省或实验数据，因数据规模小，结论难以泛化。基于语料库的间接言语行为研究多数只关注包含显性标记的语例，而忽视非规约化间接表述。有鉴于此，本研究从 2014 版“英国国家语料库”口语子库（Spoken BNC2014）中抽取了 40 个小时英语口语语料样本，人工识别出所有指令类言语行为例证，共计 2,333 条。在此基础上，本研究尝试应对经典语料库方法难以有效召回言语行为的难题，并结合言内与言外共选特征，系统考察指令类言语行为的间接性变异机制。本研究试回答三个问题：1) 在对词汇、句法和语境特征依赖程度不同的各类模型中，何种模型能更有效地识别指令类言语行为？2) 哪些言内因素可以揭示指令类言语行为的间接性变异？3) 哪些言外因素可以解释指令类言语行为的间接性变异？

为回答第一个研究问题，本研究先后采用模式匹配、特征工程、大语言模型三种识别技术，具体为：1) 基于词汇、句型的检索；2) 基于上千个词汇、句法、语义特征和浅层语境信息的支持向量机（SVM）建模；3) 基于微调后的 GPT-4o 大语言模型。结果显示，形式特征检索的召回率仅为 53.74%；最大化利用形式特征进行学习和预测的 SVM 模型召回率提高至 72.70%，而进一步引入浅层语境特征后其召回率提升至 75.00%；微调 GPT-4o 后的新模型将召回率提高至 85.51%。由此可见，语境因素在指令类言语行为的识别中不可或缺。通过深层语境建模并结合监督微调的大语言模型，能够最有效地识别指令类言语行为。

针对问题二和问题三，本研究进一步聚焦请求、建议和指示三类指令类言语行为，采用概率语境共选路径分析其间接性变异。在言内共选层面，运用凝聚层次聚类方法尝试对不同指令表达的语言形式进行归并，探索其如何反映间接程度。在言外共选层面，通过显著性检验方法考察社会语境变量对指令类言语行为表述类型及其语符长度的影响，从而揭示其在间接性制约机制中的作用。

基于聚类结果,本研究尝试构建了三维认知理论模型,以认知场景(cognitive scenario)、行事视角(illocutionary perspective)和言语行为长度(speech act length)衡量言内间接性变异。该模型的中心由认知场景的核心(CORE)阶段、第二人称行事视角和简短表达三个维度协同界定,代表最直接的指令形式。由模型中心向外扩展,间接性逐步增强,表现为认知场景变化(从核心向事前、事后转移)、行事视角转变(从“你”变为“我们”、“我”或非人称)或表达长度增加。

显著性检验结果表明,以往文献所述的九个社会语境因素对请求、建议和指示这三类言语行为的表述存在显著影响,但作用方式各异。部分因素仅在特定言语行为类型上表现显著,部分因素主要影响表述类型或长度,另一些因素在不同言语行为类型上则呈现相反的影响。仅任务难度这一变量在三类指令类言语行为中的作用方式一致,即随着任务难度增加,言语行为的表述类型和长度均指向更高的间接程度。说话人的人口学统计特征(性别、教育程度和年龄)对不同类型的指令类言语行为影响各异。相较于男性,女性在请求中更倾向于间接表述类型,而在指示中更直接,表达建议时使用的表述更长。教育程度较高的个体在建议和指示中更倾向于采用更长的表述,在请求中偏向于更间接的表述类型。年龄对请求和建议长度的影响方式大致相反。社会群体因素(对话者数量、亲密程度、熟悉程度、社会阶层和权势关系)同样影响各异。对话者数量增加时,请求和建议的表述更短,但对指示无显著影响。在关系疏远或等级分明的语境中,建议的表述更长;在亲密程度较低或不太熟悉的语境中,指示的表述更长;然而,这些因素对请求的影响并不显著。

