Complex Dynamic Systems and Interlanguage Variability: Investigating Topic, Syntactic Complexity, and Accuracy in NS-NNS Written Interaction

Anna Ciriani Dean[1]
Teachers College, Columbia University

ABSTRACT

Studies adopting a Complex Dynamic Systems Theory framework have often fallen into one of two camps: those investigating variable linguistic subsystems (namely, CAF studies) and those concerned with non-linguistic variables displaying the characteristics of dynamic systems, with little concern for their connection to linguistic development. Examining asynchronous interactional data in an 8-week email exchange between a bilingual speaker of Spanish and English and an advanced Chinese learner of English, this study attempts to reconcile these two camps by exploring the extent to which syntactic complexity and article accuracy vary situationally as a function of topic and willingness to communicate (WTC) in L2 written performance over time and by investigating any potential relationships between complexity and accuracy. A qualitative, inductive analysis reveals that topic bears the characteristics of a dynamic system. The results also show that, for both the native speaker (NS) and non-native speaker (NNS), syntactic complexity, measured as subordinate clauses per T-unit and complex nominals per T-unit, varies situationally with topic, shifting not just over time but within a single email. As might be anticipated, article accuracy exhibits the typical instability of a dynamic system for the NNS but not for the NS, whose stable performance may reflect a permanent attractor state. Both the NNS' and NS' article accuracy are subject to cross-linguistic influence. Evolving correlations between complexity and accuracy suggest that certain topics and the WTC associated with them may induce trade-offs between the two subsystems.

INTRODUCTION

Over the past few decades, *Complex Dynamic Systems Theory* (CDST) has asserted itself into the field of *Second Language Acquisition* (SLA) as a comprehensive view of *second language* (L2) development. Generally, researchers have fallen in one of two camps: those examining non-linearity in linguistic *subsystems* only (Baba & Nitta, 2014; Bulté & Housen, 2014; De Groot, 2012; Larsen-Freeman, 2006; Polat & Kim, 2014; Rosmawati, 2014; Spoelman & Verspoor, 2010; Verspoor, Schmid, & Xu, 2012; Yang & Sun, 2015), and those primarily exploring non-linguistic factors with dynamic-system characteristics, with little to no interest in how these systems may interact with specific linguistic subsystems (Cherciov, 2013; Churchill, 2007;

[1] Anna Ciriani Dean is a graduate of the M.A. Applied Linguistics programs at Teachers College, Columbia University. Her current research interests include second/foreign language acquisition, writing-to-learn, task-based language teaching, and writing center pedagogy. Correspondence should be sent to: anna.cirianidean@gmail.com.

Kang, 2004; King, 2013; MacIntyre & Legatto, 2010; Opitz, 2014; Reigel, 2008). While promising in their own right, these studies fail to connect *interlanguage* (IL) variability to external and internal variables that may impact them.

Examining interactional data in a cross-cultural email exchange, the present study attempts to fill this gap by exploring the following research questions:

1. To what extent do syntactic complexity and article accuracy vary situationally as a function of topic in L2 written performance over time?
2. Are there any trade-offs between complexity and accuracy under different topical circumstances?

Thus, this study inspects both linguistic (syntactic complexity and article accuracy) and non-linguistic (topic) variables as their own dynamic systems as well as connected components within each individual's IL development. After a survey of prior CDST literature, the study's methods and results will be discussed, followed by a conclusion and discussion of limitations and implications for future research.

LITERATURE REVIEW

In an attempt to capture the activity of fluctuating, interconnected subsystems, L2 researchers have for the most part analyzed dynamic variability within linguistic performance. Specifically, measures of *complexity, accuracy, and fluency* (CAF) have often been operationalized as separate subsystems within linguistic development (Baba & Nitta, 2014; Bulté & Housen, 2014; De Groot, 2012; Larsen-Freeman, 2006; Polat & Kim, 2014; Rosmawati, 2014; Spoelman & Verspoor, 2010; Verspoor, Schmid, & Xu, 2012; Yang & Sun, 2015). While not entirely conclusive, these studies have generally led researchers to conclude that development within CAF is non-linear, as CDST would predict, characterized by periods of high variability, or *repeller states*, as well as periods of relative stability, or *attractor states* (Spoelman & Verspoor, 2010; Zheng, 2016; De Groot, 2012). Furthermore, several studies report strong correlations between subsystems as evidence of *connected growers*, such as word complexity and noun phrase complexity (Spoelman & Verspoor, 2010), as well as negative correlations marking competitive relationships, such as between individual words and of lexical bundles (Zheng, 2016).

While the CAF framework does provide a useful tool for examining development in linguistic subsystems, Polat and Kim (2014) aptly point out that "because [C]DST highlights the importance of various factors in language learning (e.g., both cognitive and social factors) future studies are warranted that account for cultural and social aspects of language development" (p. 204). In other words, CAF studies examining language subsystems alone, with no consideration of other variables, overlook one of the tenets of CDST: the impact of internal and external resources on language development.

Beyond the study of CAF, SLA researchers have also investigated non-linguistic factors displaying the characteristics of dynamic systems. In particular, Kang (2004) and MacIntyre and Legatto (2010) apply CDST to the construct of *willingness to communicate* (WTC) in their studies of NS-NNS conversations and student performance on oral communication tasks respectively. WTC, defined as readiness or intention to speak in a communicational context, is considered its own dynamic system influenced by cognitive and affective variables, such as anxiety and extraversion (MacIntyre & Legatto, 2010), and security, excitement, and

responsibility (Kang, 2004). What distinguishes these studies from others is the idiodynamic approach they adopt: rather than trace variation across unitary data-collection points, they also observe variation within a single interaction. WTC is seen not as a stable predisposition, but rather as a situational construct influenced by several shifting variables throughout a communication event, influenced by situational variables such as topic, interlocutors, and conversational context (Kang, 2004). Through assisted recall, the researchers in both studies elicited verbal report data from participants who commented on their WTC and socio-cognitive state moment by moment.

