Lingue e Linguaggi
Lingue Linguaggi 74 (2026), 245-274
ISSN 2239-0367, e-ISSN 2239-0359
DOI 10.1285/i22390359v74p245-274
https://ese-journals.unisalento.it, © 2026 Università del Salento

THE MULTIMODAL EXPRESSION OF GRATITUDE
The Interplay of Words and Pictures in Thank-you Cards[1]

ELISA BERTOLDI, SARA GESUATO
UNIVERSITY OF PADUA

Abstract – This paper examines the linguistic and visual components of over 90 print, non-pop-up thank-you cards within the frameworks of speech act set analysis (Blum-Kulka *et al*. 1989) for the linguistic component, and Kress and van Leeuwen's (2021) social semiotic analysis, integrated with Unsworth and Cleirigh's (2017) framework on intersemiotic relations, for the visual component. The linguistic component of the texts is analysed in terms of head acts and supportive moves, identified partly inductively and partly based on the literature (Ahar and Eslami-Rasekh 2011; García and Kleifgen 2016; Eisenstein and Bodman 1986, 1993; Schauer and Adolphs 2006). The visual component is examined through the analysis of representational structures and compositional arrangements, paying special attention to the intersemiotic relations between verbal and visual elements. The cards include direct, occasionally repeated, head acts (e.g., "thank you"; "With sincere gratitude and warmest thanks"). These are often typographically enhanced through bolding or capitalisation and are sometimes characterised by plays on words. The cards may also include supportive moves, such as complimenting the benefactor ("life saver") and acknowledging the impact of the benefit ("That means everything"). The analysis reveals three primary intersemiotic relationship types: humorous engagement, whereby visual elements enable or reinforce intersemiotic wordplay; aesthetic enhancement, whereby images provide visual pleasure; and semantic support, whereby visuals clarify or emphasise verbal expressions. Food imagery most often serves a humorous function, while floral elements primarily provide aesthetic enhancement, and animal imagery demonstrates versatility across all relationship types. Overall, the study demonstrates that multimodal thank-you cards employ intersemiotic coordination strategies that transform routine politeness acts into opportunities for social bonding, suggesting that multimodal analysis can enhance our understanding of how different semiotic resources are co-deployed in positive facework.

Keywords: thank-you cards; multimodality; intersemiotic relations; verbal text; visual text.

[1] Both authors designed the study and analysed the data. They jointly wrote Sections 2.3., 3.1, 3.2, 3.3, 3.3.4, 5.1 and 5.2. Elisa Bertoldi also wrote Sections 2.2, 3.3.2, 3.3.3, 4.1.3, 4.2 and 4.3. Sara Gesuato also wrote Sections 1, 2.1, 3.3.1, 4.1.1 and 4.1.2.

1. Introduction

Pragmatics sees language, or rather language use, as a form of goal-directed social behaviour. It thus explores how people use language to perform actions and enact social relations in context. As part of purposeful communicative behaviour, speech acts (i.e., verbal actions) are privileged means for achieving goals and regulating interpersonal relationships. Expressing gratitude is one such purpose- and relation-oriented verbal action: it both expresses a psychological state and performs the interactional function of balancing out social debts and credits.

Prototypically, thanking is the verbalisation of the feeling of gratitude and indebtedness that a speaker/writer feels toward the addressee as a result of something beneficial, or at least intended as beneficial, that the addressee intentionally performed for the speaker/writer or provided him/her with. It is meant as an acknowledgement of the addressee's generous behaviour and reciprocation for it.

Thanking is a harmony-preserving interactional contribution which helps regulate interpersonal relationships. First, it indicates that the thanker is well-disposed towards the thankee; indeed, as an expressive speech act (Searle 1976), it conveys the thanker's psychological state of gratitude toward the thankee for a past action that benefited the thanker. Second, it enhances the thankee's positive face by acknowledging their previous laudable deed, thus making them feel good; that is, it functions as a positive politeness strategy (Brown and Levinson 1978). It also preserves the thanker's positive face, because their courteous behaviour makes them appear likeable. Third, in its very acknowledgement of the thanker's state of dependence on the addressee, it functions as a verbal repayment of the debt incurred in accepting the thankee's benefit. Thus, if an act of thanking is successful (i.e., if it is understood, accepted, and possibly ratified with a relevant response), the interactants' social debts and credits will have been balanced out.[2]

An act of thanking may be conventionally routinised or originally elaborate, depending on several variables. These include the actual or perceived magnitude of the benefit received, and its long-term effects; the role-relationship between the interactants (their degree of familiarity and social (a)symmetry); the context in which the interaction takes place, which

[2] However, an act of thanking is not a failsafe interactional mechanism. It threatens the thanker's negative face, since indebtedness involves an obligation to return the favour received. Also, if it is not accepted (e.g., because it is not considered adequate or because the thankee wants to break off their relationship with the thanker), the thanker's positive and negative face will be damaged: on the one hand, he/she will feel rejected, disliked, unappreciated; on the other, he/she will remain indebted.

can be private or public, informal or formal; the medium (e.g., written) and instrument (e.g., the phone) of communication, which can make the interaction synchronous or asynchronous, and ephemeral or permanent; the personality of the thanker, who can be effusive or reserved; and the conventional cultural expectations holding in the context where thanking takes place, which shape interactants' choice of content and wording.[3]

The asynchronous and monologic nature of written thanks brings about two main effects.[4] On the one hand, it removes the pressure of immediate response management, which allows writers to plan, ponder, and revise their texts before they are delivered. On the other hand, it produces potentially permanent texts, which can be stored, shared and accessed multiple times across space and time. Written thanks, therefore, both allow and call for greater elaboration and more complex face-work strategies than their synchronous spoken counterparts. This renders them suitable for formal and official occasions where they can evolve into established genre conventions (e.g., book acknowledgements).[5] One established macro genre where written thanking occurs is that of greeting cards. These are commercially produced, multimodal texts that blend verbal and visual elements, and count as culturally ritualized vehicles for expressing, among other things, gratitude.

Commercially printed greeting cards are a macro genre in which the roles of Principal, Author and Animator (Goffman 1981) shift between the card buyer/sender and the card copywriter/designer/manufacturer. At the planning/design stage, the Animator is the manufacturer, who physically produces the card; the Author is the copywriter and/or designer, who select(s) and/or create(s) verbal and visual content; the Principal is initially the brand, which chooses what stance to project (e.g., funny, sentimental, elegant). At the delivery stage, the Principal role shifts to the card sender, who appropriates the message as an expression of their own stance. That is, a ready-made resource is crafted by a professional team for potential use in a number of contexts; when it is selected, delivered and possibly personalised by a sender, it is recontextualised for a specific communicative occasion.

Greeting cards are comparable to other multimodal genres which combine words and still pictures: picture books (Moya-Guijarro 2016), traffic signs (Forceville and Kjeldsen 2018), memes (Vásquez and Aslan 2021), magazine covers (Tseronis 2015), posts (Bouko 2020), public health posters (Al-Subhi 2024) and tweets (Oakey *et al*. 2022), toiletries ads (Del Saz-Rubio

[3] Determinants of gratitude expressions may also include subtle variables like eye contact between the benefactor and the beneficiary (Okamoto and Robinson 1997).

[4] This is intended to the exclusion of texting, which is fingered speech, not writing (McWhorter 2012).

[5] Given the ease of recording speech, spoken thanks can also become permanent in both public contexts (e.g., at awards ceremonies) and private ones (e.g., at birthday parties).

2018), and postcards (Francesconi 2011). Yet, linguistic-multimodal analyses of greeting cards are scarce (Atminytė 2022). In particular, to our knowledge, thank-you cards have not been systematically explored in linguistics, particularly regarding how their multimodal nature contributes to the performance of thanking speech acts.

In the remainder of the paper, we provide a brief overview of the literature on the act of thanking and on the analysis of multimodal texts (Section 2), and then we illustrate our research method (Section 3). Next, we report our findings (Section 4). Finally, we discuss and draw implications from them (Section 5).

