International Journal of English Language Studies
ISSN: 2707-7578
DOI: 10.32996/ijels
Journal Homepage: al-kindipublisher.com/index.php/ijels/index

Genre Framing in Generated English Responses to Situational Prompts

Miguel Marcos-Bo ✉[1], **Reece Anthony Cole**[1]

Department of English Studies, University of Alicante, Alicante, Spain

Corresponding Author: Miguel Marcos-Bo**, E-mail:** mmarcos.bo@ua.es

ABSTRACT

Situational prompts specify what a speaker should accomplish without always specifying the genre of the response. This exploratory secondary analysis examines how four conversational systems frame their answers to twelve English discourse-completion scenarios. The complete existing corpus comprises 48 responses collected in April 2026. Explicit rules identify initial subject lines, greetings, square-bracket fields and final valediction–signature combinations. Prompts are classified as email, texting or message, spoken meeting, or channel-unspecified. Twenty-one responses contain 38 bracketed fields, and twelve have a signed closing. All eight responses to email scenarios contain fields and signed closings, but subject lines occur only in the four extension-request emails. None of the eight responses to texting or message scenarios has fields or signed closings. Twelve of the 28 responses to channel-unspecified scenarios contain fields; one of four responses to the spoken-meeting scenario also contains a field. Same-prompt comparisons show different combinations of address fields, conversational openings and signatures. The findings distinguish the communicative contribution represented by a response from the drafting instructions it leaves to an eventual writer. They support describing genre framing as a separate dimension of generated-text analysis, while providing no human comparison, causal estimate of channel effects, or general system ranking. The twelve situational prompts and the 48 generated responses are provided as supplementary materials. The pedagogical implication is that generated model texts require explicit decisions about channel, personalisation and whether they represent a draft or a completed turn.

KEYWORDS

Generated English; genre; discourse completion tasks; computer-mediated communication; message framing; placeholders

ARTICLE INFORMATION

ACCEPTED: 24 October 2026 **PUBLISHED:** 24 November 2026 **DOI:** 10.32996/ijels.2026.8.5.6

1. Introduction

A generated reply may answer an interlocutor and simultaneously instruct someone else how to finish the text. In the present corpus, a manager's reply to an employee can begin with “Hi [Employee Name]” and end with “Best, [Manager Name]”. Another response to the same prompt grants permission directly without those fields. Both perform the requested action, but they present different objects for use: a reusable message framework and a contribution that does not visibly require name substitution. This distinction becomes consequential when generated responses are treated as model texts for English writing.

Research on generated language increasingly examines its relationship to register and communicative purpose. Demir and Egbert (2026), for example, distinguish linguistic alignment from functional appropriateness when comparing ChatGPT-generated academic texts with human writing. Their study concerns extended academic prose. The present analysis considers a different level of organisation: the visible framing of short elicited replies. A subject line, an address field and a signature are not interchangeable with grammatical complexity or a rating of pragmatic appropriateness. They indicate how the text is packaged for a potential exchange and where further preparation is required.

The empirical question is deliberately narrow. When prompts specify an action and a prior turn, how do the preserved responses mark their genre and personalisation requirements? The study analyses the full twelve-scenario corpus rather than selecting examples according to whether a feature occurs. It distinguishes an explicitly supplied communication channel from one inferred by a reader and then examines contrasting outputs to identical prompts. The contribution is an auditable feature description connected to a linguistic account of message framing, not a new benchmark of human-like competence.

Three research questions guide the analysis. RQ1 asks how subject lines, greetings, bracketed fields and signed closings are distributed across the 48 responses. RQ2 asks how those distributions relate descriptively to the channel information in the

Copyright: © 2026 the Author(s). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) 4.0 license (https://creativecommons.org/licenses/by/4.0/). Published by Al-Kindi Centre for Research and Development, London, United Kingdom.

prompts. RQ3 asks what combinations of features are visible in same-prompt contrasts and what personalisation decisions those combinations leave to an eventual user. These questions address observed texts, not the probability that a system will choose a form in future generations.

2. Literature review

2.1 Genre framing and register

Biber and Conrad (2019) distinguish perspectives on language variation that attend to situationally associated features and to conventional text organisation. This distinction supports analysing the external structure of a response separately from its level of formality. A conversationally phrased body can still be enclosed by name fields and a signed closing. Conversely, respectful wording need not have a subject line or signature. Message framing here denotes these visible organisational choices; it is not a complete classification of each text into a genre.

Recent research comparing human-authored and AI-generated texts provides a broader context for examining such distinctions. Reinhart et al. (2025) address the identification of human and AI writing, while Ibrahim and Mahmoud (2026) compare human-authored and ChatGPT-generated research article titles. Yao and Liu (2026) examine rhetorical differences in metadiscourse between generated and human-authored academic abstracts. These studies concern different textual units and communicative settings, but collectively illustrate why generated language should be examined beyond surface fluency. The present analysis focuses more narrowly on the organisational framing of short situational responses.

Herring's (2007) classification of computer-mediated discourse distinguishes technical and social contextual facets. Accordingly, a professional relationship is not treated as evidence that an exchange occurs through email. Yus (2011) likewise foregrounds contextual interpretation in internet-mediated communication. For the present analysis, the relevant contextual question is what kind of contribution the prompt asks the system to produce. A supplied prior turn establishes a sequential context but may leave the channel open. That openness permits more than one defensible response format.

Square-bracket fields make a further distinction visible. A phrase such as [Mentor Name] is part of the generated document but is not ordinarily intended to be addressed literally to a mentor. It marks information that an eventual writer must resolve. The analysis therefore treats such fields as evidence of an unfinished drafting layer, without assuming that leaving a field is an error. A field may appropriately avoid inventing an identity that the prompt does not supply.