Despite the richness of this idiodynamic approach, these studies are limited in that they do not correlate WTC with L2 performance, and while other studies do attempt to link non-linguistic variables to L2 development, they only track general markers of proficiency level rather than L2 subsystems (Cherciov, 2013; Churchill, 2007; King, 2013; Opitz, 2014; Reigel, 2008). A comprehensive CDST view would require an analysis of the interconnected variability of both socio-cognitive factors and L2 subsystems.

CAF studies within a *task-based language teaching* (TBLT) framework somewhat reconcile these two camps. Framed around two opposing claims—Robinson's (1995, 2001, 2003) Cognition Hypothesis and Foster and Skehan's (1996; Skehan, 1998; Skehan & Foster, 1997, 1999) Limited Attentional Capacity Model—these studies have examined correlations between tasks' cognitive demands and L2 performance, testing whether higher task complexity leads to trade-offs between accuracy and complexity (Skehan and Foster's argument) or stretches learners' abilities (Robinson's argument). Due to their basis in information-processing (IP) theory, TBLT methods cannot be applied wholesale to CDST. Still, testing the existence of trade-offs seems akin to observing competitive or connected growth, and task complexity features—particularly topic—could inform researchers' understanding of cognitive factors affecting dynamic systems. By investigating complexity and accuracy through the lens of topic, the present study attempts to reconcile socio-cognitive factors with L2 performance, framing IL variability idiodynamically, as a product of changes within a single interaction.

METHODOLOGY

Data and Participants

The data consisted of an email exchange between a *non-native speaker* (NNS) of English and bilingual *native speaker* (NS) of English and Spanish. The interaction took place asynchronously in English over the course of eight weeks. A total of 24 emails were exchanged.

Methods of Analysis

Topic Analysis

The data were coded qualitatively for topic using an inductive approach. Through iterative, bottom-up analysis, 21 recurring topics were identified within and across emails and were then grouped into three categories: surface topics, connection topics, and content topics. See Table 1 for a detailed list and explanation of these topics. Once categories were determined,

the number of words for each topic occurrence was calculated. Topical patterns were then analyzed in terms of changing word count over time.

TABLE 1
List of topics discussed in the email exchange

SURF	**surface topics**	**act as general conduits through which interaction occurs; contain no in-depth content**
INT	introduction	the interlocutors introduce themselves to one another
REL	relational	remarks establishing relationships, such as compliments
GEN	general email conventions	phrases and longer sentences signaling the beginning and end of emails, as well as common email etiquette
PIC	pictures	recurring theme of sharing pictures with one another
LIF	life	general events happening in the interlocutors' lives
CONN	**connection topics**	**help interlocutors become acquainted, or "connected," with one another**
HOB	hobbies	discussion of hobbies and pastimes
EDU	education	discussion of educational goals and of the American education system
SPN	Spanish	discussion of Michelle's bilingual English-Spanish background
MND	Mandarin	discussion of Mandarin and Michelle's interest in learning it
MOV	movies	discussion of common movie interests and opinions
JAP	Japanese disaster	discussion of the 2011 Japanese nuclear disaster
MUS	music	discussion of common music interests
CONT	**content topics**	**explored in depth by interlocutors and occupy central roles in the conversation**
TXS	Texas	discussion of life in Texas—complementary with CHN
CHN	China	discussion of life in Texas—complementary with TXS
FOR	foreigners	discussion of foreigners and reactions to them in the interlocutors' area
WTH	weather	discussion of the weather in the interlocutors' respective regions of the world
FES	festivals & holidays	discussion of festivals and holidays typical in the interlocutors' hometown
COL	colors	discussion of the significance of colors in the interlocutors' cultures
TRP	trip	discussion of Belinda's recent trip to Mount Lao
GAM	games	discussion of games typical of the interlocutors' cultures

Additionally, discourse-analytic literature on topic was consulted to derive codes for measures of WTC. In her survey of research on the effect of conversational topic on NNS interactional behavior, Zuengler (1993) distinguishes between two features of topic: (1) the speaker's cognitive and affective relationship to the topic, based on Selinker and Douglas' (1985) Discourse Domain Model, which argues that ILs develop not as unitary wholes, but rather through discourse domains to which each learner has a unique socio-cognitive relationship; and (2) the speaker's conversational role within the topic, based on Giles' Speech Accommodation Theory, which claims that speakers adjust their speech depending on their relationship with their interlocutor (Street & Giles, 1982). Research within these two frameworks has included the effects of topic expertise (Cornu & Delahaye, 1987; Selinker & Douglas, 1985; Woken & Swales, 1989; Zuengler, 1989; Zuengler & Bent, 1991), topic investment (Eisentein & Starbuck, 1989), topic sensitivity (Dowd, 1984; Zuengler, 1982), or whether the interaction involves discussions of cultural differences, resulting in an inter-group encounter as opposed to an inter-

individual encounter (Zuengler, 1982). From the TBLT literature, the distinction can also be made between topics involving personal experiences and those more detached from the interlocutor, which contributes to topic familiarity (Foster & Skehan, 1996).

The deductive WTC codes derived from the literature are summarized in Table 2. The researcher then applied these codes to each topic occurrence in the dataset, inferring the participants' cognitive and affective relationship to the topic (CART) and conversational role within the topic (CRWT) based on the content of each topic segment and on indicators of tone in the participants' writing. For instance, in topic segments relating to the participants' home cultures, the codes for expert (EXP) and inter-group representative (GRP) may be applied within the category of CRWT. As for CART, the presence of emoji, exclamatives, and other aspects of tone may indicate the participants' investment in the topic (INV), and the nature of the topic itself may suggest that a topic is personal (PRS) or sensitive (SNS) for the participants.

TABLE 2
Topic features affecting WTC

CART	**Cognitive & Affective Relationship to Topic**	**the writer...**
INV	invested	demonstrates interest in the topic and therefore seems invested or engaged in it
PRS	personal	reveals personal preferences and attributes in discussing the topic
SNS	sensitive	may be sensitive to the topic, which contains upsetting or threatening issues
CRWT	**Conversational Role within Topic**	**the writer...**
EXP	expert	is the expert in the topic in relation to the other interlocutor, which confers a sense of responsibility
GRP	inter-group representative	acts as a representative of her culture in discussing the topic (as opposed to inter-individual)
WTC	**Willingness to Communicate**	sum of all CART and CRWT variables, with the exception of SNS, which instead is subtracted from the total score due to its negative impact

As sub-variables of WTC, these codes were factored into a cumulative WTC score, which was later correlated with L2 performance variables.