2. Literature review

2.1. Thanking

Studies on the speech act of thanking have identified its two core structural components: head acts and supportive moves. The head act is the gist of thanking, which minimally and explicitly conveys the speaker's/writer's illocution through conventionally recognised plain or emphatic formulae (e.g., "Thank you", "Many thanks", "I owe you one", "Much appreciated"), and which may also specify the object of the illocution (e.g., "Thank you for having me", "I really appreciate what you did for us"). A supportive move is an elaboration on the head act that is meant to enhance the effectiveness and appropriateness of the gratitude being expressed. It refers to the situational factors that give rise to, and thus motivate, the sender's gratitude, such as the positive tangible effects of the benefit (e.g., negative consequences avoided), the positive emotional impact of the benefit on the beneficiary (e.g., relief, joy), the benefactor's effort and dedication (e.g., commitment, generosity), or the actual or perceived value of the benefit (e.g., its cost, its durability, scope).

Scholars have reported the variable realisation of acts of thanking, proposing partly different taxonomies for analysing thanking strategies.

Coulmas (1981) classified speech acts of thanking along multiple categories: as ex-ante (i.e., for a potential good or service) vs. ex-post (i.e., for a good or service already provided); as relevant to material vs. immaterial goods, as originating from requested vs. non-requested benefits; and as involving indebtedness vs. non-indebtedness. Instead, Eisenstein and Bodman (1986, 1993), who examined role-play data, classified thanks into three main types: phatic, ritualised responses; expressions of thanks accompanied by subsidiary acts like jokes or questions; and speech act sets, namely expressions of gratitude qualified by additional acts like complimenting, reassuring, or promising to repay.

In DCT-elicited data, Held (1996) identified supportive moves, such as compliments to benefactors and acts of self-denigration, as well as remedial moves that emphasise the thanker's indebtedness and/or the intention to reciprocate. Similarly, in their study of DCT-elicited and corpus data, Schauer and Adolphs (2006) identified numerous supportive moves of gratitude expressions, namely complimenting the interlocutor, stating the reason, confirming the interlocutor's commitment, stating the intent to reciprocate, stating the interlocutor's non-existent obligation, and refusing.

Díaz Pérez (2005) produced a taxonomy that identifies thanking strategies differing in their register appropriateness (formal, e.g., *I give you a lot of thanks*; unmarked, e.g., *Thank you* and colloquial like *Cheers* or *Ta*), and specific supportive moves that may accompany or reinforce them. These include praise to the action (recognising the benefit received) and praise to the person (complimenting the benefactor directly). By contrast, Cheng's (2005) taxonomy comprises eight strategies: thanking, appreciation, positive feelings, apology, recognition of imposition, repayment, other expressions, and alerters.

Although most pragmatic research on thanking has focused on spoken interaction, some studies have examined written thanks such as acknowledgements in publications. The pioneering study by Ben-Ari (1987) showed that their content, structure and formulation serve to construct and manage socio-professional meanings. For example, it highlighted complex dynamics of debts and obligations in the academic community, often masking a hostile tension between respect for hierarchies and a need for independence. It also presented writers as individuals in a network of supportive interpersonal relationships.

Cronin (1995) identified the various dimensions along which acknowledgements may vary: the richness of their content, the intensity of the feelings expressed, the number of acknowledgees mentioned, the goods they are relevant to, the writers' private goals (e.g., cementing relationships).

Giannoni (1998, 2002) examined journal article acknowledgements, showing that they contain optional opening and closing moves (i.e., establishing ancestry, anticipating future interest in the article) and a compulsory credit-giving move, each expandable in specific sub-components. They were presented as manifestations of the authors' civility and membership in the scientific community, which are intended to acknowledge dependence on others for the success achieved in a competitive environment. Giannoni's (2006) analysis of book acknowledgements focused on features relevant to the texts' evaluative function, namely hyperbole, irony and emotivity, which were found to be prevalent in the soft sciences. They appeared to serve a politeness function: maximising praise of the acknowledgees, while downgrading oneself. The goal of such texts was

reported to be the furthering of the writers' careers by making explicit their network of relationships.

Hyland (2003, 2004) and Hyland and Tse (2004) studied acknowledgements in graduate dissertations, coding them for benefit categories (mostly time and academic assistance, but also technical, financial and moral support) and describing them for the sequencing of the benefactors mentioned in them (typically: academic, peers, and friends and family, in this order). The authors described their structure as consisting of three main moves: an optional Reflecting Move (i.e., a comment on the writer's research experience), an obligatory Thanking Move (i.e., mapping credit to individuals and institutions), and an optional Announcing Move (i.e., a statement of responsibility and inspiration).[6] The formulation of the acknowledgement moves was found to be restricted to a few lexico-grammatical patterns. The authors concluded that this genre allows writers to project both personal and public identities, and to promote their professional identity as competent and well-connected members of a discourse community, binding acknowledgers and their academic acknowledgees in a mechanism of long-term mutual indebtedness.

Studies relevant to thanks in general, and to written public thanks, in particular, reveal the range of strategies used to manage social relationships and negotiate face-work. Indeed, gratitude is shaped by social and situational factors, such as status and the level of imposition, as well as emotional and evaluative stance. Since each act of thanking differs from others in terms of sincerity, level of indebtedness and context, it can be variously expressed, involving multiple components, modifiers, and supportive elements.

Studies more specifically focused on acts of thanking found in publication acknowledgements involve managing conflicting social needs (face wants), and raise questions about the interactional purposes they serve, the perspectives they reflect, and the underlying messages they convey. Although they typically follow a structured sequence, they differ in content and style across languages, disciplines and individuals.

The literature reviewed above shows that rich frameworks have been put forward for analysing verbal thanking strategies and their accompanying facework, which can be usefully applied also to the analysis of thank-you cards.

[6] The structure reported by Hyland (2004) was also documented in Al-Ali's (2006) study of PhD dissertation acknowledgements written in English by Arab speakers, but with some variation in their component steps.

2.2. Multimodal texts

Research across multiple multimodal genres has demonstrated that successful multimodal texts achieve effects that are far greater than the sum of their parts through the deployment of intersemiotic relationships (e.g., Liu and O'Halloran 2009). These operate along several key dimensions.

For example, Liu (2023) showed how multimodal texts employ both convergent and divergent relationships between modes, creating elaborate communicative effects. Similarly, Yu's (2019) analysis demonstrated how visual and verbal elements can serve distinct communicative functions simultaneously, creating interpretive spaces that allow for complex meaning negotiation.

Unsworth and Cléirigh's (2017) research demonstrated that visual and verbal elements can provide complementary information units that enhance or complete one another. Thus, visual elements can provide spatial, qualitative, and compositional information, while verbal elements can provide categorical, sequential, and abstract information. Their synergistic relationship leads to the construction of meanings that neither mode could achieve independently. Furthermore, the authors' examination of multimodal reading processes demonstrated how meaning construction in image-text combinations requires readers to navigate complex interpretive processes that integrate information across multiple semiotic channels.

Extensive research across multimodal artefacts featuring brief text-image combinations has confirmed the broad applicability of key theoretical principles regarding visual-verbal coordination. In picture book studies, researchers have identified sophisticated strategies for integrating visual and verbal storytelling, with distinct modal relationships serving various narrative functions (Moya-Guijarro 2016; Nikolajeva and Scott 2006; Sipe 1998). Advertising research has similarly developed frameworks to explain how visual and verbal elements collaboratively produce persuasive effects that exceed the sum of their parts (Hidalgo-Downing and Kraljević Mujic 2015; Stöckl 2024), highlighting the interdependent nature of modes, whose roles shift across contexts (Guo and Feng 2017). In comics research, scholars have demonstrated how visual-verbal integration ranges from strict separation to full fusion, enabling both conventional representation and creative expression (Cohn *et al*. 2017; Skwarzyński 2019). Lambeens and Pint (2015) proposed a framework capturing how comics combine recognisable conventions with innovative elements that provoke engagement through sensory and emotional impact.

A study on contemporary postcards revealed how these multimodal artefacts function through coordination of written messages and images to create coherent communicative experiences that enable users to accomplish complex emotional work (Francesconi 2011). In her research on Edwardian

greeting postcards, Gillen (2013) analysed how pre-packaged visual vocabulary allowed senders to communicate intensely emotional arguments for maintaining their relationships without ever having to put these into their own words.

Hiippala's (2007) comprehensive multimodal analysis of tourist brochures provided particularly valuable insights for brief text artefact research. His study demonstrated how texts operating within a limited semiotic space require careful design when placing and planning text and images to fulfil their communicative purpose, while also adhering to aesthetic principles. Using systemic functional analysis, Hiippala showed how such artefacts imply the orchestration of multiple communicative functions (i.e., promotional, informational, and aesthetic) within the constraints of the semiotic space.

