

New technologies for feedback on, and assessment of, L2 writing

Jim Ranalli

Iowa State University

A tale of two technologies

Untitled document

that violent games will make players more easily have violent moods, words, and even movements in reality.

However, this is not the reason that violent games should be blocked for players. Firstly, both researchs is mainly focusing on students players, and their reaction could not prove to be the common influence for players that are aged 20 or more. Their minds might be easily influenced by the environment and daily activities. Thus playing the violent game frequently will make their minds aggressive in an apparent way. But for the other players that might be graduate students or workers, they have experienced more, gained enough pieces of knowledge, and had a deep understanding of our society and justice. Therefore for these

2 Correctness ⓘ Back to all suggestions X

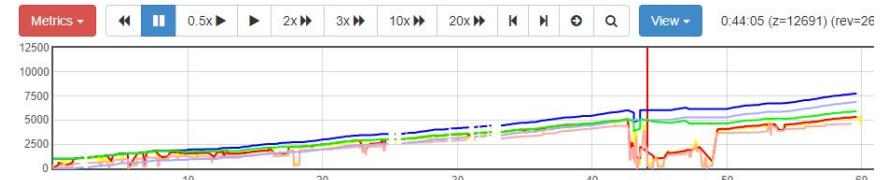
• SPELLING

researchs → research

The word **researchs** is not in our dictionary. If you're sure this spelling is correct, you can add it to your personal dictionary to prevent future alerts.

Add to dictionary ⌂ : ⌂

• they - Correct pronoun usage



necessary until you are satisfied with the final product.

Topic: Many adults become upset when young people break with traditions of the past. Do you think that these adults are justified in reacting this way? Why or why not? Support your position with evidence from your own experience or the experiences of people you know.

Automated Written Corrective Feedback

Process Tracing via Keystroke Logging and Eyetracking

Comparison of AWCF Tools, AWE Tools, and MS-NLP

	AWCF Tools	AWE Tools	MS-NLP
Examples	<i>Grammarly, Ginger, Grammar Suggestions</i>	<i>Criterion, MY Access!, Research Writing Tutor, Writing Pal</i>	The spelling and grammar checker in MS Word, Outlook, and Office 365
Access	Multiple ways (e.g., web apps, browser extensions, productivity software plugins, mobile device keyboards)	Standalone web-based interfaces	Office productivity software
Delivery Mode	Synchronous and asynchronous	Asynchronous only	Synchronous and asynchronous
Analysis	Combinations of complex techniques performed on remote servers	Combinations of complex techniques performed on remote servers	Simpler techniques performed on user's local machine
Focus	Lower-level concerns (e.g., spelling, grammar, and punctuation)	Lower- and higher-level, or only higher-level, concerns (e.g., organization, discourse, writing strategies)	Lower-level concerns

Rank	Error Type	Example
1	Content Word Choice Error	<i>We need to deliver the merchandise on a daily *base/basis.</i>
2	Preposition Error	<i>Our society is developing *in/at high speed.</i>
3	Determiner Error	<i>We must try our best to avoid *the/a shortage of fresh water.</i>
4	Comma Error	<i>However, */, I'll meet you later.</i>
5	Inflectional Morphology	<i>The women *weared/wore long dresses.</i>
6	Wrong Verb Tense	<i>I look forward to *see/seeing you.</i>
7	Derivational Morphology	<i>It has already been *arrangement/arranged.</i>
8	Pronoun	<i>I want to make *me/myself fit.</i>
9	Agreement Error	<i>I *were/was in my house.</i>
10	Run-on Sentence	<i>They deliver documents to them they provide fast service.</i>

1. How do Grammarly and MS-NLP compare in terms of their capacity to address error types common to L2 student writers?
2. How do Grammarly and MS-NLP compare in terms of the timing of error flagging?