Complexity Analysis

In CAF studies, syntactic complexity has been understood in terms of variety and sophistication of morpho-syntactic structures (Wolfe-Quintero, Inagaki, & Kim, 1998). Its operationalization has not been straightforward, though, leading to inconsistent definitions across studies (Wolfe-Quintero et al., 1998). In developing his automatic syntactic-complexity analysis tool, Lu (2010) identified fourteen measures with high levels of validity and reliability, which were then split into five distinct sub-categories: those measuring length of production unit, sentence complexity, subordination, coordination, and particular structures. Of these measures, two were chosen for the present study: dependent clauses per T-unit (DC/T) and complex nominals per T-unit (CN/T).[2] Both measures were found to have high reliability between

[2] A T-unit consists of a main clause and all of its attached or embedded subordinate clauses or nonclausal structures (Lu, 2010). A dependent clause is defined as a finite adverbial, nominal, or adjective clause dependent on a main

automated scores and manually-calculated scores (r=.95 for DC/T and r=.90 for CN/T). DC/T and CN/T were selected specifically because they both rely on the same unit of measurement (T-units) and also because they each belong to a different subcategory (subordination and particular structures), which was intended to ensure that they constitute separate subsystems of complexity. Although the two overlap at times (i.e., noun phrases with clausal postmodifiers), DC/T generally captures complexity at the clausal level while CN/T relates to noun phrase structure. Thus, in the present study, DC/T and CN/T are considered separate yet connected subsystems.

Lu's Web-based L2 Syntactic Complexity Analyzer was employed to calculate complexity scores for both participants. The analyzer was applied to the data twice: first by email exchange, for a view of development over time, and then by topic, for a fine-grained analysis of situational variation within a single email. In addition to calculating statistical correlations between the two subsystems, DC/T and CN/T were plotted on line graphs to observe each subsystem's variability and potential interactions. Cumulative complexity scores by topic were also displayed to observe potential relationships with topic.

Accuracy Analysis

While error analysis can be employed to calculate global accuracy, a target-like use (TLU) analysis, proposed by Pica (1984), may be preferable in studying IL subsystems, as it focuses on specific target forms (Ellis & Barkhuizen, 2005). In this study, the chosen target was the English article system because articles are unavoidable in production, thus providing several contexts of use, and because mastery of the article system occurs late in IL development (Master, 1997). The indefinite article *a* and definite article *the* were considered subsystems, and the definite article was split into further subsystems for its use with plural and singular nouns.

In compliance with TLU analysis procedures (Ellis & Barkhuizen, 2005), obligatory occasions for *a* and *the* were first identified in the data. Articles in fixed constructions, such as quantifiers (e.g., *a lot of* and *a little bit*) and certain invariable proper nouns (e.g., film titles) were not counted as obligatory occasions. Then, the occasions were coded for correct suppliance or oversuppliance, and left bare in cases of no suppliance. These codes were tallied both by email and by topic, and the number of correct suppliances was divided by the sum of obligatory contexts and cases of oversuppliance for each portion of text. Correlations between the subsystems were calculated, and the subsystems were plotted on line graphs to observe variability and interactions. Additionally, mean article accuracy scores were displayed by topic and placed on a bar graph to observe potential relationships between these two variables.

Trade-offs between Accuracy and Complexity

The results for both accuracy and complexity were compared by topic using statistical correlations to understand the relationship between accuracy and complexity, as well as between topical WTC and linguistic performance variables.

clause (Lu, 2010). A complex nominal consists of nouns plus any premodifiers or postmodifiers, non-clausal or clausal (Lu, 2010).

RESULTS

Topic Analysis

A detailed breakdown of the topic sequence across emails can be seen in Appendix A. This sequence reveals that several topics were addressed in each email. Even emails at the beginning of the exchange contained several topics, just with lower word counts. Figures 1 and 2 show that content topics occupy the highest number of words (2982), followed by connection topics (1420), with surface topics occupying the smallest word count (1114).

FIGURE 1
Total word count by category

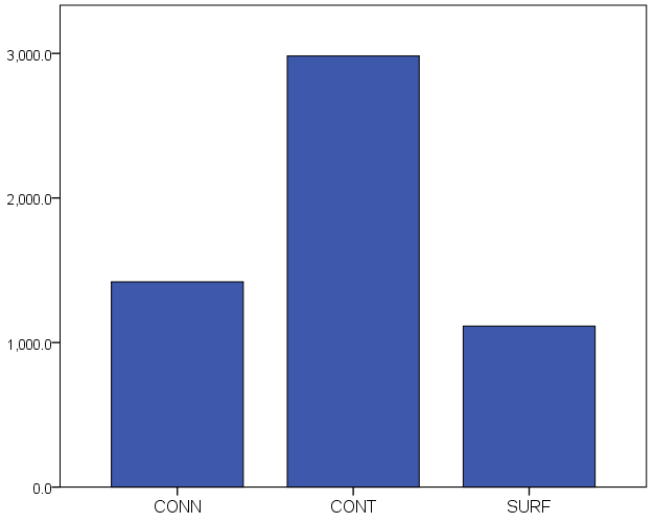
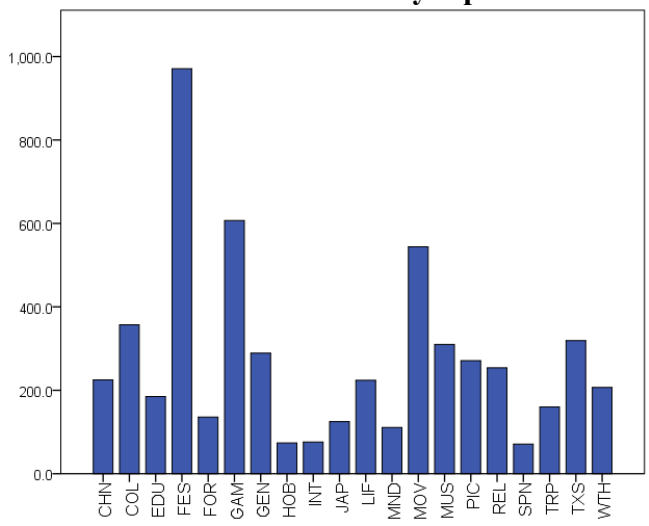