Finally, Forceville and Kjeldsen (2018) examined unusual and creative traffic signs, which, despite their coded nature, can be used as a form of visual "speech acts" to convey novel meanings. The authors demonstrated how their meaning is conveyed through verbal language and through the interaction of visual elements, genre conventions, and situational awareness.

This body of research establishes several key principles for analysing multimodal thank-you cards. First, cards can be understood as texts where meaning emerges from the coordination of visual and verbal elements. Second, the effectiveness of thank-you cards depends on their ability to create meanings that exceed what either visual or verbal elements could accomplish independently. This requires a deep understanding of how different modal relationships can be employed to achieve multiple communicative goals simultaneously.

2.3. Implications

As the above literature review demonstrates, numerous investigations have been conducted on verbal acts of thanking and various multimodal texts. However, no study has investigated how thanks are realised when expressed both verbally and visually. The present study examines how the speech act of thanking is realised multimodally in thank-you cards, private texts chosen by an addresser, but written by others, intended for exchange in private or semi-private settings (i.e., informal social gatherings). The research aims to identify the structural components of thank-you cards, to describe their verbal and visual realisation, to determine the types of relationships holding between their verbal and visual components, and to document the degree of variation in their realisation patterns. To this end, it develops and applies an analytical approach for examining intersemiotic relations in thank-you cards. By incorporating insights from speech act theory, visual semiotics and multimodal discourse analysis, it proposes a comprehensive framework for

understanding how cards accomplish their communicative work.

3. Method

3.1. Research questions

We addressed the following questions:

- RQ-1: What are the verbal head acts and verbal supportive moves that constitute the multimodal speech act of thanking in thank-you cards?
- RQ-2: How often and how widely are they used in multimodal thank-you cards?
- RQ-3: What typographic devices characterise the verbal speech act of thanking?
- RQ-4: What informational-functional role do visual components have in multimodal thank-you cards?
- RQ-5: How often and how widely are they used in multimodal thank-you cards?
- RQ-6: What kind(s) of relationship(s) is/are attested between the verbal and visual components of multimodal thank-you cards, and how often?

The questions were operationalised as follows:

- RQ-1-a: Which of the prototypical expressions of gratitude are attested in the multimodal thank-you cards considered? RQ-1-b: Which of the main supportive moves of acts of thanking described in the literature are attested in those cards?
- RQ-2: What are the frequency and dispersion values of the head acts and supportive moves identified in the multimodal thank-you cards considered?
- RQ-3-a: What different typographic resources (e.g., capitalisation, boldface, italicisation) are used in the verbal component of the speech acts of thanking? RQ-3-b: How often are they used?
- RQ-4: Of the informational-functional roles that visual components may play in multimodal texts, as reported in the literature, which ones are attested in the multimodal thank-you cards considered?
- RQ-5: What are the frequency and dispersion values of the informational-functional roles of visual components in multimodal thank-you cards?
- RQ-6-a: Of the relationship(s) that may hold between the verbal and visual components of multimodal texts, as reported in the literature, which

ones are attested in the multimodal thank-you cards considered? RQ-6-b: What are their frequency values?

3.2. Data collection

The corpus for this study consists of 97 thank-you cards. Of these, 17 thank-you cards were collected by one of the authors and a friend over time, simply out of personal interest in greeting cards. The remaining 80 thank-you cards were collected from *Etsy*, a global digital marketplace specialising in handmade and creative goods. The reason for this choice was that *Etsy* provides access to card designs that reflect current multimodal communication practices and that are produced by a diverse range of independent designers working across different cultural contexts.

Data collection on *Etsy* was conducted in early 2025, using four filters to restrict the search to cards that met specific criteria. This was intended to ensure corpus homogeneity. Cards had to display both textual and visual elements within the same spread, allowing us to examine semantic relationships between the two semiotic modes of communication. Additionally, cards had to be priced at or below 20 euro to represent accessible, mainstream designs rather than luxury or highly specialised products. Furthermore, cards had to be offered by "Star Sellers" only, a designation indicating established sellers with proven track records for quality, customer satisfaction, and reliable service; this was meant to ensure a baseline level of design competence and commercial viability. Finally, cards had to have their verbal component encoded in English, which would ensure language consistency across our data; however, by indicating a language that is widely used also by non-native speakers, cards would be collected from sellers across different geographic regions, thus possibly exemplifying the use of English as an international language. The search returned many results. From these, the first 100 cards were pre-selected, each offered by a different seller. The goal was to obtain a representative coverage of their heterogeneity, while avoiding individual designer bias. A few accidental duplicates were then discarded. The final selection consisted of 80 cards.

All 97 cards were later converted into digital format as .jpg files and uploaded to the *Padlet* platform to ensure shared access for both authors.

3.3. Data analysis

The analysis of the cards focused first on their linguistic and visual components, and then on the intersemiotic relationships between these two components. The approach integrated speech act analysis with social semiotic analysis to capture how the verbal and visual elements work together to accomplish the pragmatic functions of thanking.

Each card was assigned a numerical identifier, and both the numerical value and the verbal component of the card were entered into an *Excel* spreadsheet. Each card was then coded across multiple dimensions: linguistic features (head act types, supportive move categories, exact text transcription, typography; see Section 3.3.1), visual characteristics (content categories, functional roles, compositional features; see Section 3.3.2), intersemiotic relationships (integration levels, relationship types; see Section 3.3.3). More specifically, we applied an annotation scheme that enabled us to classify the verbal and visual components of each card, as well as the relationships between these components.[7]

3.3.1. Linguistic analysis

The verbal component of each card was linguistically analysed following the speech act set analysis approach developed by Blum-Kulka *et al*. (1989), which involved the identification of head acts (i.e., core thanking expressions) and supportive moves (i.e., ancillary speech acts favouring the successful performance of the act of thanking).

We considered core thanking expressions to be only those containing words like *thank you*, *thanks*, *thankful*, *grateful*, *gratitude*, and *appreciate*, which encode positive emotional reactions (coded as EXP for 'Explicit thanking expression'). When no such explicit, prototypical gratitude expression was present, we did not "upgrade" a supportive move to head act status, as is done in Blum-Kulka *et al*. (1989). Instead, we marked the relevant verbal speech act as lacking a head act (coded as ITE for 'Implied thanking expression'). For example, we decided that expressions like *I am indebted to you*, if attested, would be classified as realisations of supportive moves. For each head act, we also specified if it was internally modified by an intensifier (INT; e.g., *so much*) or mitigator (MIT; e.g., *a little note to say*), modulating its force; if it mentioned the benefit (BIT, i.e., the object of the illocution) and/or the benefactor (BEN), these being elements previously attested in analyses of written thanks (e.g., Hyland and Tse 2004; Al-Ali 2006); if it was characterised by humorous wordplay (HUM, e.g., *fanks* for *thanks*); and if it was repeated (DOB), thus giving rise to so-called *gratitude clusters* (Schauer and Adolphs 2006, p. 126).

We also coded the verbal speech acts for the presence (Y for 'Yes') or absence (N for 'No') of supportive moves. Then, drawing on the taxonomies by Cheng (2005) and Eisenstein and Bodman's (1986, 1993), we determined if the verbal component of the cards included these possible supportive moves: Acknowledging the imposition (ACK, e.g., *I know you're busy*),

[7] The coding scheme is available from the authors upon request.

Complimenting the addressee (PLI, e.g., *You're amazing*), Expressing indebtedness (DEB, e.g., *I owe you*), Stating the impact of the benefit (IMP, e.g., *It means a lot to me*) or Promising reciprocation (REP, e.g., *I'll make it up to you!*). The identification of the supportive moves was not based on a predetermined list of possible lexical realisations, but rather on our joint interpretation of the propositional content of the verbal speech acts. This inductive approach led to the identification of one type of supportive move, namely the quotation of a proverb or saying, or the statement of a general reflection (PRO, e.g., *It takes a village*). Additionally, each supportive move was coded for the possible presence of lexical upgraders, specifically Intensifiers (INT, e.g., *very much*) and Wordplay (PUN, e.g., *You're GRAPE*).