Ranalli & Yamashita, in press

Performance of Grammarly and MS-NLP in Addressing Common L2 Problem Areas

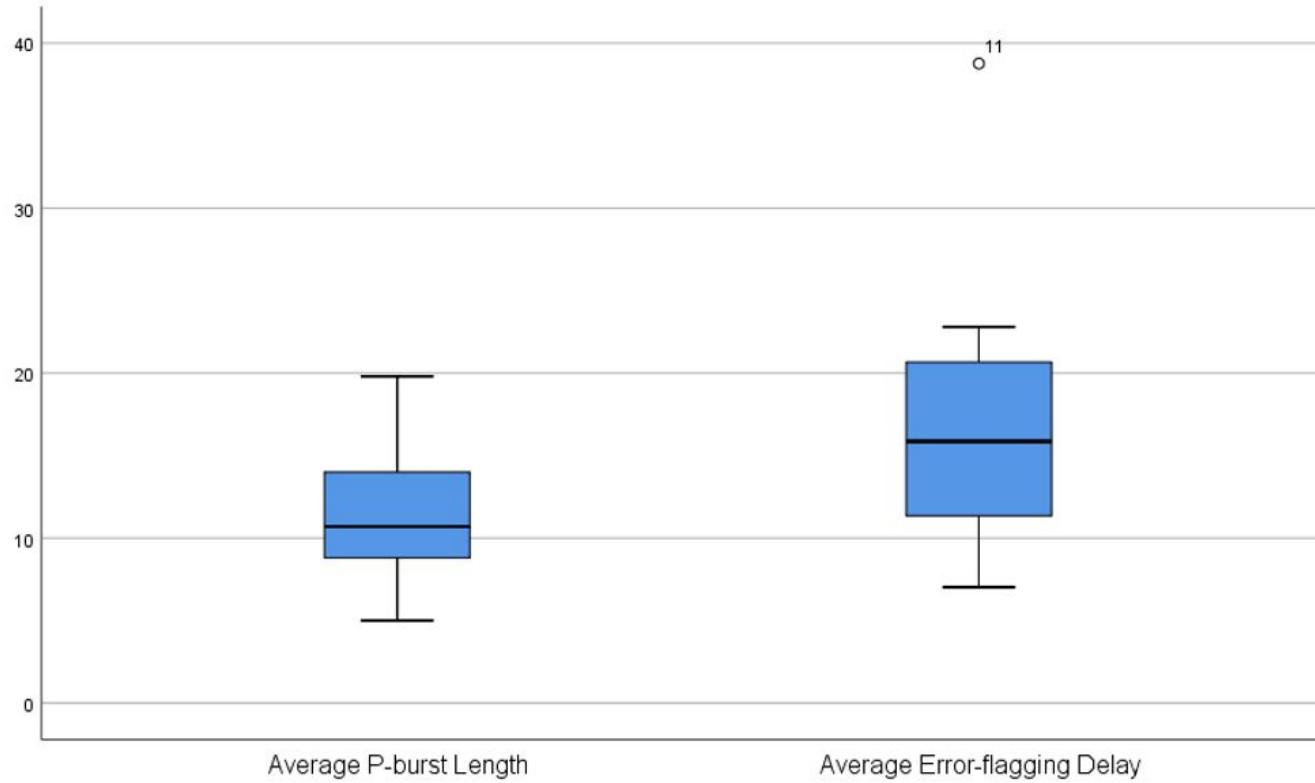
	Grammarly						MS-NLP					
	Flaggings	% of Total Flaggings	Error-types	Precision	Correction	Flaggings	% of Total Flaggings	Error-types	Precision	Correction		
Agreement Error	160	11.33%	22	0.84	0.79	61	7.80%	1	0.87	0.82		
Comma Error	141	9.99%	12	0.89	0.88	1	0.10%	1	0	0		
Content Word Choice Error	60	4.25%	5	0.83	0.8	2	0.3%	1	1	1		
Determiner Error	289	20.47%	9	0.8	0.78	8	1.03%	2	0.88	0.75		
Preposition Error	79	5.59%	5	0.94	0.94	2	0.30%	1	1	0		
Pronoun Error	30	2.12%	6	0.9	0.83	1	0.10%	1	1	1		
Run-on Sentence	2	0.14%	1	1	1	-	-	-	-	-		
Verb-form Error	95	6.73%	19	0.84	0.73	6	0.80%	1	1	0.83		
Wrong verb Tense	-	-	-	-	-	-	-	-	-	-		
Totals	856	60.62%	79	0.84	0.81	81	10.43%	8	0.88	0.79		

Frequency of Flaggings by Temporal Location

	Grammarly	MS-NLP
Point of Inscription	113 (21.1%)	475 (88.6%)
Earlier in the Text	422 (78.9%)	61 (11.4 %)
Total	535	536