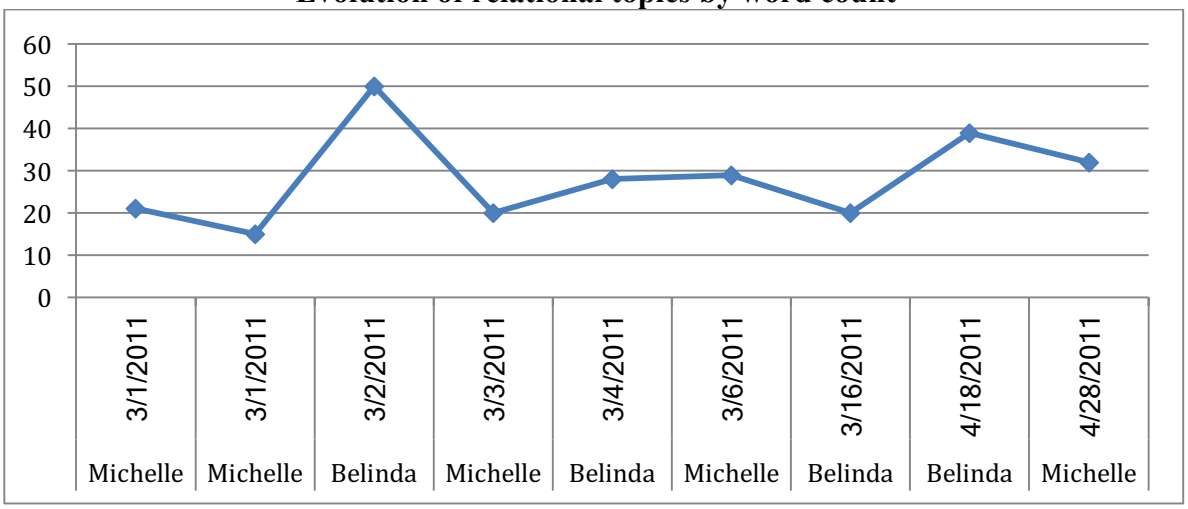
FIGURE 2
Total word count by topic



Over time, the conversation shifts from surface topics, which then branch into several connection topics. In the first two emails, for example, the discussion consists entirely of surface

topics. The participants give brief introductions of themselves, interspersed with comments intended to secure their relationship (relational topics), such as "I hope we can get along well and help each other out with this project ^_^" in Michelle's 3/1 email and "Though we may still need several e-mailed to make us acquainted with each other, I'm sure it will be nice to work with you this semester" in Belinda's 3/1 email. These early topics establish the participants' relationship, and the use of emoji, carried throughout the entire exchange, conveys a sense of familiarity and ease. As can be seen in Figure 3 and Appendix A, relational topics largely disappear from the discussion after the exchange on 3/6. At the same time, word count picks up, and the exchanges are largely taken up by connection or content topics. Relational topics may have become less frequent because the participants have become sufficiently acquainted with one another.

FIGURE 3
Evolution of relational topics by word count

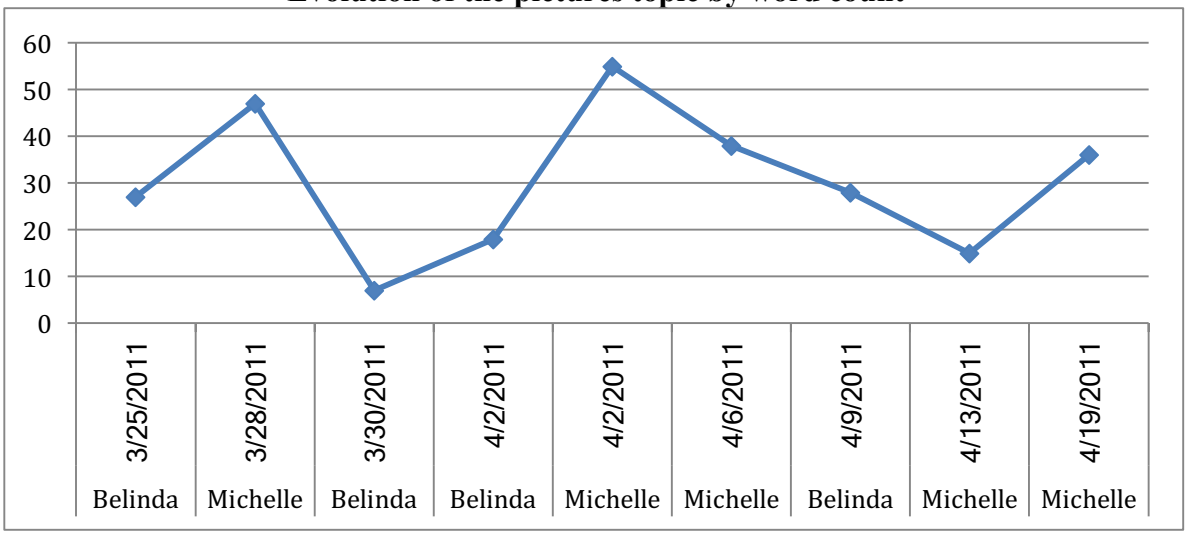


This reliance on surface topics is then mirrored in the participants' last few emails, starting with Michelle's on 4/6, when content topics and word count lose momentum and surface topics dominate the discussion once again. After Belinda's 4/9 email, the two engage only in surface topics. Relational topics return on 4/18: Belinda attempts to end the conversation, commenting "Here I want say that it's really nice to have worked with you and made a friend like you. Hope you will remember me when you teach in China.O(∩_∩)O~ Finally, good luck for you exams and happy everyday!" Michelle does not recognize Belinda's attempt to end the conversation and on 4/19 continues with an apology for not responding sooner and more surface topics. When Belinda does not respond, Michelle ends the conversation with her own relational comments: "I hope if you still want to that we can continue writing to each other. If not then I enjoyed getting to know you and I hope your future plans work out!" Relational topics seem to signal the beginning or end of the conversation.

