A failure mode and effects analysis of pragmatic errors in learner e-mails

Allan Nicholas[1*] and *John* Blake[1]

[1]Center for Language Research, University of Aizu, Aizu-Wakamatsu, Japan.

Abstract. E-mailing in the English is an important mode of communication for Japanese learners of computer science both during and after their studies. A key element of effective e-mail communication relates to pragmatic competence—understanding the relationship between social context and language choices and accordingly adapting one's language in an appropriate manner. Despite this importance, there have been few learner e-mail corpora annotated for pragmatic errors. As part of a larger ongoing study developing online tools for assessing and developing learners' e-mail writing ability, a failure mode and effects analysis was conducted on a learner corpus of e-mails, manually tagged for pragmatic errors. Tagged errors were ascribed a value between 1-5 in terms of severity, detectability and frequency, with most sever errors assigned a score of 5. Weighted priority scores were calculated, allowing for errors to be prioritized in terms of importance. Preliminary results indicate severe errors are associated with failing to address the face needs of the e-mail recipient, violating pragmatic norms, with potential negative consequences in terms of relationship maintenance. Results allow for the creation of a job queue for the software developer, and usefully inform teaching priorities in the language classroom.

1 Introduction

E-mailing in the English L2 is a key mode of communication for Japanese students in higher education institutions, due to the need to communicate effectively with non-Japanese faculty members [1]. While alternative modes are becoming increasingly common in the workplace, such as Slack [2], e-mail remains a common communication tool in academia [3] [4] and important for receiving timely feedback from faculty [5]. Furthermore, e-mail remains a frequently used tool in many workplaces; with an increasing need to operate in global networked environments and teams, the ability to communicate effectively via e-mail in future potential workplace contexts is an important skill.

Broadly, errors within e-mail texts may be grammatical-syntactical, factual, or pragmatic in nature. All three categories of error may have an adverse impact on the recipients' perceptions of the e-mail sender. Grammatical and factual errors can be relatively easily agreed upon, identified and addressed by educators. However, the relationship between context and language — the pragmatic aspect of e-mail communication is more challenging, due to the role of subjectivity and awareness of L2 pragmatic norms.

As part of a larger ongoing study of learner e-mails and pragmatic competence, this current study investigates the application of a failure mode and effects analysis of pragmatic errors in a learner corpus of e-mail text data, evaluating the severity, detectability and

frequency of errors. Implications for software development and classroom practice are discussed.

2 Background

2.1 Pragmatics and pragmatic errors

Pragmatics is "the societally necessary and consciously interactive dimension of the study of language" [6, p.315]. Pragmatics is an aspect of communicative competence addressing the complex relationship between our language choices and the social scenarios in which we interact with others. Failure to attend to the pragmatic norms of an L2 interaction can lead to negative social consequences for the learner [7], and/or perceptions of "rudeness" [4].

While the identification of pragmatic failure is, to an extent, a subjective endeavour, and the need to avoid overly prescriptive approaches to pragmatics classroom instruction important [7] [8], the potential negative consequences of pragmatic failure point to a clear need to raise learner awareness in this area [4]. Studies typically employ native English users to evaluate L2 learner e-mail texts [5] and assess faculty perceptions of students' e-mail communications [3] [4] [5] [9].

2.2 Learner corpora and pragmatic annotation

Use of corpora allows for a systematic approach to identifying patterns, common features or errors in large

---

* Corresponding author: anich@u-aizu.ac.jp

collections of text data. Corpora have been used for a wide variety of purposes, with the majority employed in the identification of features in the formal aspects of language use, such as grammatical patterns and lexis. Pragmatic annotation of corpora is less common [10] [11], often focusing on oral communication and speech acts, such as requesting, apologizing or thanking in telephone conversations [12].

