



IOWA STATE UNIVERSITY

Corpus and genre-based
automated writing evaluation
for scientific writing

Dr. Elena Cotos

International Perspectives on Corpora for Language Learning
University of Queensland
22 November 2022



Outline

- Background to genre-based Automated Writing Evaluation
- Research Writing Tutor: Genre-based AWE exemplar for scientific writing
 - A. Corpus as data
 - B. Corpus for feature design
 - C. Corpus for pedagogical implementation
- Learning potential and impact
- Future directions

Automated Writing Evaluation (AWE)

Ability of computer technology to evaluate and score written prose (Shermis & Burstein, 2003)

- 1st generation – grading of student writing
 - Project Essay Grade

A “comprehensive history of AWE has yet to be written”
(Hazelton et al., 2021)

- 2nd generation – individualized feedback on errors





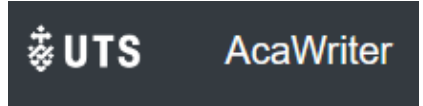
- 3rd generation – analysis of student writing across academic disciplines and writing genres

Genre-based AWE

3rd generation of AWE took a “left turn” (Burstein et al., 2016a, p. 6)

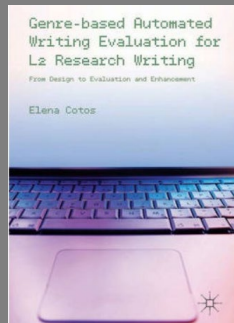
- **From** quantifiable features that, “in the aggregate, embody the meaning of writing for the assessment” (Williamson, 2013, p. 166)
- **To** “the rhetorical ability to integrate an understanding of audience, context, and purpose [...]; the ability to effectively employ multiple writing strategies; the ability to learn and use the conventions appropriate to a specific genre of writing” (Perelman, 2012, p. 129)
- Motivated by EAP teaching and learning needs
- Grounded in ESP genre theory
- Focused on the design and pedagogical uses of automated feedback tools for genre writing
 - moves (communicative goals)
 - steps (writing strategies)

Genre-based AWE

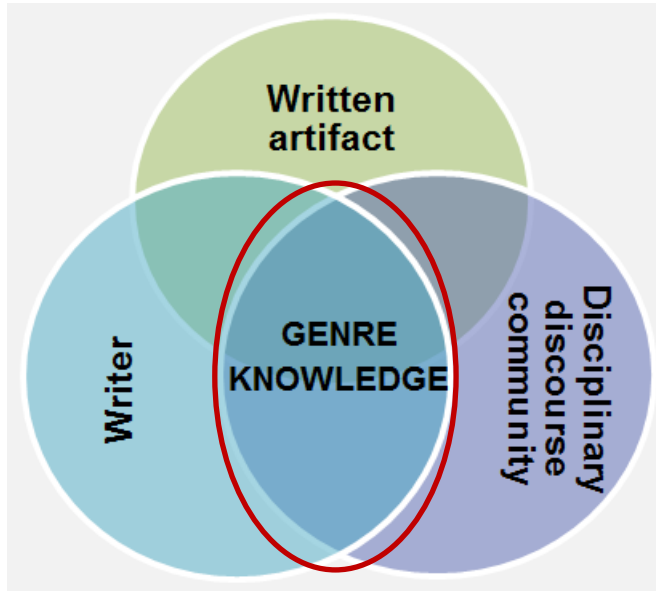
- **Mover** (Anthony & Lashkia, 2003)
 - Move structure of research article abstracts
 - Naïve Bayes classifier
- **Intelligent Academic Discourse Evaluator (IADE)** (Cotos, 2009)
 - Move feedback on research article Introduction sections
 - Support Vector Machines classifier (Pendar & Cotos, 2008)
- **Research Writing Tutor (RWT)** (Cotos, 2014, Cotos et al., 2015, 2016, 2017, 2020)
 - Move/step feedback on IMRD/C sections
 - Support Vector Machines classifier (Cotos & Pendar, 2016)
- **Academic Writing Analytics (AWA); AcaWriter** (Knight et al., 2020)
 - Feedback on rhetorically salient sentences
 - Rule-based dependency parser (Sándor, Kaplan, & Rondeau, 2006)

Outline

- Background to genre-based Automated Writing Evaluation
- Research Writing Tutor: Genre-based AWE exemplar for scientific writing
 - A. Corpus as data
 - B. Corpus for feature design
 - C. Corpus for pedagogical implementation
- Learning potential and impact
- Future directions



Genre-based AWE for scientific writing



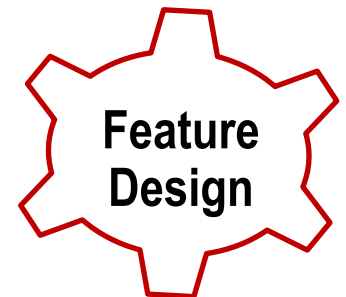
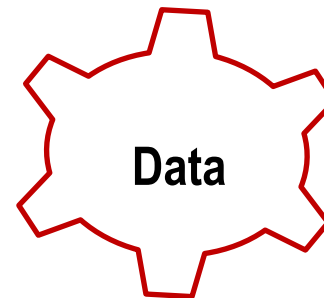
- Writer's self-awareness and metacognitive knowledge of the rhetorical task
- Socio-disciplinary awareness about the discourse community
- Metapragmatic ability to produce a research genre artifact
 - communicative action
 - genre-specific language choices
 - appropriate to the expectations of the disciplinary discourse community

Genre-based AWE for scientific writing

- Invaluable tool for genre analysis (Hyland, 2007, p. 224)
- Powerful methodology-technology for determining how disciplines use language in their major genres (Lee & Swales, 2006, p. 57)

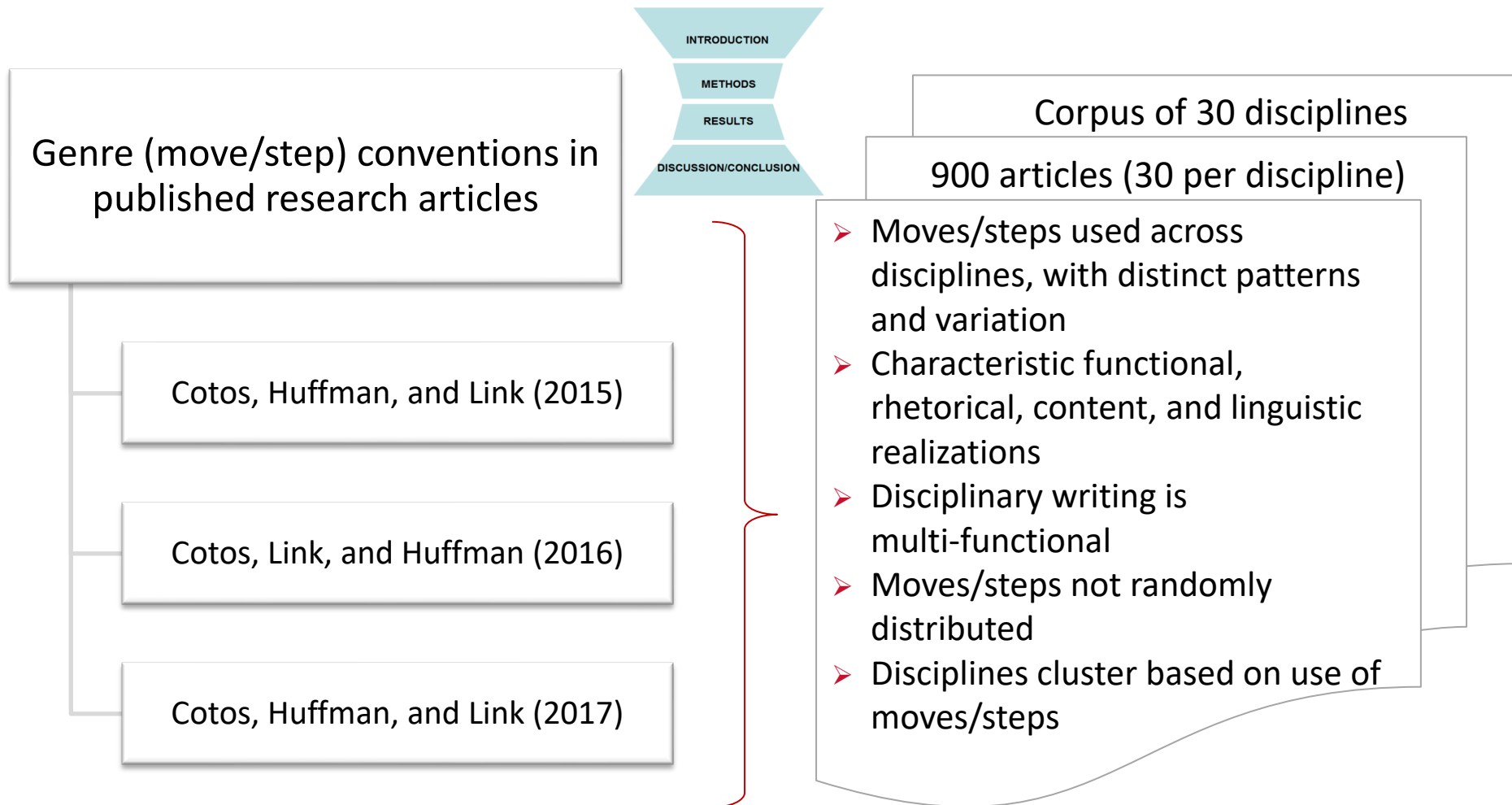


Corpus for RWT development



A. RWT: Corpus as data

Corpus analysis and annotation of research articles in the disciplines



A. RWT: Corpus as data

■ Cross-disciplinary conventions

Introduction	Methods	Results	Discussion/Conclusion
Move 1: Establishing the territory Step 1 – Claiming centrality <i>and/or</i> Step 2 – Providing general background <i>and/or</i> Step 3 – Reviewing previous research	Move 1: Contextualizing the study methods Step 1 – Referencing previous works <i>and/or</i> Step 2 – Providing general information <i>and/or</i> Step 3 – Identifying the methodological approach <i>and/or</i> Step 4 – Describing the setting <i>and/or</i> Step 5 – Introducing the subjects/participants <i>and/or</i> Step 6 – Rationalizing pre-experiment decisions	Move 1: Approaching the niche Step 1 – Providing general orientation <i>and/or</i> Step 2 – Restating study specifics <i>and/or</i> Step 3 – Justifying study specifics	Move 1: Re-establishing the territory Step 1 – Drawing on a/theoretical general background <i>and/or</i> Step 2 – Drawing on study-specific background <i>and/or</i> Step 3 – Highlighting principal findings <i>and/or</i> Step 4 – Previewing the discussion 'road map'
Move 2: Identifying a niche Step 1 – Indicating a gap <i>and/or</i> Step 2 – Highlighting a problem <i>and/or</i> Step 3 – Raising general questions <i>and/or</i> Step 4 – Proposing general hypotheses <i>and/or</i> Step 5 – Presenting justification	Move 2: Describing the study Step 1 – Acquiring the data <i>and/or</i> Step 2 – Describing the data <i>and/or</i> Step 3 – Identifying variables <i>and/or</i> Step 4 – Delineating experimental/study procedures <i>and/or</i> Step 5 – Describing tools/instruments/materials/equipment <i>and/or</i> Step 6 – Rationalizing experiment decisions <i>and/or</i> Step 7 – Reporting incrementals	Move 2: Occupying the niche Step 1 – Reporting specific results <i>and/or</i> Step 2 – Indicating alternative presentation of results	Move 2: Framing the new knowledge Step 1 – Explicating results <i>and/or</i> Step 2 – Accounting for results <i>and/or</i> Step 3 – Clarifying expectations <i>and/or</i> Step 4 – Addressing limitations
Move 3: Addressing the niche Step 1 – Introducing present research descriptively <i>and/or</i> Step 2 – Announcing present research purposefully <i>and/or</i> Step 3 – Presenting research questions <i>and/or</i> Step 4 – Presenting research hypotheses <i>and/or</i> Step 5 – Clarifying definitions <i>and/or</i> Step 6 – Summarizing methods <i>and/or</i> Step 7 – Announcing principle outcomes <i>and/or</i> Step 8 – Stating the value of present research <i>and/or</i> Step 9 – Outlining the structure of the paper	Move 3: Establishing credibility Step 1 – Preparing the data <i>and/or</i> Step 2 – Describing the data analysis <i>and/or</i> Step 3 – Rationalizing data processing/analysis	Move 3: Construing the niche Step 1 – Comparing results <i>and/or</i> Step 2 – Accounting for results <i>and/or</i> Step 3 – Explicating results <i>and/or</i> Step 4 – Clarifying expectations <i>and/or</i> Step 5 – Acknowledging limitations	Move 3: Reshaping the territory Step 1 – Supporting with evidence <i>and/or</i> Step 2 – Countering with evidence
	Move 4: Expanding the niche³ Step 1 – Generalizing results <i>and/or</i> Step 2 – Claiming the value <i>and/or</i> Step 3 – Noting implications <i>and/or</i> Step 4 – Proposing directions	Move 4: Establishing additional territory Step 1 – Generalizing results <i>and/or</i> Step 2 – Claiming the value <i>and/or</i> Step 3 – Noting implications <i>and/or</i> Step 4 – Proposing directions	