Surface topics, however, are not always concentrated at the beginning or end of the exchange. Figure 4 shows that the pictures topic, in which Belinda and Michelle discuss sharing photographs of themselves and of their environment with each other, spans a large part of the exchange, recurring in seven emails over the course of a month. Relatively few words are dispensed at each occurrence, but the topic seems to act as a conduit for maintaining the conversation. It begins in Belinda's 4/25 email as a general comment ("ps. It seems I haven't given you my picture. For better understanding, I will upload my photo and some of my school's

pictures to MSN~~") and becomes a full-fledged topic, as the participants start sending each other photos, commenting on them (Michelle on 4/2: "Aw I like your picture, too cute! I guess you are a Harry Potter fan?..."), and promising to send more photos (Michelle on 4/13: "I promise to upload the pictures this weekend, and to have a more lengthy letter."). In spite of their basic content function, surface topics are embedded throughout the exchange and support ongoing conversation.

FIGURE 4
Evolution of the pictures topic by word count



The evolution of each topic is characterized by great variability. Connection topics, while generally more substantial than surface topics, start off with relatively low word counts but later occupy longer segments of text. For instance, Figure 5 shows that the education topic, in which the participants discuss their educational goals, returns in small increments over the span of only a few days. It begins in the introduction and continues with Michelle's mention of pursuing a Masters in ESL. Belinda's curiosity about the U.S. education system (on 3/4: "By the way, is it hard to get a master in your country? […] Could you tell me more about your education system, teacher michelle? o(∩_∩)o") fuels the conversation. More substantial connection topics, such as movies in Figure 6, occupy a higher number of words but are concentrated in fewer exchanges. The movie topic in particular occurs in large blocks and is abruptly introduced by Belinda on 3/9, with no preamble: "Recently, our foreign teacher let us watch a film called 'The Kids Are All Right' and write a reaction about it…." This is uncharacteristic of most topics, which tend to stem from brief mentions and grow into larger topics.

FIGURE 5
Evolution of the education topic by word count

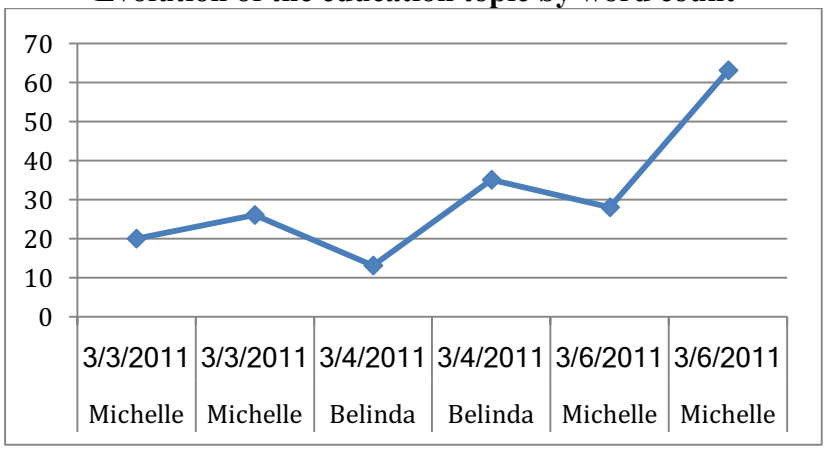
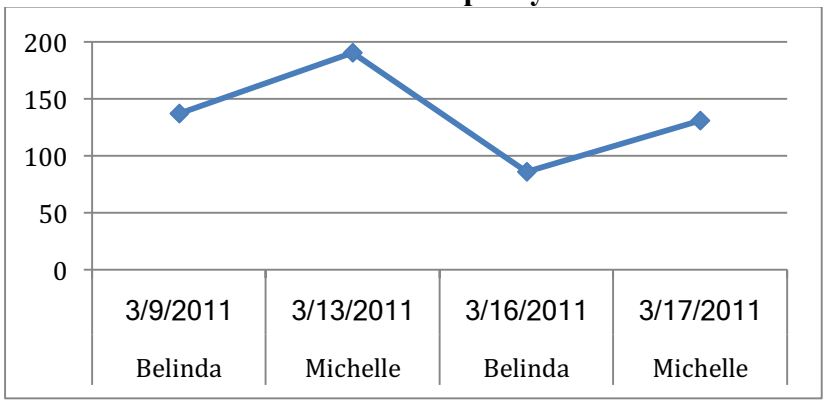
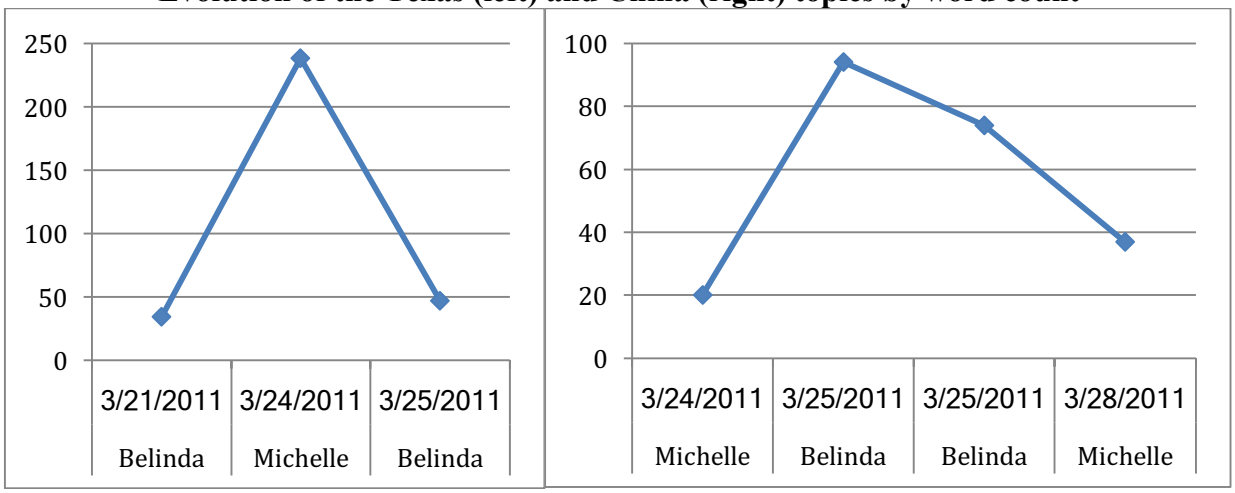