When a given verbal component or feature in a card was compatible with more than one code within a coding categories, multiple codes were specified, separated by + symbols, in order to capture the richness of these multimodal texts. For example, the head act of Card 3, namely, *Thanks for being my unpaid therapist* was coded as EXP+INT, signalling that it had an explicit expression of gratitude and an intensifier. Similarly, the supportive moves of Card 21, namely *You're otterly amazing*, *You're # 1* were coded as PLI+PLI, signalling the presence of two compliments to the benefactor.

Finally, the entire verbal speech act was coded for its typographic features, that is, the presence, on all or some of its words, of Standard text (STD), Boldface (BOL), All caps (CAP), Italics (ITL), Colour-coding of morphemes/words (COL), Glitter (GLI), Embossing (EBS) and/or Incision (INC).

Here follows an example of the complete classification of the verbal component of Card 78:

Head Act Coding:
HA_Type: EXP
HA_Text: "Thank you"

Supportive Moves Coding:
SM_Present: Y
SM_Types: PLI
SM_Text: "You're indispensable"

Typography: CAP+COL

3.3.2. Visual analysis

Visual analysis drew upon Kress and van Leeuwen's (2021) social semiotic framework and was adapted to account for representational structures (what

is depicted) and compositional arrangements (layout, prominence, spatial relationships).

For each card, a verbal description of the visual content was provided in the *Excel* file. Then, the main visual elements on the front cover (and the inside spread, when present) were classified into these types: flowers/plants (FLO), animals (ANI), food/drink (FOO), abstract/geometric patterns (ABS), typography-only designs (TYP), scenes/landscapes (SCE), objects (OBJ), and people (PEO). These labels could also be used in combinations, if applicable. For example, Card 78 was classified as OBJ for its depiction of colourful medicine bottles with pills spilling out; instead, Card 40 was classified as ABS+FLO+ANI+OBJ, since it displays heart shapes and circular patterns with butterfly and decorative swirl elements in pink and purple tones, where the swirl elements appear to be plant-like botanical features.

The visual content was also classified along these dimensions: the distinction between narrative (NAR) and conceptual (CON) representations differentiated dynamic processes (i.e., actions or transformations; e.g., *three anthropomorphised dancing pastries*, Card 1) from static displays (i.e., static situations; e.g., *watercolour floral border pattern*, Card 37). The realistic (REA) versus stylised (STY) distinction refers to the degree of photographic realism versus cartoon/simplified rendering of the pictures. Compositional positioning analysed spatial arrangement, distinguishing centrally positioned elements (CEN), which indicate high visual prominence, from marginally positioned elements (MAR), which serve decorative functions. Framing was differentiated between compositions with strong borders/frames (FRA) that clearly delineate visual spaces and open compositions (OPE) featuring flowing, unbounded elements without clear boundaries.

For cards with both front and inside visual elements, we examined how visual themes were maintained or developed across the card's interior spread. We coded interior elements by applying the same scheme, thus capturing visual continuity or contrast between front and inside presentations.

Here follows an example of the complete classification of the visual component of Card 78:

VT_Description: Colourful medicine bottles with pills spilling out
VT_F_Element(s): OBJ
VT_Content and style: CON+STY+CEN+FRA
VT_I_Element(s): – (not applicable for FRONT distribution)

3.3.3. Intersemiotic analysis

The interplay of verbal and visual modes was analysed by drawing on the theoretical framework established by Unsworth and Cléirigh (2017) for intersemiotic relations in multimodal texts, and informed by Lim's (2021)

conceptualisation of co-contextualising and re-contextualising patterns. Three types of relationships were thus identified: Aesthetic Enhancement (AES), Humorous Engagement (HUM) and Semantic Support (SEM).

AES relationships were those where visual and verbal elements combined to enhance the recipient's positive face through beauty, elegance, or sensory appeal. In these cases, visual elements do not directly relate to gratitude concepts, but communicate messages of value and investment in the relationship, making the recipient feel special through premium visual treatment. For example, formal gratitude expressions like *Thank you*, presented in elegant calligraphy and framed by watercolour botanical garland, suggest the recipient deserves beautiful things. Similarly, multisensory aesthetic features such as embossing, glitter, or raised textures applied to written text enhance the material experience of receiving gratitude beyond purely linguistic expression.

HUM relationships emerged when visual and verbal modes collaborated to create wit and playfulness, requiring cognitive effort from recipients to decode intended meanings. This category encompasses two mechanisms: visual reinforcement of existing textual puns, where wordplay already present in written gratitude is amplified through corresponding imagery (e.g., *Thank you BERRY much* accompanied by berry illustrations, Card 71); and visual pun creation, where apparently straightforward gratitude text becomes humorous only through unexpected visual elements that create double meanings (e.g., *Thanks for the support* paired with images of bras, Card 4).

SEM relationships were characterised by logical connections between visual and verbal content, with images directly illustrating, reinforcing, or clarifying gratitude messages. This category was further subdivided into literal semantic support, where images provide concrete representations of abstract gratitude concepts (e.g., *you're a life saver*, accompanied by life preserver imagery, where the visual directly illustrates the rescue concept expressed verbally, Card 26), and figurative semantic support, where visuals help decode metaphorical or figurative verbal messages (e.g., *you make the world brighter,* supported by stars and sparkles that visually demonstrate the brightening effect described in the text, Card 91).

The above analytical framework enabled examination of how different semiotic modes interact, integrate and harmonise in thank-you cards to create what Stöckl and Bateman (2022) describe as "multimodal coherence" (p. 1).

3.3.4. Validation

We each independently tested the analytical framework we developed on 17 thank-you cards. Together, we then discussed ambiguous and hard-to-classify cases until we refined the labels and definitions of the various components of

our coding scheme, which we jointly applied to the entire corpus.

4. Findings

4.1. Findings of the linguistic analysis

4.1.1. The head acts

The 97 thank-you cards considered comprise 682 words, of which 572 encode the head acts and 110 the supportive moves. Therefore, each card contains a verbal text that is, on average, 7 words long. All the cards, except one (i.e., Card 25), have an explicit head act. Supportive moves, instead, are found in only 13 cards. Therefore, the head acts are much more prominent than the supportive moves, not only because they take up the bulk of the verbal text, but also because of their higher dispersion value.

The head act of the verbal speech act of thanking may consist of a single expression of gratitude plus its optional object(s) of the thanking illocution (i.e., the benefit(s)), or it may consist of a combination of two or more such expressions (i.e., a gratitude cluster; see Section 3.3.1). In this paper, to refer to an individual *token* of a gratitude expression plus its optional benefit, we use the term *head act*; instead, the part of the speech act where the illocution and its object(s) are expressed is called *head act component*. Thus, a head act component may comprise one or more head acts.

Short Text Cards (1-10 words)		**Long Text Cards (11+ words)**	
Word count	**Number of cards**	**Word count**	**Number of cards**
1	4	11	1
2	20	12	1
3	14	13	0
4	14	14	3
5	7	15	1
6	9	16	3
7	5	17	1
8	2	18	0
9	4	19	1
10	4	20	2
		30	1

Table 1
Word count distribution of the head act components.

In our corpus, a minority of the cards (i.e., 12) contain gratitude clusters (i.e., multiple head acts) in their head act components; for example, Card 8 contains two gratitude expressions (i.e., *A little note to say thank you. Thank*

you). Most, instead, contain single expressions of gratitude (e.g., *Thank you*, Card 21). The length of the head act components is quite variable, though most are very short (see Table 1). Indeed, most are between 2 and 6 words long; the shortest and the longest are, respectively, 1 and 30 words long. We considered word strings of letters separated by blank spaces, apostrophes, commas, full stops or exclamation marks. Thus, *I'm* (Card 90) counts as two words, but *every-fin* (a wordplay on *everything*, Card 30) counts as one, and so does *my* [picture of excrement] (Card 10), since what follows the possessive *my* is a picture, not a word.

The notion of gratitude is encoded in a few lexical realisations (see Table 2): *gratitude* (3 occurrences in 3 cards), *grateful* (3 in 3 cards), *thankful* (1), *thanks* (30 in 30 cards), *thank you* (80 occurrences in 60 cards), *appreciation* (once in 1 card), and *appreciated* (twice in 2 cards), while four more realisations are humorous variations on *thanks* (i.e., *fanks*, *frank*, *thankxolotl*) or *grateful* (i.e., *gratefull*), which appear to make sense when considered in relation to the visual component of the cards (see Section 4.3). Indeed, these differently spelt gratitude expressions pick up on, and mimic phonetically, (part of) the pictures in the cards, such as a dog's fangs with *Fanks* in Card 15, frankfurters with *Frank* in Card 27, a salamander with the last part of [*thankx*]*olotl* in Card 74.