2.2 Generated language and contextual fit

Demir and Egbert (2026) investigate register alignment in journal-article and textbook excerpts. Their results motivate attention to the function of generated language rather than fluency alone. In related work, Demir (2026) addresses human perceptions of register alignment and functional appropriateness in generated and human-authored texts. The current study extends that question to message boundaries and personalisation fields, while differing in design: it has neither human reference texts nor an independent measure of functional appropriateness. Its counts cannot establish the humanlikeness findings that require such comparisons.

Wu et al. (2024) advocate examining free-form pragmatic responses rather than relying only on constrained answer formats. Ma et al. (2025) map the diversity of pragmatic evaluation resources and task designs. Sieker and Zarrieß (2026) further address listener–speaker asymmetries in the pragmatic competence of large language models, while Meizlish and Ziffo (2026) investigate LLM performance in annotating discourse strategies. Together, these studies motivate making the elicitation and measurement choices explicit. The present inventory records text features rather than scoring success. An answer can contain a recognised speech act and still leave decisions about the message format to its user; detecting those decisions does not measure their acceptability to a recipient.

Research on conversational AI in language learning provides an additional educational context. Belda-Medina and Calvo-Ferrer (2022) examine chatbots as conversational partners in language learning, while Belda-Medina and Kokošková (2023) investigate chatbot integration through the Chatbot-Human Interaction Satisfaction Model. Soriano-Gonzalez and Belda-Medina (2026) review chatbot applications for oral proficiency in English language learning. Although these studies address interaction and educational applications rather than the formal features measured here, they reinforce the relevance of considering the communicative context in which generated responses may be used.

2.3 Implications for English writing analysis

Keller et al. (2025) examine concrete genre elements in learner emails, including subject lines, salutations and closings. Those elements offer a useful point of connection between message structure and writing instruction. Their learner ratings and reliability estimates are not adopted here: the current inventory measures presence, not quality or learning.

Related research addresses AI-generated writing in educational and professional communication. Wilson and Rose (2025) compare genre, scoring and authorship in AI-generated and human-written refusal emails, while Moon (2025) examines the effects of ChatGPT on politeness in university students' email requests to faculty. Mohamed et al. (2025) investigate ChatGPT in relation to ESP writing proficiency and learner autonomy. Gabay et al. (2026) provide a broader review of generative AI applications and

challenges in academic writing. These studies situate the present analysis within discussions of AI-supported writing, although the current corpus does not measure learner outcomes, authorship judgments or pragmatic quality.

Further work in English for Specific Purposes and technology-supported writing illustrates the diversity of educational contexts in which generated language is evaluated. Abad-Bataller (2024) examines students' and teachers' evaluations of AI-generated and human-produced ESP materials for tourism, while Abad-Bataller (2026) considers teachers' and students' perspectives on integrating AI into ESP. Abad-Bataller and Soriano-Gonzalez (2025) investigate AR-enhanced AI feedback and engagement in EFL writing. These contributions provide a pedagogical context for distinguishing generated texts from texts that have been adapted to a particular writer, task and communicative setting.

Tampogao et al. (2026), studying Grammarly-corrected Philippine English tweets, also show why automated language output should be interpreted in relation to its digital context and users. Their concern is morphosyntactic correction and local norms, rather than the framing features examined in this study.

These perspectives discourage an evaluation in which more formal framing automatically means better writing. A signed closing may suit an email draft and be unnecessary in the next turn of an ongoing conversation. The analytical task is first to identify the feature and its relation to the supplied situation. Evaluating its usefulness requires information about the eventual channel and writer's purpose. The present design supplies a basis for that discussion, not an empirical test of teaching effectiveness.

3. Methodology

3.1 Corpus and secondary analysis

The corpus consists of all 48 English responses preserved from a master's dissertation: one response from each of ChatGPT, Gemini, Claude and Perplexity to each of twelve scenarios. Collection took place in April 2026. The source documentation reports identical scenario material, fresh conversations, default interface settings and retention of the first response. Those procedures are reported from the collection documentation; no independent backend or interface logs were available. System names identify the recorded sources, not verified current product versions.

Each scenario provides a speech act, social relationship, context, prior turn and production task. The prompts also contain instructions about tone and communicative strategy. They therefore elicit scaffolded production rather than spontaneous inference of all contextual requirements. The same twelve scenarios were used across the four systems, but the scenarios do not independently manipulate channel, relationship and speech act.

The complete corpus also supports a companion manuscript under review on pragmatic ratings and contextual calibration. That manuscript reports the original five-criterion scores and qualitative discussion of S10–S12. This article reuses the same texts for a new formal-feature analysis. It does not repeat the scores, inferential comparisons or interpretations of politeness, documentary stance and relational support. All twelve scenarios are included in the present counts; the detailed framing comparisons below use S3, S5 and S8. This is a secondary analysis, not independent data or a replication. The companion manuscript is identified in the accompanying editorial statement, which explicitly describes this data reuse.

The analysis was developed after inspection of the corpus and was not preregistered. All available responses are included, irrespective of their detected features. No response was newly generated or rescored.

3.2 Classification of prompt channels

Channel categories are based on the wording of the scenario, not on the response. S1 explicitly requests an email reply to a professor. S10 places the customer's reply within customer-service email correspondence. S2 specifies texting and S6 an apologetic message; these are grouped as texting or message without assuming identical platform affordances. S11 explicitly places the prior turn in a team meeting. No communication channel is fixed for the requested contribution in S3–S5, S7–S9 or S12. In S8, references to client emails do not specify the channel of the employee's reply to the mentor.