Ranalli & Yamashita, in press

Average p-burst length versus average error-flagging delay ($n = 20$)



Ranalli & Yamashita, in press

Distribution of online revisions

In timed writing, L2 learners tend to do more revision at later stages of writing.
Barkaoui, 2016; Roca de Larios et al., 2008

→ How do synchronous AWCF tools affect the distribution of revisions?
(e.g., *Do tools prompt learners to revise more in the earlier stages of writing?*)

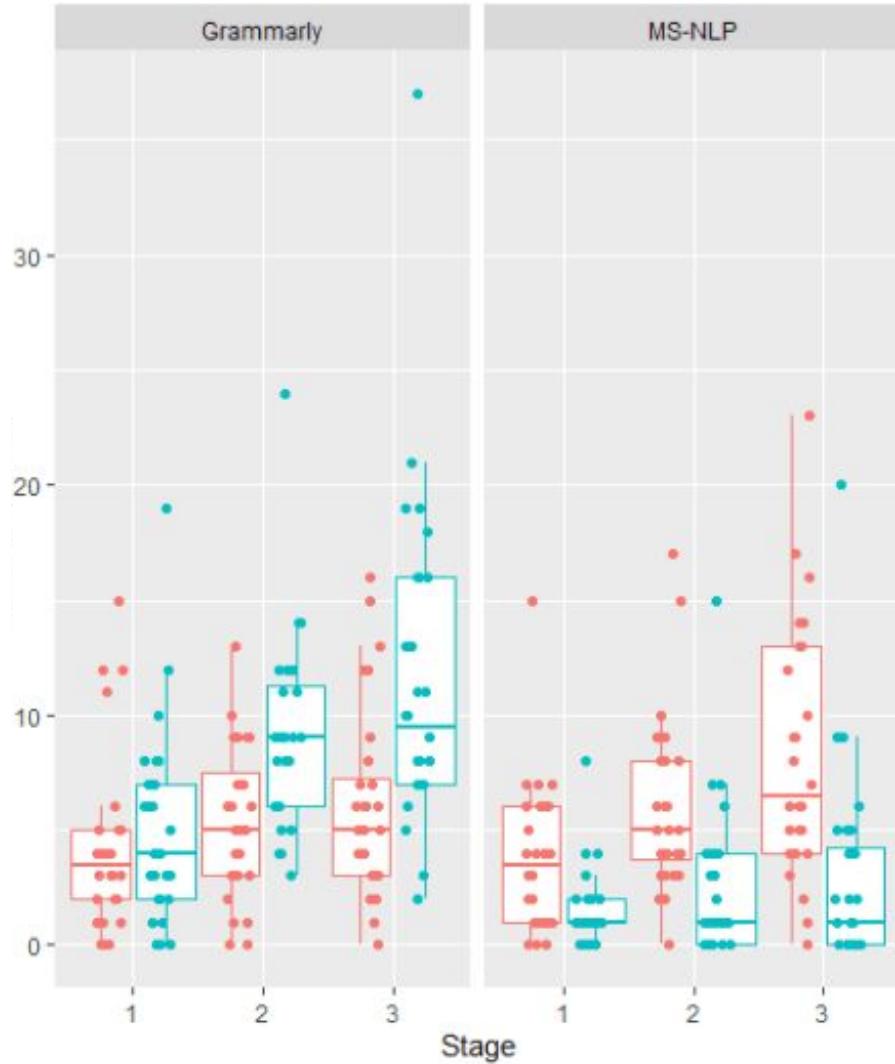
Writer-initiated revisions

Tool-initiated revisions

Ranalli & Yamashita, under review

How does choice of synchronous AWCF tool influence the frequency of revisions, both writer- and tool-initiated, in general and at different stages of the composition process?

Ranalli & Yamashita, under review



N = 28

Implications of the Grammarly evaluations

L2 student writers stand to benefit from the enhanced error-correction capabilities of AWCF tools.

But ...

Users and L2 writing instructors should be made aware of potential problems with synchronous operation.

