CHALLENGES IN CODING DISCOURSE DATA: MY EXAMPLES AND QUESTIONABLE SOLUTIONS

Jo Mackiewicz
CCCC
14 March 2013

Coding data is at the heart of being human: “Division and classification are strategies people use to organize their experience of the world” (Grant-Davie 1992, 272).

“While such definitions are generally clear, precise and easily comprehensible when presented in connection with carefully chosen examples, when it comes to applying such categories to real-life data, their application can feel somewhat arbitrary. Naturally occurring data rarely fits into these pre-formed categories neatly” (Stadler 2011, 37).

CHALLENGES

- Deciding on single or multiple codes per item
- Identifying instances versus coding discrete items in lists
- Determining code and category boundaries
- Refining codes and categories for optimal granularity

APPROACHES

Coding Scheme Type	Data Type for Coding
A: Codes and categories come from/ modified from prior research	1: Lexicon/syntax
B: Codes and categories developed during the study	2: Function

Formula	Abbreviation	Constructed Example
Noun phrase + *is/looks* + (*really*) + adjective	NP *is* ADJ	*The introduction is really good.*
Pronoun *is* (*really*) *a* adjective + noun phrase	PRO *is a* ADJ NP	*It is a really good introduction.*
You + verb + (*a*) + (*really*) + adjective + noun phrase	*You* V *a* ADJ NP	*You wrote a really good introduction.*
Adjective + noun phrase	ADJ NP	*Good introduction.*
You + verb + (noun phrase) + (*really*) + adverb	*You* V NP ADV	*You organized the introduction really well.*
I + (*really*) + *like* + noun phrase	*I like* NP	*I really like the introduction.*

A1: Syntactic codes developed from previous research

Formula	Abbreviation	Constructed Example
Noun phrase + *is/looks* + (*really*) + adjective	NP *is* ADJ	*The introduction is really good.*
Pronoun *is* (*really*) *a* adjective + noun phrase	PRO *is a* ADJ NP	*It is a really good introduction.*
You + verb + (*a*) + (*really*) + adjective + noun phrase	*You* V *a* ADJ NP	*You wrote a really good introduction.* *You delivered a stellar talk that left them wanting to discuss coding for hours after.*
Adjective + noun phrase	ADJ NP	*Good introduction.*
You + verb + (noun phrase) + (*really*) + adverb	*You* V NP ADV	*You organized the introduction really well.*
I + (*really*) + *like* + noun phrase	*I like* NP	*I really like the introduction.*

A1: Syntactic codes developed from previous research

Definition of Level of Edit	Example
Substantive: Deals with the content of the document, including the coherence and consistency of the document's parts	*Since Epinions' readers come to our site seeking information that will help them to make an informed purchasing decision, we look for both technical information and the reviewer's user experience in product reviews here.*
Format: Ensures conformity in typography, layout, and visual elements	*Just wanted to mention that dividing a review into a few paragraphs helps the eyes of those reading and also helps organise thoughts you are expressing.*
Screening: Ensures basic correctness, including absence of misspellings, complete sentences	*I would suggest that you use a spell checker and remember to capitalize your I's.*
Language: Ensures clear and effective presentation of text, including grammar and syntax, usage, conciseness, abbreviations	*I can write a very informative review of a camera without saying camera more than once, maybe twice. Using the same word repetitiously is not only distracting it can and will turn readers away.*
Policy: Ensures conformity to the organization's policies, such as absence of derogatory comments and advertisements	*While I'm not implying anything, I am saying that the review could have been written from published specs rather than from real world experience which is the essence and a requirement of the Eps TOS.*

*7 comments addressed two or three levels of edit.

A2: Functional codes developed from previous research

Levels of Edit	Example of assertion
Language + Screening	*In the future, please check your grammar, capitalization and spelling. Reviews which are "overwhelmed" by errors are supposed to be rated as Not Helpful or Don't Show. I'm giving you a break with a "show" rating today.*
Substantive + Policy	*First, as Elzora said, please delete the review under the Mid South product listing. . . .* *I wanted to rate this review helpful or better, but I think our Epinion readers would benefit from a little more information on the product. It would help if you could add some additional detail such as weight, brand of engine, fuel capacity, how long it runs on a tank of fuel, etc...*

CHALLENGE: Deciding on whether to use multiple codes

SOLUTION: Determine priority: focus on individual or sample

Boundary of Evaluation	Example of Boundary Comment
Descriptive Comment	*The first three chapters offer a survey of literature and a **detailed** description of the site and the discursive culture found there* (Miller, *JBC*, 428).
Evaluative Comment	*However, rather than dissecting and focusing on the technological side of this 'land of the mundane,' Johnson **uniquely** places his gaze upon the commonly ignored aspect of the technological everyday: the user* (Pennell, *JBC*, 277).