Table 1 Prompt channels and corpus coverage

Channel	Scenarios	n
Email	S1 and S10	8
Texting or message	S2 and S6	8
Spoken meeting	S11	4
Unspecified	S3–S5, S7–S9 and S12	28

Counts refer to the retained corpus.

These categories describe the prompts, not platforms on which the fictional exchanges actually occurred. The institutional and interpersonal situations differ in several respects besides channel, so the categories cannot be used to estimate a causal channel effect. The spoken-meeting category is retained separately even though it contains only one scenario, because a speech setting is explicitly supplied.

3.3 Feature inventory

Four binary features are extracted from the preserved response text. A subject line is the first non-empty line when it begins with Subject followed by a colon. A greeting is Dear, Hi, Hey or Hello at the start of the first non-empty line after any subject line. This restricted rule excludes discourse openings such as Thanks or Actually. A bracketed field is a non-empty single-line expression inside square brackets. Both field-bearing responses and individual fields are counted. A signed closing is a final non-empty line beginning with Yours sincerely, Yours faithfully, Sincerely, Kind regards or Best, followed by a comma and a bracketed signature field.

The extraction is case-insensitive for the lexical patterns. A multi-part expression inside one pair of brackets counts once; adjacent bracket pairs count separately. The procedure does not classify every form of address, closing move or incomplete information slot. In particular, a response without square brackets may still need editing. Absence of the detected features is not treated as proof that a response is fully ready to send.

The definitions were tailored to the observed corpus and implemented in a standard-library Python script. The closing inventory includes Yours faithfully as well as the four other observed forms. Omitting Yours faithfully returns eleven signed closings rather than twelve. This sensitivity check documents a category boundary, not statistical significance. The response-level matrix retains all texts and extracted fields so the classification can be inspected directly. No inter-rater agreement or independently validated annotation is claimed.

Table 2. Operational definitions and coding boundaries of message-framing features

Feature	Operational definition	Coding boundary
Subject line	First non-empty line beginning with Subject:	Subject lines elsewhere in the response are not counted.
Initial greeting	Dear, Hi, Hey or Hello at the beginning of the first non-empty line after any subject line.	Other openings, such as Thanks or Actually, are excluded.
Bracketed field	Non-empty, single-line expression enclosed in square brackets.	Each bracket pair counts as one field; field-bearing responses are also counted separately.
Signed closing	Final non-empty line beginning with Yours sincerely, Yours faithfully, Sincerely, Kind regards or Best, followed by a comma and a bracketed signature field.	Other closing forms and closings without a bracketed signature field are excluded.

Note. Features were extracted using explicit, case-insensitive lexical rules. The inventory captures selected observable message-framing markers rather than a comprehensive classification of genre or communicative appropriateness. Feature categories are not mutually exclusive.

3.4 Verification and analysis

The 48 response texts and identifiers were checked against the original workbook. No wording or original rating was changed. The feature script checks the number of texts and the uniqueness of each system–scenario combination and reports counts by scenario, channel and system. Illustrative excerpts retain original wording and are identified by response ID. They were selected to expose contrasts in framing, not to estimate the frequency of additional qualitative categories.

Counts are descriptive and always use the corresponding denominator. Because there is only one generation per system–scenario combination and no sampling model for communicative situations, no significance tests, general system ranking or population intervals are reported. AI-assisted pattern extraction and textual inspection were used for this secondary analysis. The extraction script and response-level feature matrix are available from the corresponding author on request. The script makes the calculations reproducible but does not constitute independent validation of the interpretation.

3.5 Descriptive statistical analysis and visualisation

The unit of analysis is the individual generated response (N = 48). Each response contributes one binary observation for each of the four message-framing features: subject line, initial greeting, presence of at least one bracketed field and signed closing. The number of individual bracketed fields is additionally retained as a separate count variable. Responses are grouped descriptively by prompt channel and recorded system, with the relevant denominators reported for each comparison.

For a given feature and subgroup, observed prevalence is calculated as the number of responses exhibiting the feature divided by the total number of responses in that subgroup, multiplied by 100. Thus, the reported percentages describe the composition of the preserved corpus rather than estimates of population-level probabilities. Feature categories are not mutually exclusive, and their percentages are therefore not expected to sum to 100%.

A channel-by-feature matrix is used to visualise these descriptive proportions. Each cell displays both the within-channel percentage and its underlying frequency and denominator, allowing differences in subgroup size to remain explicit. The visualisation is intended to complement, rather than replace, the absolute frequencies in Table 3.

The analysis does not employ hypothesis tests, regression models or population confidence intervals. The corpus comprises one retained generation per system–scenario combination, the scenarios were not sampled probabilistically, and channel categories are neither balanced nor independently manipulated. Statistical independence and generalisability beyond the observed responses are therefore not assumed. The descriptive analysis is supplemented by close comparison of selected same-prompt responses to examine combinations of features that cannot be adequately characterised by marginal frequencies alone.

The extraction rules, response-level feature matrix and original response identifiers provide an auditable basis for the calculations. The resulting statistics are treated as corpus descriptions, not measures of communicative quality, pragmatic competence or system performance.

4. Results

4.1 Overall distribution and prompt channels

Across the 48 responses, the inventory identifies four subject lines, 32 initial greetings, 21 responses containing fields and twelve signed closings. The 21 field-bearing responses contain 38 bracketed fields. Every response with a subject line or signed closing also has a field, but the converse does not hold. These features therefore describe related but non-identical aspects of framing.

