

# The pragmatics of margin comments: An empirical study

or less formally

"No apostrophe", "No conclusion", "No issues":  
The dark art of understanding margin comments

SAFeSEA Project

Debora Field

# Problem

- **Eventual Goal:** Build an automatic margin comments generator
  - For formative feedback purposes not assessment
- **Resource:**
  - Corpus of assessed Master's degree essays
  - Assessment included addition of margin comments (in English)
- **Objectives:**
  - Automatically learn human margin comment generation strategies
    - **Design a classification scheme for margin comments**
    - Manually annotate the corpus comments
    - Computationally look for relationships between a comment's category and features of the essay part to which it points
  - Implement a margin comments generator informed by that learning

# Method

- **Gain overview** of what the corpus margin comments are like
- **Consult literature** in education and linguistics
- **Engage brain**
- **Design classification scheme**, informed by the data and the theory, that:
  - Labels the **marker's intention** behind a comment
  - Requires consideration of the **semantics** of a comment
- An examination of **linguistics not assessment**
  - Is *not* concerned with how likely a comment is to 'feed forward'
    - Influence learner's future essay writing performances
  - Is *not* concerned with labelling how explicit a comment is
  - Is *not* concerned with labelling how much explanation is included

# What is a margin comment?

- A message from an assessor positioned in the 'margin' of a piece of text produced by a learner
- Graphically points to a part of the learner text

on the individual as learning object, not questioning subject.  
expectation to group conformity that it expresses.

**Comment [r5]:** Excellent point!

- Comment content typically concerns the text part to which the comment points
  - (Though the associated text part is often not accurately marked)
- The corpus margin comments were added by markers to word-processed assignments using a digital commenting tool
- Note: we are classifying the comments *without* reference to the passages to which they point
  - Ensure comments in isolation can be usefully classified
  - Can essay properties be used to predict characteristics of margin comments?

# Some more corpus comments

It ~~however~~ agree that ~~they~~ students are intuitively able to use and navigate around the internet, and that they are more visually literate than previous generations. These skills need to be taught. ~~The only~~ different ~~difference~~ being is that they have access to technology, usual at home

Comment [OE8]: Why not?

purpose, into their studies, and  
e. Evans (2008) looked at the  
podcasts after a "traditional

Comment [D3]: Why bold?

not work for those individuals.  
be located with people who  
ited or recorded in any way.

Comment [REDACTED]: No  
conclusion

educational institutions in developing countries:  
*Journal of Education and Development* using  
no. 1, pp. 5-22.

Comment [A16]: No issues.

of PLE's of fellow H800 students:

Comment [C3]: No apostrophe

# Corpus overview

- Snapshot
  - 1,408 assessed university assignments
  - Argumentative essays submitted towards a Master's in Education
  - 13 different modules
  - Official word limits ranged from 500 to 4,000
  - 20 different assessors/markers
  - 24,387 margin comments (in English natural language (NL))
- Preliminary investigations of margin comments
  - Frequency counts ->
  - Pattern matching rules for clustering similar comments
  - Initial observations...

# Comments are friendly & polite

- 9,272 Positive-sounding adjectives
  - *good* (freq. 5,177), *interesting* (954)
    - (Blue italicised examples are terms in comments not whole comments)
- Contrast with 551 negative-sounding adjectives
  - *difficult* (133), *missing* (123)
- 7,276 conditional auxiliaries
  - Used to make an instruction sound like polite suggestion
  - *you might* (882), *I would* (330)
- 3,996 softeners (to soften impact of a criticism)
  - *perhaps* (863), *rather* (422)