(Cotos et al., 2015)

A. RWT: Corpus as data

■ Description of moves and steps

Rhetorical Intent <i>[Why?]</i>	Content Realizations <i>[What content?]</i>
<ul style="list-style-type: none">– to give meaning to the results– provide logic behind arguments related to principal findings– to take a position and incline the reader to accept it	<p>Statements that express authors':</p> <ul style="list-style-type: none">– interpretations of the results– inferences based on the results– deductions from the results– evaluation of the results– hypotheses based on the results
Linguistic Realizations <i>[What language choices?]</i>	
<ul style="list-style-type: none">– <u>The results indicate that</u> incumbents do indeed react preemptively to Southwest's entry threat. [ECON]– <u>The validation results presented suggest that</u> STEMS-Air <u>can be applied to</u> both short-term and long-term modelling of PM10. [ENVE]	

Example

- Characteristic functional, rhetorical, content, and linguistic realizations

Discussion section

Move 2: Framing new knowledge

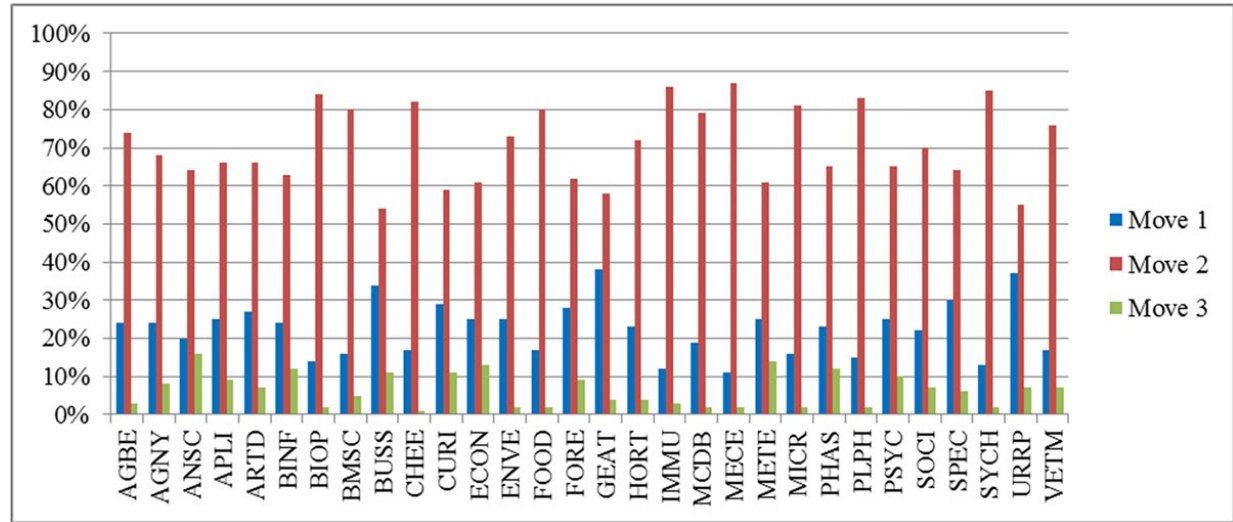
Step 1: Explicating findings

(Cotos et al., 2016)

A. RWT: Corpus as data

- Frequency of moves and steps in the disciplines

Methods section



(Cotos et al., 2017)

A. RWT: Corpus as data

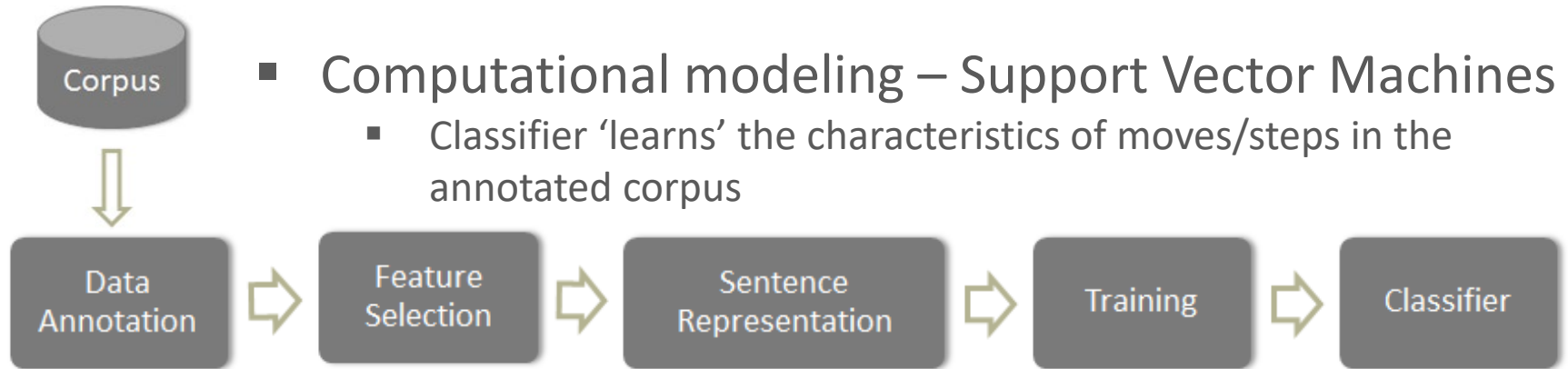
- Linguistic features for computational modeling

<i>Move 1</i>		<i>Move 2</i>		<i>Move 3</i>	
<i>Step 1-Claiming Centrality</i>		<i>Step 2-Highlighting a problem</i>		<i>Step 7-Announcing principal outcomes</i>	
n-gram	OR	n-gram	OR	n-gram	OR
recent year	20.01	as a result	6.24	found that	22.18
the past	19.53	due to the	3.52	show that	9.81
wide us	16.52	because of the	3.50	we also	8.41
interest in	15.50	the fact that	3.23	that in	5.96
the last	14.13	have not been	3.22	demonstr that	5.86
import in	11.51	larg number of	3.09	the averag	5.84
bodi of	8.43	there i a	2.44	the result	5.56
been studi	7.40	the major of	2.29	demonstr the	5.25
an import	7.28	to be a	2.28	consist with	5.07
of interest	6.95	a result of	2.24	reveal that	4.72
plai an	6.88	the number of	2.23	here we	4.70
ar import	6.70	lead to the	2.05	the membran	4.26
research have	6.43	a rang of	1.99	evid of	4.22
most import	6.40	on the other	1.92	correspond to	4.08
over the	5.92	of the most	1.89	evid that	3.96

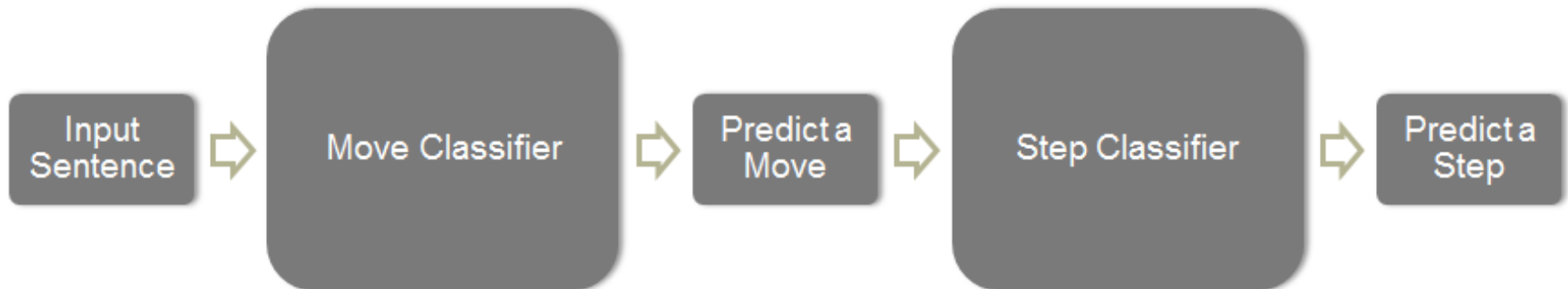
[OR - probability that n -gram will occur in a given step versus the probability that it will not]

(Cotos & Pendar, 2016)

A. RWT: Corpus as data



- Classifier predicts the moves/steps that new texts should have in order to be classified similarly to human annotation