本研究在方法论上的突出贡献在于利用大语言模型及其微调技术,突破传统语料库研究对形式检索的依赖,实现指令类言语行为在自然语料中更有效的识别。同时,研究为语境主义提供实证支撑,表明意义不可约减至词汇和句法,而是依赖更广泛的语境。基于真实语言使用数据,研究摒弃传统的直接-间接二元分类,构建衡量间接性程度的实证框架,并探讨社会语境因素对间接性变异的影响。在解决言语行为识别的关键难题后,未来研究可进一步扩展数据规模,整合多层次语境因素,深化对指令性言语行为及其间接性的理解。

关键词: 指令类言语行为, 间接性, 语料库语言学, 语用学, 大语言模型

Abstract

Indirectness is ubiquitous in everyday communication, reflecting the complexity and context-dependency of human language. In the philosophy of language and pragmatics, literalism and contextualism have shaped discussions on (in)directness. Scholars debate on whether (in)directness should be viewed as a binary or a gradable notion. These theoretical debates have also given rise to investigations into the socio-cultural factors that influence degrees of (in)directness.

The current study examines variation in (in)directness across directive speech acts. Previous research on indirect speech acts has largely relied on introspective or experimental data, making it difficult to generalize findings due to the limited sample sizes. Corpus-based studies on indirect speech acts, despite offering broader empirical coverage, have mostly focused on cases with explicit linguistic markers, overlooking the non-conventionalized ones. To address this gap, this study extracts 40 hours of spoken English data from the Spoken BNC2014, manually identifying 2,333 instances of directive speech acts. Building on this dataset, the study seeks to overcome the limitations of traditional corpus-based research, which relies on form-based retrieval and struggles to capture illocutionary force. Besides, by exploring both intra- and extra-linguistic features, this study systematically examines the mechanisms underlying the variation in (in)directness within directive speech acts. Specifically, it aims to answer the following three research questions: 1) Among models that rely on lexical, syntactic and contextual features to different degrees, which is most effective in identifying directive speech acts? 2) How can intra-linguistic features reveal the variation in (in)directness across directive speech acts in English? 3) How can extra-linguistic factors account for the variation in (in)directness across directive speech acts?

To address the first research question, this study employs three techniques: 1) lexico-grammatical pattern matching, 2) SVM modeling with feature engineering, incorporating thousands of lexical, syntactic, semantic and shallow contextual features, and 3) a fine-tuned GPT-4o model. It turns out that form-based search achieves a recall rate of only 53.74%. The SVM model, which maximally exploits formal features for learning and prediction, improves recall to 72.70%, and further incorporating shallow contextual features raises it to 75.00%. The fine-tuned GPT-4o model further increases recall to 85.51%. The results highlight an indispensable role of context in identifying

of directive speech acts. Deep contextual modeling, combined with supervised fine-tuning in large language models, proves to be the most effective approach for identifying directive speech acts.

To address the second and third research questions, this study examines three subtypes of directive speech acts – requests, suggestions and instructions, adopting a probabilistic contextual co-selection approach. Intra-linguistically, agglomerative hierarchical clustering is used to group different linguistic forms of directives, investigating how they reflect varying degrees of (in)directness. Extra-linguistically, significance tests are conducted to examine the impact of socio-contextual variables on the groups and length of directive speech acts, elucidating their role in influencing the choice of (in)directness.

Based on the clustering results, this study proposes a three-dimensional cognitive model, which integrates cognitive scenario, illocutionary perspective and speech act length, to capture intra-linguistic variation in (in)directness. At the heart of this model lies the CORE phase in the directive scenario, characterized by *you*-perspective and concise forms. Together, they represent the most direct form. Moving outward from this center, indirectness gradually increases, manifesting in: 1) a transition in cognitive scenario (from CORE to BEFORE or AFTER); 2) a shift in illocutionary perspective (from *you* to *we*, *I* or *impersonal* constructions); and 3) an increase in utterance length.