While a number of corpora have annotated for learner errors, these have been non-pragmatic (the Cambridge Learner Corpus; the Longman Learner's Corpus). Automatically annotating pragmatic features or errors has proved difficult, due to a lack of information regarding patterns of such features or errors [13]. Few corpora, therefore, have annotated for pragmatic errors, due to the time and resource-intensive nature of manually annotating large amounts of text data.

2.3 Learner e-mail and pragmatic errors

E-mails received by university faculty from students vary greatly in terms of appropriacy. Figure 1 shows an illustrative example of a learner e-mail to university faculty at a Japanese higher education institution.
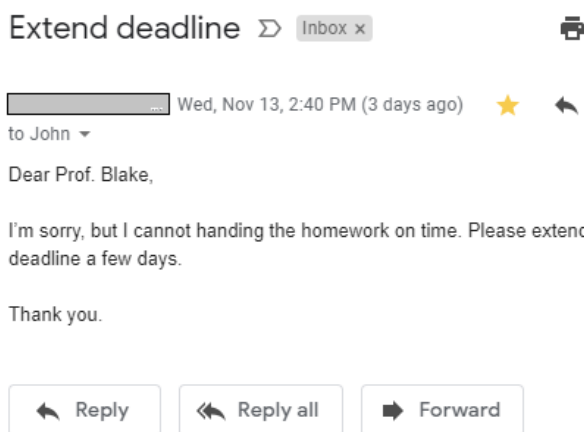

Fig. 1. Screenshot of email with sender details redacted

3 Method

3.1 Corpus specifications

The target users were undergraduate students at a computer science university in Japan, aged 18-21 years. The institution has identified e-mailing in the English L2 as a key task students should be able to perform [1], due to approximately 40% of faculty members being non-native Japanese.

Request-based e-mails were chosen as the target e-mail type, due to their prevalence in e-mail communications [4] [5] and their inherently face-threatening nature [14]. Learners of English as a Foreign Language (EFL) have been found to struggle with constructing pragmatically appropriate e-mail requests [5] [9], indicating a clear need for learners in this regard.

Target e-mail scenarios were primarily those in an academic context. However, to explore learners' abilities to adapt their e-mail writing to a variety of contexts, private social communications and e-mails to business persons outside of academia were also identified as useful task types. Rather than collect authentic e-mail data, for this study, e-mail texts were elicited via in-classroom tasks. This allowed for control of the contextual variables *power* (P; akin to relative social status between interlocutors), *distance* (D; how well the interlocutors know each other) and *imposition* (R; how potentially troublesome the request may be to the receiver), identified by [14].

3.2 Corpus creation

To create task scenarios for data elicitation, an exemplar generation questionnaire was administered to a sample of the student population (*n=108*), eliciting examples of situations in which they had needed to make a request in their daily lives, with a particular interest in academic scenarios. A count of the results was carried out, with scenarios ranked in terms of frequency. The most frequent scenarios were then used as templates for task creation, with additional scenarios added to ensure a variety of social contexts and challenges for students. The task items were then drafted, before being moderated by expert English users to ensure the tasks elicited the target (requesting), the economy of the task language, and that appropriate task responses could only be obtained through knowledge of L2 pragmatic norms. In addition, moderators were asked to evaluate the values P, D and R for each task scenario, to ensure they matched the P, D and R values assigned by the researchers. For each variable, two values could be assigned, as shown in Table 1.

Table 1. Contextual variables Power, Distance and Imposition (adapted from [15])

P	Meaning	D	Meaning	R	Meaning
+	Receiver has a higher rank, title or social position. E.g. president, supervisor	+	Sender and receiver do not know, or identify with, each other.	+	Great expenditure of goods or energy by the receiver to carry out the request.
-	Receiver has a lower rank or social position. E.g. salesperson serving a customer.	-	Sender or receiver know, or identify with, each other.	-	Small expenditure or energy by the receiver to carry out the request.

