

Speech Acts in Formal and Informal Interactions in English: Mapping the Past, Exploring the Future

Workshop convenors:

Rachele De Felice (The Open University)
Christine Elsweiler (University of Innsbruck)
Sofia Rüdiger (FU Berlin)

A corpus-pragmatic function-based investigation of questions and answers in Contemporary Spoken British and American English

Dawn Archer
(Manchester Metropolitan University)

Niall Curry
(Manchester Metropolitan University)

Questions and answers play a central role in spoken communication. As a rhetorical "talk-in interaction" device, questions can help speakers to manage turn-taking, demonstrate engagement, and give the communicative "floor" to interlocutors (Curry & Mark 2025). Questions and answers have been studied in casual as well as strategic contexts (the courtroom, classroom, etc.), in historical and modern periods, across real-world activity types, in fictional contexts, and across languages (Archer 2005, 2012; Atkinson & Drew 1979; Bolden et al. 2023; Peñarroja, 2020; Sinclair and Coulthard 1975). Questions in casual spoken communication often occur at the close of turns and, in the case of polar questions, are typically followed by mitigating devices ('well', 'I mean', 'I don't know') rather than a yes or no response (Curry & Mark 2025) as a means of softening the discourse. Taxonomically, questions are identified via their subject-verb inversion or rising intonation and are typically identifiable in transcripts through the Illocutionary force indicating device (IFID) of the question mark (Flöck & Geluykens 2015). Answers are then (normally) identified as second-pair parts to questions (but see Sinclair and Coulthard 1975). Studies of questions and answers tend to adopt a manual or semi-automatic approach to identification. Limited attention has been paid, in the extant literature, to indirect questions and answers in spoken language, however, despite their evident value, in English, as a less face-threatening approach to conversation management and interlocutor engagement. This paper is part of a small body of work interested in (in)validating an automatic approach to the effective retrieval of such phenomena (see, e.g., Landert et al 2023; Jucker 2024). We investigate questions and answers in spoken English conversation, using comparable American and British corpora. With regard to the paper's methodological contribution, we first assess the extent to which (in)direct questions can be found automatically using Curry's (2021; 2023) IFID approach on the two corpora. We then draw upon Archer's (2005) taxonomy of answer types – a taxonomy that

was applied manually for semi-automatic interrogation – to determine whether it is possible to effectively categorize responses to these IFIDs. This allows us to confirm the roles of questions and answers; including the extent to which some answers can carry “(one or more) of several illocutionary force(s) at any given time, and still function as an answer”, as Archer (ibid: 290) claims. The work is thus designed to further develop our theoretical understanding of answers (in relation to questions) as well as enabling us to assess the benefits and limitations of automatic pragmatic annotation (Archer et al. 2008; Lu 2014). An additional theoretical advantage is that the inter-varietal analysis, contrasting both American and British varieties, and the intra-varietal analysis, documenting regional variation within American and British contexts, offers insight into evident lacunae in the wider literature on casual spoken language, i.e., the shared and differing questioning and answering practices in casual spoken conversation in American and British Englishes.

References

- Archer, D. (2005). *Questions and Answers in the Historical Courtroom (1640-1760)*. Amsterdam and Philadelphia: John Benjamins.
- Archer, D. (2012). ‘Assessing Garrow’s *aggressive* questioning style’. In: G. Mazzon (ed.). *English Historical Dialogue Studies*. Milano: FrancoAngeli, pp.301-320.
- Archer, D., J. Culpeper & M. Davies (2008c). Pragmatic annotation. In A. Lüdeling & M. Kytö (eds.) *Corpus Linguistics: An International Handbook*. Volume 1. Mouton de Gruyter, pp. 613-641.
- Atkinson, J.M. and P. Drew (1979). *Order in Court: The Organisation of Verbal Interaction in Judicial Settings*. London: Macmillan.
- Bolden, G.B., J. Heritage and M-L. Sorjonen (2023). *Responding to Polar Questions Across Languages and Contexts*. Amsterdam and Philadelphia: John Benjamins.
- Curry, N. (2021). *Academic writing and reader engagement: Contrasting questions in English, French and Spanish corpora*. London: Routledge. <https://doi.org/10.4324/9780429322921>
- Curry, N. (2023). Question illocutionary force indicating devices in academic writing: A corpus-pragmatic and contrastive approach to identifying and analysing direct and indirect questions in English, French, and Spanish. *International Journal of Corpus Linguistics*, 28(1), 91-119. <https://doi.org/10.1075/ijcl.20065.cur>
- Curry, N. & G. Mark (accepted). Applications of corpus linguistics in language education: teacher, editor, and assessment developer perspectives. *Research Notes*.
- Flöck, I., & R.I.E. Geluykens (2015). Speech acts in corpus pragmatics: A quantitative contrastive study of directives in spontaneous and elicited discourse. In J. Romero-Trillo (Ed.), *Yearbook*

of corpus linguistics and pragmatics 2015 (pp. 7–37). London: Springer.
https://doi.org/10.1007/978-3-319-17948-3_2

Jucker, A. (2024). *Speech Acts: Discursive, Multimodal, Diachronic*. Cambridge: Cambridge University Press.

Landert, D., D. Dayter, T.C. Messerli & M.A. Locher (2023). *Corpus Pragmatics*. Cambridge: Cambridge University Press.

