

Determining Specificity of L2 Peer Response in Learning Management System-Based Writing

Garrett DeHond*

DeHond, Garrett. (2020). Determining specificity of L2 peer response in learning management system-based writing. *English Teaching*, 75(3), 69-91.

Second language (L2) peer response literature is defined in part by discourse research, yet there is scant research on text-specific comments, or comments that make explicit text references, thus resisting generic qualities. The purpose of this case study was to examine such peer response activities in an English writing course at a South Korean university. The data comprises two essay assignments with peer response conducted between two drafts – as accomplished during class time on the class learning management system (LMS) – as well as the subsequent revisions in second drafts. This paper expands on previous coding schemes accounting for area, nature and type commentary to account for a specificity dimension, and also links these categories to revision practices. While students entertained diverse commenting and revising options, popular practices included generic evaluating or revising local or surface-level concerns. This paper offers implications for modelling response activities as well as for how to better define specific and complex idea construction exhibited during response.

Key words: peer response, peer feedback, L2 writing, digital affordances, discourse analysis, revisions

*Author: Garrett DeHond, Professor, Department of English Linguistics and Language Technology, Hankuk University of Foreign Studies; Imun-ro 107, Dongdaemun-gu, Seoul 02450, Korea; Email: dehondg@gmail.com

Received 1 July 2020; Reviewed 26 July 2020; Accepted 5 September 2020

1. INTRODUCTION

Second language (L2) writing researchers have been investigating peer response (PR) for over thirty years (Chang, 2016; Ferris, 2003), and in that time three frequent trends have emerged. The first concerns the exploration of PR discourse characteristics to better understand how participants adopt task behaviors and react critically through the activity. A second research trend regards revision strategies adopted by the writer, which provide researchers data on potential impact. These first two trends have been prevalent since the field's early days, but the 2000s saw increased interest in computer-mediated communication (CMC) for peer response. With so many options and spaces for designing a peer response activity, there lies importance in determining how activity procedures and phenomenon correspond to communicative choices.

As researchers continue to explore these options, one notable research gap appears to be an investigation of specificity in response constructions. Ferris (1997) defined text-specific commentary as respondent's text that makes explicit reference to the target essay text, as opposed to generic comments which are often "rubber-stamping vague, prescriptive remarks" (p. 333) which can be applied to any students' essay. These language choices are worth exploring given our assumptions that articulation choices – and not just their intrinsic ideas – influence the recipient. Exploring specificity choices in L2 peer response could potentially lead to new considerations for how researchers examine such characteristics as student motivation, socializing behaviors or exhibition of critical thought.

Specificity can be an issue for any PR design and communicative mode, but presents a particular challenge in "review-only" PR activity designs. These designs favor one-way communication where the reviewer makes comments and writers choose how to respond through revisions. When practitioners employ review-only designs for digital spaces such as cloud-stored documents (i.e. Google Docs), they undercut the tool's affordances that should promote great accessibility in time, space and interaction characteristics (Breuch, 2004). Given the abundant interest in L2 PR research to investigate or compare particular mode choices (Chang, 2016), it may be worthwhile to determine how specificity characteristics reveal themselves in an activity situation where the tool appears to afford text-specific comments while the design choice does not demand them.

The current study therefore explores how specificity characteristics are exhibited in an observed practice of "review-only" L2 peer response. The literature review establishes developments in L2 PR discourse research, with specificity as an important element missing from that literature. To that end, it will highlight how a specificity dimension can interact with other established PR discourse characteristics. It also establishes the need to investigate specificity in particular mode or design contexts like the "review-only" digital option described above. Because PR research is ever interested in exploring notable trends

in discourse patterns and their impact on writers, this study will attempt to discuss the possible motivations behind specificity choices and how such choices interact with other discourse characteristics to influence revisions.

2. LITERATURE REVIEW

Most researchers trace the foundational period for fervent L2 peer response research to the late 1980s through 1990s (Ferris, 2003; Liu & Hansen, 2002), with discourse characteristics being a frequent topic. Some researchers attempted to identify discourse traits by describing the students, perhaps in how students adopted roles (Nelson & Murphy, 1992), behaviors (Nelson & Murphy, 1993) or stances (Lockhart & Ng, 1995; Mangelsdorf & Schlumberger, 1992) in the activity. Other researchers would isolate and identify discrete categories of discourse ideas, with examples including group-created “rules” (Stanley, 1992) and idea negotiations (Mendonça & Johnson, 1994). Differing interpretations aside, one quality many of these studies lacked was a sense of multi-dimensionality; ideas can indicate more than one behavior. De Guerrero and Villamil (1994) demonstrated this when they sought to investigate task episodes, cognitive regulation and social relationships all in one study.

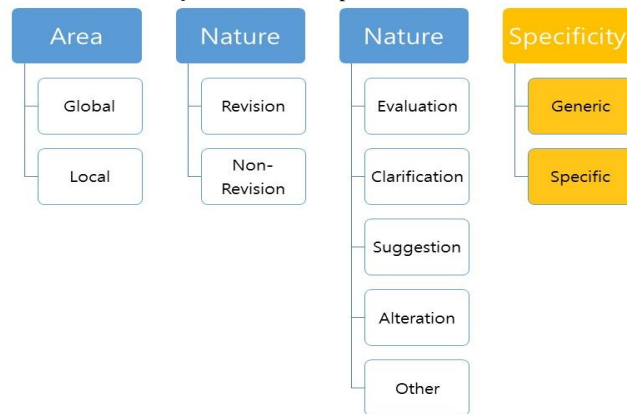
One influential study that followed this period is the Liu and Sadler (2003) study that introduced area, nature and type (ANT) characteristics as a useful construct for analyzing discourse. This paper is noteworthy for being an early comparison study of four interaction modes – pen and paper commenting, Microsoft Word commenting, oral classroom conversation, and online chat – while also employing a multidimensional analytic model consistent to all four modes.

