Int. j. adv. multidisc. res. stud. 2026; 6(6):747-752

Received: 28-09-2026
Accepted: 08-11-2026

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

Operationalising Persuasion in Multimodal Advertising: A Coding Framework for Cross-Linguistic Analysis

Thai Thi Xuan Ha
University of Transport and Communications, Hanoi, Vietnam

Corresponding Author: **Thai Thi Xuan Ha**

Abstract
Catalogues of persuasion techniques are widely cited in advertising research, but they were designed for practitioners rather than for analysts. As a result, techniques are often assigned to advertisements as impressionistic labels, and the problem is compounded in cross-linguistic research, where the same technique may be realised by different linguistic means and in different modes. This paper proposes a coding framework that operationalises fifteen persuasion techniques, drawn from the catalogue of Andrews, Van Leeuwen and Van Baaren, for the analysis of static multimodal advertisements in two or more languages. The framework rests on five design principles: indicator-based definitions, separate coding of the verbal and visual modes, language-neutral indicators paired with language-specific realisation inventories, an explicit unit of analysis, and controlled multiple coding. For each technique, the codebook specifies an operational definition together with verbal and visual indicators grounded in systemic functional grammar and the grammar of visual design. Realisation inventories are provided for English and Vietnamese, two typologically distant languages. The paper further sets out six decision rules for boundary cases and a coding procedure that includes pilot coding and reliability assessment with Krippendorff's alpha. A worked example with two constructed posters shows how mode-separated coding can reveal that two languages use the same technique through different modes, a difference that frequency counts of techniques alone would conceal.

Keywords: Advertising Discourse, Codebook, Content Analysis, Cross-Linguistic Analysis, Multimodality, Persuasion Techniques

1. Introduction

Persuasion is the defining purpose of advertising discourse (Cook, 2001; Goddard, 1998) [4, 5]. To describe how this purpose is achieved, researchers increasingly draw on catalogues of persuasion techniques from marketing and social psychology. One of the most comprehensive is that of Andrews, Van Leeuwen and Van Baaren (2013) [1], who describe 33 psychological influence techniques grouped into three levels of need (instinctive, social and personal) based on Maslow's hierarchy. Such catalogues are valuable because they link the form of an advertisement to the psychological mechanism it is intended to trigger.

These catalogues, however, were written for advertising practitioners. They define techniques in terms of the effect they aim to produce, illustrate them with memorable examples, and leave open how an analyst should decide whether a given advertisement uses a given technique. When they are applied to research data, two problems arise. First, coding becomes impressionistic: two analysts may disagree about whether a smiling face realises *trustworthiness*, *attractiveness* or neither, and there is no rule to settle the disagreement. Second, coding usually records only whether a technique is present, not *how* it is realised. In multimodal texts, the same technique may be realised in the written text, in the image or in both.

Both problems become more serious in cross-linguistic research. In earlier work on English and Vietnamese advertising (Thái, forthcoming), I argued that the two languages may perform the same function through different modes. For example, Vietnamese can position the viewer as a parent through kinship terms of address, whereas English has no ready grammatical means of doing so and must rely on the image. That study proposed a preliminary mapping between persuasion techniques and their means of realisation. If such differences are systematic, a comparison based only on the frequency of techniques would show the two languages as similar while missing a real difference in how they persuade.

This paper develops that preliminary mapping into a full coding framework. Its aim is to operationalise fifteen techniques from Andrews *et al.* (2013) [1] for the analysis of static multimodal advertisements in more than one language. Specifically, the

paper (i) sets out the design principles of the framework, (ii) presents a codebook with operational definitions and verbal and visual indicators, (iii) provides realisation inventories for English and Vietnamese, (iv) formulates decision rules for boundary cases, and (v) describes a coding and reliability procedure. A worked example illustrates how the framework is applied.

2. Background

2.1 Persuasion techniques in advertising

Andrews *et al*. (2013) [1] organise their techniques by the level of need they address. *Instinctive* techniques exploit basic cognitive tendencies, such as the preference for information that is easy to process or has been encountered before. *Social* techniques exploit our orientation towards others, such as the tendency to follow the majority or to trust authorities. *Personal* techniques exploit individual desires and plans, such as the attraction of an ideal state or of an added benefit. Because the present framework is designed for static posters, it retains fifteen techniques that can be realised without sound, movement or interaction, five from each group. Techniques that depend on sequential exposure or on the viewer's response, such as foot-in-the-door or door-in-the-face, are excluded.