Significance tests reveal that the nine socio-contextual factors mentioned in previous studies significantly influence the expressions of requests, suggestions and instructions, though their effects vary. Some factors are significant only for specific speech act types, some primarily affect expression groups or length, while others exert opposite influences across different directive types. Only imposition level exhibits a consistent effect across all three, with higher imposition levels leading to greater indirectness and increased speech act length. The demographic characteristics of the speaker (gender, education level and age) affect directive speech acts in distinct ways. Compared to males, females tend to use more indirect expressions in requests, while adopt more direct expressions in instructions, and produce longer formulations in suggestions. Individuals with higher education levels favor longer expressions in suggestions and instructions, while preferring more indirect formulations in requests. Age-related effects on suggestion and request length show opposite trends. Social interactional variables (the number of speakers, affect, familiarity, status and power asymmetry) also show distinct effects. When the number of speakers involved in

conversation increases, requests and suggestions become shorter, while instructions remain unaffected. In distant or hierarchical interactions, suggestions tend to be longer. In lower-intimacy or less-familiar contexts, instructions are expressed with greater length. However, these factors do not significantly impact requests.

This study advances speech act research by using large language models and fine-tuning techniques to overcome the limitations of traditional form-based retrieval, enabling more effective identification of directive speech acts in natural discourse. It also provides empirical support for contextualism, demonstrating that meaning cannot be reduced to lexico-grammatical features alone but relies on broader context. Drawing on authentic corpus data, this study moves beyond the traditional direct-indirect dichotomy, introducing a framework that conceptualizes (in)directness as a gradable notion, examining how socio-contextual factors influence its variation. Having addressed key challenges in speech act identification, future research can expand the dataset, integrate multi-level contextual factors, and further refine the understanding of directive speech acts and (in)directness.

Keywords: directive speech acts, indirectness, corpus linguistics, pragmatics, Large Language Models

Contents

北京外国语大学学位论文原创性声明和使用授权说明	ii
致谢	iii
摘要	v
Abstract	vii
Contents	x
Figures	xiv
Tables	xvi
Chapter 1 Introduction	1
1.1 Research background	1
1.2 Aims, objectives and research questions	5
1.3 Significance of the research	7
1.4 Organization of the dissertation	8
Chapter 2 Theoretical Framework for Speech Acts and Indirectness	12
2.1 Key notions in defining speech acts	13
2.1.1 Performatives	13
2.1.2 Illocutionary force	14
2.1.3 Felicity conditions	16
2.2 Speech act taxonomies	18
2.2.1 Austin's and Searle's classifications	18
2.2.2 Bottom-up taxonomies	20
2.2.2.1 <i>From DAMSL to ISO24617-2</i>	20
2.2.2.2 <i>SPAAC and DART</i>	28
2.2.3 Towards an operational definition of <i>directives</i>	33
2.3 Indirect speech acts	38
2.3.1 Primary and secondary illocutionary act	38
2.3.2 Literal force hypothesis and its criticisms	41
2.3.3 (In)directness as a syntax-pragmatics interface problem	44
2.3.4 Accounting for indirect speech acts	45
2.3.4.1 <i>Motivations for indirectness</i>	45
2.3.4.2 <i>Contextual variables</i>	48

2.4 Literalist and contextualist perspectives on (in)directness.....	53
2.4.1 Where does illocutionary force arise?	53
2.4.2 Indirectness: a binary or gradable notion?	57
2.5 Summary	62
Chapter 3 Corpus and Computational Approaches to Speech Act	65
3.1 Speech act and dialog act	65
3.2 Speech act analysis in corpus linguistics	70
3.2.1 Corpus pragmatics.....	70
3.2.2 Form-to-function approach	74
3.2.3 Function-to-form approach	76
3.3 Dialog act classification task in natural language processing.....	81
3.3.1 Machine learning techniques.....	82
3.3.2 Deep learning techniques	85
3.4 Summary	90
Chapter 4 Methodology	93
4.1 Research workflow.....	93
4.2 Corpus and data preparation.....	95
4.2.1 The Spoken BNC2014	95
4.2.2 Data sampling and sentence segmentation.....	98
4.2.3 Manual identification of directives	104
4.2.3.1 <i>Guidelines for recognizing directives</i>	105
4.2.3.2 <i>Inter-rater reliability</i>	111
4.2.4 The dataset dassBNC2014	112
4.3 Retrieving directives with different automated methods	115
4.3.1 Lexical-search	115
4.3.2 Support Vector Machine (SVM)	117
4.3.3 LLM-based retrieval.....	119
4.4 Analyzing directives through a co-selection approach.....	122
4.4.1 Co-selection approach to language	122
4.4.2 Clustering by intra-linguistic features	124
4.4.3 Exploring extra-linguistic influence.....	128
Chapter 5 Retrieving directive force	131