- “Area” adopts a similar definition to Zhu’s (2001) global and local comment types, which in turn were influenced by McGroarty and Zhu’s (1997); global issues concern large idea concepts like argumentation, detail, and structure, while local issues are more copy-editing in their focus (i.e., format, grammar, mechanics).
- “Nature” regards the comment’s orientation to revision or non-revision motives. Until 2003, most literature on revision examined the influence of stances or ideas on orientation (Berg, 1999; Connor & Asenavage, 1994; Nelson & Murphy, 1993). Some authors also developed similar coding schemes such as “positive” and “negative” charge (Guardado & Shi, 2007) to describe similar phenomenon.
- “Type” categories differ slightly from Mendonça and Johnson’s (1994) negotiations to account for actions of evaluation, clarification, suggestion and alteration.

In addition to being an early study on digital designs for peer response, several studies have cited and adapted the ANT coding rubric for their own purposes. Examples of ANT usage include studies focused on different communication modes (Chang, 2012; Ho, 2015), training influence (Liou & Peng, 2009), intercultural environments (Bradley, 2014), revision influence (Chang, 2012; Ho, 2015), and student perspectives (Bradley, 2014; Ho, 2015). In all studies above, the area and nature descriptions have been consistent, perhaps due to their binary treatment. Some authors also made new definitions of the “type” comments as applicable to their own studies. Liou and Peng (2009) created a “chatting” category for tangential comments (like off-task comments, or personal connections to the text), and Ho (2015) created a comprehensive “other” type category for any types not matching the main four. Largely, however, the original three-dimensional framework and descriptors appear consistent with no further dimensions developed.

While researchers could expand on the content of each dimension, no expansion appears to account for the identification of specific or generic language choice in discourse. In her paper on teacher commentary and its influence on student revisions, Ferris (1997) employed “text-specific” (p. 321) commentary as one analytic component for examining instructor feedback. She defined “generic” comments as those that “could have been written on any paper” (p. 321), while text-specific commentary was task-oriented yet not transportable, largely because such comments made specific reference to ideas within the target essay. The Ferris’s (1997) study was indeed concerned with several interactive characteristics – length by words, use of hedging language, and her own definitions of comment types. Later authors mimicked this design for their own studies on teacher commentary (Conrad & Goldstein, 1999; Ferris, Brown, Liu, & Stine, 2011; Gascoigne, 2004), though it appears few authors have sought to expand the defining traits of specificity or even explore them in peer response contexts. One noteworthy exception, the Nelson and Schunn’s (2009) study on cognitive and affective characteristics in first language (L1) peer feedback, did develop codes for problem locating, problem identification, and problem solving in connection to specificity. However, these definitions with overlap with the “type” comments in an ANT rubric (i.e., problem identification could be a revision-oriented evaluation).

FIGURE 1
The ANTS Analytic Model (Adapted from Liu & Sadler, 2003)



The evidence as presented warrants an argument for expanding the Liu and Sadler ANT framework to an “ANTS” framework, as depicted in Figure 1 above. By creating a new dimension for text-specific commentary, researchers can establish new distinctions between comments that appear to achieve the same discourse objectives under ANT. Consider the following two examples:

“your thesis statement is too factual”

“your thesis statement misinterprets Chomsky’s ideas about the critical period”

Under ANT, readers could identify both statements as global/revision/evaluation commentary. With ANTS, the first comment now appears “generic” because the language does not explicitly make reference to essay ideas; all essays should have thesis statements and may be at risk of being “too factual.” For the second comment to be “generic,” every essay must require a thesis statement about Chomsky or the critical period; without such directions, the comment is “text-specific.” Since specificity works as an interacting dimension with ANT, researchers can derive new commentaries on what it means to, for example, make an evaluation of writers’ ideas for revision purposes.

The communicative mode, however, should still be relevant to those interpretations. The appeal of digital environments for L2 PR lies in how they not only transfer but transform activity options. Breuch (2004) provides one of the strongest arguments in support of digital PR, describing it as a “remediation” of traditional peer response practices. She illustrates this transformation by describing PR activities in terms of affordances, specifically in time, space and interaction characteristics, as outlined in Table 1 below. These definitions, of course, are influenced by tool and design choices. A cloud document,

for example, is more likely than printed copy to have enhanced durability and reach, but such is not guaranteed.

TABLE 1
Characteristics of Online Peer Response (Adapted from Breuch, 2004)

Category	Characteristic	Definition
Time	Synchronicity	Responses can be immediate or delayed
	Durability	Response content can be saved and stored
	Concurrency	Response messages are not necessarily limited to “turns”
	Convenience	Response time boundaries can be shifted and expanded
Space	Social Cues	Social markers (like age and gender) are alterable
	Interpersonal presence	Respondents control the pace of relationship building
	Hyperpersonal presence	Respondents control the depth of relationship building
Interaction	Text-based	Interactions require and encourage writing skills practice
	Fixity	Discourse permanence allows for stimulated recalls
	Response structure	Response discourse can be built with several design features
	Reach	Discourse can be shared and incorporate more participants

Digital PR practices should be able to transform specificity by virtue of tool affordances, but creating specific comments is still a behavioral practice that must be modelled and guided to encourage presence (Ferris, 2003). This leads to why an ANTS investigation of “review-only” PR strategies may be vital; without checking for specificity, researchers and instructors may not well gauge the effort and influence such commentary has on PR outcomes. The following study presents one case example of an L2 writing class engaging online, review-only peer response. The research objectives concerned a quantitative and qualitative analysis of ANTS characteristics in reviewers’ commentary and how comments may have influenced revisions. In analyzing frequent trends of what students accomplished in this class, we might gain useful insights for improvements to the activity. The research agenda that follows attempts to address the following questions:

1. What ANTS characteristics do peer reviewers adopt in their commentary?
2. What revision practices (relevant to peer response) do writers complete?
3. What notable trends related to ANTS appear in commentary?