# Skills-related & advisory

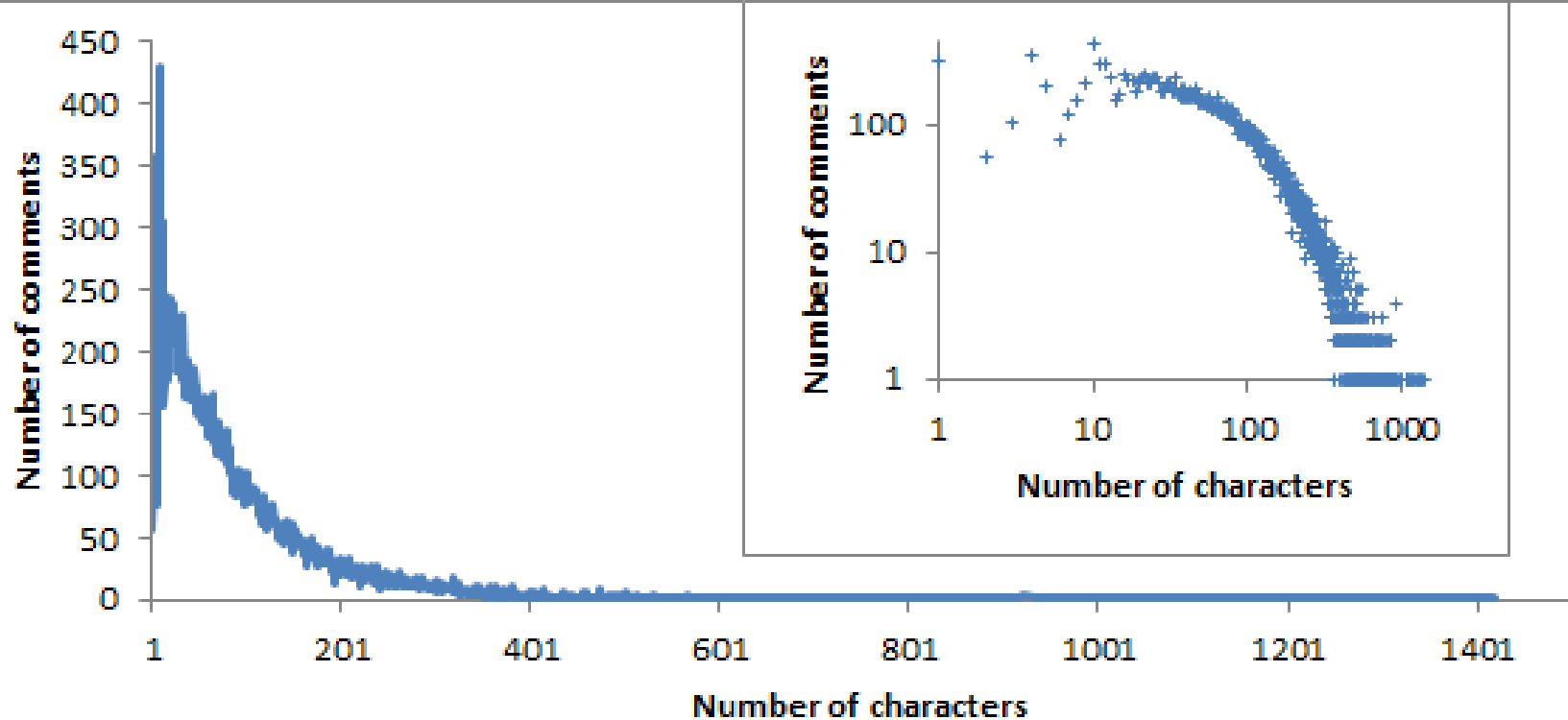
- **Argument** (14,705 comments): Content, arguing techniques, clarity
- **Referencing** (6,657): Situating work, referencing conventions
- **Essay structure** (5,243): Layout, scope, components
- **Presentation** (2,613): Grammar, punctuation, spelling, style
- 1,119 comments express **confusion or apparent uncertainty**
  - *not sure* (365), (*sure* (617))
- 1,232 comments concern **comprehensibility or clarity**
  - *clear* (908), *you mean* (500)
- 3,351 suggest something is missing that should be present
  - (E1) "Could you have developed this?"
    - (All red quoted examples are real, whole comments)
- 2,069 suggest something is present that should not be
  - (E2) "I would not leave a space."