Table 3. Feature counts by prompt channel

Channel	n	Subject	Greeting	Field texts	Signed closing
Email	8	4	8	8	8
Text or message	8	0	6	0	0
Spoken meeting	4	0	0	1	0
Unspecified	28	0	18	12	4
Total	48	4	32	21	12

Counts refer to the retained corpus. Signed closing denotes the defined valediction-signature combination.

Figure 1 presents the within-channel prevalence of the four message-framing features, displaying both percentages and observed frequencies. Email responses exhibit a consistent combination of greetings, bracketed fields and signed closings (8/8 for each feature), although subject lines occur in only half of these responses (4/8). Texting or message responses contain greetings in six of eight cases but no detected fields or signed closings. Within the channel-unspecified group, greetings occur in 18 of 28 responses (64.3%), fields in 12 (42.9%) and signed closings in four (14.3%). The spoken-meeting category contains one field-bearing response among four. These proportions illustrate differences in the observed feature profiles without implying that the channel categories are directly comparable experimental conditions.

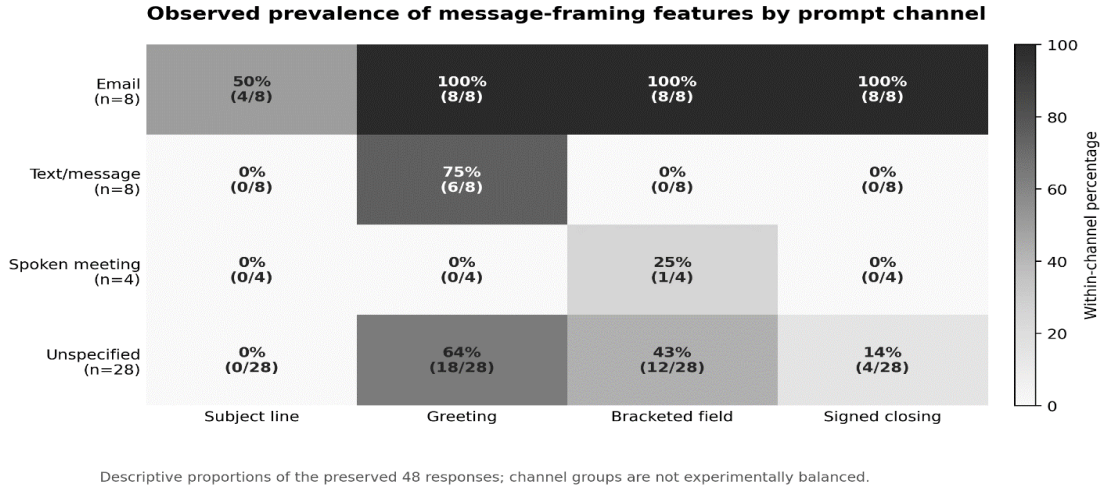


Figure 1. Observed prevalence of message-framing features by prompt channel. Cells report within-channel percentages and response frequencies (n/N). Features are not mutually exclusive. Percentages describe the retained corpus and should not be interpreted as population estimates.

The two email scenarios have a shared profile of greetings, fields and signed closings in all eight responses, but differ in subject-line provision: all four S1 responses include a subject line, whereas none of the S10 replies does. Even an explicitly email-based situation therefore does not yield one invariant set of email components. These are recorded outputs, not evidence that a subject line is required or missing in the real exchanges represented by the prompts.

The eight responses to texting or message scenarios contain no fields, subject lines or signed closings. Six begin with a greeting, demonstrating that greetings alone do not separate these contributions from email-framed responses. In the spoken-meeting scenario, no subject line, inventory greeting or signed closing occurs, but one response contains a name field. A bracketed field consequently marks a preparation requirement across more than written-message genres; it is not an exclusive diagnostic of email.

The channel-unspecified category contains 28 responses, of which twelve have fields and four have signed closings. Eighteen contain an initial greeting, and none includes a subject line. The twelve field-bearing texts comprise three responses each in S3 and S4 and two each in S5, S8 and S9. No fields occur in S7 or S12. This distribution shows variation within unspecified-channel tasks; it cannot establish that absence of a channel instruction caused the fields.

4.2 Same prompt and different framing

In S3, the manager grants permission to leave early. Gemini's R10 begins "Hi [Employee Name]," and ends "Best, [Manager Name]". Claude's R11 begins "Hi — of course, that's absolutely fine." and has no bracketed field or signed closing. The former text explicitly leaves personalisation to a writer, whereas the latter immediately continues with the permission. ChatGPT's R09 and Perplexity's R12 occupy a further combination: an addressee field without a signed closing. The set therefore contains partial combinations rather than two neatly separated genres.

S5 provides a similar contrast in an apology task. R18 begins "Hi [Colleague Name]," and ends "Best, [My Name]"; R17 begins "Hey," without a field or signature. R19 starts "Honestly, I owe you a massive apology." and proceeds directly to the requested action. R20 contains an addressee field without a signed closing. These differences concern how the text is presented for use, not whether its apology is more sincere or pragmatically successful.

S8 asks an employee to seek advice after a mentor's question about an account. R30 opens "Hi [Mentor Name]," followed by "Thanks for asking."; R31 begins "Actually, I'm glad you asked". R29 starts "Hi," with no bracketed name, while R32 provides a mentor-name field. Both a field-bearing response and one without a field can refer back to the prior turn. Thus the presence of personalisation fields does not by itself show a failure to continue the supplied exchange. It exposes an additional drafting layer alongside that continuation.