L2 writing assessment research

- Timed, impromptu writing tests are incompatible with recursiveness, extensive revision, etc. Wolcott, 1987
- Calls for assessments that measure process Cho, 2003; Deane, 2013; Graham, Harris, & Hebert, 2011; and thus expand construct coverage Deane, 2013
- Large-scale standardized tests have recently experimented with keystroke logging Almond et al., 2012; National Assessment of Educational Progress, 2011

Writing process research

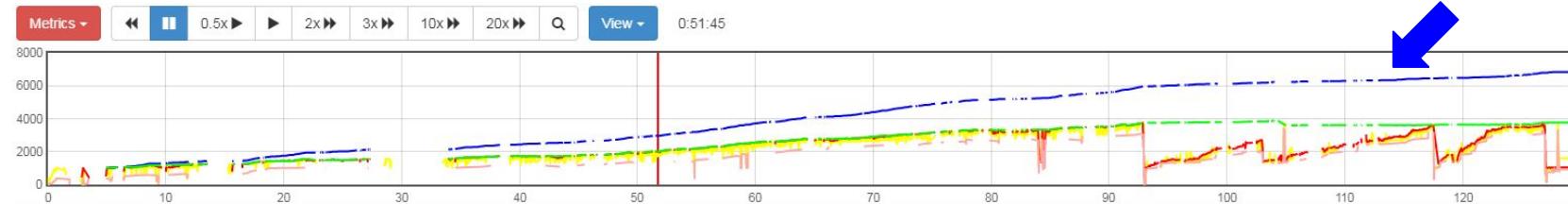
- Process visualizations used for analysis of keystroke (and in some cases eyetracking) data because of huge data sets generated by these methods Bécotte et al., 2015
- Visualizations used as formative feedback on composing behavior in small-scale, one-off studies Lindgren, Stevenson, & Sullivan, 2008; Sullivan & Lindgren, 2002

Do process-tracing technologies provide insights about L2 student writing that can enhance instructors' efforts to create more meaningful assessment for learning?

Ranalli, Feng, & Chukharev-Hudilanan, 2018

Post-session Viewer

Process graph



First reason is that with the renewal of the time everything has changed. Like the development of technology, much fierce competition, more new types of occupations come up. Base on these objective changing it is certain that the thought and action of young people change a lot and some of these changes will break the tradition. For example, in my hometown it is a tradition to visit all the relatives in their home and should bring the thought and expansive gifts. However, base on the short vacation and heavy work pressure the young people now prefer to eat in the restaurant together and go travelling instead of to visit every relatives' home everyday. |

← Playback

Case studies

Zedong*

- Male
- 19
- Freshman
- Lower-level ESL writing course
- L1 Mandarin
- Computer science
- IELTS 6.5 (TOEFL 79-93)

Mingyu*

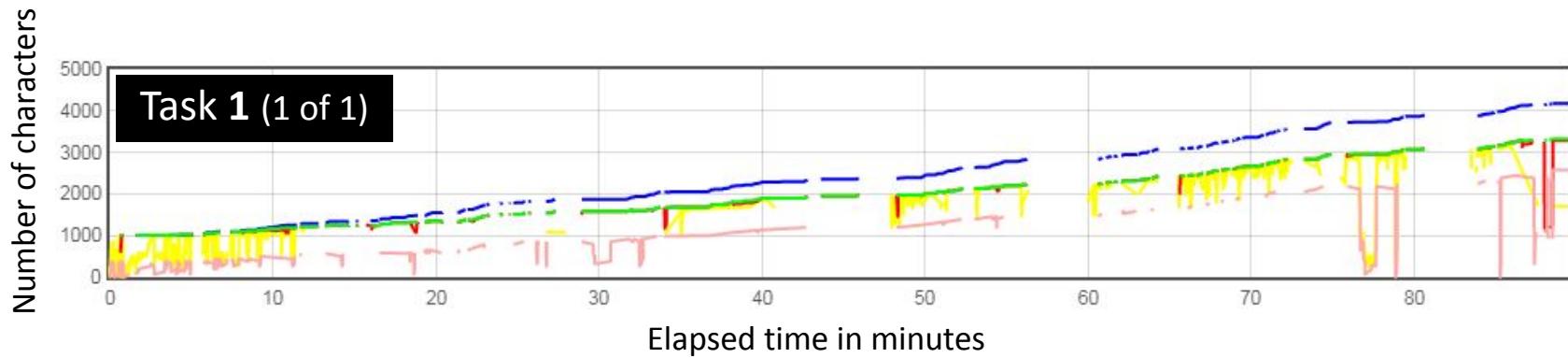
- Female
- 18
- Freshman
- Higher-level ESL writing course
- L1 Mandarin
- Computer engineering
- TOEFL 98