FIGURE 6
Evolution of the movie topic by word count



The connection topics culminate with Michelle and Belinda's discussion of movies and music, at which point they move on to content topics. These act similarly to connection topics but occupy higher word counts. Participation is not always evenly distributed for content topics: at times word count is balanced, but at others it seems one-sided. For obvious reasons, as seen in Figure 7, each participant writes more about her own home country than about the other participant's country. However, when referencing the other country, the participant only comments or asks questions. For example, when discussing the topic of Texas, Belinda is limited to remarks such as (on 3/25): "From your description, the Valley seems to be a comfortable place and to live in there is a pleasant thing. I am looking forward to being there."

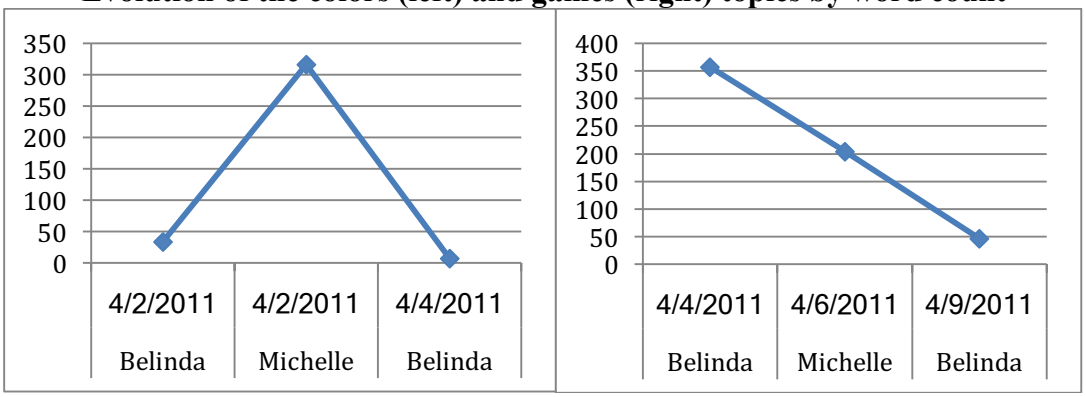
FIGURE 7
Evolution of the Texas (left) and China (right) topics by word count



However, as seen in Figure 8, the topic of colors is clearly one-sided: though Belinda initiates the topic ("Yes, I have questions for the use of color word, like use 'green' to show one is jealous. I always mix up this kind of words. Can you teach me how to use them?") and Michelle indulges her request, Belinda does not contribute to it any further ("Thanks, Michelle~~I will sum the colors up later."). This behavior may be due to Michelle's offer on 3/30: "Plus, let me know if there is any custom, behavior, word or phrase that you have any questions on and I'll try to the best of my abilities to answer you." In a way, Michelle positions herself as an authority, making the discussion unilateral.

Also from Figure 8, it can be gleaned that although Belinda writes a high number of words when discussing games, the conversation declines linearly, perhaps because the topic is of little interest to her or affords little material for further discussion. Belinda introduces the games topic as part of a story of her trip to Mount Lao on 4/4, providing a vast amount of detail regarding how to play "Who's the killer?." Michelle, on 4/6, responds by explaining a traditional Mexican game, but then Belinda merely comments on Michelle's game without going into detail about other childhood games, providing scant material for response.

FIGURE 8
Evolution of the colors (left) and games (right) topics by word count



As can be seen in Table 3, certain correlations between word count and WTC scores are somewhat correlated. Word count is moderately correlated with expertise, followed by investment and inter-group representation. The personal nature of topic is weakly correlated with word count, and topic sensitivity is slightly negatively correlated.