4.3 Profiles within the retained systems

Table 4 presents the system-level totals to make the distribution auditable, not to identify a best system. Each system contributes twelve responses to the same twelve prompts. Each system contributes twelve responses to the same twelve prompts.

Table 4. Feature counts by recorded system

System	n	Subject	Greeting	Field texts	Signed closing
ChatGPT	12	1	10	4	2
Gemini	12	1	9	8	6
Claude	12	1	4	2	2
Perplexity	12	1	9	7	2

Counts refer to the retained corpus. Signed closing denotes the defined valediction-signature combination.

All four systems contribute field-bearing responses and signed closings in the email scenarios. Outside those eight texts, the remaining thirteen field-bearing responses are distributed across the other prompt categories. The system totals are therefore partly conditioned by a common set of tasks. Without repeated generations, the data do not estimate each system's probability of choosing a field or signature. The recorded contrasts remain useful for identifying alternatives available within this corpus.

4.4 The contribution of a drafting layer

The main contribution is the distinction between a response's represented interpersonal action and the drafting instructions embedded in its form. A square-bracket field addresses the person preparing the text rather than literally addressing the fictional interlocutor. It makes a layer of user work visible. A signed closing combines such a preparation requirement with a recognisable message convention. This account explains why speech-act completion and document completion should not be collapsed into a single judgment.

The data also constrain that account. The field in the spoken-meeting scenario prevents an equation between placeholders and email. The contrasting subject-line profiles of S1 and S10 prevent treating all email replies as one identical structural template. Finally, R30 combines a field with a response to the prior turn. The framing distinction is therefore multidimensional: fields, boundaries and sequential linkage can coexist in different combinations. It is more precise to describe those combinations than to label every response either a template or a dialogue continuation.

4.5 Relation to register and digital discourse

Demir and Egbert's (2026) functional perspective directs attention to the situational demands placed on generated texts. The present results add a limited, inspectable set of message-boundary observations to that broader question. They do not establish the functional appropriateness of the outputs, since the corpus contains no recipient or human reference judgments. Herring's

(2007) separation of social and technical context further clarifies why professional roles cannot substitute for explicit channel information.

The findings also support keeping formal detection distinct from normative correction. As Tampogao et al. (2026) show in a different domain of automated writing support, linguistic decisions require attention to the context of use. A bracketed field can be useful in a draft, while a salutation may be unnecessary in a particular ongoing exchange. The present counts expose the choice without supplying a universal rule for resolving it. This is an analytical implication of the feature patterns, not evidence that the systems were designed with a particular genre model.

4.6 Applications to English writing instruction

The corpus offers material for an explicit genre-awareness task. Students could first identify the channel information actually supplied in S8, then compare R29–R32 and decide what each version would require before use as an email or chat reply. The teacher could distinguish replacing a name field, retaining or deleting a greeting, and adapting the text to a channel from evaluating the request for advice itself. This would make editing decisions visible instead of presenting a single fluent answer as a complete target norm.

Keller et al. (2025) provide a related rationale for examining concrete genre elements in writing instruction, but the current proposal is not an intervention. No learners used these materials and no improvement in writing or genre awareness was measured. The practical output is a set of traceable examples and questions, suitable for subsequent evaluation rather than a demonstrated teaching method.

5. Discussion

The findings provide a descriptive account of how conversational AI systems frame short English responses when situational prompts specify communicative actions but do not always determine the channel or final form of the response. Across the 48 retained texts, subject lines, greetings, bracketed personalisation fields and signed closings occurred in different combinations. These patterns suggest that responding to a communicative situation and producing a fully specified message are related but distinct aspects of generated language. The central contribution of the study is therefore not a judgment about which system produces better English, but the identification of an observable drafting layer that may accompany an otherwise relevant response.

The overall distribution addresses RQ1 by showing that the four features are neither equally frequent nor interchangeable. Initial greetings occurred in 32 of the 48 responses, whereas subject lines appeared in only four. Twenty-one responses contained bracketed fields, comprising 38 individual fields, and twelve included signed closings. Importantly, every response with a subject line or signed closing also contained a bracketed field, while several field-bearing responses lacked these additional features. This relationship suggests that personalisation fields can function independently of a complete email-like framework. A generated response may identify an addressee without providing a signature, or it may incorporate both elements within a more explicitly prepared message. Such distinctions are compatible with approaches to genre and register that separate linguistic choices from the conventional organisation of texts (Biber & Conrad, 2019). They also complement research on AI-generated language that emphasises the importance of functional and situational appropriateness rather than fluency alone (Demir & Egbert, 2026).

The distribution across prompt channels further clarifies the relationship between situational information and message framing. All eight responses to explicitly email-based scenarios contained greetings, bracketed fields and signed closings. However, subject lines appeared only in the four responses to S1 and were absent from the four responses to S10. This difference is noteworthy because both scenarios were classified as email-based, yet their generated outputs did not share an identical structural profile. It suggests that an explicit channel label does not necessarily determine every component of a response. In genre terms, the presence of an email setting is compatible with different assumptions about whether the generated text should constitute a complete standalone email or a reply within an existing correspondence. The study cannot determine which interpretation is more appropriate, but it demonstrates why individual framing features should be examined rather than assuming that all texts associated with a channel follow one invariant template.