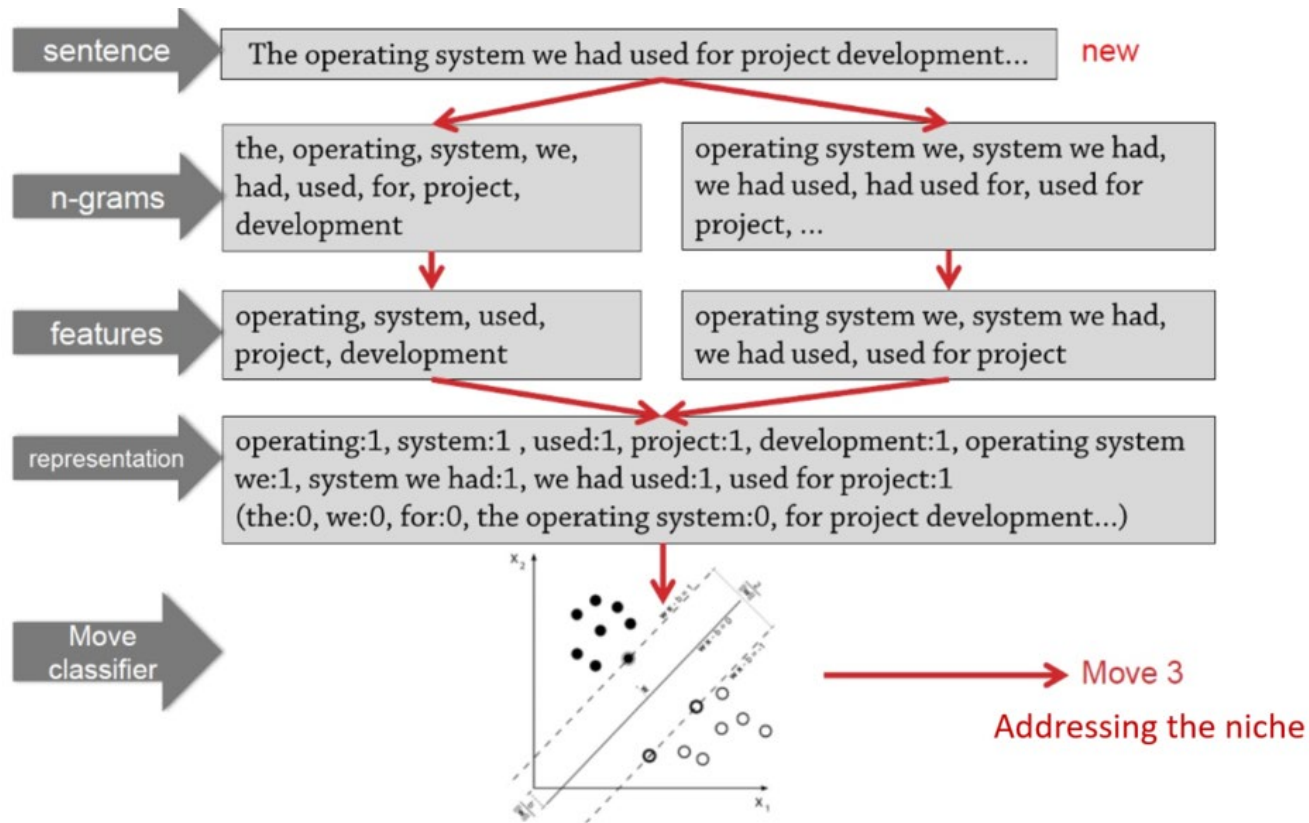


- Identified moves/steps passed on for generation of different types of feedback

(Cotos & Pendar, 2016)

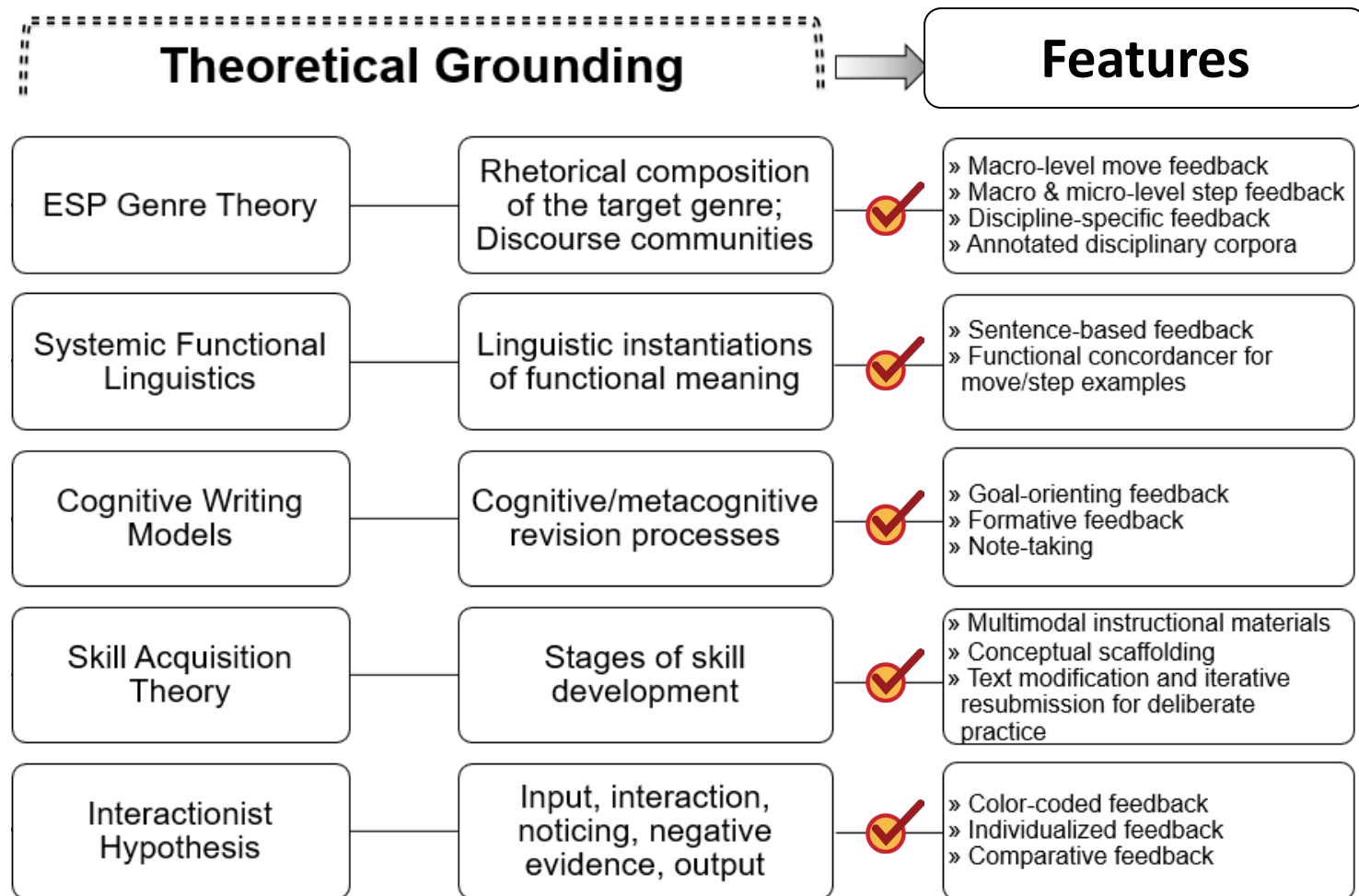
A. RWT: Corpus as language data

- Computational modeling

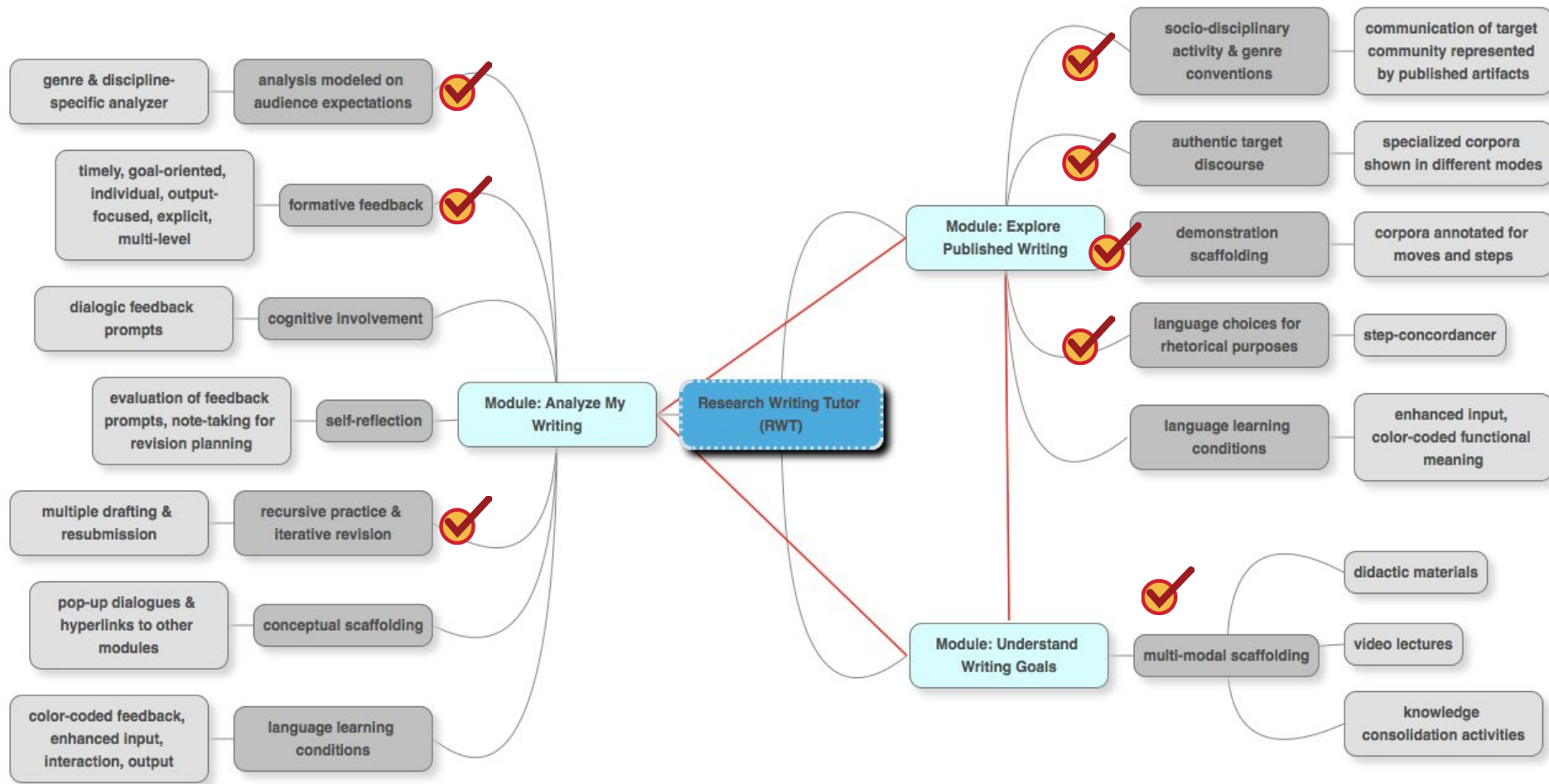


(Cotos & Pendar, 2016)

B. RWT: Corpus for feature design



B. RWT: Corpus for feature design



(Cotos, 2015)

B. RWT: Corpus for feature design

- Instruction: 'Understand Writing Goals' module

The screenshot shows a video lecture interface. On the left, a video window shows a female presenter. Below the video is a progress bar and a table of contents. On the right, a slide titled 'IOWA STATE UNIVERSITY Research Writing' displays 'Methods Move 1: Contextualizing the Study Methods'. Below the slide, two smaller panels show 'The Moves in Methods' and 'Steps in Move 1'.

