

Table of contents

List of figures	12
List of tables	14
1 Introduction	17
1.1 Studies Overview	21
1.2 Motivation	24
2 Related Work	33
2.1 Explicit and Implicit Prosody	33
2.2 Keystrokes	37
2.3 Dialogue	48
3 Research Questions	54
4 Data Collection Methodology	56
4.1 Demographics	58
4.2 Summary of Collected Conversation Data	61
4.3 Prolific Crowdsourcing	63
4.4 Experimental design details	64
4.5 Keystroke Collection	71
4.6 Engineered Features	72

Table of contents	10
5 Study 1: Keystroke patterns in dialogue acts	79
5.1 Methodology	81
5.2 Results	86
5.3 Discussion	96
5.4 Future work	103
6 Study 2: The relationship between keystrokes and sentiment	107
6.1 Related Work: Sentiment and dialogue	112
6.2 Methodology	114
6.3 Results	120
6.4 Discussion	128
6.5 Future work	135
7 Study 3: Predicting rapport levels based on keystroke patterns	138
7.1 Related Work: Rapport	141
7.2 Methodology	146
7.3 Results	154
7.4 Discussion	163
7.5 Future work	168
8 Overall Discussion and Future Work	170
References	184
Appendix A IRB Detail	205
Appendix B Prolific Advantages and Disadvantages	208
Appendix C Prosodic Parallels in Speech and Typing	210
Appendix D Experimental iterations	213

Table of contents	11
Appendix E Study 1 details	217
Appendix F Manual Sentiment Analysis Guidelines	226

List of figures

1.1	A comparison of the scope of each of my studies	21
2.1	A set of features available from extracted keystroke timing	40
4.1	A map of all 3 phases of my data collection experiment	57
4.2	The overall age distribution of participants	59
4.3	The average enjoyment rating between partners, as a function of the age difference between them.	60
4.4	Distribution of conversation lengths, as measured by the total number of messages sent by both participants.	62
4.5	Distribution of message lengths, as measured by the number of (whitespace- delimited) words.	62
4.6	The average word count for each role type, as role relates to providing or receiving recommendations.	63
4.7	A screenshot of the experimental chat interface	65
4.8	The distribution of responses to survey questions	71
5.1	The distribution of high-level dialogue act categories within my collected dataset .	82
5.2	The distribution of word counts within each dialogue act	91
5.3	Word counts within each recommendation role	106
6.1	Different turn types within a dialogue	110

6.2	The distribution of the number of turns per conversation.	115
6.3	The distribution of manual and algorithmic sentiment scores	117
6.4	The nonlinear functions that define each predictor in the optimal keystroke model .	121
6.5	The correlation of scores between each opinion question	127
7.1	Subsets of conversational data for Study 3	148
7.2	Distribution of rapport ratings for each cluster	149
7.3	PCA of 2 versus 3 clusters	150
7.4	Predictive confidence for each classifier	155
7.5	The distributions of metric scores for the full dataset versus a random subset	157
7.6	Variable importance for the full dataset and random subset	158
7.7	The distributions of metric scores for the first half subset versus the second half subset	160
7.8	Variable importance for the first half subset and second half subset	161
7.9	The distributions of metric scores for the Provider subset and Receiver subset . . .	162
7.10	Variable importance for the Provider subset and Receiver subset	163
8.1	A toy example of different typing patterns visualized with the same lexical content	177
A.1	The advertisement for the experiment, sent out by Prolific to potential participants.	206

List of tables

1.1	A comparison of keystrokes, speech, and text	28
2.1	Categories of keystroke features, along with possible parallels in speech production.	42
4.1	Employment status of participants	60
4.2	Schema for keystroke database	72
4.3	Schema for message database	73
4.4	Individual keystroke-level features	74
4.5	Features added to each word, which is a space-delimited character segment.	75
4.6	Features added to each full transmitted message	76
4.7	Features calculated for a participant overall.	77
4.8	Features calculated for an entire conversation.	78
5.1	Mapping of dialogue act classifications	84
5.2	Utterance count and examples for mapped DAs	85
5.3	Results of predicting dialogue act binaries using different timing metric covariates	89
5.4	Results of whether dialogue acts alone can predict word count	92
5.5	<i>F</i> scores measuring the influence of changes in dialogue acts and changes in word counts on keystroke timing response variables	93
5.6	Significant features from Exp. 1a	96
5.7	Significant features from Exp. 1b	97

5.8	Dialogue acts with significant keystroke timing differences as compared to the grand mean	98
6.1	The sentiment of the current turn and following turn	111
6.2	Occurrence count and features of different turn types	115
6.3	The importance of (smoothed) predictors in the optimal keystroke model	119
6.4	The model results for four prediction tasks	122
6.5	The influence of each individual predictor on the outcome of the combined models	124
6.6	The effects of each opinion rating on the timing of the respective keystroke features	125
7.1	Definitions of rapport and possible quantification by keystroke patterns	143
7.2	A list of features used in Study 3	152
7.3	Model and dataset comparison, arranged by predictive certainty	154
7.4	Full metrics for the full dataset versus random subset	156
7.5	Full metrics for the first half subset and second half subset	159
7.6	Full metrics for the Provider subset versus Receiver subset	162
C.1	A comparison of parallel features in spoken prosody and keystroke dynamics (continued in Table C.2)	211
C.2	A comparison of parallel features in spoken prosody and keystroke dynamics (continued from Table C.1)	212
E.1	Effects of adding covariates to a model predicting non-opinion vs opinion binary dialogue acts	220
E.2	Results of whether dialogue acts controlling for word count can predict the speed at which the utterance was produced.	221
E.3	Results of whether dialogue acts can predict edit counts, controlling for word count.	222
E.4	Results of whether dialogue acts predict variation in typing speed, when controlling for word count.	223

-
- E.5 Results of whether dialogue acts can predict pre-utterance gaps, when controlling
for word count, and word count + who the last sender was. 224
- E.6 Results of whether dialogue acts can predict the typing speed of word 1, word 2,
and the length of time between words 1 and 2. 225