Following these steps, four e-mail tasks were selected for administration to students (*n=426*). The tasks were chosen based on their differing combination of P, D and R values, with students needing to adapt

their language choices accordingly in order to write pragmatically appropriate texts. Table 2 shows the four task scenarios administered, and their assigned values. Tasks were administered to the participants in both the L1 and the L2 to ensure understanding, with responses required in the English L2.

Tasks were administered in the classroom using Google Forms. A total of 1,474 e-mail texts were elicited, across the four task scenarios.

Table 2. Administered task scenarios and assigned contextual values.

Task Scenario	P	D	R
You are organizing a university event in which local businesses' products are showcased to the public. To help fund the event, you are contacting local business people to ask for financial donations. Email Mr. Smith- a local business owner in Aizu-Wakamatsu- to ask for a financial donation. You do not know Mr. Smith.	+	+	+
You have a close friend who is an international student at university. You want to practice your English conversation skill next week, so you email him/her asking if they can spend five minutes of their time talking with you in English.	-	-	-
You need to go to Sendai for an academic conference next week, but the train there is too expensive. You email your friend (who has a car) asking them to drive you there. It takes about three hours to drive from your home to Sendai by car. Your friend will be busy next week, so this will be inconvenient for him/her.	-	-	+
You must submit a document to the local government office in Aizu-Wakamatsu proving that you are a student at the University of Aizu. Email the manager of the Student Affairs Office at the university asking them to provide you with the document you need.	+	+	-

3.3 Tagset and schema creation

The initial tagset was adapted from the Cross-Cultural Study of Speech Act Realization Patterns (CCSARP) framework [16], developed to identify features of speech acts in conversation, and used in adapted form by [5] to analyze learner e-mail texts. An inter-annotator reliability check was carried out, in which the tagset was further developed and adapted to suit the needs of non-expert annotators who may be unfamiliar with academic-oriented pragmatics tagsets, and to reflect the way we typically read an e-mail text, from opening to closing. An excerpt from the tagset that annotates use is given below.

H2A The marker "please"
H2B consultative devices ("would you mind" "do you think")
H2C Subjectivisers ("I'm afraid" "I wonder")

H2D Downtoners ("possibly" perhaps" "just")
H2E Hedges ("a bit" "a little" "sort of")
H2F Cajolers ("You know" " You see…")
H2G Appealers ("will you…Prof. Johnson?")

In the annotation label H indicates that this annotation relates to the sentence that contains the main request (head request). The annotations focus on content that should have been included. This type of annotation is problematic as the annotator has to judge what should be present but is not. An extract of the annotation schemata is shown in Figure 2.
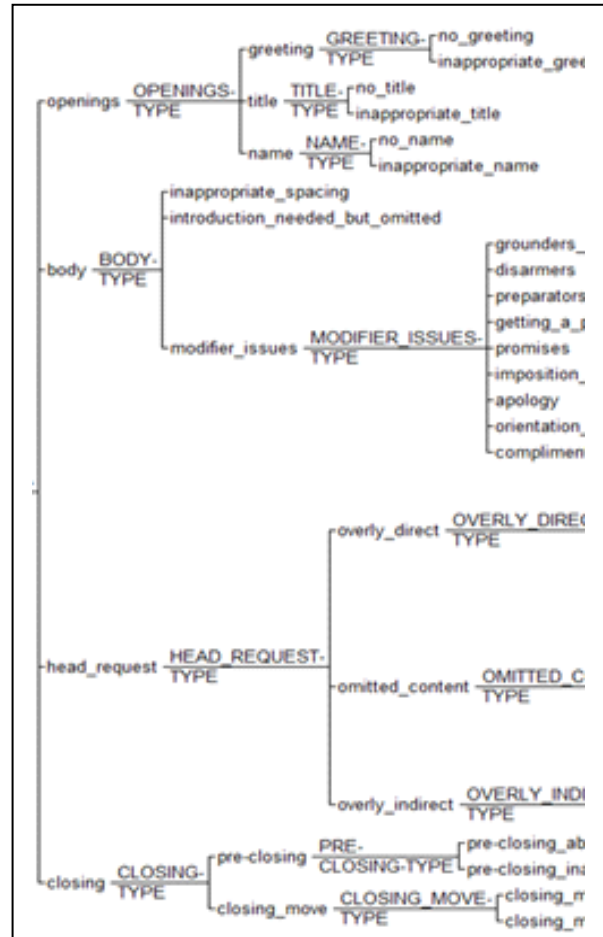