# Look like utterances

- Predominance of very short comments »
  - A very small proportion are full sentences
    - (E3) "Avoid jargon"
    - (E4) "This is unclear."
  - Vast majority are non-sententials
    - (E5) "Why not?" (Ellipticals) (Klein, 1985; Merchant, 2004)
    - (E6) "Good point" (Fragments)
    - (E7) "What a good idea." (Others)
- Contractions (*don't*, *I'd*, ...) are common (3,818)
- Fillers (CI *ok* (444), CI *yes* (1,109)) are common
  - (E8) "OK well that's a good start."
- Questions are common
  - 4,307 comments end in a question mark
  - 1,109 comments begin with a WH question word
  - 1,119 comments begin with a polar (yes/no) question

# Comment lengths



- 9.5% of comments have 11 chars or fewer
- Top 3 most freq comment lengths
  - 10 chars (freq. 430), 4 (358), 1 (316)

# Are NL margin comments conversation?

- Pedagogy has argued that margin comments are or are like conversation (Ziv, 1984; Danis, 1987; Lindemann, 1987; Anson, 1989)
  - Movement away from 'teaching product' to 'teaching process' in 1980s
  - To encourage the expression of empathy with the learner
  - Believed to make more likely that teacher comments would be read and acted upon (Hairston, 1982)
- Straub's review 'Teacher response as conversation' (1996) concludes
  - Margin comments have a conversational style
  - They are not conversational utterances, real or imaginary
- Schegloff's (1999) definition of 'ordinary conversation' also excludes margin comments
  - Yes: 'Talk-in-interaction' (All talk with intention to communicate messages)
  - Yes: 'Speech exchange system' (Sacks, 1974) (lectures, classroom discourse, courts, meetings...)
  - No: Turn-taking, sequence organisation, repair organisation are missing
  - No: They {are} subject to functionally specific or context-specific restrictions

# And from pragmatics

- Conversation requires **common ground** (Stalnaker, 1972; Thomason, 1990)
  - A 'conversational record' that contains only public objects that have been explicitly entered into it
  - Each utterance results in the addition of new information to the common ground
- Conversation requires **accommodation** (Clark and Haviland, 1974; Lewis, 1979; Kamp, 1981)
  - New information entered into the common ground is accommodated by appropriate adjustment of belief state
- Conversation requires **grounding** (Clark and Schaefer, 1989)
  - Grounding is necessary to ensure that speaker and hearer's views of the common ground do not diverge
  - Grounding is achieved through a process of **presentation** (by S) and **acceptance** (by H)
  - Acceptance or non-acceptance demonstrated by H taking the next turn

- **Yes, there is common ground** in the transmission of margin comments
  - Marker M's comments are eventually perceived by the essay author A
  - But note that only M gets to 'speak'
- **Yes, there is accommodation** in the transmission of margin comments
  - A will read, try to understand, and to accommodate the comments (hopefully)
  - But note that only A gets to 'hear'
- **But no, there is no grounding**
  - The context constraints demand that A must accept the evidence
  - A does not get a turn to speak and demonstrate acceptance
  - There's no opportunity for clarification
- NL margin comments are not conversation, even though they look like it
- So perhaps a dialogue act taxonomy (DIT (Bunt, 1990), DAMSL (Core and Allen, 1997)) is not appropriate for classifying margin comments
- **Key question: What are margin comments 'doing'?**
  - (speaker's communicative intention, communicative function, illocutionary force, ...)