Tanaka (1994) [15] showed that the persuasive strategies of advertisements vary across cultures, and that this variation can be studied linguistically. The present framework builds on this insight, but approaches it through systematic coding rather than through the interpretation of individual cases.

2.2 Multimodal analysis as a source of indicators

To be coded reliably, a technique must be tied to observable features of the text. Systemic functional grammar (Halliday & Matthiessen, 2004) [6] and the grammar of visual design (Kress & Van Leeuwen, 2006) [9] provide a vocabulary for such features. On the verbal side, relevant features include process types, mood, modality and forms of address. On the visual side, they include participant types, narrative and conceptual processes, gaze (demand or offer), social distance, modality and information value. Because the two grammars share the same metafunctional architecture, features from both modes can be described in comparable terms (Jewitt *et al*., 2016) [8]. The framework uses these features as the building blocks of indicators.

2.3 Content analysis and reliability

The framework adopts the methodological standards of content analysis. Krippendorff (2004) [10] and Neuendorf (2017) [12] stress that categories must be defined explicitly enough for independent coders to apply them consistently, and that the consistency of coding must be measured rather than assumed. Bell (2001) [3] shows how these principles apply to visual content, and Bateman (2008) [2] argues that multimodal analysis needs explicit, corpus-oriented annotation schemes if its claims are to be tested. Following Hayes and Krippendorff (2007) [7], the framework uses Krippendorff's alpha as its reliability measure, because it accommodates any number of coders, missing data and different levels of measurement.

3. Design principles

The framework is built on five principles.

Principle 1: indicator-based definitions. Each technique is defined by observable indicators rather than by the effect the advertiser presumably intended. Coders do not ask whether the poster *aims to* build trust, but whether it contains a specified feature, such as a human participant in close-up with a demand gaze. Intention is inferred from form, not the other way round.

Principle 2: separate coding of modes. For each technique, presence is coded separately for the verbal mode and the visual mode. A technique may therefore be coded as verbal only (V), visual only (I) or both (V+I). This is the principle that allows the framework to detect cross-linguistic differences in the division of labour between modes.

Principle 3: language-neutral indicators with language-specific inventories. Indicators are formulated at a level that applies to any language, for example *a directive softened by a mitigating device*. The forms that realise the indicator in each language are listed in separate realisation inventories (Section 5). This keeps the categories comparable across languages, in line with the requirement of a *tertium comparationis* in contrastive analysis (Krzeszowski, 1990) [11], while respecting typological differences.

Principle 4: an explicit unit of analysis. The unit of analysis is the single static poster. Within it, the verbal mode comprises the brand name, headline, slogan, body copy and any text on packaging or badges; the visual mode comprises represented participants (human, animal, mascot), the product, the setting and the layout.

Principle 5: controlled multiple coding. A poster may realise several techniques, and all are coded. However, a single element may count towards more than one technique only when it meets the full definition of each. Rule R1 (Section 6) settles the remaining cases.

4. The codebook

Tables 1 to 3 present the codebook for the three need groups. For each technique, the table gives an operational definition and the verbal and visual indicators. A technique is coded as present in a mode when at least one indicator for that mode is found, unless a decision rule in Section 6 specifies otherwise.

Table 1: Codebook: Techniques addressing instinctive needs

Technique	Operational definition	Verbal indicators	Visual indicators
Fluency	The message is designed to require minimal processing effort	Headline of one clause or less Rhyme, alliteration or parallelism Numerals in place of words	One clear focal point Three or fewer salient elements Plain or blurred background
Mere exposure	A brand element is repeated within the poster beyond the genre default	Brand name appears two or more times Brand name embedded in the slogan	Logo appears two or more times Brand colour dominates the whole layout
Anchoring	The product's value is presented relative to an explicit reference point	Original and reduced price Comparative quantities (*twice as much*)	Struck-through price Product placed next to an object of comparison
Trustworthiness	A source presented as sincere vouches for the product	First-person guarantee by the brand (*we promise*) Claims of transparency (*no added sugar*)	Human face with a demand gaze Close or medium shot High naturalistic modality
Primacy & recency	A product claim, not merely the brand name, occupies the first or last position	Claim as the first or last element of the verbal sequence	Claim placed in the top (Ideal) or bottom (Real) zone and made salient

Source: Developed by the author from Andrews *et al*. (2013) [1] and Kress & Van Leeuwen (2006) [9]