This observation is particularly relevant to research on email writing. Keller et al. (2025) analyse subject lines, salutations and closings as distinguishable components of learner-produced emails, showing the value of examining genre-specific elements individually. Moon (2025) investigates politeness in university students' email requests to faculty, while Wilson and Rose (2025) consider genre and authorship in AI-generated and human-written refusal emails. These studies address different research questions, but they help situate the present findings within a broader concern with how communicative purposes are realised through the conventions of written interaction. The present corpus adds a narrower observation: the visible presence of an email component should not automatically be interpreted as evidence of greater communicative success, and its absence should not automatically be classified as a deficiency. Such evaluations would require the intended message format, the expectations of the recipient and the communicative purpose to be established independently.

The texting or message scenarios provide a contrasting pattern. None of their eight responses contained subject lines, bracketed fields or signed closings, although six began with greetings. This combination suggests that greetings are not exclusive to email framing and that their presence alone cannot establish whether a generated text is intended as a formal document or a conversational contribution. The absence of bracketed fields and signed closings in these scenarios is consistent with a more

immediately usable message format, but it does not prove that the texts would require no further editing. These results support Herring's (2007) distinction between the social and technical dimensions of computer-mediated discourse and Yus's (2011) emphasis on contextual interpretation. The relevant issue is not simply the social relationship represented in a prompt, but also the communicative channel and the form of participation that the prompt makes available.

The channel-unspecified scenarios are especially informative for RQ2. Of the 28 responses in this category, twelve contained bracketed fields and four included signed closings, while eighteen began with greetings. The field-bearing responses were distributed across S3, S4, S5, S8 and S9, with none in S7 or S12. These differences indicate that the absence of an explicitly stated channel does not lead to one consistent framing strategy. Some responses adopt features associated with editable written messages, whereas others proceed more directly with the requested interpersonal action. Nevertheless, the scenarios differ in their speech acts, social relationships and task instructions. The findings therefore cannot establish that channel underspecification caused the observed variation. Instead, they identify a pattern that would justify more controlled investigation of channel instructions in future research.

The spoken-meeting scenario provides an additional qualification. Although none of its four responses contained an inventory greeting, subject line or signed closing, one included a bracketed name field. This observation prevents bracketed personalisation fields from being treated as exclusive markers of email or other written-message genres. A field may also occur when a response is framed as something to be prepared or adapted for a spoken situation. Consequently, the most defensible interpretation of bracketed fields is that they indicate information left for an eventual user to resolve, rather than that they identify a particular communication channel. This distinction is important because it separates the detection of a textual feature from an assumption about its communicative function.

The same-prompt comparisons address RQ3 by illustrating how different framing choices can accompany responses to an identical communicative task. In S3, Gemini's R10 includes both an employee-name field and a signed closing, while Claude's R11 grants permission without these elements. ChatGPT's R09 and Perplexity's R12 provide intermediate combinations, incorporating an addressee field without a signed closing. The contrast is therefore not a simple opposition between email templates and conversational responses. Instead, the four outputs display different degrees of explicit message preparation. A similar pattern appears in S5, where the apology responses range from a message with addressee and signature fields to a response that begins directly with the apology. These differences concern the presentation of the response for subsequent use; they do not establish differences in sincerity, politeness or pragmatic effectiveness.

S8 further demonstrates that personalisation and conversational continuity can coexist. R30 combines a mentor-name field with an explicit acknowledgement of the preceding turn, whereas R31 continues the interaction without a bracketed field. R29 and R32 offer additional variations in the treatment of the addressee. The presence of a placeholder therefore does not necessarily mean that a response fails to engage with its conversational context. A text may simultaneously perform the requested communicative action and signal that certain details remain to be completed. This finding helps distinguish message framing from broader evaluations of pragmatic competence. Research by Wu et al. (2024) and Ma et al. (2025) emphasises the complexity of evaluating pragmatic behaviour in large language models, while Sieker and Zarrieß (2026) examine differences between pragmatic production and evaluation. The present analysis does not assess these broader abilities. Instead, it offers a restricted, reproducible description of how generated responses are packaged for possible use.

The system-level totals should also be interpreted within these boundaries. Each of the four systems produced one response to each scenario, and all contributed field-bearing responses and signed closings in the email category. The observed differences in the remaining responses show variation within the retained corpus, but they cannot support a general ranking of systems or predictions about future outputs. Repeated generations, a larger range of scenarios and controlled comparisons would be required for such claims. This distinction between descriptive feature analysis and broader performance evaluation is relevant to research comparing AI-generated and human-authored texts. Studies of register alignment (Demir & Egbert, 2026), research article titles (Ibrahim & Mahmoud, 2026) and academic metadiscourse (Yao & Liu, 2026) examine different linguistic dimensions and textual settings. The present study complements these approaches by identifying a formal dimension that can be observed without assigning an overall quality score or assuming that a response should resemble human-authored text in every respect.

A further implication concerns the relationship between automated feature detection and linguistic interpretation. The inventory identifies specific textual markers through explicit rules, making the reported counts inspectable and reproducible. However, a detected bracketed field does not reveal why it was generated, whether it is appropriate for the situation or how much editing an actual user would undertake. Likewise, a response without a detected field may still contain information that requires personalisation. The analysis therefore illustrates both the usefulness and the limits of formal annotation. Research on the annotation of discourse strategies by large language models (Meizlish & Ziffo, 2026) addresses a related methodological concern, although the present study uses rule-based extraction rather than evaluating an LLM as an annotator. The distinction between observable form and inferred communicative function remains essential when interpreting the results.