Feature

Learning materials
(each move/step explained/exemplified)

Emphasis

Multi-modal scaffolding
(video and textual)

Table of Contents:

Time	Topic
0:00	Methods Move 1: Contextualizing the Study Methods
0:14	The Moves in Methods
1:54	Steps in Move 1
2:43	Referencing previous works
4:38	Providing general information
6:48	Identifying the methodological approach
8:10	Describing the setting
10:17	Introducing the subjects/participants

The Moves in Methods

- Move 1: Contextualizing the study methods
- Move 2: Describing the study
- Move 3: Analyzing the data

Steps in Move 1

- Referencing previous works and
- Providing general information and
- Identifying the methodological approach and/or
- Describing the setting and/or
- Introducing the subjects/participants
- Rationalizing pre-experiment decisions



B. RWT: Corpus for feature design

- Pedagogical mediation: 'Explore Published Writing' module

Discipline: Agricultural and Bio-Systems Engineering Section: Introduction Search

1 2 3 Next ->

► Ni, J.-Q., Heber, A. J., Sutton, A. L., Kelly, D. T., "Mechanisms of Gas Releases from Swine Wastes", *Transactions of the ASABE* 52(6):2013-2025, 2009

▼ Xu, H., He, P., Wang, G., Yu, G., Shao, L., "Enhanced storage stability of aerobic granules seeded with pellets", *Bioresour*

1. Introduction
Aerobic granulation is a novel self-immobilization process of microorganisms. Compared with conventional activated sludge of better settleability, stronger microbial structure, higher biomass retention, and better ability to handle toxic compound (Liu and Tay, 2002; Morgenroth et al., 1997). The application of aerobic granules was regarded as one of the promising bi (Adav et al., 2008b). Unfortunately, aerobic granules could easily lose stability and activity during storage, which would be application (Adav et al., 2008a; Zhang et al., 2005). Tay et al. (2002a) and Ng (2002) noted that the granules became more soluble organic materials after storage for 8 weeks. Tay et al. (2002b) reported that aerobic granules lost about 60% metabolic activity and acetate-fed granules lost about 90% metabolic activity after storage for 8 weeks. Lee et al. (2000) reported the endogenous respiration in the core of granules after starved for one month. Some storage methods were shown to improve the stability of granules. Storage at subfreezing temperatures (-20 degrees Celsius) was an ideal method for preserving granule stability and activity. The addition of a stabilizing substance (phenol) in the storing solution was beneficial to preserve the granule stability. However, this method is not quite suitable for large-scale storage. The information about the storage stability of aerobic granules seeded with different inocula is hardly found in the literatures so far. In most studies, aerobic granules were used directly (Adav et al., 2008a). It is widely known that activated sludge flocs are mainly composed of cells and extracellular polymeric substance (EPS). Zhang et al. (2008a, b) employed a novel EPS fractionation approach to obtain the pellets (cells) by extracting the EPS matrix. In this study, activated sludge flocs and pellets, respectively, were cultivated in two sequencing batch reactors (SBRs), then these two were stored at +/- 1 degrees Celsius for 3 weeks. The main purposes of this study were to investigate their responses to storage and expansion. The EPS concentration variation during storage was monitored, and scanning electron microscopy (SEM) and confocal laser scanning microscopy (CLSM) were applied to investigate the microstructure and distribution patterns of EPS and cells. Increased knowledge on this issue would be useful for the application of aerobic granules.

Move 2. Identifying a niche: Highlighting a problem

Klocke, N., Currie, R., Tomsicek, D., Koehn, J., "Corn yield response to deficit irrigation", *Trans. ASABE* 54(3):931-940, 2011

Feature

Annotated corpus
(shows each move/step glossed and color-coded)

Emphasis

Representation of target disciplinary writing
(presents authentic texts as enhanced input)



B. RWT: Corpus for feature design

- Pedagogical mediation: 'Explore Published Writing' module

Discipline: Agricultural and Bio-Systems Engineering

Section: Introduction

Moves: Move 2-Identifying a niche

Step: Indicating a gap

SEARCH

Move/Step	Percentage
Highlighting a problem	51%
Presenting justification	19%
Raising general questions	1%
Indicating a gap	13%
Proposing general hypotheses	16%

Feature

Functional concordancer
(move/step search engine)

Emphasis

Examples of "Indicating a gap"

Dufault, N., Isard, S., "A portable rainfall simulator for evaluating the wet deposition of plant pathogens", *Applied Engineering in Agriculture* 26(1):71-78, 2010

Ecological and environmental factors that influence organisms during each stage of the dispersal process are important for understanding the movement and development of plant disease epidemics (Aylor, 1986; Madden, 1992). However, there is limited information about some of these processes, especially for the deposition of plant pathogen propagules from the atmosphere onto susceptible host plant tissues. Rainfall has been identified as an important mechanism in the spread of plant pathogen propagules (Aylor, 1986; Madden, 1992; Isard and Gage, 2001; Isard et al., 2005).

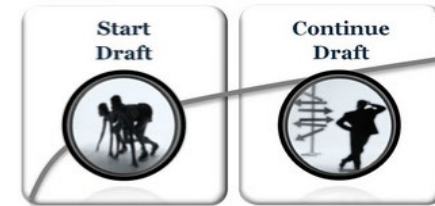
Wong, K., Sugumaran, M., Lee, D., Zahid, M., "Ecological aspects of endemic plant populations on Klang Gates quartz ridge, a habitat island in Peninsular Malaysia", *Biodiversity and Conservation* 17(1):1-12, 2008

This ridge is the most prominent of a series of such ridges in this part of Peninsular Malaysia. Current collecting indicates that inventory of the plant species is the earlier listings by Ridley (1922) (21 species) and Henderson (1928) (265 species), and a later enumeration by Kiew (1982) (175 species). Many records do not distinguish between an occurrence on the quartz ridge proper and one on the abutting (non-quartz) slopes.

Line, D., Shaffer, M., Blackwell, J., "SEDIMENT EXPORT FROM A HIGHWAY CONSTRUCTION SITE IN CENTRAL NORTH CAROLINA", *Transactions of the ASABE* 54(1):105-111, 2011

Historically, construction-related activities were cited by the state as a major source of degradation to lakes (NC DENR, 1992), and the perception that highways typically make up a significant proportion of urban construction, it is unknown what the contribution of highway construction is to the overall sediment development. In fact, few studies have documented the effects of highway construction on the water quality of receiving streams.

Language choices for rhetorical purposes
(all examples extracted from the corpus)



B. RWT: Corpus for feature design

- Feedback: 'Analyze My Writing' module

Emphasis

multiple drafting and continued revision

Analysis Results

Pharmaceuticals and wastewater treatment plants

Editing Introduce Add Methods Add Results Add Discussion

Your text

Pharmaceuticals are only partially removed in wastewater treatment plants (WWTPs) and have been detected in water bodies. For some, negative consequences on aquatic ecosystems have been established, but often cause-effect relationships are lacking [1]. X-ray contrast agents give rise to concern because of their persistence and usage in highest quantities [2]. Additionally, aquatic ecosystems must cope with mixtures of different compounds that interact or show concentration additivity [3]. Many countries are discussing precautionary measures; Swiss authorities propose upgrading the 100 largest WWTPs [4] by ozonation or activated carbon [5,6].

Upgrading WWTPs to deal with water pollution follows the ruling paradigm of centralized wastewater treatment, which is increasingly challenged [7]; source control is an alternative [8]. There are several source control options to decrease water pollution from pharmaceuticals. The most radical, strict prohibition, is hardly conceivable in the case of pharmaceuticals. Designing biodegradable pharmaceuticals and labeling of environmentally friendly medicaments are options [9]. Partial source control measures capture some micropollutants. Urine separation (NoMix-technology [8]); is one such possibility and could decrease 60-70% of the pharmaceutical load [10], and 50% of the ecotoxicological risk [11]. Another possibility, separate treatment of polluted waste-water known from

Words: 658 Goal: 426 to 1049

Iterative revision and submission Draft4 2019/02/27 14:11

ANALYZE

Export Options

Select Format: ☒ PDF ☐ Email

Include: ☒ Analysis and feedback ☒ Your notes

EXPORT

Macro-level color-coded move feedback

pathogens and multiantibiotic resistant bacteria give rise to concern [16]. To determine whether the hospital point source is relevant, several projects have quantified the pharmaceutical fraction from hospitals at WWTPs. Most found the hospitals' contribution low, ranging from 10 to 18% [17-19], because pharmaceuticals are widely used throughout the population. Moreover, although cytostatics or X-ray contrast agents are administered in hospitals, only ca. 50% of the latter are found in hospital sewers, because half the patients go home after X-ray-treatment [18]. Although the above studies indicate that hospitals are not always major sources of pharmaceuticals, it is premature to reject separate treatment of hospital wastewater. In a related study we found that the hospitals' contribution to

You are likely highlighting a problem.

Micro-level sentence feedback

potential are not necessarily those administered in large amounts [1]. Hence, further research to determine the relevance of hospitals as point-sources is needed, and this study intends to contribute. In complex decision situations with multiple, conflicting objectives and large uncertainty, Multiple-Criteria Decision Analysis (MCDA) offers support

Although the above studies indicate that hospitals are not always major sources of pharmaceuticals, it is premature to reject separate treatment of hospital wastewater.

You are likely highlighting a problem.

Macro-level discipline-specific step feedback

MOVE 1. ESTABLISHING A TERRITORY

You: 77%

46% - 58% 58% - 83% 83% - 95%

not enough goal too much

1 step(s) needs work | 2 step(s) good work >>

MOVE 2. IDENTIFYING A NICHE

You: 3%

0% - 7% 7% - 25% 25% - 33%

not enough goal too much

4 step(s) needs work | 1 step(s) good work >>

MOVE 3. ADDRESSING THE NICHE

You: 20%

0% - 5% 5% - 20% 20% - 28%

not enough goal too much

5 step(s) needs work | 4 step(s) good work >>

Lacking clarifying definitions, which is typically used in Agricultural and Bio-Systems Engineering papers. Needs more work. [Learn More](#) | [Examples](#)

Good work on introducing present research descriptively. Very similar to Agricultural and Bio-Systems Engineering papers.

B. RWT: Corpus for feature design

- Feedback: ‘Analyze My Writing’ module

Feature

Color-coded feedback
(entire text is shown in
colors representing
particular moves)

Emphasis

Textual context
(to depict the move
structure of the draft)

Tredici, 1987). Georgia plume grows on sand ridges, dry oak ridges, evergreen hammocks, and sandstone outcrops (Patrick et al., 1995). The apparently wide variety of habitats, but limited number of locations in which georgia plume occurs, has made it difficult to draw any definitive conclusions as to causes for rarity of the species. Georgia plume is reported to be found in less than three dozen

You are likely highlighting a problem.

Positioning System/Geographic Information System-based

Feature

Functional feedback
(comments or clarifying
questions about the step
function of individual
sentences)

Emphasis

Monitoring of rhetorical
intent
(to draw attention to
functional meaning)

B. RWT: Corpus for feature design

■ Feedback: 'Analyze My Writing' module

Feature

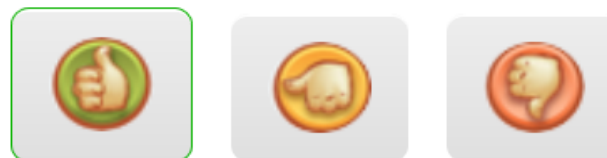
Note-taking
(comments or clarifying
questions about the step
function of individual
sentences)

Emphasis

Intra-personal interaction
(to foster cognitive
processing)

Critical information on fundamental aspects of floral development are needed to develop an understanding of what factors are causing lack of sexual reproduction in native habitats.

You are likely presenting justification.

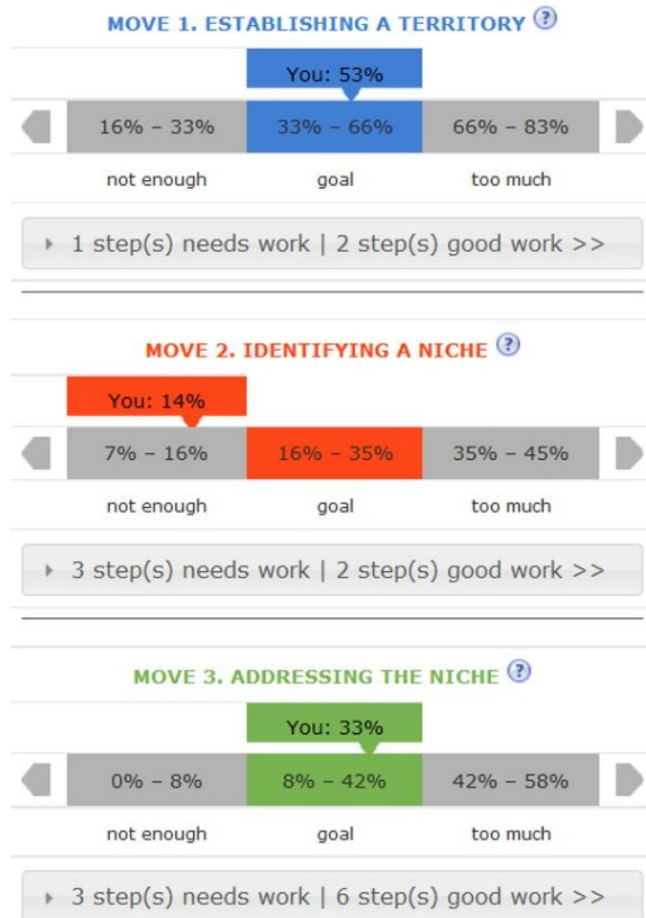


Your revision notes:

Yes, but maybe I should say that new research is needed because it's the research that would provide the critical information needed.