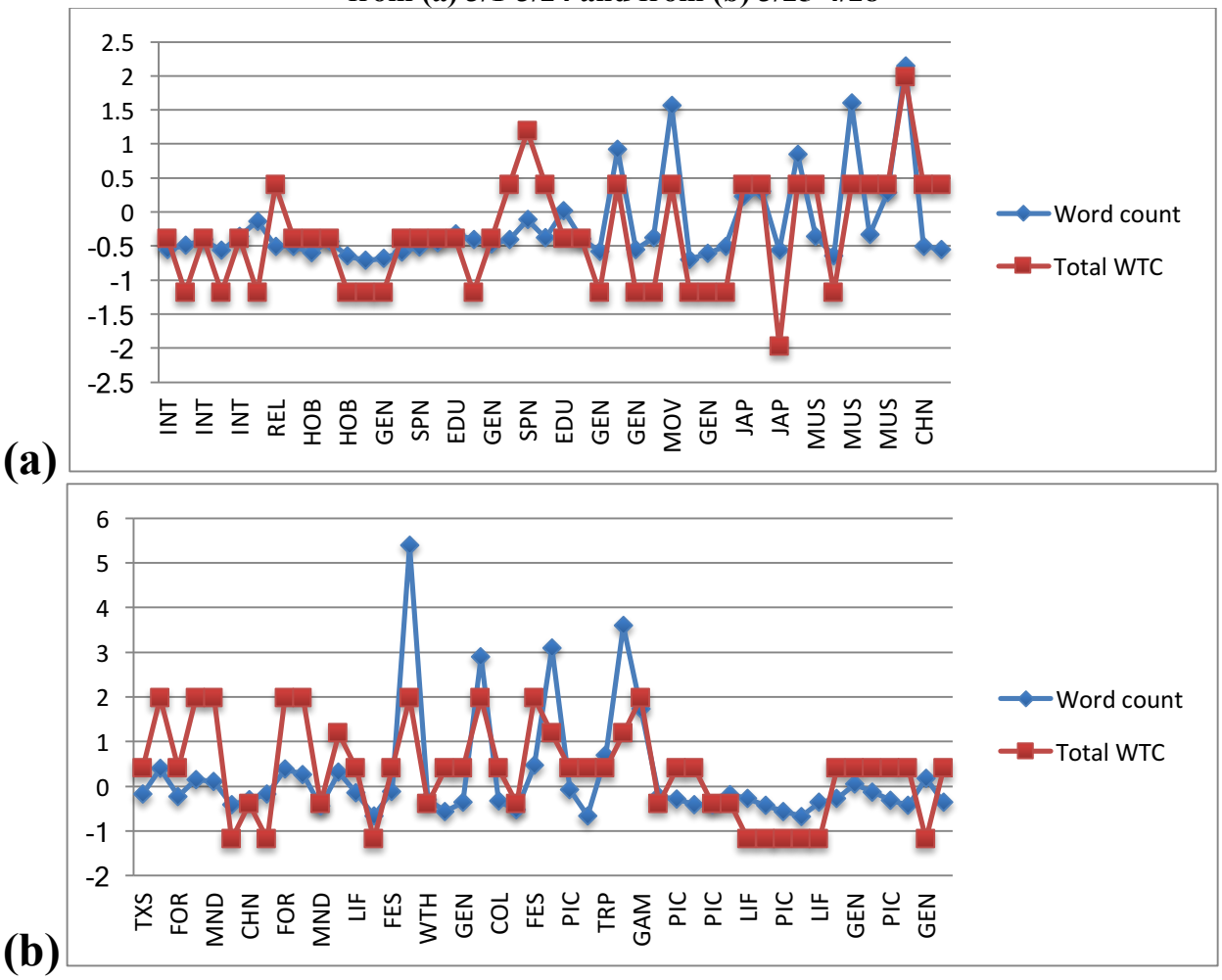
TABLE 3
Correlations between word count and features of WTC

	INV	PRS	SNS	EXP	GRP
Word count	.411**	.278**	-.045	.561**	.402**

**. Correlation is significant at the 0.01 level.
INV=invested, PRS=personal, SNS=sensitive, EXP=expert, GRP=inter-group representative

It may therefore be possible that the five WTC features operationalized in this study have some relationship with topic evolution. Indeed, Figure 9 shows that word count and WTC often coincide. See Appendix A for a detailed breakdown of WTC features associated with each topic.

FIGURE 9
Topic evolution by standardized word count and WTC scores from (a) 3/1-3/24 and from (b) 3/25-4/28



Complexity

Syntactic complexity data for Michelle reveals both variability and connected use. Figure 10 shows substantial amounts of variability for both subsystems from email to email, with no clear upward or downward trend. Although some divergence is evident and the correlation between the two is weak (r=.32), nominal complexity and amount of subordination do sometimes evolve in a parallel manner, suggesting some amount of connected use. This variability is amplified when visualizing the data by topic, as in Figure 11. However, CN/T and DC/T seem more correlated when split by topic (r=.62) than by email, providing slightly stronger evidence for connected use, possibly due to the larger number of data points. This finding is not surprising, as several types of nominal phrases are in fact clausal (i.e., clausal postmodifiers) and thus also count as dependent clauses.

FIGURE 10
Michelle's variability in syntactic complexity organized chronologically by email

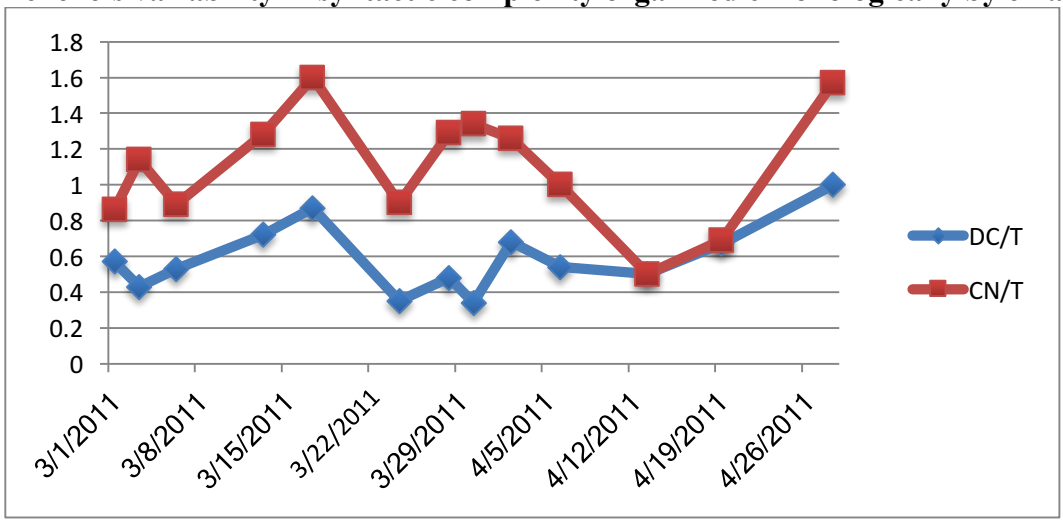


FIGURE 11
Michelle's variability in syntactic complexity organized chronologically by topic