The educational relevance of these findings lies in the possibility of treating generated responses as materials for critical genre analysis rather than as automatically complete writing models. Learners could compare alternative responses to the same prompt, identify which elements are supplied by the system and decide which details would need to be replaced, retained or removed for

a particular channel. The S3, S5 and S8 contrasts are especially suitable for illustrating how a response can perform an interpersonal action while also containing editable message components. This approach is consistent with research examining conversational AI in language learning (Belda-Medina & Calvo-Ferrer, 2022), chatbot integration and user interaction (Belda-Medina & Kokošková, 2023), and the use of chatbots to support English oral proficiency (Soriano-Gonzalez & Belda-Medina, 2026). These studies provide a broader educational context, although they do not directly validate the proposed classroom use of the present corpus.

The findings also connect with research on AI-supported writing and English for Specific Purposes. Mohamed et al. (2025) investigate ChatGPT in relation to ESP writing proficiency and learner autonomy, while Gabay et al. (2026) review applications and challenges associated with generative AI in academic writing. Abad-Bataller (2024) examines evaluations of AI-generated and human-produced ESP materials, and Abad-Bataller (2026) considers teachers' and students' perspectives on AI integration in ESP. Abad-Bataller and Soriano-Gonzalez (2025) investigate AI-supported feedback and engagement in EFL writing. Together, these areas of research underline the importance of considering how generated texts are selected, interpreted and adapted in educational settings. The present results suggest one specific instructional focus: helping learners recognise the difference between a response that fulfils a communicative purpose and a draft that still requires decisions about personalisation and genre conventions.

This pedagogical interpretation should remain proportionate to the evidence. No learners participated in the present analysis, and no intervention assessed whether identifying framing features improves writing quality, pragmatic awareness or learner autonomy. The proposed classroom applications are therefore possibilities derived from the observed texts rather than demonstrated educational outcomes. Similarly, the corpus does not contain recipient judgments or human-authored comparison responses, and its channel categories are not experimentally balanced. The findings should consequently be understood as a bounded description of one preserved set of outputs. Future studies could examine repeated generations, independently varied channel instructions, alternative placeholder conventions and human evaluations of the usefulness of different response formats.

Overall, the discussion reinforces the value of separating the interpersonal action represented by generated language from the drafting layer through which that action is presented. The observed combinations of greetings, fields, subject lines and signed closings show that these dimensions do not necessarily coincide. Genre framing can therefore be examined as an additional descriptive dimension alongside register, pragmatic interpretation and writing quality. Within the limits of the present corpus, this distinction offers a transparent way to describe how generated English responses may function both as contributions to an imagined exchange and as texts awaiting adaptation by an eventual writer.

6. Conclusion

The complete corpus shows systematic combinations of message-framing features within explicitly bounded tasks: greetings, fields and signed closings occur in all eight email responses, while fields and signed closings are absent from the eight texting or message responses. The channel-unspecified category and the spoken-meeting case reveal further variation, including partial frames and a field outside email. Subject-line differences between the two email scenarios show that channel labels alone do not determine all components of generated message framing.

Several limitations constrain the interpretation of these findings. The corpus contains only one response from each of four systems per scenario, and including every available response does not make the twelve-scenario battery representative of English genres or current system behaviour. Channel categories are unbalanced and confounded with relationship, communicative action and task wording. In particular, one spoken-meeting scenario cannot establish a general pattern of spoken-response framing.

The feature-extraction rules are transparent but deliberately restricted. Closing forms outside the predefined inventory and other placeholder conventions would be missed. The sensitivity of the signed-closing count to the inclusion of Yours faithfully illustrates the importance of documenting category boundaries. Although the response-level feature matrix permits inspection of the classifications, it does not provide independent annotation validation. Counts of visible fields also cannot establish actual editing time or all the modifications required before a response can be used.

The absence of human-authored comparison texts, repeated generations and recipient judgments further limits the conclusions to the preserved outputs. The same empirical corpus is also used in a companion manuscript addressing different analytical variables, and the two datasets should not be treated as independent evidence. No causal explanation of system behaviour, learner benefit or general hierarchy of pragmatic competence is supported by the present design. Future research could investigate larger and more varied corpora, experimentally controlled channel instructions, repeated generations and human evaluations of how different framing choices affect the perceived usability of generated responses.

The principal linguistic contribution is the identification of a drafting layer made visible through personalisation fields and message boundaries. Same-prompt contrasts demonstrate that this layer can accompany a contribution that also responds to the preceding conversational turn. The distinction between communicative action and textual preparation therefore provides a useful analytical perspective on generated English. For research and teaching, generated responses should be examined not only for the action they represent, but also for the decisions about genre, channel and personalisation that their form leaves to an eventual writer.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

ORCID iD: https://orcid.org/0009-0006-7254-1342 Miguel Marcos Bo

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Abad-Bataller, S. (2024). Students' and teachers' evaluation of the interestingness of AI-generated vs. human-produced ESP materials for tourism. In J. Belda-Medina, C. Holgado-Sáez, & J. Gutiérrez-Artacho (Eds.), *Digital language learning: New approaches and methods* (pp. 105–126). Peter Lang. https://doi.org/10.3726/b21639

[2] Abad-Bataller, S. (2026). Integrating artificial intelligence in English for Specific Purposes (ESP): Teachers' and students' perspectives. In R. Casañ Pitarch (Ed.), *Emerging digital technologies in language education: Virtual simulation, games and artificial intelligence* (pp. 261–282). Tirant Humanidades.