Table 2: Codebook: Techniques addressing social needs

Technique	Operational definition	Verbal indicators	Visual indicators
Social proof	The product is presented as chosen by many people	User numbers, sales figures, votes Majority expressions (*everyone's favourite*)	Groups of people using the product Images presented as user-generated content
Attractiveness	A physically attractive participant is foregrounded	Evaluation of appearance (*glowing, slim*)	Professional model as the most salient participant Close-up with idealised lighting
Fleeting attraction	The poster signals similarity between the source and the target group	In-group register (slang, youth language) Inclusive or kinship terms of address	Participants matching the target group in age or life situation Everyday, recognisable settings
Anthropomorphism	The product or brand is given human qualities	Product as Senser or Sayer Product speaks in the first person	Mascot or product with face or limbs Action vector originating from the product Speech bubble attached to the product
Authority	An expert, institution or celebrity endorses the product	Expert quotation Certification, award, institutional name	Participant in expert attire (e.g. white coat) Recognisable celebrity presented as endorser

Source: Developed by the author from Andrews *et al*. (2013) [1] and Kress & Van Leeuwen (2006) [9]

Table 3: Codebook: Techniques addressing personal needs

Technique	Operational definition	Verbal indicators	Visual indicators
Promised land	The product is linked to an ideal state that goes beyond its attributes	Mental or relational processes construing an ideal state (happiness, success, freedom)	Symbolic process Idealised setting Ideal zone featuring a state rather than the product
Humour	The poster contains an incongruity intended to amuse	Pun or wordplay Unexpected collocation	Incongruous narrative situation Exaggerated proportions
That's not all	A benefit is added on top of the product	Free gift, bonus, combo, extra quantity	Bonus item shown alongside the product Starburst or badge announcing the extra
Implementation intentions	The viewer is told when, where or on what occasion to use the product	Directive accompanied by a circumstance of time, place or occasion	A concrete usage situation with cues of time or place (breakfast table, beach)
God terms	High-value evaluative words are attached to the product	Positive evaluative lexis of high intensity (*new, natural, perfect, pure*)	Not applicable (verbal-only technique)

Source: Developed by the author from Andrews *et al*. (2013) [1] and Kress & Van Leeuwen (2006) [9]

5. Realisation inventories for English and Vietnamese

Several verbal indicators in the codebook are defined functionally, for example *a directive softened by a mitigating device* or *inclusive or kinship terms of address*. Their concrete forms differ between languages. English marks mood mainly through the order of Subject and Finite, and its imperative uses the base form of the verb. Vietnamese, an isolating language, signals mood largely through adverbs and sentence-final particles, which also convey intimacy, urgency or mitigation (Nguyễn, 2008; Palmer, 2001) [13, 14]. Vietnamese also offers a rich system of kinship terms used as forms of address, which has no grammatical counterpart in English.

Table 4 provides starting inventories for the indicators where the two languages differ most. The inventories are open lists: they are intended to be extended during pilot coding, and any additions should be documented in the codebook.

Table 4: Realisation inventories for selected verbal indicators

Indicator	English	Vietnamese
Inclusive address	*we, us, let's, our*	*chúng mình, mình, chúng ta, cả nhà*
Kinship address positioning the viewer	No grammatical equivalent; realised visually or through nouns (*mums, kids*)	*mẹ* (mother), *con* (child), *bố/ba* (father), *anh/chị/em* (sibling terms)
Softened directive	*why not, just, go ahead, treat yourself*	Sentence-final particles *nhé, nha, đi, thôi*; preverbal *hãy*
Immediacy	*now, today, instantly*	*ngay, liền, ngay hôm nay*
Added value	*free, plus, bonus, extra*	*tặng, kèm, thêm, miễn phí*
God terms	*new, natural, pure, fresh, perfect*	*mới, tự nhiên, nguyên chất, tươi, tuyệt hảo*

Source: Compiled by the author

One further feature is specific to the Vietnamese data: the insertion of English words or slogans into Vietnamese advertisements. Such insertions should not be coded as a technique in their own right. Instead, each inserted item is coded under the technique whose indicator it meets, for example *fresh* under God terms, and is additionally flagged as an English insertion so that its contribution can be analysed separately.

6. Decision rules

The following rules settle the boundary cases most likely to cause disagreement between coders.