Fig. 2. Annotation schemata

4 Failure mode effects analysis

4.1 Introduction to FMEA

Failure Mode Effects Analysis (FMEA) is a quality management process to investigate points of failure and evaluate the effects of failure [17]. In this project, FMEA analysis is applied to pragmatic errors, and so the severity, frequency and detectability are evaluated.

The sum of three values of severity (S), detectability (D) and frequency (F) gives a priority score for an error. Each score category is multiplied by a coefficient to create a weighted priority score (WP), as shown in equation (1). The weighted priority provides a mechanism to take into account the relative importance of each category.

$$\alpha S \cdot \beta D \cdot \gamma F = WP \qquad (1)$$

where:

S= severity α=coefficienct of severity
D= detectability β=coefficient of detectabilty
F= frequency γ=coefficient of frequency
WP= weighted priority

To conduct the FMEA, annotated errors were extracted from the corpus and collated into an error bank. Each error was ascribed a value of between 1 and 5 for the three criteria of severity, detectability and frequency. The coefficients of each criterion were initially set at 1, meaning that the range of weighted priority score was from 0 to125.

4.2 Severity

Severity was judged by the researchers based on the perceived effect on the reader of the email. Errors judged to be more likely to create negativity in the recipient were judged as more severe. Errors that were judged as severe received a score of 5 while non-serious errors received 1.

4.3 Detectability

Detectability was judged by a software developer who based the value on the difficulty of automatically identifying the error with precision [18]. Precision is calculated by dividing the number of true positive results by the sum of the number of true and false positive results as shown in Equation (2).

$$P = \frac{TP}{TP+FP} \qquad (2)$$

where:

P = precision
TP = number of true positive results
FP = number of false positive results

False positive results are those that are detected incorrectly while false negative results are those that should have been detected but were not. In short, the more likely the occurrence of false positive and false negative results, the lower the precision score and the lower the detectability score. The detectability score correlates with the precision score but also takes account of the difficulty in creating the expression to match the error. In this category, a score of 5 means that the error is easy to automatically identify.

4.4 Frequency

The frequency of each error type was counted in the pilot corpus of 50 email texts. The likelihood of the same error type occurring in the full corpus was estimated. A score of 5 was awarded for errors that were frequent and 1 for errors that occurred rarely.

As an illustrative example, professors frequently receive requests to extend the deadline for assignments from their students. Should an email begin with "I want you to extend the deadline," it would be unlikely to put the professor in a favorable mood. In this example, the student failed to employ strategies to soften the request, such as providing a reason. In this example, the key problematic phrase is "I want you". When an FMEA analysis is applied to this pragmatic error, the resultant values can be tabulated as shown in Table 3.

Table 3: Example FMEA score for pragmatic error

Error type	S	D	F	WP*
I want you	4	5	3	60

* The coefficient weighting is set at 1

5 Results

The expected main outcome was achieved. The job queue for the software developer can be created by sorting the pragmatic errors in descending order by weighted priority. The software developer needs to write the program to automatically match the errors using rule-based parsing. The cost-benefit of the task decreases as the weighted priority decreases, and so this is an effective way to maximize productivity.