Lu, X. (2014). *Computational Methods for Corpus Annotation and Analysis*. Netherlands: Springer.

Peñarroja, M.R. (2020). *Analysing the Pragmatics of Speech Acts in Sitcom and Drama Audiovisual Genres*. Newcastle: Cambridge Scholars.

Sinclair, J. & M. Coulthard (1975). *Towards and Analysis of Discourse: The English Used by Teachers and Pupils*. Oxford: Oxford University Press.

Automatic Prediction of Speech Acts

Sophia Conrad
(University of Zurich)

Tilia Ellendorff
(University of Zurich)

Gerold Schneider
(University of Zurich)

The corpus of spoken Irish English, SPICE Ireland [1] has been manually annotated for the five speech act categories of illocutionary speech acts, following [4], namely *representatives*, *directives*, *commissives*, *expressives*, and *declaratives*. Using this annotated dataset as training data, we use state-of-the-art machine learning approaches to predict speech acts.

We address the following four research questions (RQ):

1. How well does a classical document classifier perform on the task of speechact classification?
2. To what extent does class imbalance affect classification performance, and what strategies can mitigate the overrepresentation of large classes (e.g. *representatives*) compared to smaller classes (e.g. *declaratives*)?
3. Do distributional semantics approaches, such as word and sentence embeddings, improve speech act classification compared to traditional wordbased methods?
4. Can large language models (LLMs) outperform traditional machine learning approaches in speech act classification?

Our level of annotation are the turns, as given in SPICE Ireland. The key features used in our approach include the vectorized text of the current turn, as well as the preceding and following turns for context. Additionally, we incorporate register information at three levels as given by the corpus. For some experiments, we also include the gold label of the previous speech act.

Concerning RQ1, we report preliminary results on two baselines, namely logistic regression (given in Table 1, “Baseline F-Score”) and a rule-based system.

Performance on small classes is considerably worse than on large classes (see RQ2). In order to boost underrepresented classes, we thus test oversampling strategies. To this end, we employ GPT-4 [2] as a data augmentation tool to create further instances of small classes (marked with * in Table 1), providing the model with some examples from the corpus and prompting it to generate

Class	Class Description	Frequency	Baseline Score	F- Score	Augmented F-Score
rep	Representatives	35246		0.84	0.84
dir	Directives	10522		0.57	0.44
icu	Indet.communicativeunits	3782		0.76	0.27
xpa	Notanalysable	2513		0	0.26
exp*	Expressives	1242		0.01	0.17
com*	Commissives	685		0.04	0.56
soc*	Socialgreetings	173		0.39	0.08
dec*	Declaratives	57		0	0.02
P	Micro-AverageAccuracy	54220		0.72	0.73

Table 1: Comparison of Baseline System to Augmented System

similar turns. The performance of a system using a more balanced sample and parameter tuning is given in Table 1, “Augmented F-Score”). There is only improvement in some classes while partially impairing the performance of the larger classes.

Speech acts are pragmatic categories and thus often expressed very indirectly. Accordingly, turning to RQ3, we can expect on the one hand that approaches including distributional semantics (ranging from word embeddings via document embeddings, e.g. Roberta or snowflake, to GPT-4 and other large language models) leverage contextual information. On the other hand, pragmatics as the art of reading between the lines and understanding situational knowledge may still remain a challenging task for automatic systems which cannot understand texts in a human sense [3]. Preliminary results comparing a parameter-tuned (but not data augmented) system, show that we could increase accuracy from 0.76 to 0.79 with *word2vec* word embedding, indicating that embeddings partly manage to include situational knowledge.

References

- [1] Jeffrey L Kallen and John Kirk. *SPICE-Ireland: A user’s guide*. Cl’o Ollscoil na Banr’iona, 2012.
- [2] OpenAI et al. *GPT-4 Technical Report*. 2024. arXiv: 2303.08774[cs.CL]. url: <https://arxiv.org/abs/2303.08774>.
- [3] Gerold Schneider. “Automatically detecting Directives with SPICE-Ireland”. In: *Sociopragmatic Variation in Ireland: Using Pragmatic Variation to Construct Social Identity*. Trends in Linguistics. Studies and Monographs. Berlin: De Gruyter Mouton, 2024.
- [4] John R Searle. *Toward a linguistic theory of speech acts*. 1976.

Apology conventions in business-related correspondence from the 18th century to the present day

Christine Elswailer
(University of Innsbruck)

Rachele De Felice
(The Open University)

PDE apologies are highly routinised speech acts which are typically realised explicitly by the IFID *sorry* (Deutschmann 2003). However, it is also possible to apologise indirectly through other strategies such as taking on responsibility for an offense or by combining apology strategies (Holmes 1990: 167; Aijmer 1996: 82–84). In contrast to Present-Day English, the speaker-oriented IFID *sorry* was less prominent than addressee-oriented forms such as (*I pray you*) *pardon me* in Late Modern English and combinations of indirect apology strategies are often found. Apologies were, moreover, less routinised and more complex (Jacobsson 2004; Jucker & Taavitsainen 2008; Jucker 2018). Nevertheless, in genres such as letter-writing, certain conventional apology patterns are found, for instance, the recurrent *trouble* formula, e.g., *yet I cannot but giue you this trouble*. In addition to these general trends, the choice of apology strategies also depends on the type of offense that is being apologised for, e.g., time offenses such as being late or talk offenses such as interrupting someone (Aijmer 1996: 109; Deutschmann 2003: 62, Thaler & Elswailer 2023: 234) as well as the communicative setting (Thaler & Elswailer 2023).