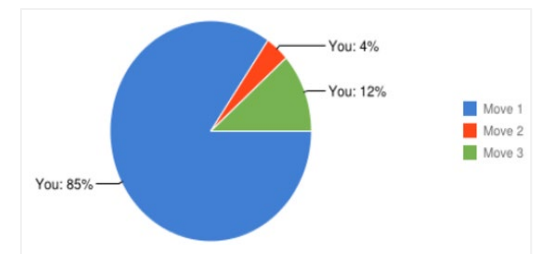
B. RWT: Corpus for feature design

Feedback: 'Analyze My Writing' module



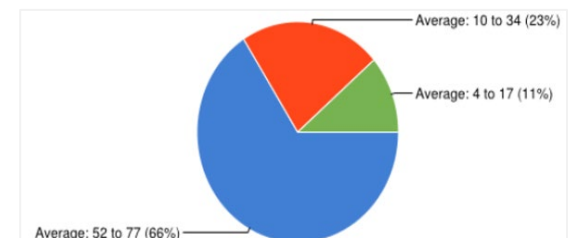
Feature

Numerical, comparative feedback
(range bars and pie charts summarize the distribution of moves in the draft compared with disciplinary corpus)



Emphasis

Socio-disciplinary context
(to compare the draft with the target discipline)



B. RWT: Corpus for feature design

■ Feedback: 'Analyze My Writing' module

Feature

Goal-orienting feedback
(range bars for moves expand to show which step/s are addressed well and which may be lacking or needing more work)

Emphasis

Planning and revision
(to set goals based on textual and socio-disciplinary patterns)

▼ 2 step(s) needs work | 1 step(s) good work >>

😊 Good work on providing general background. Very similar to Horticulture papers.

[Learn More](#) | [Examples](#)

⚠ Not enough focus on reviewing previous research compared to Horticulture papers. Needs more work.

[Learn More](#) | [Examples](#)

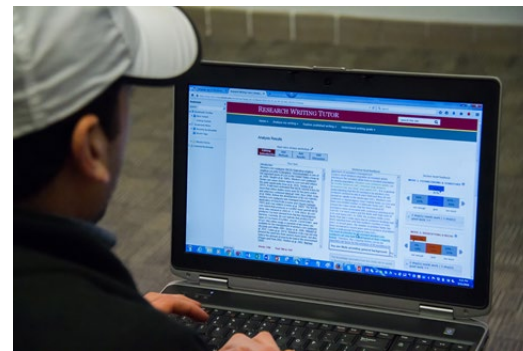
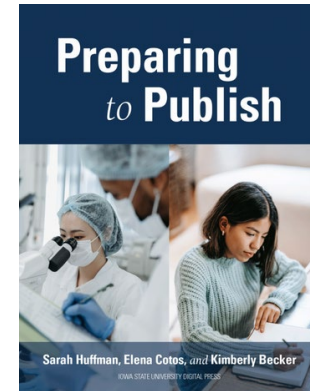
⚠ Lacking claiming centrality, which is typically used in Horticulture papers. Needs more work.

[Learn More](#) | [Examples](#)

C. RWT: Corpus for pedagogical implementation

Course: Preparing Publishable Thesis Chapters

- Knowledge and comprehension
 - Instructional videos (RWT)
 - Readings - genre writing conventions (RWT) (Huffman, Cotos, & Becker, 2021; open access <https://iastate.pressbooks.pub/preparingtopublish>)
 - Knowledge consolidation activity
- Corpus exploration
 - Rhetorical composition (RWT)
 - Language use (RWT)
- Analysis and application
 - Drafting and revision (RWT)
 - Peer review
 - Final draft submission



- Corpus exploration: Top-town analysis of rhetorical composition

- How is [Move] distributed in the Introduction?
- Do all the three steps appear in [Move]? If yes, in what order? Does the order matter?
- Are there any steps in [Move] that are not used or are very rare? Why do you think they are not typical of the research articles from your discipline?
- Is there any communicative overlap? That is, are there sentences that represent more than one step in a given move or more than one move?

Discipline: Section:

1 2 3 Next ->

Trefan, L., Bünger, L., Bloom-Hansen, J., Rooke, J., Salmi, B., Larzul, C., Terlouw, C., Doeschl-Wilson, A., "Meta-analysis of the effects of dietary vitamin E supplementation on α -tocopherol concentration and lipid oxidation in pork", *Meat science* 87(4):305-314, 2011

1. Introduction

Oxidation of lipids is a major cause of deterioration in the quality of muscle foods. Once oxidation is initiated, it continues as free radicals catalyse additional free-radical generating reactions (Abuja & Albertini, 2001; Jadhav, Nimbalkar, Kulkarni, & Madhavi, 1996). These oxidative processes lead to the deterioration of flavour, odour and colour of meat (Faustman & Cassens, 1990; Liu, Lanari, & Schaefer, 1995). Many studies have therefore aimed to delay the initiation of oxidation and subsequent loss of quality. In contrast to most synthetic antioxidants, dietary vitamin E can be directly fed to the animals and absorbed by the tissues (Buckley & Conolly, 1980). The role of vitamin E in particular, as a major lipid-soluble antioxidant, has been investigated for several decades and shown to stabilize animal products as far back as 1946 (Tsai, Wellington, & Pond, 1978). The stabilizing effect of vitamin E on flavor, color, texture and nutritive value has since been confirmed for chicken meat, beef and pork (Buckley & Conolly, 1980; Gray, Gomaa, & Buckley, 1996). While there is a substantial amount of literature that investigates the effects of dietary vitamin E on pork quality traits and on subcellular membranes (mitochondria, microsomes) (Asghar et al., 1991; Buckley et al., 1989; Monahan et al., 1994) under specific experimental conditions, there have been few attempts (Apple, 2007) to unify diverse experimental findings in a statistically robust way. Review articles (Buckley, Morisson, & Gray, 1995; Dikeman, 2007; Dunshea, 2000; Monahan, 1994; Nisbet, 1994; Nisbet, 1995; Nisbet, 1996; Nisbet, 1997; Nisbet, 1998; Nisbet, 1999; Nisbet, 2000; Nisbet, 2001; Nisbet, 2002; Nisbet, 2003; Nisbet, 2004; Nisbet, 2005; Nisbet, 2006; Nisbet, 2007; Nisbet, 2008; Nisbet, 2009; Nisbet, 2010; Nisbet, 2011; Nisbet, 2012; Nisbet, 2013; Nisbet, 2014; Nisbet, 2015; Nisbet, 2016; Nisbet, 2017; Nisbet, 2018; Nisbet, 2019; Nisbet, 2020; Nisbet, 2021; Nisbet, 2022; Nisbet, 2023; Nisbet, 2024; Nisbet, 2025) have compiled information on the effects of dietary vitamin E on pork quality traits, but do not offer quantitative relationships. Meta-analysis has been established as a useful

Move 2. Identifying a niche: Indicating a gap



RWT: Corpus exploration

■ Corpus exploration : Bottom-up analysis of language use

- Identify the linguistic choices indicative of certain step functions
- Compile findings into a list of move/step examples of functional language use

* Note grammatical features (e.g., present vs past, active vs passive, modals, etc.)

Claiming centrality

These vegetables are widely accepted by consumers because they are easy to prepare for eating. In addition, there is an increasing demand by consumers for safe and nutritious foods that improve physical performance, reduce risks of diseases, and increase the life span. Enhancing the nutritional levels of vegetables would improve nutrient intake without requiring an increase in consumption.

However, in some cases the terms QTL and gene will be used interchangeably for ease in explanation. Because of the wide use of PI 88788 as a resistance source, it has been the focus of a number of studies. These mapping studies have revealed the presence of only two resistance QTL that can be traced to this source (Concibido et al., 1997; Glover et al., 2004).

Since a major role of transpiration is leaf cooling, canopy temperature (T_c) and its reduction relative to ambient temperature is an indication of the role of transpiration in cooling the leaves. Thus, interest is increasing in using canopy temperature when breeding for drought tolerance. This involves selection of genotypes that maintain lower canopy temperature as compared with other genotypes under the same field conditions.

The effects were particularly large in the arils. Considering the increasing demand for intensely red pomegranates, further understanding of climate effects on the fruit coloration is needed, especially in face of "global warming" that already affects the climate in traditional cultivation regions of pomegranates. Knowledge of the factors and processes involved in anthocyanin

RWT: Effects of corpus & genre-based AWE

■ Writing task: Self-analysis and revision

- When you receive your color-coded text in the "Sentence-level feedback" text box, click on each sentence and check if you agree with the feedback prompt. Analyze all your sentences and take clarification/revision notes as follows:
 - If you agree, click on "thumbs up";
 - If you don't agree, click on "thumbs down" and write the step/strategy you think the sentence represents;
 - If you partially agree (e.g., the move is colored as you intended, but the step meaning you are trying to convey is different; or, your sentence carries two or more step functions, but the feedback prompt contains only one), click on "neutral thumbs" and take a note of what you agree and what you disagree with in "Your revision notes" box below.
 - If the feedback prompt is "The function of this sentence is not clear", clarify it with a revision note as well.
- Make revisions based on your notes in the textbox called "Your text".
 - You can make revisions concurrently with your analysis of individual color-coded sentences. Make the changes you want and then re-submit your modified draft for automated analysis and feedback by clicking on ANALYZE.
 - After you identify the parts of your draft that might need to be improved, export your notes as PDF and begin revising in the textbox called "Your text".
- Revise and resubmit your revised draft for sentence-level feedback as many times as you need.
- If your discipline is represented in RWT, consider the section-level feedback in the form of bar-graphs as well.
 - See whether or not each of your moves is within the "goal" range compared to your discipline.
 - Under each move, click to expand the feedback on steps and see which steps may be lacking or may need more work.
 - Determine which step to address first, second, next, etc. If you need a reminder of what a step means, hover over or click on "Learn More" to see its definition, or click on "Examples" to see sentences from the RWT corpus representing that particular step.
- Revise and resubmit your revised draft for section-level feedback as many times as you need.
- When you finish revising with RWT, export your draft as PDF and save it as "Your name_RWT Introduction revised draft".

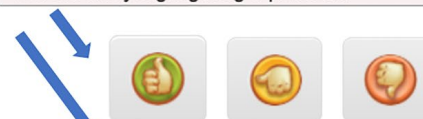
Tredici, 1987). Georgia plume grows on sand ridges, dry oak ridges, evergreen hammocks, and sandstone outcrops (Patrick et al., 1995). The apparently wide variety of habitats, but limited number of locations in which georgia plume occurs, has made it difficult to draw any definitive conclusions as to causes for rarity of the species. Georgia plume is reported to be found in less than three dozen

You are likely highlighting a problem.

Positioning System/Geographic Information System-based conservation management tool for georgia plume indicates that many recorded populations no longer exist with human

The apparently wide variety of habitats, but limited number of locations in which georgia plume occurs, has made it difficult to draw any definitive conclusions as to causes for rarity of the species.

You are likely highlighting a problem.