3. METHODOLOGY

3.1. Participants

Study participants comprised 27 undergraduate students in an English writing course at a South Korean university. These students self-selected as participatory volunteers following a first-week survey with study information and questions on student demographics. Of these, 24 students were of Korean descent (the others were of Chinese, Japanese and French nationality), and the gender split was 13 females and 14 males. Though this course is intended for second-year students, class levels included 6 freshmen, 4 sophomores, 8 juniors and 9 seniors. Students' department information included five students in the course's home department, and 22 students double-majoring in an English-language department with another department. 20 students had attended some form of schooling outside Korea, 9 of whom having done so in the United States, and 13 had completed some form of English writing study in either private academies or with tutors.

Students were asked questions about their academic experiences with English, learning management systems (LMS), and peer response activities. Only 2 students were taking their first university class where English writing was required, whereas 8 were taking their second and 17 having already taken at least two such courses previously. All had taken some number of courses where the professor primarily spoke in English, with 18 having taken at least four courses taught in English. Regarding LMS, two students were new to using such programs for submitting work, with 10 having done at least once and 15 at least twice. As for peer response, 9 students claimed to have never done peer response in English, and 15 claimed the same for Korean language.

3.2. Data Collection Procedure

Data collection consisted of all essay drafts and peer response submissions. For their class, participants met in a computer-networked lab. The instructor required students to submit two essays' first drafts before specific class meetings, and peer review of writers' drafts during those class meetings, all through the instructor's Canvas LMS course site. Following peer review, the instructor expected students to submit revisions of first drafts to Canvas LMS one week later. The first essay's instructions required students to select one of the course readings from a unit of famous essays about language and write a critical response. The second essay's instructions required students to develop an argument that directly references one or more of the course readings from a unit on famous essays about cognition. For both essays, the instructor expected students to write approximately 1,000 words in an academic English essay style taught in the class.

For the peer review activities, the instructor randomly sorted students into triads and had students read and respond to essay drafts during class time. The LMS provided a preview screen for peers' drafts and allow paired readers to create both marginal annotations and summary comments, without altering the original text. While the instructor graded students for their participation during peer response, the instructor allowed for flexibility in how students addressed 'higher order' (i.e., idea quality) and 'lower order' (i.e., copy quality) concerns, referencing the Purdue OWL lab descriptions in peer review instructions (Purdue Online Writing Lab, n.d.). The LMS would save each confirmed comment online and could then be visible to the original writer. For the second drafts, the instructor graded students not only for overall qualities but also in part for their revisions as they pertained to students' feedback.

Table 2 below adapts Breuch's digital PR affordances to summarize the specific communicative mode and activity design observed in this study. In most respects, these activity features are defined by the specific LMS tool affordances (such as the observed durability or reach). In some cases, the instructors' activity choices (such as the observed convenience and social cues) influence activity procedures.

TABLE 2
L2 Peer Response Design Choices (Modelled from Breuch, 2004)

Category	Characteristic	Activity Features
Time	Synchronicity	Asynchronous (via document annotations)
	Durability	Recorded (all comments are saved and downloadable)
	Concurrency	Monologic (no required writer-response impetus)
	Convenience	Teacher-defined (limited to 45-minute sessions)
Space	Social Cues	High identity visibility (student names appear with comment)
	Interpersonal presence	One-way (participants post comments, must refresh to see replies)
	Hyperpersonal presence	Limited (writer responses optional and visible to peer, reviewers cannot view other reviewers' comments)
Interaction	Text-based	Typewritten (options for highlighting)
	Fixity	Digital (exists as long as course is open, comments could be altered until deadline)
	Response structure	Annotation-style (comment bubbles, summary paragraphs)
	Reach	Writer/Teacher (only writers and instructor can download)

3.3. Data Analysis Procedure

Peer response coding consisted of the ANTS dimensions described in the literature review: area (global, local), nature (revision and non-revision oriented), type (evaluation, clarification, suggestion, alteration and other) and specificity (generic and text-specific). "Other" type levels were given to ideas that were purely informational or off-task. Ideas

units were identified by location (marginal or summary comments) and constructed at the clausal level (Chafe, 1980; Halliday, 1967). A visual representation of the ANTS coding scheme with examples can be found in the appendix.

For analyzing revisions, Hall's (1990) revision taxonomy was adapted to account for relation, level, and action characteristics. Relation characteristics replace Hall's "stimulus" descriptor to approach the likelihood that visible revisions are traceable to peer comments (likely, unclear or unlikely), or if the target area is entirely unchanged (none). Level characteristics consist of text changes within one sentence (mechanics, text and ideas), within one paragraph or beyond (global idea changes or surface level changes throughout the paper). Revision actions regard the precise practice taken by the writer, as described by Hall: additions (more text with new ideas), deletions (less text with fewer ideas), substitutions (changed text where one idea replaces another), expansions (additions and substitutions) and consolidations (deletions and substitutions). As this study is concerned with PR influence, data on revision levels and actions for either unclear or unlikely relations are not presented in the findings.

4. FINDINGS AND DISCUSSION

The following section summarizes the main findings of ANTS characteristics and their potential relation to revision practices. Except where otherwise noted, the tables in this section represent findings in both assignments combined, and parenthetical numbers in the tables represent an increase or decrease (- symbol) in values found in the second assignment. The first two sections will highlight noteworthy observations from the quantitative analysis, while the third will attend to important trends exhibited in response.

4.1. ANTS Characteristics

Overall, reviewers combined for 1,447 expressed ideas in PR discourse, with 123 ideas more in the second essay activity than in the first. Of these ideas, 1,163 come from marginal comments while 459 come from summary comments. Table 3 above displays the 1,163 ideas found in marginal comments. With 654 global/revision and 311 local/revision ideas, reviewers certainly opted for more revision commentary than non-revision. While generic evaluations (203) and generic suggestions (156) were more frequent than their specific counterparts, specific other/revisions (123) appear more often than generic (38). Many of these can be explained by reviewers' habits to make an evaluation or suggestion and also support this with specific contextual detail (e.g., factual statements like "your topic sentence is about national identity"). Total clarifications (117) and alterations (39) of

the specific variety also appear more commonly than generic constructions. Readers may note that in the second essay, PR ideas shifted more to generic evaluations (178).