* Pseudonyms

Ranalli, Feng, & Chukharev-Hudilanan, 2018

Finding 1

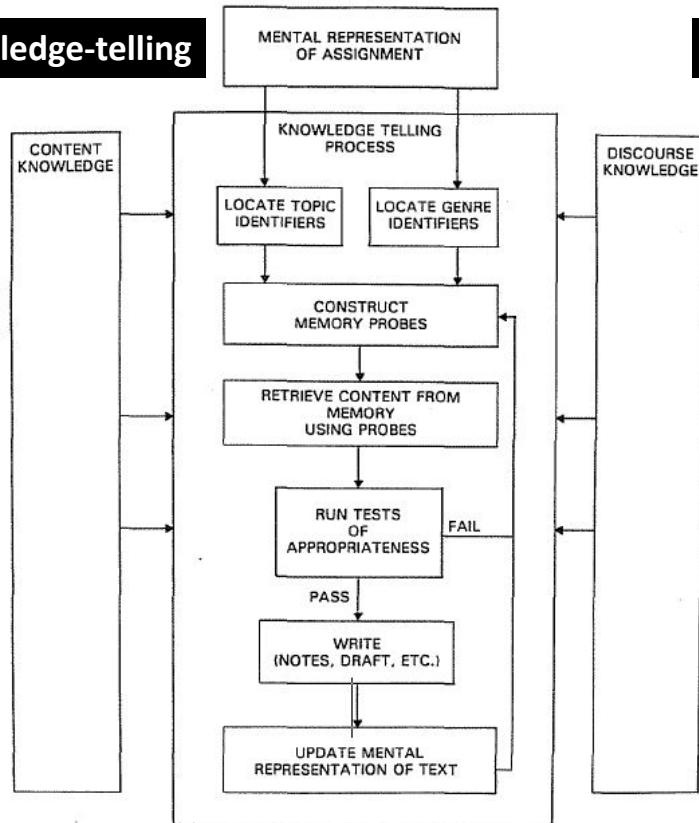
Diagnosing case studies with respect to developmental models