Lexis	**Occurrences**
Gratitude	3
Grateful	3
thanks	30
Thank you	80
Appreciation	1
Appreciated	2
Fanks	1
Frank	1
thankxolotl	1
gratefull	1

Table 2
Lexical realisations of the notion of gratitude.

Lexical upgraders of gratitude expressions, coded as Intensifiers in our coding scheme, are found in 43 head acts, and only Cards 84 and 88 contain two intensifiers each.[8] They are mostly realised with varied expressions of quantity (i.e., adverbials denoting the amount of gratitude being conveyed; e.g., *very much* in *Thank you very much*, Card 60), but also adverbs denoting the degree of intensity (e.g., *so* in *I'm so grateful*, Card 90). Infrequent

[8] We did not consider other types of upgraders like repetitions of the gratitude expression, which we classified instead as a gratitude cluster, or the use of exclamation marks. For the use of typographic devices like all caps or bolding, see Section 4.1.3.

realisations include a noun (e.g., *an abundance of* in *An abundance of gratitude*, Card 1), three adjectival phrases denoting a great amount (e.g., *a very big* in *A little card to say a very big thank you*, Card 47), and a bound morpheme (i.e., *-olotl*, part of *Thankxolotl*, Card 74).

It is mainly in these internal modifiers of the head acts that one can notice the playful nature of many of these cards. Indeed, they provide an opportunity for plays on words such as *berry much* (Card 7), *so matcha* (Card 13), and *donut […] enough* (Card 38). As in the case of funny gratitude expressions, these playful upgraders often rely on the accompanying pictures for a humorous effect. Thus, *berry much* is visually sustained by drawings of strawberries (Card 7); *so matcha* (Card 13) is accompanied by the drawing of a hand pouring green tea from a teapot into a cup; and *donut […] enough* (Card 38) is paired with the drawing of an anthropomorphised, smiling doughnut soaking in a cup.

The complete list of the lexical intensifiers we identified is the following (see Table 3): *an abundance of*; *a brunch; a bunch* (4 times); *a huge*; *a latte*; *a melon* (twice); *a ton*; *a very big*; *beary much*; *berry much* (twice); *can't […] enough*; *carrot […] enough*; *doesn't […] enough*; *donut* […] *enough*; *ever so*; *for the record*;[9] *from my head tomatoes* (with the letter strings *to*, *ma* and *toes* differently coloured); *many*; […]*olotl*; *much* (once); *oh-so-much*; *so matcha*; *so* (3 times); *so much* (twice); *soup-er*; *soy much* (twice); *so very much*; *this much*; *very much* (3 times).

Intensifiers	**Occurrences**	**Intensifiers**	**Occurrences**
An abundance of	1	*Ever so*	1
A brunch	1	*For the record*	1
A bunch	4	*from my head tomatoes*	1
A huge	1	*many*	1
A latte	1	*[…]olotl*	1
A melon	2	*much*	1
A ton	1	*oh-so-much*	1
A very big	1	*so matcha*	1
Beary much	1	*so*	3
Berry much	2	*so much*	2
Can't ... enough	1	*soup-er*	1
Carrot ... enough	1	*soy much*	2
Doesn't... enough	1	*so very much*	1
Donut ... enough	1	*this much*	1
		very much	1

Table 3
Lexical intensifiers of the notion of gratitude.

[9] We considered *for the record* an upgrader because it is an idiom that is used to emphasise a point that the addressee should notice.

References to the benefits occur 34 times in 31 cards, and therefore exhibit similar frequency and dispersion values. Benefits include actions performed by the benefactor (17 occurrences; e.g., *and still you make time to show you care*, Card 96), qualities of, or roles played by, the benefactor (13 occurrences; e.g., *for simply being you*, Card 95), attitudes or experiences characteristic of the benefactor (1 occurrence; e.g., *for believing in me*, Card 12), or generic goods or services (3 occurrences; e.g., *for everything*, Card 35).

Excluding the formulation *Thank you*, and the presence of the possessive *your*, explicit references to the benefactor are rare, appearing 11 times in 8 cards. It may show up as the object of a preposition (e.g., *because of you*, Card 16), the subject of a clause (e.g., *all that you do*, Card 42), or a subject complement (e.g., *simply being you*, Card 95).

Only Card 28 contains an alerter, the humorous attention-getter *Hay* (a play on *Hey*).

The syntactic formulation of the head acts is varied but mostly conforms to the combination of the interjections *Thank you* or *Thanks* with a *to*-headed prepositional phrase denoting the benefactor or a *for*-headed prepositional phrase denoting the benefit, optionally enriched by an adverbial. More elaborate formulations are also attested, for instance when the benefit is encoded in a sentence or incomplete sentence. The various realisations are listed in Tables 4, 5 and 6.

Formulations	**Frequency**
Thank you	19
Thank you + *to* PP	1
Thank you + *for* PP	9
Thank you + other PP	1
Thank you + adverbial	12
Thank you + adverbial + *for* PP	1
Thank you + adverbial + PP	1
Thank you + adverbial + AP + *for* PP	2
Thank you + sentence	7
Thank you + incomplete sentence + NP	1
Thank you + sentence + *for* PP	2
Thank you + sentence + adverbial	1

Table 4
Formulations of head acts with the interjection *Thank you*.

Formulations	Frequency
Thanks	2
Thanks + *to*_PP	0
Thanks + *for*_PP	7
Thanks + other PP	0
Thanks + adverbial	12
Thanks + adverbial + PP	1
Thanks + alerter	1
Thanks + incomplete sentence	1

Table 5
Formulations of head acts with the interjection *Thanks*.

Other formulations	Frequency
Other interjection	1
Other interjection + adverbial	1
NP	1
Sentence	4
Sentence + NP	1
NP + to_PP + incomplete sentence	1
Incomplete sentence	1
Incomplete sentence + sentence	1
NP + PP	2
AP + PP	1

Table 6
Other formulations of the head acts.

4.1.2. The supportive moves

Only 13 speech acts include single (10) or, more rarely, combinations (3) of supportive moves (see Table 7). These exemplify compliments to the thankee (7 PLI in 6 cards; e.g., *You're a life saver*, Card 26), sayings (4 PRO in 4 cards; e.g., *Life is crammed with things to take care of*, Card 96) and expressions of emotional impact (5 IMP in 5 cards; e.g., *Just thinking about your kindness brings a smile*, Card 94). Only Card 21 contains a humorous word-play (*You're otterly amazing*). These moves either enhance the perceived value of the benefit or positively evaluate the benefactor's behaviour or qualities. No instances were found of Acknowledgement of the imposition (ACK), Expression of indebtedness (DEB) or Promise of reciprocation (REP). Of the moves, one is encoded as a noun phrase (i.e., *Life saver*, Card 87), one as a coordinate clause (i.e., *and so are you*, Card 83), two as combinations of sentences (i.e., *Life is crammed with things to take care of. That means everything*, Card 96), and the rest as single sentences (e.g., *You make the world a better place*, Card 81).

Supportive moves	Frequency
ACK	0
DEB	0
IMP	5
PLI	7
PRO	4
REP	0

Table 7
Frequency of different types of supportive moves.

4.1.3. Typographic characteristics of the verbal speech acts

The use of typographic devices is summarised in Table 8. About half of the cards have verbal text characterised by a single typographic feature. Of the remaining half, 30.9% display two features, and 17.6% display three or more. More specifically, among the single typographic features (see Table 9), CAP is the most frequent one (almost 40%), followed by ITL, BOL, and STD, which display similar frequency values (i.e., around 20%). Among the double features (see Table 10), the most frequent is BOL+CAP (30%), while only three additional combinations (i.e., BOL+ITL, CAP+ITL, CAP+STD) account for just over 13% of the data each. Among the combinations of triple features (Table 11), only three occur more than once, namely BOL+CAP+COL, BOL+CAP+ITL and CAP+COL+STD. Finally, only one card shows a combination of four typographic features.

No. of typographic features	No. of cards
1	48 (49.5%)
2	31 (31.9%)
3	17 (17.6%)
4	1 (1.0%)
Total	97 (100%)

Table 8
Summary of typographic features of the verbal text.