5.1 A landscape of directives in spoken English	132
5.2 Retrieval via lexical and syntactic search (baseline 1).....	137
5.2.1 Feature set for lexical and syntactic search.....	137
5.2.2 Precision, recall, accuracy and F2-score	140
5.2.3 Analysis of retrieval failures	144
5.3 Retrieval through feature engineering (baseline 2).....	149
5.3.1 Feature set for SVM classifiers	149
5.3.1.1 <i>Lexical features</i>	149
5.3.1.2 <i>Syntactic features</i>	152
5.3.1.3 <i>Semantic feature</i>	154
5.3.1.4 <i>Contextual features</i>	155
5.3.2 Precision, recall, accuracy and F2-score	156
5.3.2.1 <i>Baseline 2-1</i>	157
5.3.2.2 <i>Baseline 2-2</i>	161
5.3.3 Analysis of retrieval failures	165
5.4 Retrieval based on contextual representation (baseline 3).....	171
5.4.1 Prompt and format of training data	171
5.4.2 Precision, recall, accuracy and F2-score	176
5.4.2.1 <i>Baseline 3-1</i>	176
5.4.2.2 <i>Baseline 3-2</i>	179
5.4.3 Analysis of retrieval failures	183
5.5 Benchmark analysis of different retrieval methods.....	190
Chapter 6 Variation across situational directives.....	193
6.1 The case of <i>Request</i>	194
6.1.1 Definition of <i>request</i>	194
6.1.2 Intra-linguistic variation.....	195
6.1.2.1 <i>Agglomerative hierarchical clustering</i>	195
6.1.2.2 <i>Patterns of request</i>	204
6.1.3 Extra-linguistic variation.....	207
6.1.3.1 <i>Sixfold classification and extra-linguistic variables</i>	207
6.1.3.2 <i>Speech act length and extra-linguistic variables</i>	212
6.2 The case of <i>Suggest</i>	215
6.2.1 Definition of <i>suggest</i>	215

6.2.2 Intra-linguistic variation.....	216
6.2.2.1 Agglomerative hierarchical clustering.....	216
6.2.2.2 Patterns of suggestion.....	225
6.2.3 Extra-linguistic variation.....	229
6.2.3.1 Sixfold classification and extra-linguistic variables.....	229
6.2.3.2 Speech act length and extra-linguistic variables.....	232
6.3 The case of <i>Instruct</i>	239
6.3.1 Definition of <i>instruct</i>	239
6.3.2 Intra-linguistic variation.....	240
6.3.2.1 Agglomerative hierarchical clustering.....	240
6.3.2.2 Patterns of instruction.....	246
6.3.3 Extra-linguistic variation.....	249
6.3.3.1 Fourfold classification and extra-linguistic variables.....	249
6.3.3.2 Speech act length and extra-linguistic variables.....	251
6.4 Summary.....	255
Chapter 7 Discussion.....	258
7.1 Retrieving illocutionary force between form and function.....	258
7.2 Recategorizing (in)directness based on intra-linguistic features.....	264
7.3 Profiling (in)directness with extra-linguistic factors.....	269
Chapter 8 Conclusion.....	277
8.1 Major findings.....	277
8.1.1 LLM as a holistic approach for directive retrieval.....	277
8.1.2 A three-dimensional model for intra-linguistic variation in (in)directness.....	279
8.1.3 Multi-level contextual constraints on (in)directness across directives	281
8.2 Implications.....	284
8.3 Limitations.....	286
8.4 Outlook.....	287
References.....	289
Appendices.....	309