Your revision notes:

Your text

Elliottia racemosa, commonly known as georgia plume, is a threatened, woody plant that occurs only in the Coastal Plain region of Georgia in the southeastern United States. It was first discovered by the explorer William Bartram in 1773, documented by botanist Stephen Elliott in 1807, and so rare that the species was thought to be extinct until its rediscovery in 1901. Georgia plume is a striking deciduous tree reaching heights of 9 m and can have a single or multitrunked form. Plume-like racemes of attractive, white flowers appear in early summer above lustrous, green foliage. Despite all of its positive horticultural attributes, georgia plume has a limited range not only in the wild, but also in cultivation. It is difficult to propagate and characteristically exhibits extremely rare fruit production, scarcity of ripe seed, and difficulty with transplanting (Del Tredici, 1987).

Georgia plume grows on sand ridges, dry oak ridges, evergreen hammocks, and sandstone outcrops (Patrick et al., 1995). The apparently wide variety of habitats, but limited number of locations in which georgia plume occurs, has made it difficult to draw any definitive conclusions as to causes for rarity of the species. Georgia plume is reported to be found in less than three dozen populations (Chafin, 2007). However, our recent work conducting population surveys to develop a Global Positioning System/Geographic Information System-based

Words: 713 Goal: 511 to 895

Change Draft:

Draft10 2019/03/29 13:41

ANALYZE

Export Options

Select Format: ☒ PDF ☐ Email
 Include: ☒ Analysis and feedback ☒ Your notes

EXPORT

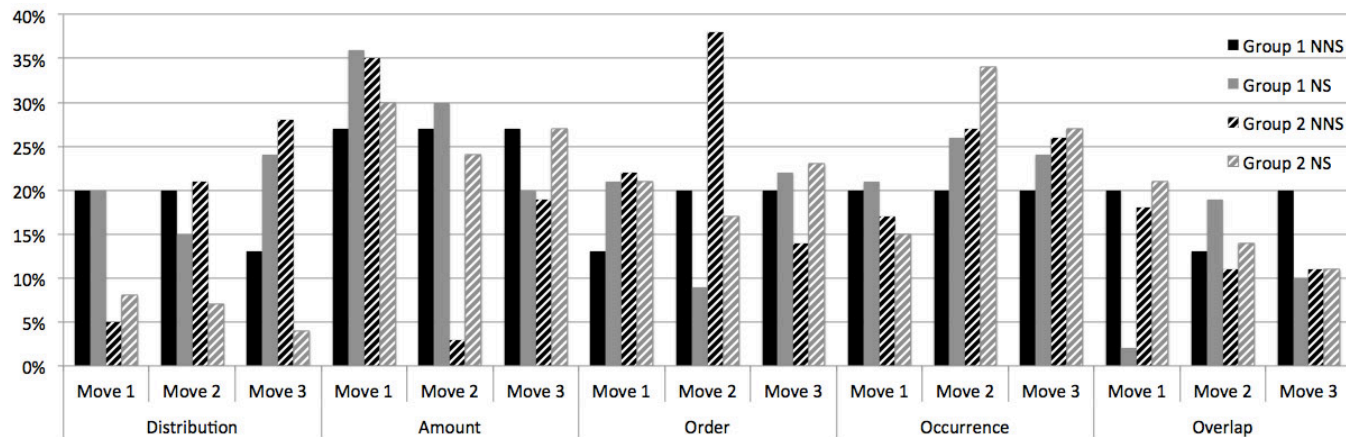
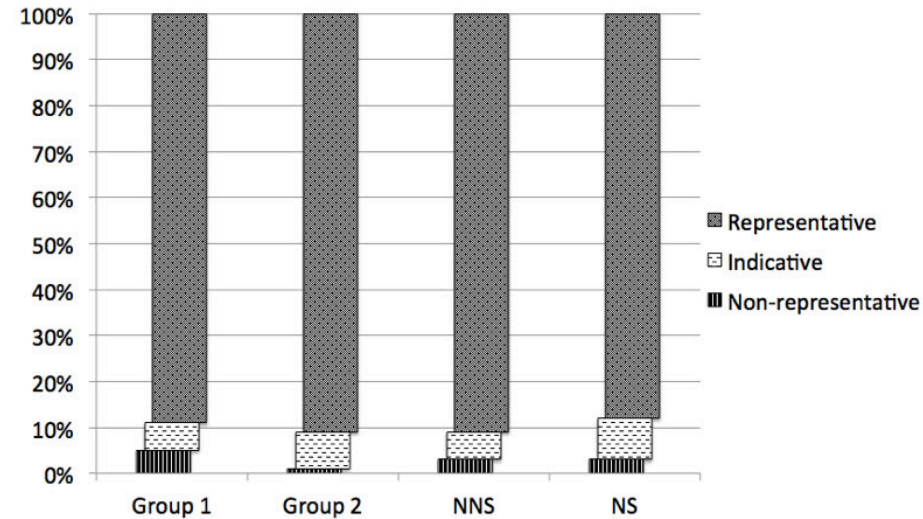
Outline

- Background to genre-based Automated Writing Evaluation
- Research Writing Tutor: Genre-based AWE exemplar for scientific writing
 - A. Corpus as discourse data
 - B. Corpus for feature design
 - C. Corpus for pedagogical implementation
- Learning potential and impact
- Future directions



RWT: Effects of corpus exploration

- Corpus-based features can
 - raise genre awareness
 - increase understanding of disciplinary conventions
 - foster noticing of
 - discipline-specific patterns in rhetorical composition
 - functional language expressing specific rhetorical intent

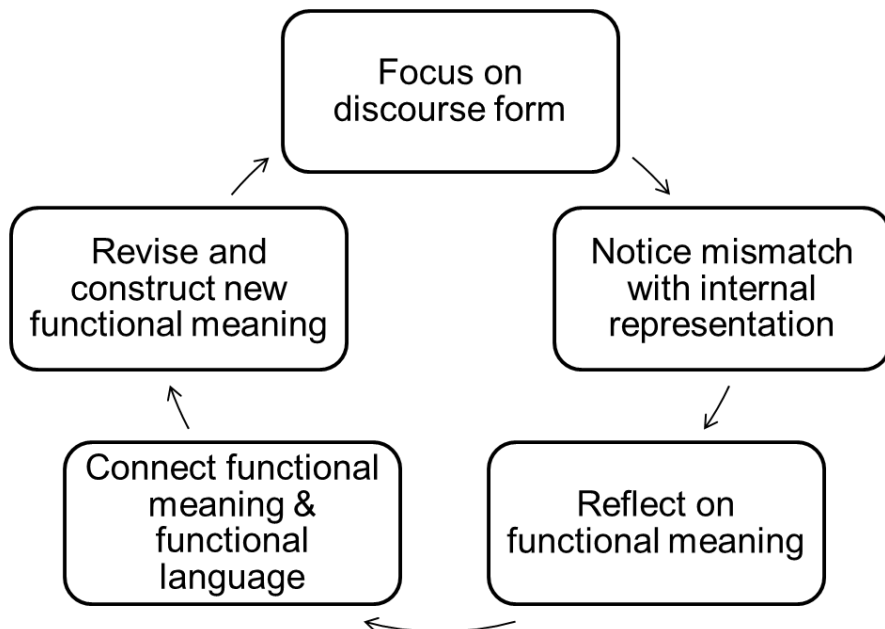


- Distribution [move placement]
- Amount [step extent, quantity]
- Order [step succession]
- Occurrence [step presence, absence]
- Overlap [multiple step functions in the same stretch of text]

(Cotos, Link, & Huffman, 2017)

RWT: Effects of corpus & genre-based AWE

- Corpus-based automated feedback can
 - foster intra-personal interaction
 - enable learning-revision cycle
 - develop a meaning-oriented internal representation



(Cotos, 2011, 2014)

Noticing meaning mismatch

So this part I want it to be shortly describe my research, but uh it displays in m1.

Reflecting on functional meaning

See here, this part. [highlights a blue part of the text] in my mind I wanted to write about method. So I think I describe the method, but it's m1. I don't know why. So, see... [highlights a blue sentence] this sentence may be purpose and reason, but in this model ... [highlights a blue sentence] this sentence I think I want to describe the model, how it works. So this, uh, some parameters and ... the method. Actually I don't know which move it will be. I just want to first ... in my mind, I think first I describe the field, uh, the technology in this field. Second, I do some reviews, and later I um... from the review I find some um... maybe not perfect in the previous research, so I can do something more. And in this part... so I describe what I want to do in this paper, and briefly, briefly introduce my method.

Connecting function and lexical choice

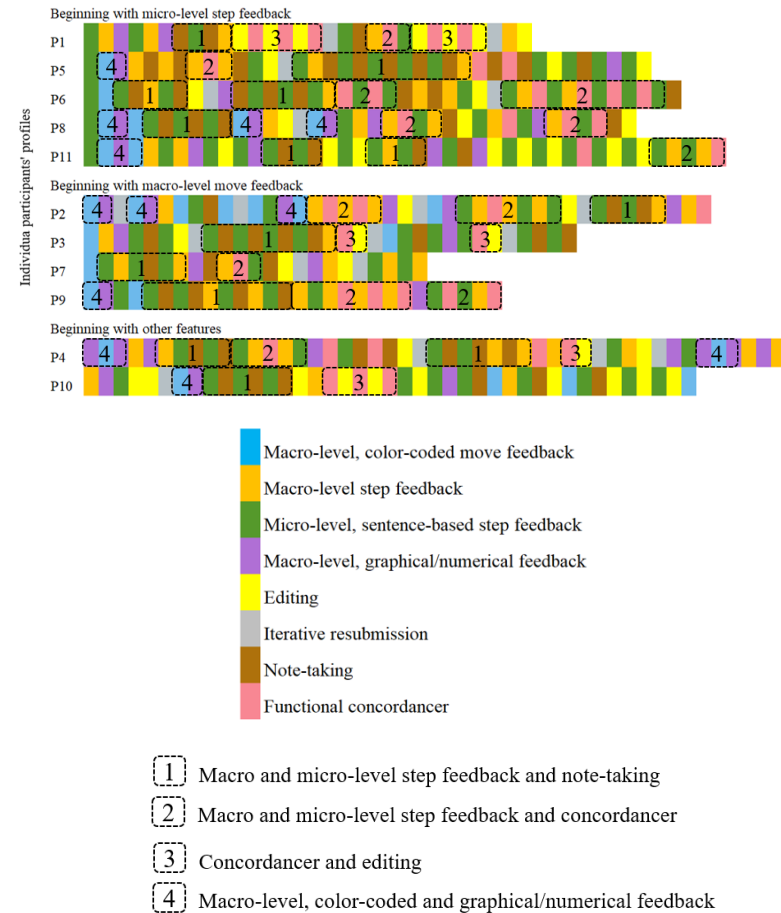
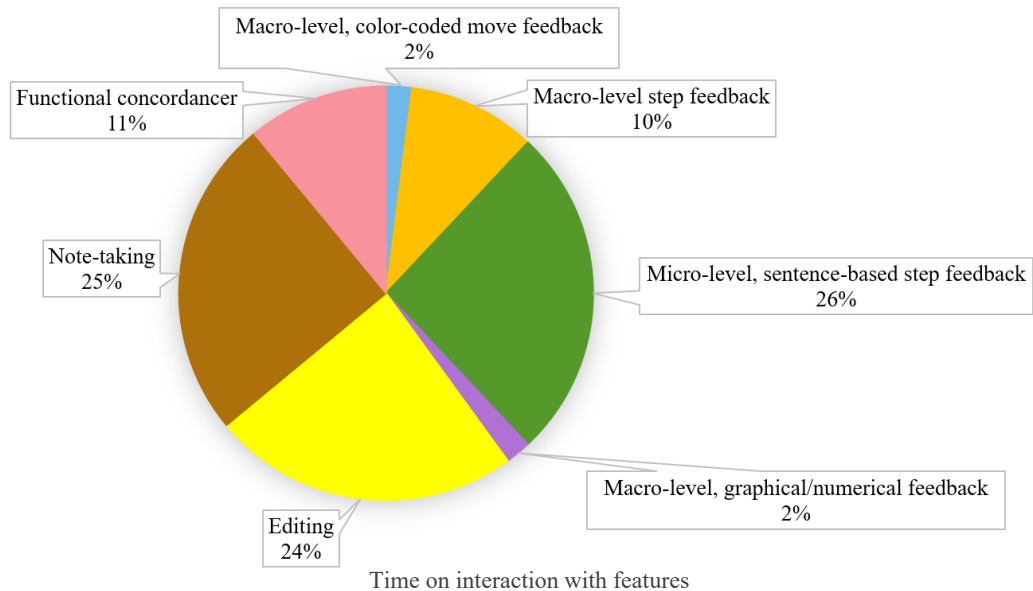
Let me see some examples [opens AC in the HO. looks at green sentences and their steps. highlights some.] Hmm. Let me find some more examples to describe the model. [opens the AC, scrolls down to the green text in an annotated introduction, checks the steps by clicking, highlights parts of the green text. opens another annotated text and does the same] maybe I can write it as description. [speaks aloud and emphasizes] "our work", "of our proposed system"... Let's try this first.