Single typographic features	No. of cards
BOL	10 (20.8%)
CAP	18 (37.5)
COL	0 (0.0%)
EBS	0 (0.0%)
GLI	0 (0.0%)
INC	0 (0.0%)
ITL	11 (22.9%)
STD	9 (18.8%)
Total	48 (100%)

Table 9
Frequency of single typographic features in the verbal text.

Double typographic features	No. of cards
BOL + CAP	9 (29.0%)
BOL + COL	1 (3.2%)
BOL + ITL	4 (12.9%)
BOL + STD	1 (3.2.%)
CAP + COL	2 (6.5.%)
CAP + ITL	5 (16.1%)
CAP + STD	4 (12.9%)
COL + ITL	2 (6.5.%)
COL + STD	1 (3.2.%)
ITL + STD	2 (6.5%)
Total	31 (100%)

Table 10
Summary of double typographic features of the verbal text.

Triple typographic features	No. of cards
BOL + CAP + COL	4 (23.5%)
BOL + CAP + ITL	4 (23.5%)
BOL + CAP + STD	1 (5.9%)
BOL + COL + ITL	1 (5.9%)
BOL + ITL + STD	1 (5.9%)
CAP + COL + ITL	1 (5.9%)
CAP + COL + STD	3 (17.6%)
CAP + ITL + STD	1 (5.9%)
COL + ITL + STD	1 (5.9%)
Total	17 (100%)

Table 11
Summary of triple typographic features of the verbal text.

4.2. Findings of the visual analysis

The analysis of visual and verbal content distribution showed that front-only distribution (FRONT) dominates with 80 cards, where images and written text appear exclusively on the front cover, leaving the interior of the card blank. Extended distribution (EXTENDED) accounts for 16 cards, where content spans both the front and the interior of the card. Only one card (Card 93) employs progressive distribution (PROGRESSIVE), where visual and textual elements continue a narrative from the front to the interior.

The visual analysis of the 97 thank-you cards revealed a variety of subjects depicted both as single visual elements and in combinations (see Table 12). Cards with single visual elements display food imagery as the most prominent category (21 occurrences), followed by floral/botanical elements (20 occurrences) and animals (18 occurrences). Abstract/geometric patterns appear in 8 cards, objects in 11, typography-only designs in 2, scenes in 2, and people in Card 60 only. Additionally, 14 cards display more elaborate compositions combining multiple visual elements. Compositions of objects and floral/botanical elements (vase and flowers) appear in 5 cards;

combinations of landscapes and background scenes with flowers or animals or objects are found in 6 cards; 2 cards combine abstract elements with animals or flowers, and Card 40 displays a complex combination of abstract elements, flowers, animals, and objects.

Subjects depicted	Frequency
FOO	21
FLO	27
ANI	24
OBJ	20
ABS	10
TYP	2
SCE	8
PEO	1

Table 12
Frequency of subjects depicted in the cards.

The food category encompasses a diverse range of items, including anthropomorphised pastries (Card 1), ice cream (Card 2), fruits like grapes (Card 5) and strawberries (Card 7), beverages such as coffee and tea (Cards 13, 55), and other prepared foods. Animal imagery spans domestic pets including cats (Cards 19, 35, 86) and dogs (Cards 15, 29), farm animals including pigs (Card 31) and horses (Card 28), wildlife including bears (Cards 24, 79), sea otters (Card 21), and other marine life (Cards 11, 30), and bees (Cards 10, 11). Botanical elements range from simple decorative borders (Cards 37, 46) to elaborate watercolour arrangements (Cards 39, 83) and to wildflower compositions (Cards 45, 48).

The analysis of content revealed that 76 cards display conceptual representations (CON) showing static objects, symbols, or decorative elements. Narrative actions (NAR) appear in 21 cards, depicting dynamic scenes where characters (mainly anthropomorphized objects, animals or food items) perform actions or engage in human-like behaviours.

Stylistic rendering analysis showed a strong preference for stylized/cartoon/simplified imagery (STY) appearing in 83 cards, while realistic/photographic representation (REA) occurs in 14 cards. This stylistic choice suggests a tendency toward playful, accessible visual communication rather than formal photographic presentation.

Front-page visual analysis revealed that central positioning with high visual importance (CEN) appears in 89 cards, making it the preferred layout choice. Marginal or decorative positioning (MAR) occurs in 8 cards, typically for subtle accent elements. Frame characteristics showed strong borders or frames (FRA) in 94 cards, while open, boundary-free compositions (OPE) appear in only 3 cards.

Interior pages analysis showed a marked preference for marginal

positioning (MAR) of visual elements with 16 occurrences versus 1 occurrence for central positioning (CEN) in Card 94. This reversal of compositional priorities indicates that interior visual elements serve supporting functions, mainly displaying decorative frames that enclose the written text and echo the pattern, colour palette, or visual elements presented on the front cover.

4.3. Findings of the intersemiotic analysis

The intersemiotic analysis revealed three primary relationship types between visual and verbal elements: HUM, AES and SEM (see Table 13).

Intersemiotic relationship	Frequency
HUM	44
AES	32
SEM	11
HUM+SEM	2
AES+SEM	8
Total	**97**

Table 13
Frequency of intersemiotic relationships.

HUM appears in 44 cards, making it the most frequent relationship type. This category encompasses visual reinforcement of textual puns, such as Card 7's *thank you berry much* accompanied by anthropomorphised strawberries with happy facial expressions, or Card 13's *Thanks so MATCHA* paired with a tea preparation scene showing matcha being poured into a cup. Intersemiotic wordplay occurs where written text becomes humorous through unexpected imagery, as in Card 4's *Thanks for all the support* depicted with various colourful bras.

AES occurs in 32 cards, where visual elements enhance the reader's positive face through beauty and elegance without additional semantic layers. These cards feature refined presentations such as Card 37's simple *Thank you* framed by watercolour floral borders in purple, blue, and orange tones, or Card 41's *Thank you for your kindness* accompanied by a wildflower bouquet in a glass jar.

SEM appears in 11 cards as a standalone function. However, semantic interplay combines with other relationship types in 10 additional cases: 2 cards display both SEM and HUM, while 8 cards combine SEM with AES.

Standalone SEM includes literal semantic support, for instance, Card 20's *Thanks a million* illustrated with a red wheelbarrow filled with green paper money, directly representing the large quantity concept. Figurative SEM appears in Card 16: *The wheels go round because of you. Thank you!*

combines with the image of a yellow school bus, connecting the figurative "wheels" of organization to literal vehicle wheels.

Combined semantic functions demonstrate how visual elements can serve multiple purposes simultaneously. Card 87's *Life saver* accompanies an orange and white life ring on blue ocean waves. This exemplifies the combination of SEM+HUM where the life ring imagery clarifies the rescue metaphor (SEM) and creates wordplay humour through the visual reference to life ring (HUM). Card 52's *Thanks a bunch* pairs with a watercolour bouquet exemplifying SEM+AES interplay, where the floral imagery both clarifies the double meaning of *bunch* (SEM), while providing aesthetic enhancement through botanical illustration.

The distribution of these relationship types correlates with visual element categories: food imagery most frequently serves humorous functions, floral elements predominantly provide aesthetic enhancement, while animal imagery demonstrates versatility across all three relationship types, reflecting their adaptability in multimodal meaning construction.

5. Discussion and conclusion

5.1. Discussion of the findings

The linguistic analysis shows that the verbal component of the cards is characterised by the very frequent presence of explicit and prototypical head acts of gratitude. Indeed, most of the gratitude-related expressions are based on the morpheme *thank*. This suggests that the spatial constraints and face-enhancing function of thank-you cards may favour verbal expressions that are immediately clear as to the sentiment being conveyed. Only a few creative, playful adaptations (e.g., *fanks*, *thankxolotl*) are attested, which can be easily understood when anchored to the visual mode. This shows, on one hand, that the verbal and visual modes of expression are complementary and co-dependent, and on the other, that humour operates as an auxiliary resource for affective engagement.

Lexical upgraders modify over 40% of the head acts. They include conventional intensifiers (*very much*, *so*), and inventive ones derived from plays on words (*berry much*, *so matcha*), reinforced by the visual component. This suggests a synergistic relationship between lexis and image, where the former hints at humour or affection, which can only be fully processed when considered together with the latter. The frequency of playful engagement as the primary visual-verbal relationship reveals humour's central role in transforming gratitude into relational bonding. The cards achieve this through intersemiotic wordplay, which requires cognitive collaboration between sender and recipient to decode visual-verbal puns. This shared interpretive

work points to an investment in the relationship beyond mere acknowledgement of benefits received.