[3] Abad-Bataller, S., & Soriano-Gonzalez, L. (2025). AR-enhanced AI feedback and EFL writing engagement. *International Journal of English Language Studies*, *7*(6), 1–12. https://doi.org/10.32996/ijels.2025.7.6.1

[4] Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as AI conversational partners in language learning. *Applied Sciences*, *12*(17), Article 8427. https://doi.org/10.3390/app12178427

[5] Belda-Medina, J., & Kokošková, V. (2023). Integrating chatbots in education: Insights from the Chatbot-Human Interaction Satisfaction Model (CHISM). *International Journal of Educational Technology in Higher Education*, *20*, Article 62. https://doi.org/10.1186/s41239-023-00432-3

[6] Biber, D., & Conrad, S. (2019). *Register, genre, and style* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/9781108686136

[7] Demir, N. Y. (2026). *Human perceptions of register alignment and functional appropriateness: ChatGPT-generated vs. human-authored texts*. Digital Scholarship in the Humanities, Article fqag084. https://doi.org/10.1093/llc/fqag084

[8] Demir, N. Y., & Egbert, J. (2026). Register alignment of ChatGPT-generated academic texts. *Applied Corpus Linguistics*, *6*(1), Article 100174. https://doi.org/10.1016/j.acorp.2025.100174

[9] Gabay, R. A. E., Funa, A. A., & Ricafort, J. D. (2026). Generative artificial intelligence (GenAI) for academic writing in higher education: A scoping review of applications, challenges, and implications. *International Journal of Education in Mathematics, Science and Technology*, *14*(1), 200–232. https://doi.org/10.46328/ijemst.5682

[10] Herring, S. C. (2007). A faceted classification scheme for computer-mediated discourse. *Language@Internet*, *4*. https://scholarworks.iu.edu/journals/index.php/li/article/view/37562

[11] Ibrahim, S. K. M., & Mahmoud, Z. A. Z. (2026). Generative AI in academic writing: A comparison of human-authored and ChatGPT-generated research article titles. *Humanities and Social Sciences Communications*, *13*, Article 394. https://doi.org/10.1057/s41599-026-06956-z

[12] Keller, S. D., Trüb, R., Horbach, A., Jansen, T., & Fleckenstein, J. (2025). Mastering the genre of English emails: A linguistic analysis of learner uptake in an on-line learning environment. *Research in Subject-matter Teaching and Learning*, *8*(2), 27–46. https://doi.org/10.2478/ristal-2025-0005

[13] Ma, B., Li, Y., Zhou, W., Gong, Z., Liu, Y. J., Jasinskaja, K., Friedrich, A., Hirschberg, J., Kreuter, F., & Plank, B. (2025). *Pragmatics in the era of large language models: A survey on datasets, evaluation, opportunities and challenges*. In Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 8679–8696). Association for Computational Linguistics. https://doi.org/10.18653/v1/2025.acl-long.425

[14] Meizlish, T., & Ziffo, C. (2026). Evaluating an LLM's performance in annotating discourse strategies. *Corpus Pragmatics*, *10*, Article 23. https://doi.org/10.1007/s41701-025-00224-2

[15] Mohamed, A. M., El Deen, A. A. M. M., Abukhait, R. O., Dahleb, F., & Khan, Y. (2025). ChatGPT's impact on ESP writing proficiency and learner autonomy: An experimental study. *Technology in Language Teaching & Learning*, *7*(3), Article 102964. https://doi.org/10.29140/tltl.v7n3.102964

[16] Moon, Y. (2025). The effects of ChatGPT on politeness in L2 university students' email requests to faculty. *Modern English Education*, *26*, 352–363. https://doi.org/10.18095/meeso.2025.26.1.352

[17] Reinhart, A., Markey, B., Laudenbach, M., Pantusen, K., Yurko, R., Weinberg, G., & Brown, D. W. (2025). Do LLMs write like humans? Variation in grammatical and rhetorical styles. *Proceedings of the National Academy of Sciences*, *122*(8), Article e2422455122. https://doi.org/10.1073/pnas.2422455122

[18] Sieker, J., & Zarrieß, S. (2026). How hypocritical is your LLM judge? Listener-speaker asymmetries in the pragmatic competence of large language models. In M. Liakata, V. P. Moreira, J. Zhang, & D. Jurgens (Eds.), *Findings of the Association*

for Computational Linguistics: ACL 2026 (pp. 7551–7565). Association for Computational Linguistics. https://doi.org/10.18653/v1/2026.findings-acl.372

[19] Soriano-Gonzalez, L., & Belda-Medina, J. (2026). Chatbots for oral proficiency in English language learning: A systematic review. *Telematics and Informatics Reports*, *23*, Article 100346. https://doi.org/10.1016/j.teler.2026.100346

[20] Tampogao, A. R., Cosain, S. A., & Sieras, S. (2026). A corpus-based morphosyntactic analysis of Grammarly-corrected Philippine English tweets. *International Journal of Linguistics Studies*, *6*(1), 33–44. https://doi.org/10.32996/ijls.2026.6.1.3

[21] Wilson, W., & Rose, H. (2025). *A genre, scoring, and authorship analysis of AI-generated and human-written refusal emails*. Business and Professional Communication Quarterly. https://doi.org/10.1177/23294906251322890

[22] Wu, S., Yang, S., Chen, Z., & Su, Q. (2024). *Rethinking pragmatics in large language models: Towards open-ended evaluation and preference tuning. In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (pp. 22583–22599)*. Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.emnlp-main.1258

[23] Yao, G., & Liu, Z. (2026). Rhetorical divergence of AI: Comparing interactive and interactional metadiscourse in GPT-generated and human-authored research abstracts across disciplines. *Journal of English for Academic Purposes*, *80*, Article 101640. https://doi.org/10.1016/j.jeap.2026.101640

[24] Yus, F. (2011). *Cyberpragmatics: Internet-mediated communication in context*. John Benjamins. https://doi.org/10.1075/pbns.213