Constructing new functional meaning

[reads his text, then makes changes: "In this model, tungsten cones with different sizes are used to enhance the optical field." --> In our model, several tungsten cones with different sizes were chosen to enhance the optical field." [submits] (Student 40, think-aloud/Camtasia transcript)

RWT: Effects of corpus & genre-based AWE

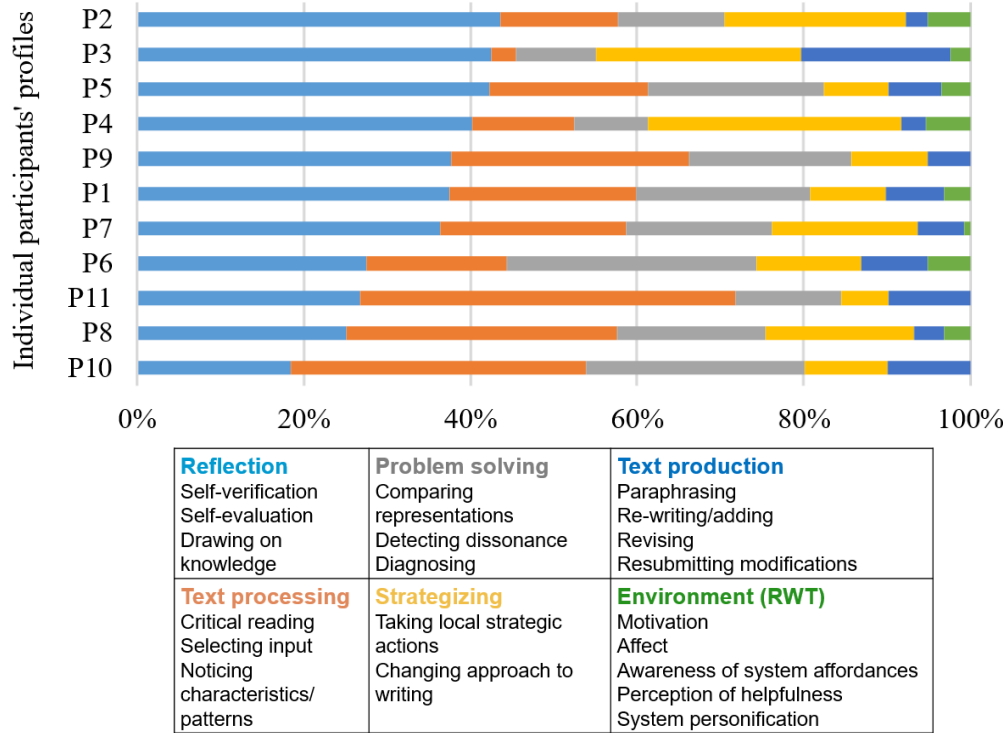
- Corpus-based automated feedback and features can
 - enable students to tailor their own interaction strategies
 - allow the necessary degree of learner control



(Cotos et al, 2020)

RWT: Effects of corpus & genre-based AWE

- Corpus-based automated feedback can
 - foster metacognition
 - enhance revision processes



"I agreed with the feedback of the system because I could also see that, yes, my sentence has this function but not as a primary goal."

Process: Reflection (self-verification)

"I didn't have this in mind."

Process: Problem solving (detecting dissonance)

"And maybe those, if I figure out those, maybe that's the blind point that I have in my writing."

Process: Problem solving (diagnosing)

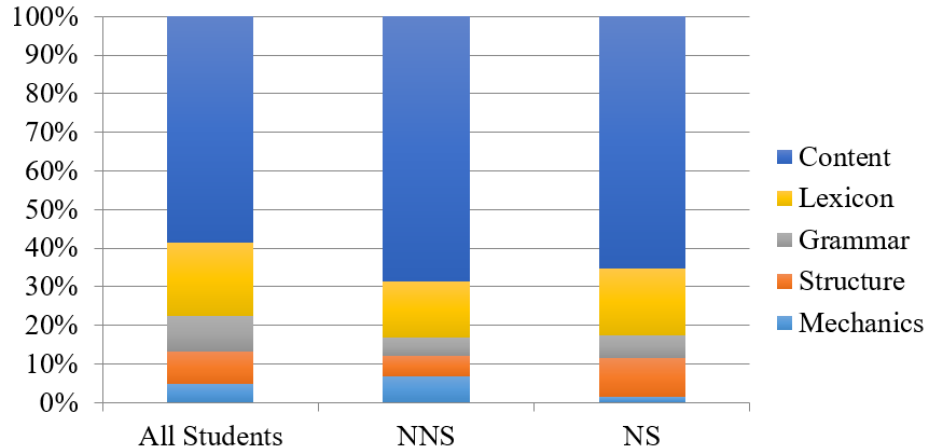
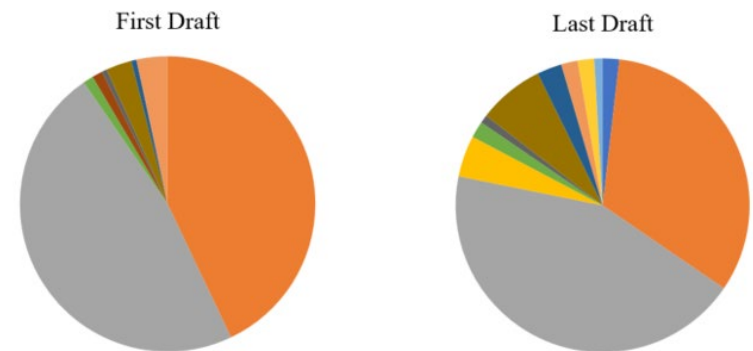
"I focused on the ones actually that needed more work."

Strategizing: Reflection (sequencing actions)

(Cotos et al., 2020)

RWT: Effects of corpus & genre-based AWE

- Corpus-based automated feedback can
 - help transfer knowledge from corpus exploration to drafting and revising
 - contribute to improving genre writing quality

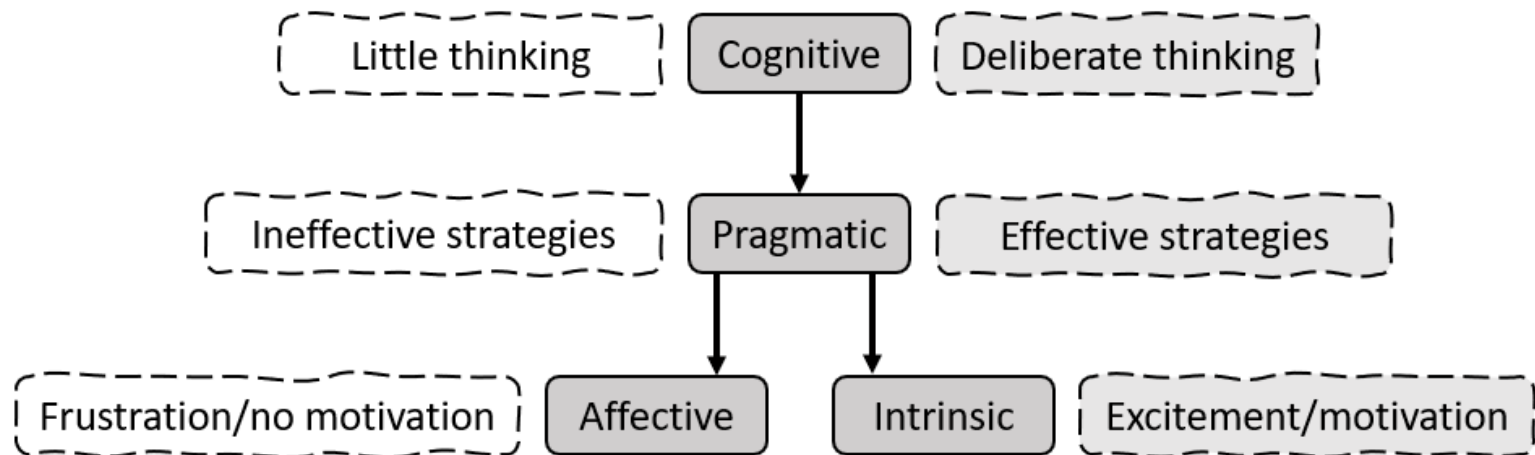
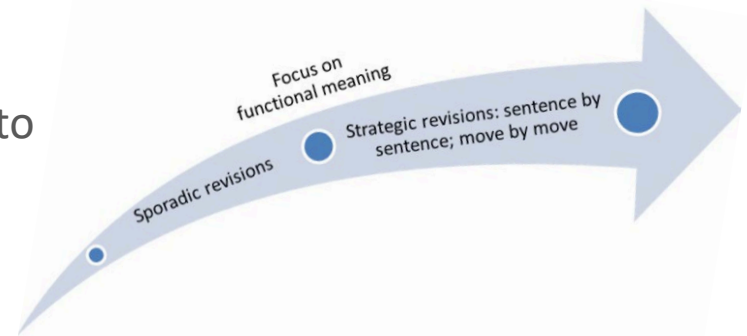


■ Claiming centrality
■ Providing general background
■ Reviewing previous research
■ Indicating a gap
■ Highlighting a problem
■ Raising general questions
■ Proposing general hypotheses
■ Presenting justification
■ Introducing the present research descriptively
■ Introducing the present research purposefully
■ Presenting research questions
■ Presenting research hypotheses
■ Clarifying definitions
■ Summarizing methods
■ Announcing principal outcomes
■ Stating the value of present research
■ Outlining the structure of the paper

(Cotos et al., 2017, 2020)

RWT: Effects of corpus & genre-based AWE

- Corpus-based automated feedback can
 - exert impact at multiple levels
 - focus on functional meaning is key to positive impact



(Cotos, 2012, 2014; Cotos & Huffman 2013; Cotos et al, 2017, 2020)

Outline

- Background to genre-based Automated Writing Evaluation
- Research Writing Tutor: Genre-based AWE exemplar for scientific writing
 - A. Corpus as discourse data
 - B. Corpus for feature design
 - C. Corpus for pedagogical implementation
- Learning potential and impact
- Future directions