Supportive moves are not only infrequent, but also restricted to compliments, general reflections, and statements of emotional impact. The cards thus appear to be verbally focused on positive affect and celebration rather than the negotiation of social obligations, serving more to boost than to redress social relations. This verbal restraint, combined with the predominance of AES and HUM over SEM, suggests that the cards serve as a means for managing social relationships.

Typographic features are a pervasive feature of the cards, used with subtle variations, both as single formatting choices and in combinations. The most common is CAP, followed by BOL and ITL. They add visual salience to the textual component, and may be considered a form of visual emphasis. The frequent combinations of multiple typographic resources not only reinforce the multimodal nature of the card in general, but also specifically draw the reader's attention to the *verbal* component of the cards, which might otherwise go unnoticed. This typographic strategy works in concert with compositional choices that position visual elements centrally for immediate impact, while ensuring the verbal gratitude expressions remain prominent. This reflects the cards' dual function of providing visual pleasure while maintaining clear communicative intent.

While the mutual reinforcement of words and images is frequently attested, the actual replacement of words with pictures occurs only twice (Cards 22 and 51) as a euphemistic device to avoid using four-letter words. This gives rise to multimodal sentences (cf. D'Ulizia *et al*. 2007) in which the subject and predicator are expressed verbally, while the direct object or adverbial are rendered visually.

In general, the visual components in multimodal thank-you cards serve informational-functional roles that extend far beyond mere decoration or illustration. They operate as active meaning-making resources that disambiguate verbal wordplay, amplify abstract gratitude concepts through concrete imagery, and intensify emotional engagement through aesthetic appeal and anthropomorphised characters.

Overall, the description of the verbal component suggests that the cards are lexically and typographically optimised for immediacy and clarity, while intersemiotic relationships and visual aesthetics serve as supplementary resources for emotional engagement, producing humour and enhancing visual appeal.

The analytical framework recognises that thank-you cards employ multiple strategies for coordinating visual and verbal elements, ranging from simple reinforcement to complex interplay, that create sophisticated pragmatic effects. Our analysis reveals that multimodal texts cannot be fully understood through additive approaches that merely combine verbal and

visual meaning. Rather, intersemiotic relationships create emergent pragmatic effects that exceed the sum of individual modal contributions; as a result, the cards' effectiveness depends on readers' ability to integrate multiple semiotic resources into coherent communicative experiences.

5.2. Concluding remarks

The integrated linguistic, visual and intersemiotic analysis has provided a holistic representation of a small corpus of thank-you cards, showing that typographic prominence, concise explicit gratitude, and visually anchored humour are key communicative strategies for engaging audiences.

The brevity of the texts, the conventionality of their gratitude expressions, and the limited number of their supportive moves suggest that the primary goal of such cards is to convey the message at a glance, rather than elaborate on it. Also, the occasional recourse to humour indicates that an additional function of such a message is to provide entertainment (i.e., aesthetic appreciation), possibly as a form of gift-giving or reciprocation meant to please the addressee further.

Our analysis also reveals that politeness strategies operate through intermodal relations. While verbal elements maintain clarity and immediacy through explicit expressions of gratitude, visual elements serve affective and relational goals through aesthetic enhancement and humour creation. The affordances of individual modes and their co-deployment allow the cards to ensure comprehensible illocutionary force and to maximise emotional engagement.

This study presents a few limitations. First, the corpus we analysed comprises exclusively English-language cards, limiting our understanding of how the multimodal properties identified might manifest across different linguistic and cultural contexts.

Second, the data collection process introduced potential biases that may have influenced the representativeness of the corpus. The cards collected through personal networks reflect the aesthetic preferences and social demographics of the collectors, while the *Etsy* selection process, guided by platform algorithms and search parameters, probably favoured commercially popular designs over niche or experimental approaches. Additionally, the time gap between the personal and online collections of the cards precluded the examination of a unified time period, potentially introducing variations related to changing design trends or seasonal factors.

Third, the corpus size of 97 cards, while sufficient for identifying prominent patterns, does not allow us to make definitive claims about frequency distributions or rare phenomena within the genre. A larger dataset might reveal different proportions of intersemiotic relationship types or uncover additional multimodal strategies not represented in this sample.

Future research could examine thank-you cards from additional sources, in other languages, and across a broader price range to determine their degree of genre-internal variation. Additionally, in line with Abdullah and Nor (2022), investigations could be conducted into the digital manifestations of gratitude, such as e-cards and memes, to examine whether and how they replicate or adapt the multimodal strategies observed in print cards, including verbal-visual wordplay and expressive typography. Furthermore, reception studies could be conducted to investigate whether readers notice the content, humour, typographic emphasis, and visual characteristics of thank-you cards, and how they cognitively and emotionally respond to them. Finally, researchers could analyse cards associated with other speech acts, such as apologies, congratulations, or condolences, to identify both their shared and distinctive features.

Bionotes: Elisa Bertoldi is a postdoctoral researcher at the University of Padua.
Sara Gesuato researches and teaches English linguistics at the University of Padua.

Authors' addresses: elisa.bertoldi.1@unipd.it, sara.gesuato@unipd.it

Acknowledgements: We thank two anonymous reviewers for helpful comments on our paper.

References

Abdullah S. and Nor J.M. 2022, *Exploring Non-linguistic Features of Malay Hari Raya E-cards,* in "International Journal of Academic Research in Business and Social Sciences" 12 [1], pp. 1076-1083.

Ahar V. and Eslami-Rasekh A. 2011, *The Effect of Social Status and Size of Imposition on the Gratitude Strategies of Persian and English Speakers*, in "Journal of Language Teaching & Research" 2 [1], pp. 120-128.

Al-Ali M.N. 2006, *Conveying Academic and Social Identity in Graduate Dissertation Acknowledgements*, in Neumann C.-P., Pérez-Llantada Auría C. and Alastruér R.P. (ed.) *Proceedings of the Fifth International AELFE Conference*, Prensas Universitarias de Zaragoza, University of Zaragoza, pp. 35-42.

Al-Subhi A.S. 2024, *Interactional Multimodal Metadiscourse in Public Health Posters During The COVID-19 Pandemic*, in "Pragmatics & Society" 16 [2], pp. 255-281.

Atminytė E. 2022, *Soviet Lithuanian New Year Greeting Card: Semiotics of Recto*, in "Semiotika" 17, pp. 162-185.

Ben-Ari E. 1987, *On Acknowledgements in Ethnographies*, in "Journal of Anthropological Research" 43 [1], pp. 63-84.

Blum-Kulka S., House J. and Kasper G. 1989, *Cross-cultural Pragmatics: Requests and Apologies*, Ablex Publishing Corporation, Norwood (NJ).

Bouko C. 2020, *Emotions through Texts and Images: A Multimodal Analysis of Reactions to the Brexit Vote on Flickr*, in "Pragmatics" 30 [2], pp. 222-246.

Brown P. and Levinson S.C. 1978, *Universals in Language Use: Politeness Phenomena,* in Goody E.N. (ed.), *Questions and Politeness: Strategies in Social Interaction*, Cambridge University Press, Cambridge, pp. 56-289.

Cheng S.W. 2005, *An Exploratory Cross-sectional Study of Interlanguage Pragmatic Development of Expressions of Gratitude by Chinese Learners of English*, PhD dissertation, University of Iowa.

Cohn N., Murthy B. and Foulsham T. 2017, *A Picture is Worth More Words Over Time: Multimodality and Narrative Structure Across Eight Decades of American Superhero Comics*, in "Multimodal Communication" 6 [1], pp. 1-18.

Coulmas F. 1981, *"Poison to your soul". Thanks and Apologies Contrastively Viewed*, in Coulmas F. (ed.), *Conversational Routine. Explorations in Standardized Communication Situations and Prepatterned Speech*, Mouton, The Hague/Paris/New York, pp. 69-91.

Cronin B. 1995, *The Scholar's Courtesy. The Role of Acknowledgement in the Primary Communication Process*, Taylor Graham, London/Los Angeles.