In this paper, we address apologies in the communicative setting of British English business-related correspondence from a diachronic perspective. Our goals are

(1) to examine if offense types in business-related correspondence, for instance, slowness in replying or length of the letter/e-mail, are stable from the Late Modern period through to the present day,

(2) to explore which apology strategies and combinations of strategies are chosen for the different offense types and how these are formally realised across the periods under investigation.

Our data are drawn from different correspondence corpora and editions spanning the 18th to the 21st century, totaling between 100 and 150 letters per century. For the 18th century, they comprise Scottish letters on estate business from the *Helsinki Corpus of Scottish Correspondence, 1540–1750* (ScotsCorr), the correspondence of the Scottish philosophers Adam Ferguson and Adam Smith with their book-sellers, printers and agents, as well as business-related letters from the *Corpus of Early English Correspondence Extension Sampler* part 1 and 2 (CEECES1 & CEECES2). The 19th- and 20th-century data include letters from the *British Telecom Correspondence Corpus* (BTCC; Morton and Nesi 2019) and the professional letters sub-component of the 1994 *British National Corpus* (BNC1994; Burnard 2000). For the 21st century, we look at the publicly available emails recently released as part of the Post Office Horizon Inquiry

(<https://www.postofficehorizoninquiry.org.uk/>). The apologies are retrieved combining a form-to-function approach with close reading. Each instance is categorised for both formal features (i.e. strategies used) and offence types (cf. Ancarno 2005, De Felice 2024, Elswailer 2024). Previous research on other datasets has found that minor offences centred around correspondence- and communication-related issues are the most frequent categories in emails and that they do not typically require complex apology strategies (Harrison and Allton 2013, Marsden 2019, De Felice 2024). Our work will compare whether these findings persist across different datasets and time periods and can therefore be considered a constant feature of correspondence.

References

- Aijmer, Karin. 1996. *Conversational Routines in English: Convention and Creativity*. London/New York: Longman.
- Ancarno, Clyde. 2005. “The Style of Academic E-mails and Conventional Letters: Contrastive Analysis of Four Conversational Routines. *IBÉRICA* 9: 103–122.
- Burnard, Lou. 2000. *Reference Guide for the British National Corpus*. Oxford: Oxford University Press.
- CEECES 1 = *Corpus of Early English Correspondence Extension Sampler* part 1. 2021. Compiled by Terttu Nevalainen, Helena Raumolin-Brunberg, Samuli Kaislaniemi, Mikko Laitinen, Minna Nevala, Arja Nurmi, Minna Palander-Collin, Tanja Säily and Anni Sairio at the Department of Languages, University of Helsinki. XML conversion by Lassi Saario. Helsinki: VARIENG.
- CEECES 2 = *Corpus of Early English Correspondence Extension Sampler* part 2. 2022. Compiled by Terttu Nevalainen, Helena Raumolin-Brunberg, Samuli Kaislaniemi, Mikko Laitinen, Minna Nevala, Arja Nurmi, Minna Palander-Collin, Tanja Säily and Anni Sairio at the Department of Languages, University of Helsinki. XML conversion and encoding by Lassi Saario. Helsinki: VARIENG.
- De Felice, Rachele. 2024. “Stability of Pragmatic Markers: The Case of *Sorry* in Organizational Emails from the Clinton Email Corpus”. Paper presented at the *Sociopragmatic Variation in Late Modern English Workshop*, ICAME45, Vigo, Spain.
- Deutschmann, Mats. 2003. *Apologising in British English*. Umeå: Skrifter Från Moderna Språk.
- Elswailer, Christine. 2024. “The Organisation of Macro-requests in Early 18th-century Scottish and English Letters”. Paper presented at the *Sociopragmatic Variation in Late Modern English Workshop*, ICAME45, Vigo, Spain.
- Harrison, Sandra, and Diane Allton. 2013. Apologies in Email Discussions. Susan C. Herring,

- Dieter Stein, Tuija Virtanen (eds.). *Pragmatics of Computer-Mediated Communication*. Berlin: Walter de Gruyter. 315–337.
- Holmes, Janet. 1990. “Apologies in New Zealand English.” *Language in Society* 19: 155–99.
- Jacobsson, Mattias. 2004. “Apologies and Apologetic Attitude in Early Modern English”. *Nordic Journal of English Studies* 3(3): 187–204.
- Jucker, Andreas H., and Irma Taavitsainen. 2008. “Apologies in the History of English: Routinized and Lexicalized Expressions of Responsibility and Regret”. Andreas H. Jucker and Irma Taavitsainen (eds.). *Speech Acts in the History of English*. Amsterdam/Philadelphia: Benjamins. 229–244.
- Jucker, Andreas H. 2018. “Apologies in the History of English: Evidence from the *Corpus of Historical American English* (COHA)”. *Corpus Pragmatics* 2(4): 375–398.
- Marsden, Elizabeth. 2019. *Relationship Management in Intercultural Business Emails*. Doctoral thesis, University of Huddersfield, available at <http://eprints.hud.ac.uk/id/eprint/34902/>.
- Morton, Ralph and Hilary Nesi. 2019. “Institutional Collaboration in the Creation of Digital Linguistic Resources: The Case of the *British Telecom Correspondence Corpus*”. Simon Pople, Andrew Prescott A. and Daniel H. Mutibwa, (eds.). *Communities, Archives and New Collaborative Practices*. Bristol: Policy Press.
- ScotsCorr = *The Helsinki Corpus of Scottish Correspondence 1540–1750*. 2017. Ed. Anneli Meurman-Solin. Helsinki: University of Helsinki. <http://urn.fi/urn:nbn:fi:lb-201411071>.
- Thaler, Marion, and Christine Elsweiler. 2023. “The Role of Gender in the Realisation of Apologies in Local Council Meetings: A Variational Pragmatic Approach in British and New Zealand English”. Special Issue on Gender and Language. Ed. by Anna Islentyeva and Anatol Stefanowitsch. *Zeitschrift für Anglistik und Amerikanistik* 71(3): 217–239.