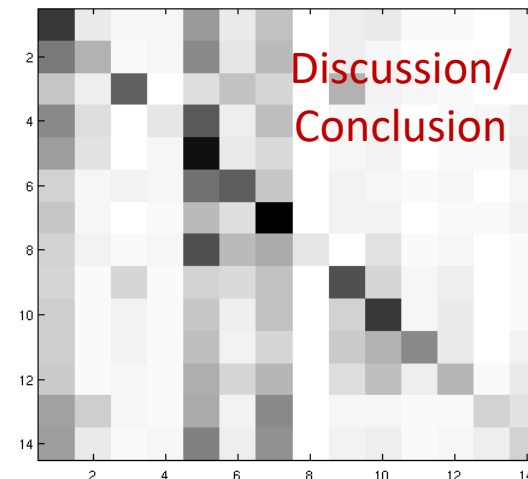
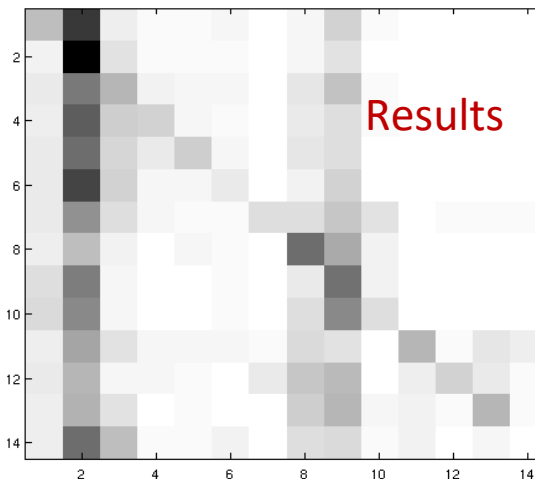
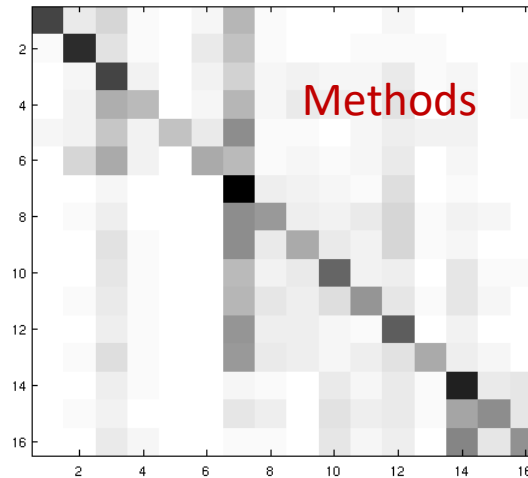
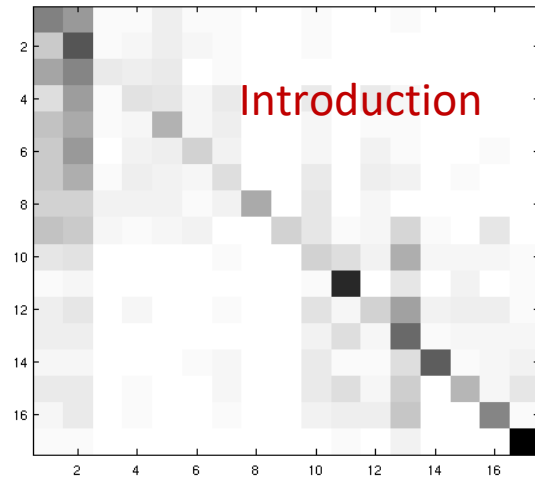


Future directions: Development

- New feedback on move sequencing



Future directions: Development



- New feedback on step sequencing

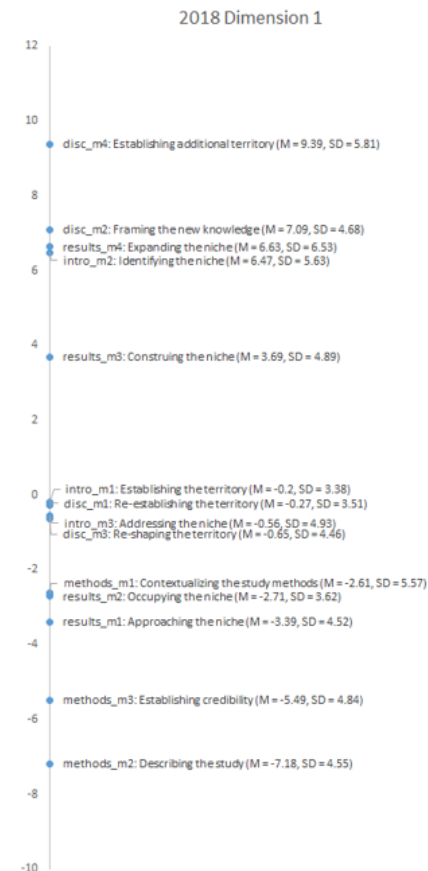
Matrices represent the probability of transitioning from one step to another: black --> high probability; white --> low probability

(Cotos et al., 2015)

Future directions: Development

- Future feedback on linguistic features
 - (Multi-Dimensional Analysis of moves)

Dimension 1	Interpretation and Expansion vs. Simple Reportage
Dimension 2	Abstraction / Overt Empiricism
Dimension 3	Procedural Narration
Dimension 4	Interpreting Results vs. Informational Density



Interpretation & Expansion

Positive Features

Verbs: verb be (.59), present tense (.33)

Modals: possibility (.64), prediction (.41) and necessity (.35) modals

Adverbs: linking adverbials (.48), adverbs (.47)

Stance: adjective-based stance (lexical and grammatical) (.49)

Adjectives: predicative adjectives (.47), attributive adjectives (.36)

Other: pronoun it (.45)

(Gray, Cotos, & Smith, 2020)

Future directions: Evaluation research

- Optimize teachers' understanding of automated feedback
- Improve feedback generation

Student:
Why?

RWT: ***"You are likely
highlighting a problem"***
*[but I'm not entirely
sure]*



Teacher: **How**
*do I explain
why? I can't
just say that
it makes
errors!*

Me: *What is going
on in the "black
box"?*

(Cotos, in press)

Future directions: Evaluation research

- To what extent can GBAWE feedback accurately reflect students' genre writing competence?
- How helpful is GBAWE for revision practice and genre writing improvement?
- How do students use GBAWE, and what strategies are most effective?
- What kind of and how much training do students and teachers need to use GBAWE effectively?
- How can teachers assess the effectiveness of GBAWE implementations in their classrooms?
- What are the strengths and limitations of GBAWE compared to other digital writing technologies, and how can different strands of research on feedback, usefulness, and impact address limitations and further inform the advancement of GBAWE?
- How can we most appropriately operationalize genre constructs to design 'actionable' GBAWE feedback; i.e., feedback that provides the guidance needed to improve?
- How can the functionality and output of GBAWE engines be evaluated and interpreted in meaningful ways for teachers and students?
- How can different theories, research results, and practical needs be best integrated in the design of new GBAWE for a range of target contexts?
- What principles should be developed to scale GBAWE from individual genres to genre systems spanning different contexts and discourse communities?

(Cotos, 2022)

References

- Anthony, L., & Lashkia, G. (2003). Mover: A machine learning tool to assist in the reading and writing of technical papers. *IEEE Transactions on Professional Communication*, 46 (3), 185-193
- Cotos, E. (2009). Designing an intelligent discourse evaluation tool: Theoretical, empirical, and technological considerations. In C. A. Chapelle, H.- S. Jun, & I. Katz (Eds.), *Developing and evaluating language learning materials* (pp. 103–127). Ames, IA.
- Cotos, E. (2011). Potential of automated writing evaluation feedback. *CALICO Journal*, 28(2), 420-459.
- Cotos, E. (2012). Towards effective integration and positive impact of automated writing evaluation in L2 Writing. In G. Kessler, A. Oskoz & I. Elola (Eds.), *Technology across writing contexts and tasks, CALICO Monograph Series* (Vol. 10, pp. 81–112). CALICO: San Marcos, TX.
- Cotos, E. (2014). *Genre-based automated writing evaluation for L2 research writing: From design to evaluation and enhancement*. Basingstoke, UK: Palgrave Macmillan.
- Cotos, E. (2015). Automated writing analysis for writing pedagogy: From healthy tension to tangible prospects. *Writing & Pedagogy*, 7, 197–231.
- Cotos, E. (2022). Genre-based automated writing evaluation. In H. Mohebbi & C. Coombe (Eds.), *Research Questions in Language Education: A Reference Guide for Teachers* (pp. 645–650). Springer Nature Switzerland AG.
- Cotos, E. (in press). Towards a validity argument for genre-based AWE. In J. Xu & G. Yu (Eds.), *Language test validation in a digital age*. Cambridge Assessment English and Cambridge University Press.
- Cotos, E., & Huffman, S. (2013). Learner fit in scaling up automated writing evaluation. *International Journal of Computer-Assisted Language Learning and Teaching*, 3(3), 77–98.
- Cotos, E., Huffman, S., & Link, S. (2015). Furthering and applying move/step constructs: Technology-driven marshalling of Swalesian genre theory for EAP pedagogy. *Journal of English for Academic Purposes*, 19, 52–72.
- Cotos, E., Huffman, S., & Link, S. (2017). A move/step model for Methods sections: Demonstrating rigour and credibility. *English for Specific Purposes*, 46, 90–106.
- Cotos, E., Huffman, S., & Link, S. (2020). Understanding graduate writers' interaction with and impact of RWT during revision. *Journal of Writing Research*, 12(1), 187-232.
- Cotos, E., Link, S., & Huffman, S. (2017). Effects of DDL technology on genre learning. *Language Learning and Technology*, 21(3), 104–130.
- Cotos, E., Link, S., & Huffman, S. (2016). Studying disciplinary corpora to teach the craft of Discussion. *Writing & Pedagogy*, 8, 33–64.
- Cotos, E., & Pendar, N. (2016). Discourse classification into rhetorical functions for AWE feedback. *CALICO Journal*, 33(1), 92-116.

References

- Gray, B., Cotos, E., & Smith, J. (2020). Combining rhetorical move analysis with multi-dimensional analysis: Research writing across disciplines. In U. Römer, V. Cortes, & E. Friginal (Eds.), *Advances in corpus-based research on academic writing: Effects of discipline, register, and writer expertise* (pp. 138–168). John Benjamins .
- Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing*, 16, 148-164.
- Knight, S., Shibani, A., Abel, S., Gibson, A., Ryan, P., Sutton, N., ..., Buckingham Shum, S. (2020). AcaWriter: A learning analytics tool for formative feedback on academic writing. *Journal of Writing Research*, 12(1), 141-186.
- Lee, D., & Swales, J. M. (2006). A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes*, 25(1), 56-75.
- Perelman, L. (2012). Construct validity, length, score, and time in holistically graded writing assessments: The case against automated essay scoring (AES). In C. Bazerman, C. Dean, J. Early, K. Lunsford, S. Null, P. Rogers, & A. Stansell (Eds.), *International advances in writing research: Cultures, places, measures* (pp. 121-131). Fort Collins, Colorado: WAC Clearinghouse/Anderson.
- Sándor, Á., Kaplan, A., & Rondeau, G. (2006). Discourse and citation analysis with concept-matching. In *Proceedings of the International Symposium: Discourse and Document*, 15–16 June 2006, Caen, France (pp. 147-152). Retrieved from https://www.researchgate.net/publication/240828357_Discourse_and_citation_analysis_with_concept-matching
- Shermis, M.D. & Burstein, J.C. (2003). *Automated essay scoring: A cross-disciplinary perspective*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Williamson, D. M. (2013). Probable cause: Developing warrants for automated scoring. In M. D. Shermis & J. C. Burstein (Eds.), *Handbook of automated essay evaluation: Current applications and new directions* (pp. 153–180). Routledge, New York.

Thank you!

(ecotos@iastate.edu)