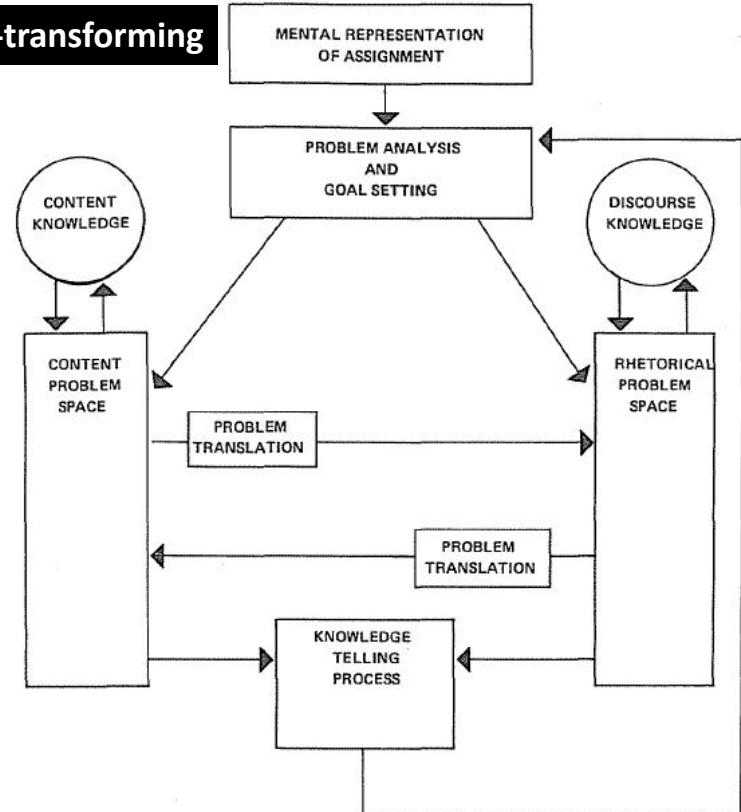
Zedong: Knowledge telling

Ranalli, Feng, & Chukharev-Hudilanan, 2018

Knowledge-telling



Knowledge-transforming

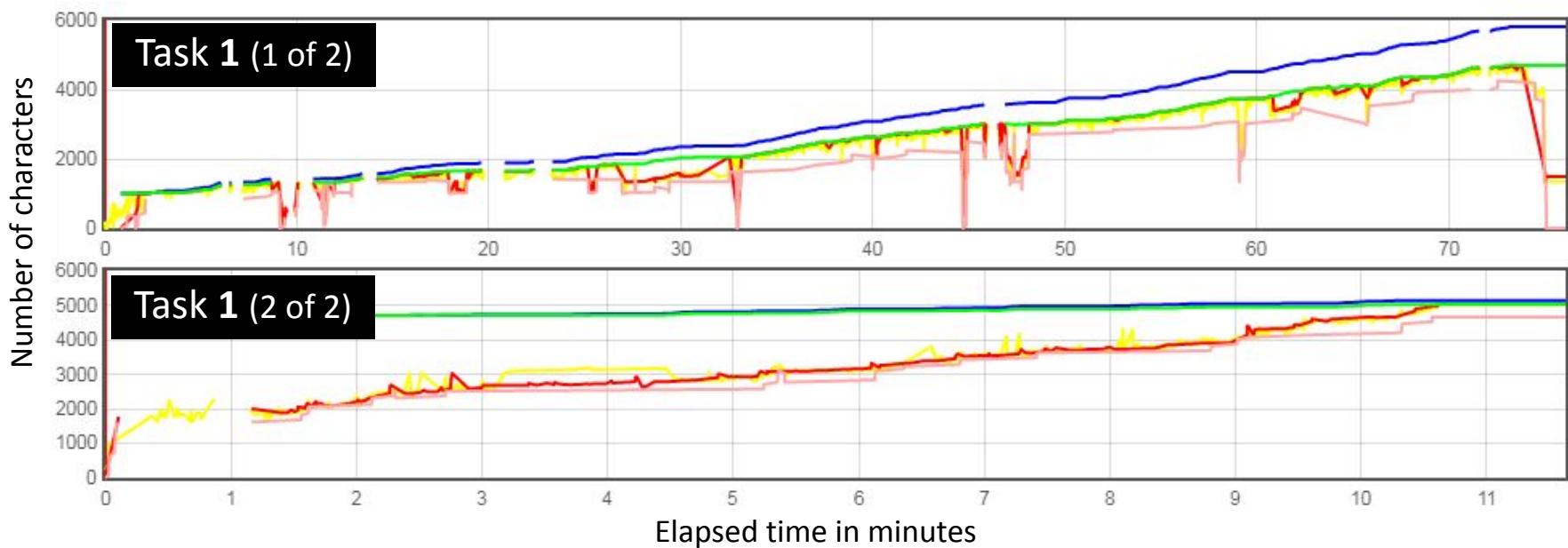


Developmental models of the writing process

Bereiter & Scardamalia, 1987, pp. 28-32

Finding 1

Diagnosing case studies with respect to developmental models

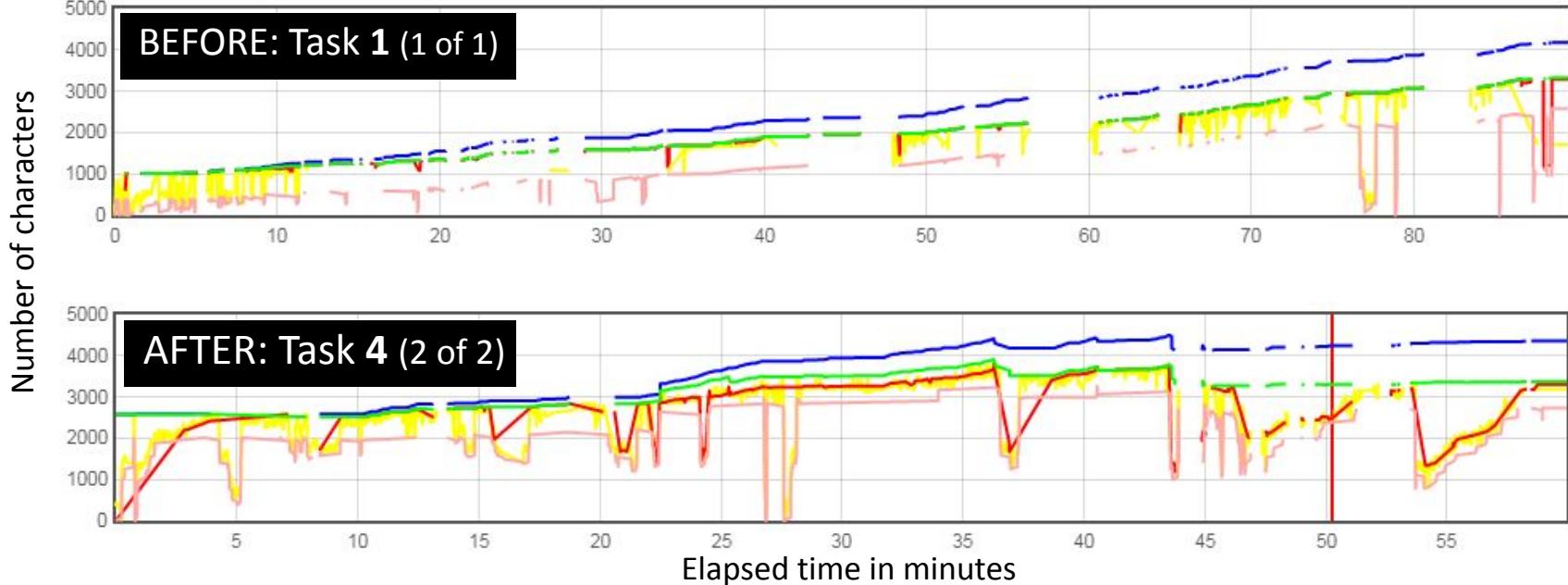


Mingyu: knowledge transforming

Ranalli, Feng, & Chukharev-Hudilanan, 2018

Finding 2

Identifying and addressing specific process/strategy issues

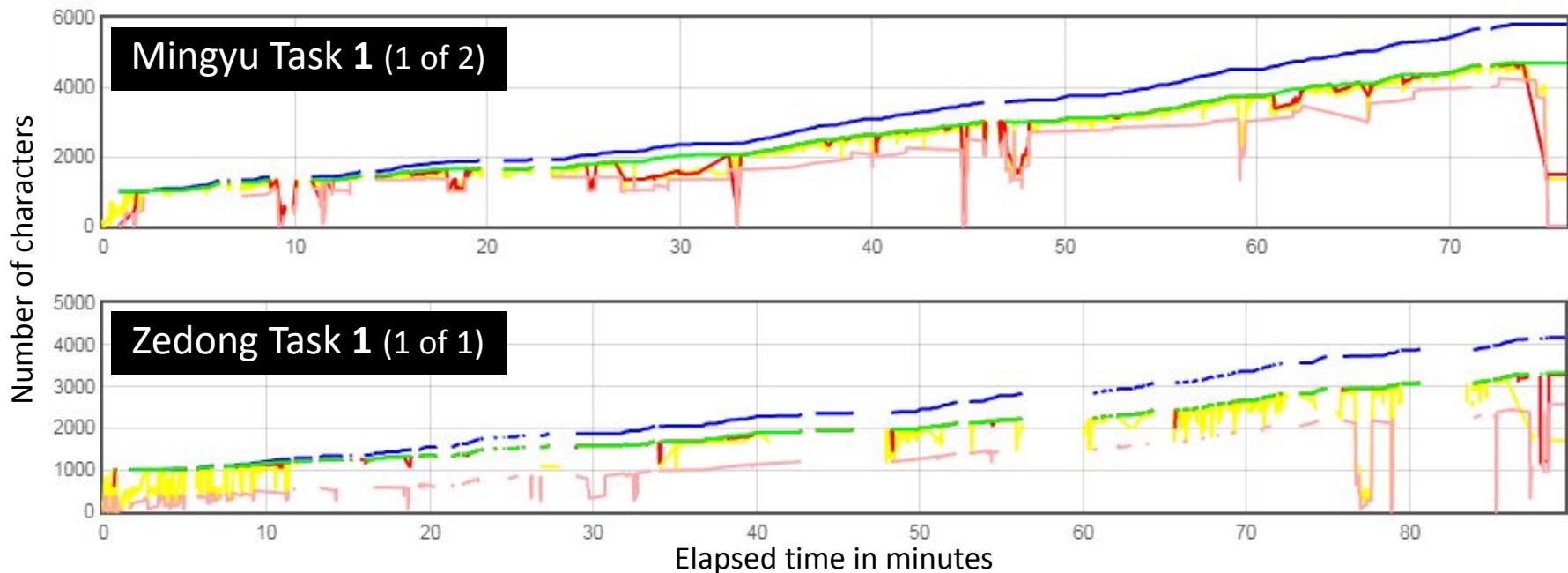


Zedong: inefficient formulation

Ranalli, Feng, & Chukharev-Hudilanan, 2018

Finding 2

Identifying and addressing specific process/strategy issues



Mingyu: idiosyncratic task definition

Ranalli, Feng, & Chukharev-Hudilanan, 2018

Summing up

Processing tracing supported more meaningful AfL by making it possible to:

- position participants in relation to developmental models of writing
- identify and address issues with specific processes and strategies

More generally, process tracing allows for ...

- an expansion of the construct of writing used in L2 classroom assessment
- writing-process research to inform instruction

Challenges to adoption of process tracing

- Making space for it among L2 writing instructors' already burdensome feedback duties
- Supporting students' abilities to self-assess and self-regulate their process engagement
- Providing teaching materials that support a data-driven (not generic) process approach