Letter-Writing and Speech act Analysis: Tracing (Im)politeness Patterns in Historical Irish English

David Sotoca Fernández
(University of Extremadura)

Carolina Amador-Moreno
(University of Bergen)

This paper presents an overview of three different case studies focusing on the use of (im)politeness (Brown & Levinson 1987; Haugh & Culpeper 2018) strategies pertaining to the encoding of requests, apologies and reproaches in historical Irish English. It resorts to a subsection of CORIECOR (Corpus of Irish English Correspondence) (Amador-Moreno 2021) comprising missives exchanged by Irish emigrants that moved to the US and their relatives and close ones from 1700 to 1940. This subsection contains a total of 596 letters that had been annotated manually in order to extract these speech acts systematically. All these investigations focus on intimate discourse as described by Clancy (2005) and use different theoretical frameworks (Archer 2017; Blum-Kulka 1984) within (im)politeness studies to tackle the data at hand. The first case study deals with the use of the mental verb “hope” in three speech acts and observes its historical value as a mitigator among all of them. The case study focuses on reproaches, a speech act that has been so far overlooked and categorized within the notion of critique. The analysis shows the existence of two different kinds of reproaches in the data employing Corpus Linguistics tools to shed light on the relevant lemmas that play a key function in the encoding of either of them. The third case study takes a more traditional approach to the analysis of requests within the data. It uses a combination of Blum-Kulka’s (1984) and Ackermann’s (2023) taxonomy for the analysis of requestive speech acts and categorizes each instance to determine the level of directness and indirectness appearing in this subcorpus when performing this speech act.

The focus of the paper will be on methodological issues regarding the analysis of these phenomena in historical data as illustrated in the three case studies. It concentrates on the value of their results for Irish English as a linguistic variety as well as their limitations and the caveats they present for further research. The paper discusses their potential comparability with previous studies using the same set of data and/or similar corpora, focusing on the value of these pieces for the area of (im)politeness studies and their validity within this field.

References

Ackermann, T. (2023). Mitigating strategies and politeness in German requests. *Journal of Politeness Research*, 19(2), 355–389.

- Amador-Moreno, C. P. (2022). Contact, Variation and Change: Mapping the History of Irish English through CORIECOR. *Nexus*, 2, 49–53.
- Blum-Kulka, S. & Olshain, E. (1984). Requests and apologies: A cross-cultural study of speech act realization patterns (CCSARP). *Applied Linguistics*, 5(3), 196–213.
- Brown, P. and Stephen C. L. (1987) *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
- Clancy, Brian. (2015). *Investigating Intimate Discourse: Exploring the spoken interactions of families, couples and friends*. Routledge Taylor and Francis Group
- Haugh, M. and Culpeper, J. (2018). Integrative Pragmatics and (Im)politeness Theory. In Ilie C. & Norrick N. R. (Eds.) *Pragmatics and its Interfaces* (pp. 213–239). John Benjamins.

**Speech acts, communicative setting and discursive context:
Problems of data and methodology in historical speech act analysis**

Alexander Haselow
(University of Wuppertal)

Julian Häde
(University of Wuppertal)

Johannes Lässig
(University of Wuppertal)

This talk discusses findings deriving from the *Diachrony of Communicative Actions Project* (DiCAP), which is devoted to the empirical, largely qualitative study of the long-term diachrony of – first and foremost – expressive, and commissive speech acts in the history of English, relating historical-linguistic findings to major transformations in the socio-cultural history of Britain. In this project, verbal actions are identified manually through close inspection of historical corpora of English and collected in a database, paying close attention to their sequential context, if available. As a second step, these actions are coded for contextual variables by which they are conditioned, such as the relationship between the intra- or extradiegetic co-participants, or the setting (e.g. *religious, military, private*), and according to functional and formal aspects.

In this talk, we will outline the discursive approach underlying the project (see also e.g. Jucker 2016), which studies speech acts not with reference to isolated utterances taken out of their concrete discourse contexts. Rather, it links linguistic observation to information about the way in which specific forms of discourse unfold, to the setting, and to the relation between the interlocutors. We will illustrate how this approach may tackle the following notorious problems in historical speech act analysis:

- i. Speech acts are licensed by the discursive context and the setting. Hence, the communicative function of an utterance is not transdiscursively generalisable and, in theory, each utterance requires case-by-case treatment.
- ii. Many utterances in historical sources are not embedded in dialogues, yet we have to assume that they carried some pragmatic function, which has to be identified somehow.
- iii. Speech acts may occur in sequence and often serve as vehicles for other speech acts.
- iv. Speech acts are often hybrid entities, combining different illocutions.