TABLE 3
ANTS Codes, Marginal Comments

	Global		Local		Total
	Revision	Non-Revision	Revision	Non-Revision	
Evaluations					
Generic	153 (57)	156 (98)	85 (19)	12 (4)	406 (178)
Specific	73 (-3)	10 (4)	15 (-7)	3 (3)	101 (-3)
Clarifications					
Generic	55 (-9)	0 (0)	27 (-11)	0 (0)	82 (-20)
Specific	86 (4)	1 (1)	30 (-18)	0 (0)	117 (-13)
Suggestions					
Generic	98 (0)	0 (0)	58 (-8)	0 (0)	156 (-8)
Specific	56 (18)	0 (0)	9 (1)	0 (0)	65 (19)
Alterations					
Generic	0 (0)	0 (0)	30 (-16)	0 (0)	30 (-16)
Specific	0 (0)	0 (0)	39 (-9)	0 (0)	39 (-9)
Statements					
Generic	31 (3)	7 (7)	7 (-1)	1 (-1)	46 (8)
Specific	102 (36)	8 (6)	11 (-3)	0 (0)	121 (39)
Total	654 (106)	182 (116)	311 (-53)	16 (6)	1,163 (175)

It is important to note, however, that the ideas of Table 3 are presented in isolation rather than interaction, and we know multiple ideas were expressed in marginal comments. Table 4 above summarizes coding for marginal comment grouping; if a comment bubble contained only generic or specific ideas, those labels signify that group, whereas a minimum of one idea of both types are “combined.” Here we find that 561 of 1,163 ideas were found in generic-only marginal comment groups, whereas earlier it was stated there were 720 total generic ideas. Considering specific-only (236) and combined (366) together, 51% of comment constructions contained at least one specific idea. While non-revision commentary is more often generic-only (138 global, 13 local), 34% of global/revision ideas (225) and 59% of local/revision ideas (185) are generic-only. Looking over data examples, there are many instances where students sought to directly identify essay text when constructing marginal comment bubbles (e.g., “Isn’t the case of Victor and Helen Keller different?”), though there are instances where they relied on comment bubble highlighting (for example, highlighting some essay text and typing, “I like this line” in the bubble). As for the lack of specificity in non-revision instances, much of those examples can be described as generic praise, like the example in the previous sentence.

TABLE 4
ANS Codes, Marginal Comment Grouping

	Global		Local		Total
	Revision	Non-Revision	Revision	Non-Revision	
All Generic	225 (43)	138 (84)	185 (-7)	13 (3)	561 (123)
All Specific	151 (27)	8 (4)	75 (-23)	2 (2)	236 (10)
Combined	278 (36)	36 (28)	51 (-23)	1 (1)	366 (42)
Total	654 (106)	182 (116)	311 (-53)	16 (6)	1,163 (175)

Table 5 presents the ANTS findings for summary comments. Of 284 summary comment ideas, 83% contained generic language, and generic comments also outnumber specific comments across all four columns. Summary comments were dominated by generic evaluations (163) followed by generic suggestions (59), which together make up more than 78% of all summary ideas. Also worth noting would be the drop in local ideas (26 revision, 2 non-revision), which make up less than 10% of all summary ideas. Certainly, summary ideas do not appear as varied as with marginal ideas, though these findings are consistent with observations by Ferris (1997) and others that summary comments tend to focus on general reactions to the text as a whole, along with ideas for what may be the most vital improvements.

TABLE 5
ANTS Codes, Summary Comments

	Global		Local		Total
	Revision	Non-Revision	Revision	Non-Revision	
Evaluations					
Generic	49 (-17)	102 (-2)	10 (0)	2 (0)	163 (-19)
Specific	9 (-9)	6 (4)	1 (-1)	0 (0)	16 (-6)
Clarifications					
Generic	2 (0)	0 (0)	0 (0)	0 (0)	2 (0)
Specific	6 (-6)	0 (0)	0 (0)	0 (0)	6 (-6)
Suggestions					
Generic	50 (-2)	0 (0)	9 (-3)	0 (0)	59 (-5)
Specific	6 (-4)	0 (0)	2 (0)	0 (0)	8 (-4)
Alterations					
Generic	0 (0)	0 (0)	3 (-1)	0 (0)	3 (-1)
Specific	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Statements					
Generic	8 (2)	1 (-1)	1 (1)	0 (0)	10 (2)
Specific	14 (-14)	3 (1)	0 (0)	0 (0)	17 (-13)
Total	144 (-50)	112 (2)	26 (-4)	2 (0)	284 (-52)

4.2. Revision Practices

Of the 1,447 comment ideas, 1,135 of these were revision-oriented, including 965 marginal ideas and 170 summary ideas. Revision-oriented PR ideas placed near eventual writers' revisions comprise 703 total PR ideas (61%), including 572 marginal ideas (60%) and 131 summary ideas (77%). The revision percentages suggest most students in class indeed attempted to revise their work before the second draft submission. Table 6 below breaks down these 703 ideas into their relation to revision practices. The chart organization here orders results by number of likely relations.