Ranalli, Feng, & Chukharev-Hudilainen, 2019

References

- Barkaoui, K. (2016). What and when second-language learners revise when responding to timed writing tasks on the computer: The roles of task type, second language proficiency, and keyboarding skills. *The Modern Language Journal*, 100(1), 320-340. doi:10.1111/modl.12316
- Bécotte, H.-S., Caporossi, G., Hertz, A., & Leblay, C. (2015). Writing and rewriting: Keystroke logging's colored numerical visualization.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cho, Y. (2003). Assessing writing: Are we bound by only one method? *Assessing Writing*, 8(3), 165-191. doi:[http://dx.doi.org/10.1016/S1075-2935\(03\)00018-7](http://dx.doi.org/10.1016/S1075-2935(03)00018-7)
- Deane, P. (2013). On the relation between automated essay scoring and modern views of the writing construct. *Assessing Writing*, 18(1), 7-24. doi:<http://dx.doi.org/10.1016/j.asw.2012.10.002>
- Graham, S., Harris, K., & Hebert, M. (2011). *Informing writing: The benefits of formative assessment*. In. Washington, DC: Carnegie Corporation of New York.
- Lindgren, E., Stevenson, M., & Sullivan, K. P. H. (2008). Supporting the reflective language learner with computer keystroke logging. In F. Zhang & B. Barber (Eds.), *Handbook of research on computer-enhanced language acquisition and learning* (pp. 189-204). Hershey, NY: Information Science Reference.