We utilize the discursive approach and, where applicable, the next-action proof-procedure as a supplementary analytical tool for objectively identifying potential illocutions. Using EXPRESSIONS OF GRATITUDE, OF FUTURE COMMITMENT and INVOCATION OF HARM, we will show how access to the communicative context helps to understand important, yet underexplored aspects of the long-term developments of individual verbal actions, focusing on:

- a) how the setting by which an utterance may be licensed has changed (e.g. Which factors condition the expression of gratitude, or make it normatively required?, Haselow 2024),
- b) how cultural perceptions of certain utterances have changed (manifesting themselves, e.g. in the responses to or in metapragmatic comments on an action, e.g. Brinton 2021),
- c) how utterances that once indicated the illocutionary point of one action have to come to serve other actions and thus shifted the domain (e.g. when utterances originally used to perform commissive acts shifted into the domain of expressive acts serving politeness; or when thanking shifted from an act of deference to an act that enhances one's own face in the 16th century),
- d) how the utterance forms associated with one speech act have changed over time.

References

- Brinton, Laurel. 2021. Responding to thanks. From *you're welcome* to *you bet*. *Journal of Historical Pragmatics* 22.2, 180–201.
- Jucker, Andreas H. 2016. Politeness in eighteenth-century drama: A discursive approach. *Journal of Politeness Research* 12.1, 95–115.
- Haselow, Alexander. 2024. Politeness, speech acts and socio-cultural change. The expression of gratitude in the history of English. *Journal of Historical Pragmatics* 25.3, 419–449

Speech Act Annotation and the Description of Spoken Registers: A Corpus Study

John Kirk
(University of Vienna)

This paper on pragmatic corpus annotation explores the challenges of speech-act annotation and offers solutions as well as pragmatic characterisations of spoken registers.

Towards this end, the SPICE-Ireland Corpus (Systems of Pragmatic Annotation in the Spoken Component of the ICE-Ireland Corpus) (Kirk et al. 2011; Kallen & Kirk 2012; Kirk 2012; 2016; 2019; Kirk & Andersen 2016) encodes the speech act status of each utterance in the spoken component of the corpus, using a system that is developed from the well-known work of Searle (1976). Searle's taxonomy is designed to illustrate systemic aspects of language use, not to encode actual examples of language in use. Nevertheless, it does provide a realistic basis on which it was possible to build a system of pragmatic annotation that provides for an exhaustive and explicit categorisation of all the diverse material in a corpus.

To Searle's five categories SPICE-Ireland adds four of its own: '**indeterminate conversationally-relevant units**' (ICU), such as feedback responses or signals such as *right*, *yes*, or *ok* which provide conversational coherence but are not uttered with an intended pragmatic function or with any other commitment in the unfolding conversation or discourse, but which are crucial to the development of the ongoing discourse; '**incomplete utterances or fragments**' which are pragmatically indecipherable; '**social expressions**' such as greetings or leave takings; and '**keyings**', following Goffman (1974) for utterances involving humour or irony where speakers are not being literal or felicitous, and where normal conditions of language use do not apply.

No simple algorithm exists for determining the speech act status of an utterance; annotation is made on the basis of detailed and, it must be stressed, manual analysis of language in use. Further annotations identify discourse markers, quotative constructions, utterance tags, intonational units, and stressed syllables, all of which contribute to the identification of any particular speech act in question. The usefulness of the corpus is presented in Aijmer 2018; Archer & Culpeper 2018: §3.1.2; Culpeper & Haugh 2014; Kirk & Ronan 2022 and O'Keeffe 2018.

Among its 300 texts each of 2,000 words and comprising 15 discourse situations, both formal and informal, many spontaneous, mostly interactive but with a few monologues, distributed equally between Northern Ireland and the Republic of Ireland, the SPICE-Ireland corpus has 54,612 speech acts. No study has analysed so many Speech Acts so consistently, across so many types of spoken data, and across two geo-political jurisdictions. The results merit attention and wider application, notwithstanding the few studies specifically on speech acts that are known to us: Jansen & Flöck 2024; Rehbein et al. 2016; Ronan 2015, 2022, Ronan & Elsweiler 2024) (whereas Kirk, 2015, 2018a deal with pragmatic discourse markers, and Kirk 2018b deals with the pragmatics of intonation).

The first research question concerns how just such a speech act annotation scheme is devised. The next research question concerns the distribution across the 15 text types as well as across the two geo-political zones, in terms of raw as well as relativized (or normalized) frequencies and percentages. Exhaustive frequencies will be presented. And there will be some critical discussion about speech acts as a marker of register in spoken interaction.