R1. Specificity. When one element meets the indicators of two techniques, it is coded under the technique whose definition it matches more specifically. For example, a doctor in a white coat looking at the viewer meets indicators of both Trustworthiness and Authority; because expert attire is specific to Authority, it is coded as Authority only. The element counts for both techniques only if it independently meets the full definition of each.

R2. Genre default. Features present in virtually every poster of the genre are not coded unless they are intensified. A single logo in a corner does not count as Mere exposure; a brand name at the bottom of the poster does not count as Primacy & recency.

R3. Co-occurrence for Fluency. Because most posters are designed to be easy to read, Fluency is coded in a mode only when at least two of its indicators co-occur in that mode.

R4. Independence of modes. The verbal and visual modes are coded separately, even when they realise the same technique. A poster that announces a free gift in writing and also shows the gift is coded as That's not all (V+I), not as That's not all (V).

R5. Evidence. Every positive code must be supported by a reference to the element that triggered it, recorded in the coding sheet. For techniques that require interpretation, notably Humour and Attractiveness, the coder must also describe the incongruity or the feature judged attractive.

R6. Original language. Verbal elements are coded in their original language, using the relevant realisation inventory. Translations may be used to report results but not as the basis for coding.

7. Coding procedure and reliability

The framework is applied in five steps.

Step 1: Recording. Each poster is assigned an identifier. All verbal elements are transcribed in the original language and labelled by type (brand name, headline, slogan, body copy, badge). The visual elements are described in a standard template covering participants, product, setting, gaze, size of frame and layout zones. Recording before coding ensures that all coders work from the same description of the text.

Step 2: Coder training and pilot coding. At least two coders proficient in both languages are trained on the codebook. They then independently code a pilot set of posters that is not part of the final sample. Disagreements are discussed, and definitions, indicators, inventories and decision rules are revised where needed. All revisions are recorded.

Step 3: Independent coding. Each coder codes the full sample, or a reliability subsample of it. For each of the fifteen techniques, the coder records presence or absence in the verbal mode and in the visual mode, together with the evidence required by R5. To reduce bias, each coder codes posters from both languages, presented in randomised order.

Step 4: Reliability assessment. Agreement is calculated with Krippendorff's alpha, separately for each technique and each mode. Following Krippendorff (2004) [10], $\alpha \geq .800$ is taken as acceptable, and values between .667 and .800 as sufficient only for tentative conclusions. A technique that falls below .667 is either revised and re-tested or excluded from the analysis, and the exclusion is reported.

Step 5: Resolution and analysis. Remaining disagreements are resolved by discussion, and the resolutions are documented. The final data allow two kinds of comparison between languages: the frequency of each technique, and the distribution of each technique across modes (V, I, V+I).

8. A worked example

To illustrate the framework, this section codes two posters. They are *constructed for illustrative purposes* and do not describe real advertisements. Both advertise a children's milk drink.

(1) *Poster E (English).* Headline: *Grow strong, starting today.* Badge: *No added sugar.* Image: a close-up of a girl holding a glass of milk and smiling directly at the viewer; plain background; the product pack in the bottom right corner.

(2) *Poster V (Vietnamese).* Headline: *Mẹ yên tâm, con khỏe mạnh mỗi ngày nhé!* ('Mum, rest assured, [your] child will be healthy every day PRT!') Badge: *Tặng kèm bình nước* ('Free water bottle included'). Image: a medium shot of a mother and child at a breakfast table, looking at each other rather than at the viewer; a water bottle shown beside the product pack.

Table 5 presents the coding of the two posters.

Table 5: Coding of the two constructed posters

Technique	Poster E	Poster V
Trustworthiness	I: child's face, demand gaze, close-up. V: transparency claim (*no added sugar*)	Not coded
Fleeting attraction	Not coded	V: kinship address (*mẹ, con*). I: participants matching the target group (parent and child)
Implementation intentions	V: directive with circumstance of time (*starting today*)	V: directive (*Mẹ yên tâm*) with circumstance of time (*mỗi ngày*). I: breakfast setting
That's not all	Not coded	V: *tặng kèm*. I: bonus item shown (R4)
Promised land	Not coded	I: idealised family scene

Source: Coded by the author

Three observations follow. First, both posters use Implementation intentions, but Poster E realises it in writing only, whereas Poster V realises it in both modes. Second, both posters establish a relation with the viewer, but by different means. Poster E does so visually, through the child's demand gaze, which is coded as Trustworthiness. Poster V does so verbally: the kinship terms *mẹ* and *con* position the viewer as a parent, while the participants in the image look at each other rather than out of the frame. Third, the decision rules shape the outcome. Under R4, That's not all in Poster V is coded as V+I rather than V. Under R1, the child's gaze in Poster E is coded as Trustworthiness and not also as Attractiveness, because the child is not presented as a professional model.