Preliminary results from the pilot study (see Table 4) show that errors allocated the highest scores in severity failed to address the face needs of the e-mail recipient by, for example, failing to include the recipient's title in the opening, or the sender's name in the closing. Pragmatic failure relating to the head act are also categorized as high in severity, with a lack of internal modifiers or use of 'want' or 'need' statements being perceived as overly direct. This lack of adherence to pragmatic conventions has the potential to lead to difficulties in the sender-receiver relationship.

Among the most frequent errors, many also scored highly for detectability. Examples of these include a lack of the recipient's name in the opening or a lack of a closing salutation. On the other hand, other frequent errors, such as a lack of external modifiers in the email body, scored lower for detectability. Errors that received the highest scores for detectability leaned towards the lexical side of the lexical-grammatical cline.

Grammatical errors tended to receive medium scores for detectability while functional errors were considered the most difficult due to the complexity of the form-function relationship. Reader intuition is needed to deduce the intended function. This implies that there is a degree of subjectivity and that two readers with differing world knowledge and expectations can interpret the same language form in different ways.

The lack of a one-to-one correlation between form and function makes automatic detection considerably more difficult. A particularly difficult error to detect is the lack of presence of an expected item. Errors of omission vary in detectability depending on whether the words omitted could be considered a closed or open set. Closed sets comprise of a fixed number of items and so can be stored as lists in a detection system, while this is

not possible for open sets. For example, the omission of the term 'please' is simple to detect; however, this does not take into account whether a writer used the slightly archaic 'kindly' instead. The difficulty is exacerbated when there are multiple functional exponents that can be used. Checking the presence or absence of language features using a computer script is therefore a non-trivial task.

Currently, the frequency of error types in the full corpus is estimated based on the count in the pilot study. The errors in the full corpus will be tallied using a tailormade script once the corpus is fully annotated.

Table 4. FMEA scores for pragmatic errors, ranked for severity

Error type	S	D	F	Tally	WP*
No recipient name	5	5	5	28	125
Imperative used	5	4	3	15	60
No title	5	5	2	11	50
Head act- 'want' statement	5	5	2	10	50
No closing salutation /name	4	5	5	36	100
No greeting	4	5	3	14	60
No self-introduction when needed	4	4	2	5	32
No self-introduction because no name in closing	4	4	2	4	32
Head act- 'would like' statement	4	5	1	2	20
Head act- 'need' statement	4	5	1	1	20
No pre-closing	3	3	5	37	45
No or insufficient external modifiers	3	1	5	25	15
Head act- 'would/ could /can'	3	5	2	13	30
Greeting- 'nice to meet…'	3	5	1	2	15
Inappropriate closing	3	5	1	2	15
Inappropriate pre-closing	3	5	1	1	15
Body runs on from greeting	2	5	4	24	40
Greeting- 'hello'	2	5	3	17	30
Head act- performative	2	5	2	4	20
Opening- inappropriate name, office	2	4	1	1	8
No or insufficient internal modifiers	1	2	4	19	8
Greeting- 'dear my friend'	1	5	1	2	5
Greeting- given name	1	5	1	1	5

6 Discussion

Our findings provide useful insights into the types of pragmatic failure in learners' email texts, their frequency, perceived severity, and the potential detectability of such failure types by software. Such instances of failing to adhere to pragmatic norms may potentially lead to perceptions of impoliteness by the recipient of an e-mail text. This is in line with research on the perceptions of grammatical versus pragmatic errors in the business community showing that pragmatic errors in politeness by non-native users of English were more troublesome than grammatical errors [19].

Preliminary findings from this initial pilot study are in line with studies in other contexts [5] that found students' failure to attend to the recipients' face needs, such as employing overly direct requesting strategies, or failing to modify a request with an appropriate e-mail opening or closing, can lead to perceptions of 'rudeness' by the receiver of an email text. It is these errors- ones that most baldly fail to reflect the degree of imposition of a request- that rate highest in severity in the FMEA analysis.