References

- Aijmer, Karin, 2018. 'Corpus Pragmatics: From Form to Function', in Jucker, Andreas H., Schneider, K.P. & W. Bublitz, eds. *Methods in Pragmatics*. Berlin: de Gruyter Mouton. 555-585.
- Archer, Dawn & Jonathan Culpeper, 2018. 'Corpus Annotation', in Jucker, Andreas H., Schneider, K.P. & W. Bublitz, eds. *Methods in Pragmatics*. Berlin: de Gruyter Mouton. 495-525.
- Culpeper; Jonathan & Michael Haugh. 2014. *Pragmatics and the English Language*. Houndmills: Palgrave.
- Goffman, Erving. 1974. *Frame Analysis: An Essay on the Organization of Experience*. Boston: Northeastern University Press.
- Jansen, Alena & Ilka Flöck. 2024. 'Apologies and Corpus Pragmatics: Comparing a Form-to-Function and Function-to-Form Approach in SPICE-Ireland', in Ronald Geluykens & Ilka Flöck, eds. *Analyzing Pragmatic Variation in English: New Developments in Contrastive, Cross-Cultural and Interlanguage Pragmatics*. München: LINCOM. 247-281.
- Kallen, Jeffrey L. & John M. Kirk, 2012. *SPICE-Ireland: A User's Guide*. Belfast: Cló Ollscoil na Banríona.
- Kirk, John M. 2012. 'Beyond the Structural Levels of Language: An Introduction to the SPICE-Ireland Corpus and its Uses', in Janet Cruickshank and Robert McColl Millar, eds. *After the Storm: Papers from the Forum for Research on the Languages of Scotland and Ulster Triennial Meeting, Aberdeen 2012*. Aberdeen: Forum for Research on the Languages of Scotland and Ireland. 207-232. [available for download at: <http://www.abdn.ac.uk/pfrlsu/volumes/vol4/>]
- Kirk, John M. 2015. 'Kind of and sort of: Pragmatic Discourse Markers in the SPICE-Ireland Corpus', in Carolina P. Amador-Moreno, Kevin McCafferty & Elaine Vaighan, eds. *Pragmatic Markers in Irish English*. Amsterdam & Philadelphia: John Benjamins. 89-113.
- Kirk, John M. 2016. 'The Pragmatic Annotation Scheme of the SPICE-Ireland Corpus'. *International Journal of Corpus Linguistics* 21:3: 299-322.
- Kirk, John M. 2018a. 'The Pragmatics of *Well* as a Discourse Marker in Broadcast Discussions', in Sebastian Hoffmann, Andrea Sand, Sabine Arndt-Lippe & Lisa Maria Dillmann, eds. *Corpora and Lexis*. Leiden: Brill. 140-172.

- Kirk, John M. 2018b. 'The Pragmatics of Intonation: Fall-rise Tones in the SPICE-Ireland Corpus', Paper presented at ICAME39, Tampere, 30 May-3 June.
- Kirk, John M. 2019. 'The SPICE-Ireland Corpus as a Research Resource for the Pragmatics of Irish English', Symposium on New Approaches to Irish English (Corpus) Pragmatics, Technische-Universität Dortmund, 26-27 April.
- Kirk, John M. & Gisle Andersen, 2016. 'Compilation, Transcription, Markup and Annotation of Spoken Corpora'. *International Journal of Corpus Linguistics* 21:3: 291-298.
- Kirk, John M., Jeffrey L. Kallen, Orla Lowry, Anne Rooney & Margaret Mannion. 2011. *The SPICE-Ireland Corpus: Systems of Pragmatic Annotation for the Spoken Component of ICE-Ireland*. Version 1.2.2. Belfast: Queen's University Belfast and Dublin: Trinity College Dublin.
- Kirk, John & Patricia Ronan. 2022. 'Researching Corpus Pragmatics in Irish English', *International Journal of Corpus Linguistics and Pragmatics*, 6.2: 95-99.
- O'Keeffe, Anne, 2018. 'Corpus-based function-to-form approaches', in Andreas H. Jucker, Klaus P. Schneider & Wolfram Bublitz, eds. *Methods in Pragmatics*. Berlin: Mouton de Gruyter. 587-618.
- Rehbein, Ines, Merel Scholman, & Vera Demberg. 2016. 'Annotating Discourse Relations in Spoken Language: A Comparison of the PDTB and CCR Frameworks'. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*: 1039-1046. Portorož, Slovenia: ELRA. Refers to Disco-SPICE (Spoken conversations from the SPICE-Ireland corpus annotated with discourse relations)
- Ronan, Patricia. 2015. 'Categorizing expressive speech acts in the pragmatically annotated SPICE Ireland corpus' *ICAME Journal* 39: 24-45.
- Ronan, Patricia. 2022. 'Directives and Politeness in SPICE-Ireland'. *International Journal of Corpus Linguistics and Pragmatics*, 6.2: 175-199.
- Ronan, Patricia & Christine Elswiler. 2024. 'Directives in SPICE-Ireland and ICE-Scotland from a Cross-Varietal Perspective', paper presented at FRLSU14, Belfast, 29.11.24.
- Searle, J. 1969. *Speech Acts*. Cambridge: Cambridge University Press.
- Searle, J. 1976. 'A Classification of Illocutionary Acts'. *Language in Society* 5: 1-23.