If these two posters were analysed only by counting techniques, they would appear to share one technique and differ on four. Mode-separated coding shows something more specific: in this pair, the interpersonal work that the English poster gives to the image is done by the grammar in the Vietnamese poster. Whether this pattern holds across a corpus is an empirical question, and the framework is designed to answer it.

9. Discussion

The framework makes three contributions. It turns a practitioner-oriented catalogue into a set of operational categories that independent coders can apply consistently. By coding modes separately, it allows researchers to compare not only *whether* languages use the same persuasion techniques but *how* they distribute them across writing and image. Through language-neutral indicators and language-specific inventories, it keeps categories comparable across typologically distant languages while respecting their differences.

The framework also has limitations. First, it inherits the scope of its source: the catalogue of Andrews *et al*. (2013) [1] is a synthesis of findings from psychology written for practitioners, not an empirically validated taxonomy, and the boundaries between some techniques are fuzzy. The decision rules mitigate this but cannot eliminate it. Second, some techniques, notably Humour and Attractiveness, still require interpretive judgement, and their reliability is likely to be lower than that of more formal categories. Third, coding Vietnamese and English requires coders with cultural as well as linguistic competence, which may limit the pool of available coders. Fourth, the framework is restricted to static posters; extending it to video or interactive advertising would require additional indicators for sound, movement and sequence. Finally, the framework has been presented here at the design stage; its categories and rules should be refined through pilot studies, and its reliability should be reported for each application.

Although the inventories in this paper cover English and Vietnamese, the architecture of the framework is not specific to these languages. Applying it to another language requires only a new realisation inventory, built from that language's grammatical resources for mood, modality, address and evaluation.

10. Conclusion

This paper has proposed a coding framework that operationalises fifteen persuasion techniques for the cross-linguistic analysis of static multimodal advertisements. The framework combines indicator-based definitions derived from systemic functional grammar and the grammar of visual design with separate coding of the verbal and visual modes, realisation inventories for English and Vietnamese, six decision rules and a reliability procedure based on Krippendorff's alpha. A worked example has shown that mode-separated coding can reveal differences in how two languages persuade that would remain invisible in counts of techniques alone. It is hoped that the framework will help make the analysis of persuasion in multimodal advertising more transparent, replicable and comparable across languages.

11. References

1. Andrews M, Van Leeuwen M, Van Baaren R. Hidden persuasion: 33 psychological influence techniques in advertising. BIS Publishers, 2013.
2. Bateman JA. Multimodality and genre: A foundation for the systematic analysis of multimodal documents. Palgrave Macmillan, 2008.
3. Bell P. Content analysis of visual images. In T. Van Leeuwen & C. Jewitt (Eds.), Handbook of visual analysis. Sage, 2001, 10-34.
4. Cook G. The discourse of advertising (2nd ed.). Routledge, 2001.
5. Goddard A. The language of advertising: Written texts. Routledge, 1998.
6. Halliday MAK, Matthiessen CMIM. An introduction to functional grammar (3rd ed.). Arnold, 2004.
7. Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. Communication Methods and Measures. 2007; 1(1):77-89.
8. Jewitt C, Bezemer J, O'Halloran K. Introducing multimodality. Routledge, 2016.
9. Kress G, Van Leeuwen T. Reading images: The grammar of visual design (2nd ed.). Routledge, 2006.
10. Krippendorff K. Content analysis: An introduction to its methodology (2nd ed.). Sage, 2004.
11. Krzeszowski TP. Contrasting languages: The scope of contrastive linguistics. Mouton de Gruyter, 1990.

12. Neuendorf KA. The content analysis guidebook (2nd ed.). Sage, 2017.
13. Nguyễn VH. Cơ sở ngữ nghĩa phân tích cú pháp [Semantic foundations of syntactic analysis]. Nhà xuất bản Giáo dục, 2008.
14. Palmer FR. Mood and modality (2nd ed.). Cambridge University Press, 2001.
15. Tanaka K. Advertising language: A pragmatic approach to advertisements in Britain and Japan. Routledge, 1994.