TABLE 6
Common Revision Codes by Response Types, All Comments

Code Type	Likely Relation	Unclear Relation	Unlikely Relation	Total
Local/Evaluations/Generic*	58 (2)	4 (-2)	5 (1)	67 (1)
Global/Evaluations/Generic	57 (9)	39 (1)	27 (1)	123 (11)
Local/Suggestions/Generic**	43 (-13)	4 (-4)	2 (-2)	49 (-19)
Global/Suggestions/Generic	42 (4)	29 (-3)	21 (-13)	92 (-12)
Global/Clarifications/Specific**	28 (-14)	7 (-3)	7 (-1)	42 (-18)
Local/Alterations/Generic*	27 (-19)	1 (1)	0 (0)	28 (-18)
Local/Alterations/Specific*	27 (-5)	2 (0)	1 (-1)	30 (-6)
Local/Clarifications/Generic*	21 (-7)	0 (0)	0 (0)	21 (-7)
Local/Clarifications/Specific*	21 (-17)	3 (-3)	1 (1)	25 (-19)
Global/Statements/Specific	21 (-5)	30 (8)	8 (-4)	59 (-1)
Global/Suggestions/Specific**	18 (6)	11 (3)	6 (-2)	35 (7)
Global/Clarifications/Generic**	17 (-7)	6 (0)	6 (-2)	29 (-9)
Local/Evaluations/Specific*	13 (-9)	0 (0)	0 (0)	13 (-9)
Global/Evaluations/Specific	12 (-4)	20 (2)	15 (-15)	47 (-17)
Local/Statements/Specific*	9 (-5)	0 (0)	0 (0)	9 (-5)
Global/Statements/Generic	9 (3)	5 (5)	8 (-6)	22 (2)
Local/Suggestions/Specific*	5 (1)	1 (-1)	0 (0)	6 (0)
Local/Statements/Generic*	5 (1)	0 (0)	1 (-1)	6 (0)
Total	433 (-79)	162 (4)	108 (-44)	703 (-119)

* likely relations at least 80% of total comments in code type

** likely relations at least 50% of total comments in code type

At 433 likely related ideas, 61% of these revision actions appear to have some relation to PR ideas; this number also represents 38% of total PR ideas. 162 were labelled unclear because while the revisions attend to the peer comments' ideas, other localized revisions at the target area indicate the revision could have been for other reasons. This appeared to most often happen when a writer's paragraph had multiple comment bubbles with differing demands, like attending to a main idea and a supporting detail. 108 had no discernable

relationship to the PR idea; for example, one reviewer commented on a grammar error, but the writer changed many ideas within one paragraph for other reasons, and thus eliminated the grammar error. Both generic evaluations (115) and generic suggestions (85) top the chart. Some combinations indicate quite high levels of response; nine rows indicate a “likely response” above 80% (*), and four rows above 50% (**) per their respective code type totals. It may be noteworthy that all nine * rows indicate local area, four of which are specific. This suggests that writers may have favored ideas pointing to copy-editing issues, as local ideas are more likely than global ideas to require quick or undebatable changes. Of the four ** categories – these indicating between 50% to 80% “likely response” – three are global and specific, and all are clarifications or suggestions. It is noteworthy that while global and generic evaluations (58) and suggestions (42) are high in the “likely response” column, they do not break 50% of their own code type revision opportunities. Clearly, while writers valued clarifications and suggestions, generic ideas about global features were not often appreciated for revision considerations.

TABLE 7
Revisions by Action and Level

Codes (ANS)	Action					Level				Total
	Add	Del	Sub	Exp	Con	Sen	Par	Sur	Glo	
Loc/Eva/Gen	11	11	28	4	4	36	4	16	0	58
Glo/Eva/Gen	14	0	8	27	8	13	31	2	11	57
Loc/Sug/Gen	13	1	20	4	5	16	4	20	2	43
Glo/Sug/Gen	10	0	6	22	4	8	19	0	15	42
Glo/Cla/Spe	5	0	2	19	2	6	15	2	5	28
Loc/Alt/Gen	13	5	7	0	2	10	2	15	0	27
Loc/Alt/Spe	3	4	17	1	2	24	0	2	1	27
Loc/Cla/Gen	5	2	10	1	3	8	2	11	0	21
Loc/Cla/Spe	2	0	14	4	1	17	2	0	2	21
Glo/Sta/Spe	5	1	1	11	3	7	13	0	1	21
Glo/Sug/Spe	5	0	1	12	0	4	11	1	2	18
Glo/Cla/Gen	3	1	2	10	1	4	11	1	1	17
Loc/Eva/Spe	2	4	6	1	0	8	0	5	0	13
Glo/Eva/Spe	4	3	0	4	1	4	5	1	2	12
Loc/Sta/Spe	2	0	2	5	0	2	5	0	2	9
Glo/Sta/Gen	1	2	5	1	0	7	0	2	0	9
Loc/Sug/Spe	2	1	1	0	1	2	1	2	0	5
Loc/Sta/Gen	0	0	4	1	0	4	1	0	0	5
Total	100	35	134	127	37	183	126	80	44	433

With revision-relations established, Table 7 presents the level and action influence likely-related ideas had on revision practices. Here, substitutions (134) and expansions

(127) make up 60% of revisions, followed by additions (100), consolidations (37) and deletions (35). Highest counts per row in both the action and level columns indicate local ideas more often led to substitutions while global ideas led to expansions. In terms of revision level, 48% of revisions are at the paragraph (126) and global (80) level, suggesting that while multiple ideas were frequently revised in these papers, more focus was given either to ideas regarding sentence and surface issues. The “sentence or below” column (183) accounts for 42% of level revisions. Among other things, writers appeared to most often use commentary to facilitate opportunities at expanding on global ideas at the paragraph level and making substitutions on local issues within sentences or across the paper. One noteworthy exception, the 13 additions responding to local/generic/alterations, can be explained by missing elements such as titles, citations or a reference list.

4.3. Notable Trends

These findings suggest a variety of complex activities taking place in this response environment, though certainly some trends are noteworthy. Just as students favored generic evaluations in their PR, writers appeared to favor local or generic ideas when revising after response. These tendencies provide much to consider regarding students’ preferences in how to show critical thoughts and respond to PR work in ways that might limit workload or challenge. The following discussion attends to some qualitative trends within the comment that reveal more about these complexities, with each trend accompanied by examples. Later in this section, the discussion will also attend to some limitations experienced when applying ANTS to the analysis. Please note that all presented comments are unedited and may contain language errors, and each comment is labeled to signify distinct students and drafts (for example, “S1D1” = “Student 1, Essay 1”).