7 Conclusion

This current study describes the process of creating a corpus of English L2 learner e-mails, annotated for pragmatic failure, and a subsequent FMEA analysis of pragmatic errors, evaluating severity, detectability and frequency. This analysis allows the researchers to create a job queue for programming of future error-detecting tools, and provides useful insight into how learners' pragmatic errors are perceived by e-mail recipients.

Results of the pilot study and FMEA analysis also provide useful information not only for the software programmer, but also for classroom practice. Priority weighting scores for error types provide guidance for teachers working under classroom time constraints, who wish to address learners' needs for developing pragmatic competence in their e-mail writing.

Acknowledgments

References

[1] E. Kaneko, M. Park, I. Wilson, Y. Heo, D. Roy, T. Yasuta, A. Nicholas, J. Blake, *English curriculum innovation for computer science majors in the Japanese EFL context: From needs to tasks,* in Proceedings of IEEE International Professional Communication Conference (ProComm), Toronto, ON, pp.84-89 (2018)

[2] C. Durscheid, C. Frehner, *Email communication,* in S. Herring, D. Stein, T. Virtanen (eds.) Handbook of Pragmatics of Computer-Mediated Communication. De Gruyter Mouton, Berlin, pp.35-52 (2013)

[3] C. Felix-Brasdefer, *Email openings and closings: pragmalinguistic and gender variation in learner-instructor cyber consultations.* in E.A. Soler and M. Safont-Jorda (eds.), Discourse and Language Learning Across L2 Instructional Settings. Brill Academic Publishers, Amsterdam/New York (2012)

[4] M. Savic, J. Prag, **124,** 52-72 (2018)

[5] M. Economidou-Kogetsidis, J. Prag. **43**, 13, 3193-3215 (2011)

[6] J. Mey, Pragmatics: An Introduction. Blackwell, Oxford (1993)

[7] V. LoCastro, Pragmatics for language educators: A sociolinguistic perspective, Routledge (2013)

[8] N. Ishihara, A. Cohen, Teaching and Learning Pragmatics: Where Language and Culture Meet. Longman, London (2010)

[9] S. Biesenbach-Lucas, Lang. Learn. Tech. **11**, 2, 59-81 (2007)

[10] C. Maynard, S. Leicher, *Pragmatic annotation of an academic spoken purposes*, in E. Fitzpatrick (ed.), Corpus linguistics beyond the word: Corpus research from phrase to discourse, Rodopi, Amsterdam (2007)

[11] J. Kallen, J. Kirk, SPICE-Ireland: A User's Guide. Clo Ollscoil na Banriona, Belfast (2012)

[12] G. Leech, M. Weisser, in *Generic speech act annotation for task-orientated dialogues'*, in D. Archer, P. Rayson, A. Wilson, and A. McCenery (eds.), Proceedings of Corpus Linguistics, 28-31 March 2003, Lancaster University, United Kingdom, pp.441-446 (2003)

[13] Y. Tono, *Learner Corpora: Design, development and applications.* in D. Archer, P. Rayson, A. Wilson, and A. McCenery (eds.), Proceedings of Corpus Linguistics 28-31 March 2003, Lancaster University, United Kingdom, pp.800-809 (2003)

[14] P. Brown, S. Levinson, Politeness: Some Universals in Language Use, Cambridge University Press, Cambridge (1987)

[15] T. Hudson, E. Detmer, J.D Brown, Developing Prototypic Measures of Cross-Cultural Pragmatics, University of Hawai'i Press, Honolulu (1995)

[16] S. Blum-Kulka, E. Olshtain, App. Ling. **5**, 3, 196-213 (1984)

[17] D.H. Stamatis, Failure mode and effect analysis: FMEA from theory to execution. ASQ Quality press (2003)

[18] D. Olson, D. Dursun, Advanced data mining techniques, Springer Science & Business Media (2008)

[19] J. Wolfe, N. Shanmugaraj, J. Sipe., Bus. Prof. Comm. Quart. **79**, 4, 397-415 (2016)