CHALLENGE: Determining code boundaries

SOLUTION: Choose a context for interpretation: dictionary or data set

<I3> do you think that it would have been (.) easier to believe if it had been kind of more of a commonplace incident, like they just said, </**I3>** **<OMISSION, I2>** (3s) I can't think of a creation myth that has something that's normal. Because even in Genesis it's like, and... dust. **</I2>** **<M3>** You know? **</M3>**

A2: Functional types developed from previous research

CHALLENGE: Achieving agreement in identifying coded units and marking unit boundaries in transcripts

SOLUTION: Set rules to follow syntactic units, intonation units, t-units...

Definition of Tutoring Strategy	Example of Tutoring Strategy
Leading Question (Instruction): A question with very limited set of answers. The tutor knows the answer.	• *You're not actually using that name are you?* • *Where does the comma go in this sentence?*
Pump Question (Cognitive Scaffold): A question that allows multiple possible answers. The tutor does not know the answer.	• *How would you summarize this whole paragraph?* • *So what's the connection between that?*

A2: Functional types developed from previous research

CHALLENGE: Developing a code (validity) and achieving reliable identification

Definition of Tutoring Strategy	Example of Tutoring Strategy
Pump Question (Cognitive Scaffold): Tutors withhold their advice or part of the answer. Pumping can be constraining, or it can be open ended. We included leading questions in this category because they can act as pumps for thinking and require at least minimal responses from students.	• *You think that would be kind of a good progression?* • *What does the poem mean to you?*

SOLUTION: Refining and testing types and categories

Assertion of Expertise Type	Example
Regular experience with (use of) the product, (i.e., regularly using the product over some duration)	*My husband and I purchased this camera almost 4 years ago and it's still working great. It has taken thousand of pictures over the years.*
Testing of the product (i.e., using it to see what it can do, how well it performs)	*I have put my PowerShot SD850 through its paces. I've experimented with a wide range of conditions (light, temperatures, distance) and have been quite happy with my SD850.*

B2: Functional codes developed during the study

CHALLENGE: Creating useful codes and categories

Category of Assertion	Type
Relevant Role	Received or receiving formal training or education relevant to the product
	Relevant experience from a hobby (i.e., nonprofessional experience)
	Employment in a profession relevant to the product
	Association with someone who has expertise relevant to the product
	Conducted research on the product (e.g., online research)
Relevant Familiarity	Familiarity/ownership of previous versions of the specific product under review
	Familiarity/ownership of comparable product model
	Familiarity with the brand and brand's products (as opposed to the specific product under review)
Specific Experience	Regular experience with (use of) the product, (i.e., regularly using the product over some duration)
	Testing of the product (i.e., using it to see what it can do, how well it performs)

SOLUTION: Grouping codes into more useful categories and testing categories for reliability

“Ishikawa went into further detail, and although one may not agree with all of her decisions, the important point is that anyone reading her study has a good sense of how she handled correctness…. [I]t was helpful to learn from Ishikawa that others were grappling with problems of identifying an error, something not evident in most of the other articles” (Polio and Gass 1997, 504–505).

CONCLUSIONS AND IMPLICATIONS

- Describe coding methods in more detail (appendices and supplementary materials)
- Provide prototypical examples and boundary cases
- Describe procedure for determining inter-rater reliability in more detail (discrete list items or identified in context?)
- Consider coding scheme's usability in other contexts

THANK YOU.
QUESTIONS?

Jo Mackiewicz
mackiewicz@auburn.edu

Definition of Evaluative Comment	Example of Evaluative Comment
Compliment: A speech act that . . . attributes credit . . . usually the person addressed, for some 'good' (possession, characteristic, skill, etc.) that is positively valued by the speaker and the hearer (Brown & Levinson 1987, 446).	*Scholars and researchers interested in genre analysis, intercultural communication, and e-mail communication will find [book author's] book fascinating and rewarding.*
Criticism: The expression of dissatisfaction or negative comment on the volume [the book] (Hyland 2000, 44; Hyland & Hyland 2001, 186).	*The difficulty here is that in so short an overview, these snapshots of complex national cultures seem a bit simplistic, bordering on the stereotypic.*

A2: Functional codes developed from prior research

Ambivalent Evaluation	Example of Ambivalent Evaluation
(Criticism?) + Compliment	*Even though the demand/supply equation may seem simplistic,* ***Headrick carefully shows the complexity of innovation in communication*** (Bruckner, *JBTC,* 343).

CHALLENGE: Deciding on whether to use multiple codes

SOLUTION: Define unit of measure (e.g., independent clause)

<I4> So **maybe we could** even make a list of everything **we** know about their target audience. **</I4>**

<I1> You **could**, yeah, **I think**, that's **kind of** what you could mention, is, **you** [**know, maybe** not necessarily affect his party in any way … **</I1>**

A1: Lexical codes developed from previous research

CHALLENGE: Treating multiple occurrences within unit of measure

SOLUTION: Treat the unit as unit binary (or not)