“Pray pardon the hast this is ended in”
A corpus-based analysis of apologies in Early Modern English correspondence

Daniela Landert
(Heidelberg University)

Andreas H. Jucker
(Heidelberg University)

Apologies are usually defined as expressions of regret for a past event for which the apologizer accepts at least some responsibility, but recent work on the history of apologies in English has shown that such definitions need to be treated with care. In the course of time not only the manifestations of apologies have been subject to change but at least to some extent also their functional profile (see Jucker and Taavitsainen 2008; Williams 2018; Jucker 2019). Jucker (2019) proposed the term “attenuation” to account for the weakening of the illocutionary force of apologies over time. Haselow (2024), working on expressions of gratitude, put this into a larger context and suggested the four related processes of recontextualization, functional expansion, attenuation and routinisation to account for the typical development of speech acts across time.

In this contribution, we want to put Haselow’s concepts to the test by a focused analysis of apologies in a corpus of relatively informal interactions in Early English. For this purpose, we use the *Parsed Corpus of Early English Correspondence*, consisting of about 2.1 million words spread over five time periods from 1350 to 1710. In a first step, we extract a carefully stratified sample from the entire corpus and manually annotate all attested examples of apologies. This is expected to retrieve all manifestations of apologies that appear with a reasonable frequency in the entire corpus. In a second step, these manifestations are then searched for with appropriate corpus tools to ascertain their distribution in the entire corpus.

Preliminary results indicate that there is an overall increase of apologies from the late fourteenth to the early eighteenth century with clear evidence for Haselow’s concepts of functional expansion, attenuation and routinisation. While early instances of apologies consist largely of implicit expressions of regret by the apologizer for a past event and their own responsibility for it, later instances show an increasing level of attenuation and routinized formulations including the illocutionary point indicating devices *pardon* and *sorry*.

References

Haselow, Alexander. (2024) Politeness, speech acts and socio-cultural change. The expression of gratitude in the history of English. *Journal of Historical Pragmatics*, <https://doi.org/10.1075/jhp.21005.has>

Jucker, Andreas H. (2019) Speech act attenuation in the history of English: The case of apologies. *Glossa: A Journal of General Linguistics* 4.1, 45. 1-25. DOI: <https://doi.org/10.5334/gjgl.878>

Jucker, Andreas H., and Irma Taavitsainen (2008) Apologies in the history of English: Routinized and lexicalized expressions of responsibility and regret. In: Andreas H. Jucker and Irma Taavitsainen (eds.). *Speech Acts in the History of English*. (Pragmatics & Beyond New Series 176). Amsterdam/Philadelphia: John Benjamins, 229-244.

Williams, Graham. (2018) *Sincerity in Medieval English Language and Literature*. London: Palgrave Macmillan.

Grumbling Through Time: The Metadiscourse of Complaints in Historical Letters and Emails

Sofia Rüdiger
(FU Berlin)

Rachele De Felice
(Open University)

Complaints are ever-present in our interactions, be it when making small talk about the weather (which is too cold), chatting about politics (which we disagree with), or engaging with the waiter in a restaurant about our food (which was too salty). These examples illustrate two main complaint types: 1) other-addressed complaints (the first two cases), which, from a present-day perspective have been shown to be essential in building rapport (Boxer 1993) and eliciting ‘emotional reciprocity’ (Günthner 1997), while being socially-stigmatized at the same time (e.g., Heinemann & Traverso 2009: 2381); and 2) complainee-addressed complaints (the third case), which are considered inherently face-threatening to the hearer (Olshtain & Weinbach 1987: 196). The complex interaction between this speech act and both own- and other-face considerations makes it an interesting case study from the perspective of metadiscourse (i.e., displays of reflective awareness (Haugh 2018) as performed in people’s speech or writing (cf. Jucker 2020)), as speakers navigate the balance between explicitly expressing their stance and maintaining interpersonal harmony. In this article, we take a historical corpus linguistic perspective on complaint metadiscourse, to shed light on the variation of this function at different points in time.

We do this by investigating three datasets: 1) American English letters from the Late Modern English period (a subset of the *Corpus of Early American Literature*, CEAL; Höglund & Syrjänen 2016), 2) British English letters from the Late Modern English period (*British Telecom Correspondence Corpus*; Morton & Nesi 2020), and 3) present-day American English emails (*Clinton Email Corpus*; De Felice & Garretson 2018). We search these corpora for lexical items (and related lemmas and word forms) which were identified as salient in previous research on complaint metadiscourse in personal historical letters (Rüdiger in prep), for example, *complain*, *dissatisfaction*, *grudge*, *grumble*, *murmur*, and *remonstrate*. Based on the patterns of use of these terms, we answer the following research questions: 1) How are these lexical items used - are they IFIDs or do they perform other functions? 2) Are there any differences between the two historical varieties (i.e., British English and American English)? 3) How does this compare to data from the 20th/early 21st century? The study thus allows us to observe variation in the use of these metadiscursive items over time and space. In addition, the article makes methodological contributions by employing a number of under-used corpus resources and by attempting a first comparison of complaint-metadiscourse across the genres of handwritten and electronic correspondence from different time periods.