In many comments, students employed liberal use of class language when sharing their ideas. “Class language” or shared language (Reid, 1994) refers to common terms and expressions frequently employed in target environments. Non-revision/evaluations, which are typically praise, were a prolific source of class language as students attempted to compliment some aspect of the passage. Many students might attach a reaction (“interesting,” “good”) to class language to construct an idea. Some ideas, like counter-arguments and “F&T” (the professor’s “formatting and tips” sheet), have additional salience given how the instructor stressed them in lectures, course materials and PR modeling. Shared language also frequently occurs when reviewers detect a missing element, as shown in generic clarifications and suggestions. Class language is one example of trends where the idea is often transformed based on generic, specific or combined-idea organization.

- S1E1: "I think showing Martin Luther King's speech as a example is good counter argument."
- S11E1: "Is this supporting detail supports the topic sentence?"
- S12E1: "Interesting Hook! It grabbed my attention right away."
- S3E2: "You should add more examples and supporting details in each paragraph."
- S6E2: "Check capitalization for your title."
- S8E2: "I like the part that you actually reminding the readers the summary of the story..."

Many single-idea marginal comment bubbles expressed either generic praise or clarification requests; all of the examples below target non-revised essay text. As with many marginal comments, readers need rely on certain contextual features such as how the reviewer highlighted targets and content to construct meaning. The example "this one was good" hit at one particular sentence within a paragraph, but whatever "this one" means was unclear. "I like this line" and "clever" are clearer evaluations, but an explanation for the chosen evaluation is missing. The clarification questions, meanwhile, sometimes challenge the essayist without stated reason. Some questions like "What do you mean?" may express confusion, while others like "Who is he?" may be leading or rhetorical, though of course the writer is not compelled to answer in this activity design.

- S12E1: "This one is good."
- S14E1: "What do you mean?"
- S11E2: "I like this line."
- S14E2: "Who is he?"
- S15E2: "Clever (:"
- S16E2: "Are you sure?"

Many reviewers adopted hedging strategies, concessions and idea comparisons as they expressed multiple ideas together. The expression "may be better" or its variants frequently appears in reviews, as do "but" and "however" to express concessions with criticism. A common practice by reviewers were combined-idea marginal posts, with examples including combined evaluations and suggestions. The two-type combinations provide a sense of layered ideas, though that does not guarantee ideas are more cognitively constructive than expressing simple dislike and desiring change. Some students may also be indicating politeness strategies ("it looks fancy") to soften criticisms or save face. Word choices like "guess" or "maybe" show in many comments, and could indicate politeness or trepidation. Another issue with many "better" comments is they can be generic evaluations

disguised as suggestions. The comment “guess you need to develop your idea into better way” does not express what is problematic with the target item.

- S6E1: “It looks facny, but considering the audiences are english learning students, how about change some of the words to easier words.”
 S7E2: “guess you need to explain more to help readers to understand”
 S7E2: “guess you need to develop your idea into better way”
 S8E2: “it would be better if you make clear diffenrece between topic sentence 2 and topic sentence 3”

Most questions are labelled as clarifications, but some have the appearance of rhetorical questions hinting at other ideas. “Did you choose your title?” and “so?” are more obvious evidence, though with differing demands. Many comments like the “Plato” and “sleeping” examples have questions connected to another idea type. This raises one issue with the ANT framework in that much coding relies on explicit formations even as there appears room for interpretation. Everyone in this class should know who (Bill) Bryson is because he was in the required reading. If the reviewer had written “explain who Bryson is” this is no longer a clarification request.

- S6E1: “Did you choose your title?”
 S17E1: “so?”
 S18E1: “Who is Bryson? There should be a citation to the language unit of the book.”
 S8E2: “You didn't refer the essay from our textbook?”
 S11E2: “Isn't that what sleeping does? A lot of people don't dream when sleeping.”
 S14E2: “Plato were? unhappy? I can't understand this part.”

The discussion thus far has skirted two common concerns about PR quality: vague generalization and faux specificity, both found here in the data. Vague generalizations are generic comments that have inaccessible or inoperable meaning. Indirect requests like “try to write it differently” or just “F&T” may either be not useful or unrelatable to revisions (this happened with “try to write it differently”). This can happen with specific commentary as well, but as the review activity does not promote negotiation, a critical analysis of comment influence via revision is somewhat harder to come by. Asking a writer to write a more specific thesis, and the writer revising the thesis, does not engender clear indications of successful peer response.

S6E1: "I think you can write more interesting hook like more specific situation."

S13E1: "Try to write it differently"

S3E2 & S15E2: "F&T"

The second concern, "faux specificity," regards the deliberate dressing of an otherwise generic comment to give a specific appearance. While this analysis treated generic and specific comments as having a binary nature, specificity might range under new definitions. The impactful difference between a comment of "nice repetition" and "Nice repetition of 'five senses'" (the real comment) might require further debate. Likewise, the "true meaning" example was coded specific because it used that text ("understand") from the essay, but without explanation on why that makes for an "impressive hook". These observations have important implications for determining cognition and affect in the activity.

S4E2: "Impressive hook. Everyone knows the word 'understand', but they does not really think about the true meaning of this word."

S8E2: "Another interesting example story, Hamlet, to start a paragraph."

S26E2: "Nice repetition of 'Five senses.'"

Given the L2 context, it is appropriate to make note of clear language issues, which could include misspellings, word choice issues or problematic sentence structures. Some errors can be explained away as typos ("linual" below should be "lingual") or mental mistakes unrelated to L2. Still, there is also opportunity to mistake ideas in communication, and review-only practices do little to mitigate such possibilities. The comments themselves also might reveal students' difficulties or frustrations with constructing PR, as evidenced by the request for "easier words."