# What are margin comments 'doing'?

- (E9) "Why bold?"
  - **Looks like** a WH question (DAMSL: *Wh-question*, DIT++: *Set question*)
  - **But** marker M is not desiring or expecting addressee A to supply the requested information to M
  - A will never take a conversational turn in response to that comment
- (E10) "Explain what they do."
  - **Looks like** an instruction (DAMSL: *Action-directive*, DIT++: *Instruct*)
  - **But** M's comments were added to final, submitted version of essay
  - M did not desire or expect that A would revise the essay in response to M's comments.
- Conclusion: **NL margin comments communicate M's opinion to A about the essay parts to which they point**
  - Perhaps uncontroversial
  - All margin comments do this, even non-NL coded schemes
- Dialogue act taxonomies are beginning to look unsuitable

# What kinds of opinion do NL margin comments express?

- NL margin comment communicates 2 main messages
  - i.* **Whether or not essay part P** to which a comment points **attained the required standard** (in M's opinion)
  - ii.* **How P attained (or did not attain) the required standard** (in M's opinion)
    - The required standard is defined by some set of principles or instructions of which M and A are typically mutually aware
- Message **(i)** is rarely explicitly stated in NL comments and needs to be inferred
  - (E11) "A strong argument" (Was the standard attained?)
  - (E12) "A very strong adjective and claim?"
  - (E13) "Very strong supporting quote."
  - (E14) "A very long sentence."
  - (E15) "This section is a bit short."
- To understand whether the standard was attained or not:
  - Addressee needs to be sensitive to English compositional semantics
    - Non-native English speakers are likely to have difficulties
  - Addressee needs to possess expert knowledge about essay writing

# How was the standard met (or not met)?

- Consider:
  - (E17) "Why not?"
  - (E18) "Why bold?"
- These are statements of opinion masquerading as questions
- Two quite different criticisms (suggested meanings):
  - (E17') *The argument here would have been improved by including an explanation of why not*
  - (E18') *The use of bold font here is questionable*
- How can the addressee tell they mean such very different things?
  1. Identify the **targeted skill area**
    - E17: Argument; E18: Formatting
  2. Consider what that skill area is *like* (how it can be good/bad)
    - E17: An argument is bad if key points are missing
    - E18: Formatting is bad if it is applied in the wrong place
- Our classification scheme labels a comment's targeted skill area
  - First of three layers [»](#) slide 8 (Total of 11 Target categories)



# Second layer: Marker attitude

- M believes something present is of questionable value
    - (E19) "I got a bit confused here!" DOUBT
  - M believes something is missing that should be present
    - (E20) "No conclusion" MISS
  - M believes something is present that should not be
    - (E21) "No apostrophe" REJECT
  - M believes something is present that needed amending
    - (E22) "No italics" CONDEMN
  - M considers that something in the essay has attained or exceeded the required standard, or is pleasing or interesting to M
    - (E23) "No issues" COMMEND
  - M holds views that are in opposition to some proposition
    - (E24) "Not necessarily." DISPUTE
  - M believes that A would benefit from reading some source
    - (E25) "Ditto." REFER
- (Total of 10 Attitude categories)

# Third layer: Linguistic act

- Strongly based on **surface form**
- **Interrogative-like** acts
  - **WH Question:** (E26) "Why not?"
  - **Polar Question:** (E27) "Is this a word?"
- **Declarative-like** acts
  - **Assertion:** (E28) "I don't understand"
    - Reserved for assertions of propositions in response to argument and explicit expressions concerning understanding, agreement, verification or certainty
  - **Description:** (E29) "Too many references."
    - Description of a non-propositional object in or quality of essay part P or author action evidenced by P
- **Imperative-like** acts
  - **Instruction:** (E30) "I would not leave a space."
  - Loose definition which allows for politeness and non-sententials
  - About what should be or have been done rather than what was done

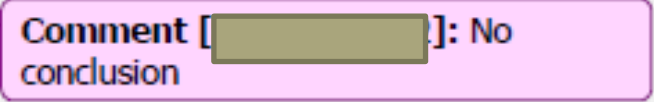
# Passing observation

- Working out what an utterance means is usually straightforward
  - Speakers send very clear signals (as per Grice's CP)
    - About what part of the dialogue history their new utterance is responding to and
    - About which entities the new utterance is referring to
    - About what their new utterance is about
- Working out what a NL margin comment means is much harder
  - The signals the marker sends about which part of the essay the comment refers to are incomplete
    - Yes, they graphically point to an essay part (though often inaccurately)
    - But they often do not say which features of that part are being commented on
  - The informal language that comment writers use is by nature fragmented, ambiguous, and open to a wide range of interpretations