Figures

Figure 2.1 Evolution of the dialog act annotation schemes	24
Figure 2.2 The ISO24617-2 taxonomy.....	27
Figure 2.3 The SPAAC taxonomy	30
Figure 2.4 Positions from literalist to contextualist according to Recanati (2003)...	54
Figure 3.1 DA classification architecture as elaborated by Durran, Battle & Smith (2021).....	86
Figure 4.1 Research workflow for the dissertation	93
Figure 4.2 Metadata in the transcription for the audio recording SHSD	96
Figure 4.3 Screenshot of transcribed utterances for the audio recording SHSD.....	98
Figure 4.4 A boxplot for average token per file across three different units	103
Figure 4.5 ISO 24617-2 classification of general-purpose communicative functions in conversation	106
Figure 4.6 Directive ratios across different relationship groups	113
Figure 4.7 Workflow for prompt-based learning and full-model fine-tuning	120
Figure 4.8 The probabilistic contextual co-selection approach to meaning.....	123
Figure 5.1 An overview of different methods for directive retrieval	131
Figure 5.2 Categorization of retrieved directive speech acts in dassBNC2014	133
Figure 5.3 Heatmap of performance metrics and directive ratio across files in test set (lexical search)	143
Figure 5.4 Precision-recall curve for SVM classifier 1	158
Figure 5.5 Heatmap of performance metrics and directive ratio across files in test set (SVM classifier 1)	160
Figure 5.6 Precision-recall curve for SVM classifier 2	162
Figure 5.7 Heatmap of performance metrics and directive ratio across files in test set (SVM classifier 2)	164
Figure 5.8 Heatmap of performance metrics and directive ratio across files in test set (GPT-4o).....	178
Figure 5.9 Heatmap of performance metrics and directive ratio across files in test set (GPT-4o-Directives).....	182
Figure 6.1 Plot of Silhouette Score curve in request-clustering.....	196
Figure 6.2 Hierarchical clustering dendrogram of requests	197

Figure 6.3 Heatmap of request groups and gender of speakers.....	208
Figure 6.4 Heatmap of request groups and age range of speakers	209
Figure 6.5 Heatmap of request groups and education level of speakers	210
Figure 6.6 Heatmap of request groups and imposition level.....	211
Figure 6.7 Scatter plot of logged word count of requests and number of speakers	213
Figure 6.8 Boxplot of logged word count of requests and age group	214
Figure 6.9 Boxplot of logged word count of requests and imposition level	215
Figure 6.10 Plot of Silhouette Score curve in suggestion-clustering	217
Figure 6.11 Hierarchical clustering dendrogram of suggestion	218
Figure 6.12 Violin plot of suggestion groups and number of speakers.....	230
Figure 6.13 Heatmap of suggestion groups and imposition level	231
Figure 6.14 Scatter plot of logged word count of suggestions and number of speakers	233
Figure 6.15 Boxplot of logged word count of suggestions and gender.....	234
Figure 6.16 Boxplot of logged word count of suggestions and age group.....	234
Figure 6.17 Boxplot of logged word count of suggestions and education level	235
Figure 6.18 Boxplot of logged word count of suggestions and social status	236
Figure 6.19 Boxplot of logged word count of suggestions and affect level.....	237
Figure 6.20 Boxplot of logged word count of suggestions and power asymmetry.	237
Figure 6.21 Boxplot of logged word count of suggestions and imposition	238
Figure 6.22 Plot of Silhouette Score curve in instruction-clustering	240
Figure 6.23 Hierarchical clustering dendrogram of instructions.....	242
Figure 6.24 Heatmap of instruction groups and speaker gender.....	250
Figure 6.25 Heatmap of instruction groups and imposition level	251
Figure 6.26 Boxplot of logged word count of instructions and education level	253
Figure 6.27 Boxplot of logged word count of instructions and familiarity.....	253
Figure 6.28 Boxplot of logged word count of instructions and relationship affect.	254
Figure 6.29 Boxplot of logged word count of instructions and imposition level....	255
Figure 7.1 Summary of the retrieval experiments.....	260
Figure 7.2 A three-dimensional cognitive model of indirectness in directive speech acts.....	265