S12E1: "It looks facny, but considering the audiences are english learning students, how about change some of the words to easier words."

S12E1: "Those sources were paraphrased well."

"does it mean English has altaic linual feature? If not, this can just be removed to clarify the misunderstanding."

S12E1: "Do Korean government really regulate official formant of Korean?"

The findings above should legitimize specificity as a complementary new dimension to the original Liu and Sadler ANT framework. There do, however, still exist some limitations with its current design. An obvious place to start would be the treatment of

specific and generic comments as a binary product, as done with area and nature. There is evidence in this study that specific comments not only vary in their distance from the generic definition, they become specific in diverse ways, such as referring to the text or to the professor's comments. Ferris (1997) had noted in her original study the difficulties with separating some text-specific comments from generic ones. Clearly, the definition of generic is much stronger than that of specific and needs attention.

Another matter regards the complexity of idea constructions that neither ANT nor ANTS is designed to account for. All examples below can be coded as generic/global/non-revision/evaluations, or in other words, generic praise. They vary, though, in number of ideas, target areas and language constructions, and ANTS has little regard for such nuance. This has implications for other matters, like genuine commentary or student workload.

S12E1: "Interesting Hook! It grabbed my attention right away."

S24E1: "Nice rebuttal."

S25E1: "This last paragraph contains all your opinion and covers the whole essay well."

S25E1: "Nice hook and the connecting is good. Also your thesis statement is clear and controversial so it is very good."

Ideas are not only complex in their isolated constructions, but in their possible interactions. This study brings to attention the necessity for understand how isolated ideas interact. Multiple comment bubbles can share ideas, and summary statements can repeat or refer back to marginal ideas. The following examples highlight this likelihood:

S2E2: "You're simply criticizing Park's attitudes. I think you should add some details that say citizen should not be deceived by those attitude, since it IS your thesis idea, if I understood your first paragraph well."

S2E2: "Same here. Is your essay about criticizing Park, or insisting that citizens should be enlightened?"

S2E2: "again merely criticizing Park."

S2E2: "Hmm. now I'm getting confused what your essay is about. If your essay is simply about criticizing Park, I think you should revise your thesis statements."

All four statements are from the same reviewer examining the same paper, but are isolated comment bubbles for different parts of the text. Here, "Park" refers to "Park Geun-Hye", a former Korean president, and one of the important subjects of the essayist's paper. Clearly, these comments have the potential to draw strength from each other, and can

interact when influencing revisions, but the ANTS framework as designed cannot code for such occurrences. This fact carries an important implication for specificity analysis, as the reviewer could have chosen alternate paths to sharing these ideas that would alter their eventual ANTS coding.

ANTS is also limited at identifying cognitive or affective features, which could be also be important for understanding reviewers' behaviors or indicating further idea complexities. While it can account for 'evaluations,' it is reliant on explicit statement, and concepts of tone or mood are altered in the written word. Consider the following examples:

S14E1: "How is Black English racist?"

S14E1: "That sounds super racist."

These are two separate marginal comments taken from the same reviewer for the same paper. The first question challenges a stance taken by the writer, while the second statement appears an emotional response to a writer's claim. A determination regarding emotional reaction here relies on implication rather than direct statement – for example, "I hate how super racist that sounds" – and there are also logical or ethical challenges the reviewer could be making as well. These observations hint at some of the difficulty regarding the coding of a statement as being cognitively or affectively different from others. Those challenges aside, there might be some possibility that the expressive nature of a comment influences or interacts with specificity in unique ways.

5. CONCLUSION

This study investigated a South Korean university's English writing course for the instructor's implementation of L2 peer response through review-only online annotation practices. The objectives were to investigate how the review-only practice style influenced students to express ideas as they pertain to area, nature, type and specificity dimensions, and the degree to which they may have influenced revisions. The findings suggest complexity in reviewers' interests in their commenting, though with some preference for evaluations, revision-orientation, global and generic commentary. While revision practices do not indicate peer response was more influential than otherwise, revisers did attend to many characteristic types. The dominant trends carry important implications for both PR response practice and the ANTS framework. Students in this activity valued class language, generic text-referencing, diverse questioning strategies, hedging and politeness strategies, and comments that indicate vague generalizations and faux specificity. While such behaviors can happen in any PR design, the review-only design does not encourage mitigation by other factors (like

partners or writers' responses) and instructors should model in ways that maximize their likely influence. As for ANTS, this study's findings suggest further investigations are needed to define specificity while also determining idea complexities.

While the ANTS framework should be useful beyond this learning context, there is potential for diverse results. The particulars about the observed class's activity design and practices – review-only style, computer lab with limited oral interactions, limited in-class time – certainly inhibited or discouraged potential negotiations that could have enhanced idea qualities. In many cases, an alternate or additional response activity could present new opportunities in the discourse (Guardado & Shi, 2007). However, as there are methodological reasons for this activity planning that do not necessarily invalidate its use, there might be merit in conducting and researching specific modeling practices that promote deliberate use of generic and specific comments in context to course objectives. Examinations of how students visually construct ideas can also be useful, especially when there are so many tools – typing, marginal or summary fields, highlights – at the reviewer's disposal. Some authors have advocated the merits of keystroke logging (Hyland, 2016; Van Waes, Van Weijen, & Leijten, 2014) in L1 and L2 writing research, and for a review-style peer response activity, logging may be one useful tool for determining how students perceive the time and space characteristics of their options (Breuch, 2004).