Del Saz-Rubio M.M. 2018, *Female Identities in TV Toiletries Ads: A Pragmatic and Multimodal Analysis of Implied Meanings,* in *"*Journal of Pragmatics" 136, pp. 54-78.

Díaz Pérez F.J. 2005, *The Speech Act of Thanking in English: Differences Between Native and Non-native Speakers'* B*ehaviour*, in "Revista de Filología Inglesa" 26, pp. 91-102.

D'Ulizia A., Ferri F. and Grifoni P. 2007, *A Hybrid Grammar-based Approach to Multimodal Languages Specification*, in Meersman R. *et al*. (ed.), *On the Move to Meaningful Internet Systems 2007: OTM 2007 Workshops. OTM 2007. Lecture Notes in Computer Science*, Springer-Verlag, Berlin/Heidelberg, pp. 367-376.

Eisenstein M. and Bodman J.W. 1986, *"I very appreciate": Expressions of Gratitude by Native and Non-native Speakers of American English*, in "Applied Linguistics" 7

[2], pp. 167-185.
Eisenstein M. and Bodman J.W. 1993, *Expressing Gratitude in American English*, in Kasper G. and Blum-Kulka S. (ed.), *Interlanguage Pragmatics*, Oxford University Press, New York/Oxford, pp. 64-81.
Etsy. 2025, *Etsy – Global Digital Marketplace*, https://www.etsy.com/ (last accessed January 2025).
Forceville C. and Kjeldsen J.E. 2018, *The Affordances and Constraints of Situation and Genre: Visual and Multimodal Rhetoric in Unusual Traffic Signs*, in "International Review of Pragmatics" 10 [2], pp. 158-178.
Francesconi S. 2011, *Multimodally Expressed Humor Shaping Scottishness in Tourist Postcards*, in "Journal of Tourism and Cultural Change" 9 [1], pp. 1-17.
García O. and Kleifgen J.A. 2016, *Translanguaging and Literacies*, in "Reading Research Journal" 0 [0], pp. 1-19.
Giannoni D.S. 1998, *The Genre of Journal Acknowledgements: Findings of a Crossdisciplinary Investigation*, in "Linguistica e Filologia" 6, pp. 61-84.
Giannoni D.S. 2002, *Worlds of Gratitude: A Contrastive Study of Acknowledgement Texts in English and Italian Research Articles*, in "Applied Linguistics" 23 [1], pp. 1-31.
Giannoni D.S. 2006, *Book Acknowledgements Across Disciplines and Texts*, in Hyland K. and Bondi M. (ed.), *Academic Discourse Across Disciplines*, Peter Lang, Berlin, pp. 151-175.
Gillen, J. 2013, *Stoic and Sentimental: The Emotional Work of the Edwardian Greetings Postcard*, in "Journal of Social History" 47 [2], pp. 313-341.
Goffman E. 1981, *Forms of Talk*, University of Pennsylvania Press, Philadelphia.
Guo F.B. and Feng X. 2017, *A Multimodal Discourse Analysis of Advertisements-based on Visual Grammar*, in "Journal of Arts and Humanities" 6 [3], pp. 59-69.
Held G. 1996, *Two Polite Speech Acts in Contrastive View: Aspects of the Realization of Requesting and Thanking in French and Italian*, in Hellinger M. and Ammon U. (ed.), *Contrastive Sociolinguistics*, Mouton de Gruyter, Berlin/New York, pp. 363-384.
Hidalgo-Downing L. and Kraljević Mujic B. 2015, *Multimodal Metaphor and Fictional Storytelling in Printed and Internet Ads: Recontextualizing Social Practices and Globalization*, in "Revista Brasileira de Linguística Aplicada" 15, pp. 377-402.
Hiippala T., 2007, *Linguistic and Multimodal Insights into Tourist Brochures,* in "Finnish Journal of Tourism Research", 10 [2], pp. 62-68.
Hyland K. 2003, *Dissertation Acknowledgements: The Anatomy of a Cinderella Genre*, in "Written Communication" 20 [3], pp. 242-268.
Hyland K. 2004, *Graduates' Gratitude: The Generic Structure of Dissertation Acknowledgements*, in "English for Specific Purposes" 23, pp. 303-324.
Hyland K. and Tse P. 2004, *'I would like to thank my supervisor'. Acknowledgements in Graduate Dissertations*, in "International Journal of Applied Linguistics" 14 [2], pp. 259-275.
Kress G. and van Leeuwen T. 2021, *Reading Images. The Grammar of Visual Design*, 3rd ed., Routledge, London/New York.
Lambeens T. and Pint K. 2015, *The Interaction of Image and Text in Modern Comics*, in Lardinois A., Levie S., Hoeken H. and Lüthy Ch. (ed.), *Texts, Transmissions, Receptions: Modern Approaches to Narratives*, Brill, Leiden, pp. 240-257.
Lim F.V. 2021, *Investigating Intersemiosis: A Systemic Functional Multimodal Discourse Analysis of the Relationship Between Language and Gesture in Classroom Discourse*, in "Visual Communication" 20 [1], pp. 34-58.
Liu Y. 2023, *Analysing Tension Between Language and Images: A Socialsemiotic View*, in

"Social Semiotics" 33 [1], pp. 132-150.
Liu Y. and O'Halloran K. L. 2009, *Intersemiotic Texture: Analyzing Cohesive Devices Between Language and Images*, in "Social Semiotics" 19 [4], pp. 367-388.
McWhorter J. 2012, *A Surprising New Language – Texting* https://www.youtube.com/watch?v=yoF2vdLxsVQ (8.8.2025).
Moya-Guijarro A.J. 2016, *A Multimodal Analysis of Picture Books for Children: A Systemic Functional Approach*, Equinox, Sheffield.
Nikolajeva M. and Scott C. 2006, *How Picturebooks Work*, Routledge, London and New York.
Oakey D., Jones C. and O'Halloran K.L. 2022. *Phraseology and Imagery in UK Public Health Agency COVID-19 Tweets,* in O'Halloran K.L., Jones C. and Oakey D. (ed.), *Discourses, Modes, Media and Meaning in an Era of Pandemic,* Routledge, London, pp. 89-114.
Okamoto S. and Robinson W.P. 1997, *Determinants of Gratitude Expressions in England*, in "Journal of Language and Social Psychology" 16 [4], pp. 411-433.
Padlet. 2025, *Padlet – Online Collaboration Platform*, https://it.padlet.com (last accessed January 2025).
Schauer G.A., Adolphs S. 2006, *Expressions of Gratitude in Corpus and DCT Data: Vocabulary, Formulaic Sequences, and Pedagogy*, in "System" 34, pp. 119-134.
Searle J.R. 1976, *A Classification of Illocutionary Acts*, in "Language in Society" 5 [1], pp. 1-23.
Sipe L.R. 1998, *How Picture Books Work: A Semiotically Framed Theory of Text-Picture Relationship*, in "Children's Literature in Education" 29 [2], pp. 97-108.
Skwarzyński J. 2019, *Reading Images. Comics and Its Multimodality in Cultural Communication, Interpretation and Translation*, in "New Horizons in English Studies" 4, pp. 102-117.
Stöckl H. 2024, *Detecting Generic Patterns in Multimodal Argumentation: A Corpus-based Study of Environmental Protection Print-Ads*, in "Journal of Argumentation in Context" 13 [2], pp. 260-291.
Stöckl H. and Bateman, J.A. 2022, *Editorial: Multimodal Coherence Across Media and Genres*, in "Front. Commun" 7 [1104128], pp. 1-2.
Tseronis A. 2015, *Multimodal Argumentation in News Magazine Covers: A Case Study of Front Covers Putting Greece on the Spot of the European Economic Crisis*, in "Discourse, Context & Media" 7, pp. 18-27.
Unsworth L. and Cléirigh C. 2017, *Multimodality and Reading: The Construction of Meaning Through Image-Text Interaction*, in Jewitt C. (ed.), *The Routledge Handbook of Multimodal Analysis,* 2nd edn., Routledge, London, pp. 176-188.
Vásquez C. and Aslan E. 2021, *"Cats be outside, how about meow": Multimodal Humor and Creativity in an Internet Meme*, in "Journal of Pragmatics" 171, pp. 101-117.
Yu H. 2019, *One Page, Two Stories: Intersemiotic Dissonance in a Comic Adaptation of Journey to the West*, in "Social Semiotics" 31 [2], pp.1-21.