Tables

Table 2.1 Austin’s and Searle’s analyses of a complete speech act into its constituent acts.....	15
Table 2.2 A comparative overview of Austin’s and Searle’s taxonomies.....	18
Table 2.3 A comparative overview of dialog act annotation schemes.....	21
Table 2.4 Direct and indirect requests to <i>pass the salt</i>	39
Table 3.1 A summary of research on speech acts using the function-to-form approach	77
Table 4.1 Recording metadata in the Spoken BNC2014	96
Table 4.2 Speaker metadata in the Spoken BNC2014.....	97
Table 4.3 An overview of the sampled data from the Spoken BNC2014	100
Table 4.4 An overview of three units based on different segmentation approaches	103
Table 4.5 Inter-rater reliability between chief annotator and 12 assistant annotators	112
Table 4.6 Distribution of annotated sub-utterances as directives in dassBNC2014	113
Table 4.7 Overview of training, test and validation set of dassBNC2014.....	114
Table 4.8 Intra-linguistic features for clustering	125
Table 4.9 Extra-linguistic variables	128
Table 5.1 Distribution of directive speech acts across subtypes.....	136
Table 5.2 Feature set for lexical and syntactic search	138
Table 5.3 Confusion matrix for lexical-search-based retrieval on test set.....	140
Table 5.4 Precision, recall, accuracy and F2-score for lexical search across files in test set	142
Table 5.5 Confusion matrix for SVM classifier 1 on test set	157
Table 5.6 Precision, recall, accuracy and F2-score for SVM classifier 1 across files in test set.....	159
Table 5.7 Confusion matrix for SVM classifier 2 on test set	161
Table 5.8 Precision, recall, accuracy and F2-score for SVM classifier 2 across files in test set.....	163
Table 5.9 Number of misclassified sub-utterances by SVM classifier 1 and 2	165
Table 5.10 Confusion matrix for the performance of GPT-4o base model on test set	176

Table 5.11 Precision, recall, accuracy and F2-score for GPT-4o base model across files in test set.....	177
Table 5.12 Fine-tuning configuration	179
Table 5.13 Confusion matrix for the performance of fine-tuned GPT-4o model on test set	180
Table 5.14 Precision, recall, accuracy and F2-score for GPT-4o-Directives across files in test set.....	181
Table 5.15 Number of misclassified sub-utterances by GPT-4o and GPT-4o-Directives	184
Table 5.16 Performance of different methods on the directive retrieval task.....	190
Table 6.1 Subcategorization of imperatives in requestive speech acts.....	198
Table 6.2 Subcategorization of <i>you</i> -expressions in requestive speech acts.....	200
Table 6.3 Subcategorization of <i>I</i> -expressions in requestive speech acts	202
Table 6.4 Subcategorization of <i>we</i> -expressions in requestive speech acts	203
Table 6.5 Patterns of requests	204
Table 6.6 Statistical tests for request groups and extra-linguistic variables.....	207
Table 6.7 Statistical tests for requests length and extra-linguistic variables	212
Table 6.8 Subcategorization of <i>you</i> -imperatives in suggestive speech acts	219
Table 6.9 Subcategorization of <i>let's</i> -imperatives in suggestive speech acts.....	220
Table 6.10 Subcategorization of <i>you</i> -expressions in suggestive speech acts	221
Table 6.11 Subcategorization of <i>we</i> -expressions in suggestive speech acts.....	223
Table 6.12 Patterns of suggestions	225
Table 6.13 Statistical tests for suggestion groups and extra-linguistic variables	229
Table 6.14 Statistical tests for suggestion length and extra-linguistic variables	232
Table 6.15 Patterns of instructions	246
Table 6.16 Statistical tests for instruction groups and extra-linguistic variables	249
Table 6.17 Statistical tests for instruction length and extra-linguistic variables	252
Table 7.1 An overview of extra-linguistic influence on forms of requests, suggestions and instructions	271
Table 8.1 Summary of extra-linguistic influence on instructions, suggestions and requests.....	282