- Ranalli, J., Feng, H.-H., & Chukarev-Khudilaynen, E. (2019). The affordances of process-tracing technologies for supporting L2 writing instruction. *Language Learning & Technology*, 23(2), 1-11. doi:<https://doi.org/10125/44678>
- Ranalli, J., Feng, H.-H., & Chukharev-Hudilainen, E. (2018). Exploring the potential of process-tracing technologies to support assessment for learning of L2 writing. *Assessing Writing*, 36, 77-89. doi:<https://doi.org/10.1016/j.asw.2018.03.007>
- Ranalli, J., & Yamashita, T. (in press). Automated written corrective feedback: Error-correction performance and timing of delivery. *Language Learning & Technology*.
- Ranalli, J., & Yamashita, T. (under review). Synchronous automated written corrective feedback in timed L2 writing: Effects on the distribution and focus of revisions and writing quality.
- Roca de Larios, J., Manchón, R. M., Murphy, L., & Marín, J. (2008). The foreign language writer's strategic behaviour in the allocation of time to writing processes. *Journal of Second Language Writing*, 17(1), 30-47. doi:<http://dx.doi.org/10.1016/j.jslw.2007.08.005>
- Sullivan, K. P. H., & Lindgren, E. (2002). Self-assessment in autonomous computer-aided second language writing. *ELT Journal*, 56(3), 258-266. doi:10.1093/elt/56.3.258
- Wolcott, W. (1987). Writing instruction and assessment: The need for interplay between process and product. *College Composition and Communication*, 38(1), 40-46. doi:10.2307/357585

For discussion

What are some of the main challenges and opportunities that new technologies present to the field of L2 writing assessment?