Applicable levels: Tertiary

REFERENCES

- Berg, E. C. (1999). The effects of trained peer response on ESL students' revision types and writing quality. *Journal of Second Language Writing*, 8(3), 215–241.
- Bradley, L. (2014). Peer-reviewing in an intercultural wiki environment - student interaction and reflections. *Computers and Composition*, 34, 80–95.
- Breuch, L. A. K. (2004). *Virtual peer review: Teaching and learning about writing in online environments*. Albany, NY: State University of New York Press.
- Chafe, W. L. (1980). *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood, NJ: Ablex
- Chang, C. F. (2012). Peer review via three modes in an EFL writing course. *Computers and Composition*, 29(1), 63–78.
- Chang, C. Y. H. (2016). Two decades of research in L2 peer review. *Journal of Writing*

- Research*, 8(1), 81–117.
- Connor, U., & Asenavage, K. (1994). Peer response groups in ESL writing classes: How much impact on revision? *Journal of Second Language Writing*, 3(3), 257–276.
- Conrad, S., & Goldstein, L. M. (1999). ESL student revision after teacher-written comments: Text, contexts, and individuals. *Journal of Second Language Writing*, 8(2), 147–179.
- de Guerrero, M., & Villamil, O. S. (1994). Social-cognitive dimensions of interaction in L2 peer revision. *The Modern Language Journal*, 78(4), 484–496.
- Ferris, D. (1997). The influence of teacher commentary on student revision. *TESOL Quarterly*, 31(2), 315–339.
- Ferris, D.R. (2003). *Response to student writing: Implications for second language students*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Ferris, D., Brown, J., Liu, H., & Stine, M. E. A. (2011). Responding to L2 students in college writing classes: Teacher perspectives. *TESOL Quarterly*, 45(2), 207–234.
- Gascoigne, C. (2004). Examining the effect of feedback in beginning L2 composition. *Foreign Language Annals*, 37(1), 71–76.
- Guardado, M., & Shi, L. (2007). ESL students' experiences of online peer feedback. *Computers and Composition*, 24(4), 443–461.
- Hall, C. (1990). Managing the complexity of revising across languages. *TESOL Quarterly*, 24(1), 43–60.
- Halliday, M. A. K. (1967). Notes on transitivity and theme in English Part I. *Journal of Linguistics*, 3(1), 37–81.
- Ho, M. (2015). The effects of face-to-face and computer-mediated peer review on EFL writers' comments and revisions. *Australasian Journal of Educational Technology*, 31(1), 1–15.
- Hyland, K. (2016). Methods and methodologies in second language writing research. *System*, 59, 116–125.
- Liou, H. C., & Peng, Z. Y. Y. (2009). Training effects on computer-mediated peer review. *System*, 37(3), 514–525.
- Liu, J., & Sadler, R. W. (2003). The effect and affect of peer review in electronic versus traditional modes on L2 writing. *Journal of English for Academic Purposes*, 2(3) 193-227.
- Lockhart, C., & Ng, P. (1995). Analyzing talk in ESL peer response groups: Stances, functions, and content. *Language Learning*, 45(4), 605–651.
- Mangelsdorf, K., & Schlumberger, A. (1992). ESL student response stances in a peer-review task. *Journal of Second Language Writing*, 1(3), 235–254.
- McGroarty, M. E., & Zhu, W. (1997). Triangulation in classroom research: A study of peer revision. *Language Learning*, 47(1), 1–43.

- Mendonça, C. O., & Johnson, K. E. (1994). Peer review negotiations: Revision activities in ESL writing instruction. *TESOL Quarterly*, 28(4), 745–769.
- Nelson, G. L., & Murphy, J. M. (1992). An L2 writing group: Task and social dimensions. *Journal of Second Language Writing*, 1(3), 171–193.
- Nelson, G. L., & Murphy, J. M. (1993). Peer response groups: Do L2 writers use peer comments in revising their drafts? *TESOL Quarterly*, 27(1), 135–141.
- Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science*, 37(4), 375–401.
- Purdue Online Writing Lab. (n.d.). Higher order concerns (hocs) and lower order concerns (locs). Retrieved August 2, 2020, from https://owl.purdue.edu/owl/general_writing/mechanics/hocs_and_locs.html
- Reid, J. (1994). Responding to ESL students' texts: The myths of appropriation. *TESOL Quarterly*, 28(2), 273–292.
- Stanley, J. (1992). Coaching student writers to be effective peer evaluators. *Journal of Second Language Writing*, 1(3), 217–233.
- Van Waes, L., Van Weijen, D., & Leijten, M. (2014). Learning to write in an online writing center: The effect of learning styles on the writing process. *Computers & Education*, 73, 60–71.
- Zhu, W. (2001). Interaction and feedback in mixed peer response groups. *Journal of Second Language Writing*, 10(4), 251–276.

APPENDIX

ANTS Coding Scheme With Examples

Area Type Specificity	Global Revision	Global Non-Revision	Local Revision	Local Non-Revision
Evaluations Generic	Not sure this is a good count-argument to your thesis.	This is a very nice thesis statement.	I don't really understand.	The reference page is good.
Evaluations Specific	again merely criticizing Park.	This enhances readers' understanding innate morality	'people of the language' seems vague.	...thesis statement is explain the factors of crises...
Clarifications Generic	Is there proof/evidence of this?	(No examples)	I am not sure but is this correct APA style?	(No examples)
Clarifications Specific	Not appropriate for what reasons?	You think the mayor is not a sincere politician	What do you mean by 'distinct features?'	(No examples)

Suggestions Generic	It would be better if you sum up the body paragraphs.	(No examples)	...need a different way to cite something on your draft.	(No examples)
Suggestions Specific	I think you should mention horizontal structure more...	(No examples)	...citing the part where Baldwin claims such thing is better.	(No examples)
Alterations Generic	(No examples)	(No examples)	“Put a comma”	(No examples)
Alterations Specific	(No examples)	(No examples)	“Change it into dynamic classroom”	(No examples)
Statements Generic	...considering the audiences are english learning students...	I believe this is your thesis statement.	If you have in-text citation...	...I could understand the meaning of words...
Statements Specific	There are people saying they rather not...	I did not realize that Black English is that much different...	For people who didn't read the essay of Langer...	(No examples)