- An utterance

A says	"Where were you last night?"
B thinks	<i>A wants to know where I was last night</i>
B says	"No comment"
A thinks	<i>B is refusing to tell me where she was last night</i>

- A margin comment

A writes	evidence to be developed as to what works and what does not work for those individuals. The difficulty is capturing this tacit information as it tends to be located with people who have worked with those individuals in the past and not collated or recorded in any way.
B thinks	This essay part has not met the required standard because A has not written a conclusion to his essay
B writes	<p>not work for those individuals. be located with people who ited or recorded in any way.</p> 
A thinks	I'm not sure what aspect of my work B is commenting on. I'll have to guess from the semantics of the comment together with the part of my essay the comment is pointing to. Perhaps I have missed something out, or included something that's wrong. It's pointing to the last sentence in the essay and mentioning a conclusion. I think an essay is supposed to end with a summing up section called a conclusion. I thought I did that. Maybe I didn't do it right.

# Putting it all together

- The 3-layered scheme enables the intended evaluative meanings of margin comments to be captured despite their conversational style:
  - **Target (typically skill aspect being targeted)**
  - **Attitude of the marker towards a feature of the related essay part**
  - **Linguistic act (close to surface form)**
- (E33) "No conclusion"
  - Attitude: Miss      Target: Structure      Act: Description
- (E34) "No apostrophe"
  - Attitude: Reject      Target: Punctuation      Act: Instruction
- (E35) "No issues"
  - Attitude: Commend      Target: Context-Dependent      Act: Assertion
- (E36) "I know how you feel."
  - Attitude: Engage      Target: Author      Act: Misc

# Scheme evaluation

- 2 annotators, 313 sample comments

Layer	Cohen's Kappa
Attitude	.874
Target	.791
Act	.822

Percentage agreement across all 3 layers was 72.1 %

- Mean no. comments per essay per marker: 4.83--47.00
  - Only for some tutors was there a correlation between essay length and the number of margin comments.
- To avoid potential bias towards prolific marker styles, the same number of essays and comments sampled per marker
- Handful of cross-layer dependencies
- Some linguistic acts are unlikely to combine with some attitudes
- Estimated no. possible combinations of the 3 layers is 155
- Some categories from a given layer are more frequent than other categories from same layer
- Weighted coefficient method may be more suitable

# Other work

- Other categorisation schemes have been devised or re-used to analyse written feedback (Ferris et al., 1997; Hyland, 2001; Perpignan, 2003; Whitelock et al., 2004; Brown & Glover, 2006; Nelson & Schunn, 2009)
- Mainly interested in whether the marker was writing comments that would 'feed forward'
- Measures for deciding this tended to revolve around:
  - The power of a comment to motivate its addressee
  - Whether the comment contained explanatory text that would make it clear how to do things better in future
- Not found any feedback categorisation schemes primarily concerned with
  - How opinion in comments is conveyed through the medium of NL
  - How to understand a margin comment

# Final comments

- We have rules for identifying 'principal segment':
  - (E37) "Yes<sup>1</sup>, good recommendation<sup>2</sup>, **but pin down exactly WHO will do it, HOW they will do it and ideally also by WHEN they will achieve it<sup>3</sup>** (although it may be a stab in the dark it's worth putting a date on recommendations to act as a goal<sup>4</sup> - it can always be moved later if proved to be unrealistic)<sup>5</sup>"
- Arguably suitable for all NL margin comments for all disciplines
  - Possibly all NL assessor feedback (SCD, viva, driving test, film review,...)
  - But targeted skills are domain-specific
- Planned machine learning investigations will
  - Attempt to categorise appropriate opportunities for feedback comments
  - Look for associations between margin comment and essay part
    - syntactic structure, n-gram features, other semantic similarity measures
- Comment generator will be informed by the ML investigations
- No plans to generate fragmented English