References

- Boxer, Diana. 1993. "Social Distance and Speech Behavior: The Case of Indirect Complaints." *Journal of Pragmatics* 19: 103-125.
- De Felice, Rachele & Gregory Garretson. 2018. "Politeness at Work in the Clinton Email Corpus: A First Look at the Effects of Status and Gender." *Corpus Pragmatics* 2: 221-242.
- Günthner, Susanne. 1997. "Complaint Stories: Constructing Emotional Reciprocity among Women." In: Helga Kotthoff & Ruth Wodak, eds. *Communicating Gender in Context*. Amsterdam: John Benjamins. 179-218.
- Haugh, Michael. 2018. "Corpus-Based Metapragmatics." In: Andreas H. Jucker, Klaus P. Schneider & Wolfram Bublitz, eds. *Methods in Pragmatics*. Berlin: De Gruyter. 619-644.
- Heinemann, Trine & Véronique Traverso. 2009. "Complaining in Interaction." *Journal of Pragmatics* 41(12): 2381-2384.
- Höglund, Mikko & Kaj Syrjänen. 2016. "Corpus of Early American Literature." *ICAME Journal* 40: 17-38.
- Jucker, Andreas H. 2020. *Politeness in the History of English. From the Middle Ages to the Present Day*. Cambridge: Cambridge University Press.
- Morton, Ralph & Hilary Nesi. 2020. "Institutional Collaboration in the Creation of Digital Linguistic Resources: The Case of the *British Telecom Correspondence Corpus*." In Simon Pople, Andrew Prescott & Daniel H. Mutibwa (eds.), *Communities, Archives and New Collaborative Practices*. Bristol: Policy Press. 153-164.
- Olshtain, Elite & Liora Weinbach. 1987. "Complaints: A Study of Speech Act Behavior among Native and Nonnative Speakers of Hebrew." In Jef Verschueren & Marcella Bertuccelli-Papi (eds.), *The Pragmatic Perspective: Selected Papers from the 1985 International Pragmatics Conference*. Amsterdam: John Benjamins. 195-209.
- Rüdiger, Sofia (in prep). *Historical Perspectives on the Speech Act of Other-Addressed Complaint – Other-Addressed Complaint Realization, Strategies, and Responses*.

- *Hello, are you kidding?* A study of speech acts realized by the discourse markers *hello* in English and *hallo* in Norwegian

Jan Svennevig
(University of Agder)

Ingrid Kristine Hasund
(University of Agder)

This paper explores the methodological challenges involved in a cross-linguistic corpus-based analysis of the discourse markers *hello* in English and *hallo* in Norwegian. *Hello/hallo* are interesting in a speech act perspective as they can form an utterance alone and perform a range of different speech acts depending on the context. While the traditional vocative uses as a greeting and a summons (Schegloff, 1968) are well established, their evolving non-vocative usages—such as expressing reproach or surprise—remain underexplored (Andersen, 2014). To the best of our knowledge, there exists no in-depth study of *hello* as a discourse marker in English, and Svennevig (2012) is to date the only study of *hallo* in Norwegian. Using data from three corpora of informal spoken Norwegian with audio-linked transcripts—the UNO corpus of teenage language (1997-1998), the Big Brother corpus of young adult language (2001) and the NoTa Oslo corpus (different age groups, 2005) —Svennevig found that *hallo* may be used with three main non-vocative functions: as a reproach to an addressee for having said or done something inappropriate or incorrect, as a negative evaluation of some event, and as an announcement of a newsworthy or interesting event.

To investigate potential cross-linguistic pragmatic borrowing from English to Norwegian, we initially examined several English corpora from around the same time period as the Norwegian data: the CABNC corpus of informal conversations (1980s-1990s), the CallHome and CallFriend corpora of unscripted telephone conversations (1997) and the Santa Barbara Corpus of Spoken American English (1990-1997). Surprisingly, few instances of non-vocative *hello* were found within these datasets. While the reproach and announcement functions appeared, the expected negative evaluation function was absent. This led us to expand our methodological approach. We complemented the corpus searches with a diachronic survey of English dictionaries (1964-2024), revealing that only the reproach function aligns with Norwegian *hallo*. Contrary to the Norwegian data, the English dictionaries frequently list *hello* as ‘expressing surprise’; without context, however, it is hard to interpret this speech act.

Recognizing the importance of non-verbal cues—such as gaze and gesture—in interpreting *hello/hallo*, we incorporated multimodal data. Analysis of the US TV series *Seinfeld* scripts (1989-1998) uncovered a few instances of reproach and negative evaluation, but no cases of

announcements. Further, using a subset from 2016 of the NewsScape corpus of US TV news, which integrates video with transcripts, we identified approximately 80 cases of non-vocative *hello*. In addition to the same three functions described by Svennevig (2012), we observed *hello* used with three other functions not previously documented in research: as an epistemic marker for self-evident truths, as positive evaluation, and as a metaphorical summons.

This study underscores the complexities of using older data for diachronic analysis and highlights the necessity for multimodal resources to capture the full communicative spectrum of discourse markers. Further, it discusses issues relating to metadata, annotation and sound/transcript quality and alignment across different corpora. Aligning with the workshop's focus on methodological rigor, we contribute to broader discussions on dataset comparability and corpus-based speech act research.

References

- Andersen, G. (2014). Pragmatic borrowing. *Journal of Pragmatics*, 67, 17-33. <http://dx.doi.org/10.1016/j.pragma.2014.03.005>
- Schegloff, Emanuel A. (1968). "Sequencing in conversational openings". *American Anthropologist*, 70, 1075–1095.
- Svennevig, J. (2012). "og jeg bare hallo liksom" Diskursmarkøren hallo i samtale. *Språk och interaktion* 3, 157 – 174. <http://hdl.handle.net/10138/37523>