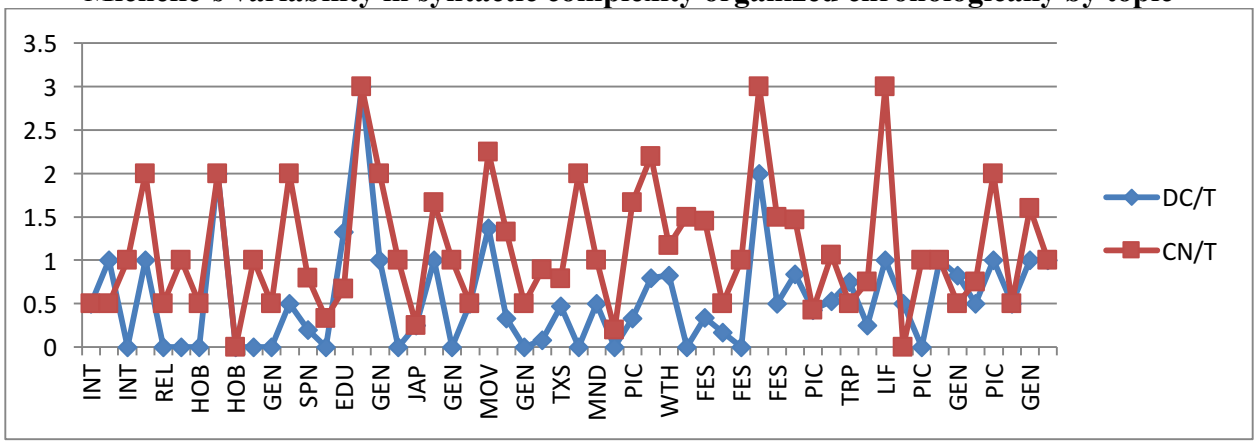


Figure 11 shows that complexity measures—particularly CN/T—peak for three surface topics: a relational topic on 3/6, a general topic on 3/30, and a life topic on 4/6. Each of these topics lasts only a short number of words, as follows:

> **Excerpt 1.1 (REL, 3/6)**: And thank you, I do hope to be a good teacher, I still have a long way to go, but that is what school is getting me ready for.
> (29 words, DC/T 3, CN/T 3)

> **Excerpt 1.2 (GEN, 3/30)**: Plus, let me know if there is any custom, behavior, word or phrase that you have any questions on and I'll try to the best of my abilities to answer you.
> (32 words, DC/T 2, CN/T 3)

> **Excerpt 1.3 (LIF, 4/6)**: We're taking my Mom to the beach on Sunday for her birthday so I'll upload some pictures of our beach and some sight-seeing stuff too!
> (27 words, DC/T 1, CN/T 3)

A look at the raw data indicates that the DC/T and CN/T scores produced by the Web-based L2 Syntactic Complexity Analyzer are not entirely precise. In excerpt 1.1, the Analyzer seems to consider each T-unit as a dependent clause, likely due to a misinterpretation of "thank you" as an independent clause and of the function of each comma, which Michelle often employs to separate independent clauses. Similarly, in excerpt 1.2, the Analyzer fails to recognize the presence of two T-units rather than one. Still, the score for excerpt 1.3 seem accurate. The reliability of the tool would likely be greater for higher word counts.

When aggregating the data by topic (Figures 12 and 13), higher syntactic complexity seems somewhat sporadic. Surface topics slightly supersede other categories, a finding which, while largely negligible, seems counter-intuitive, as surface topics generally consist of less complicated content.

FIGURE 12
Michelle's overall syntactic complexity organized collectively by topic

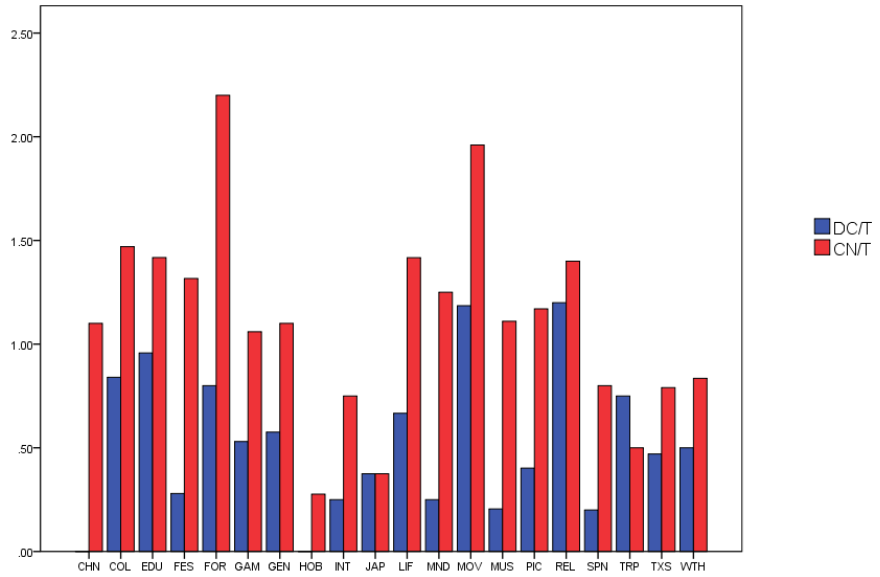
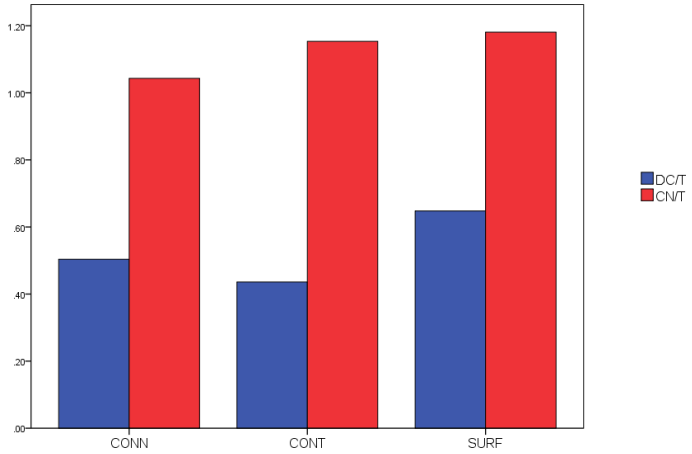


FIGURE 13
Michelle's overall syntactic complexity organized collectively by topic category



One topic that stands out is Michelle's discussion of foreigners on 3/28, with a particularly high amount of complex nominalization (CN/T 2.2), possibly amplified by the singular mention and low word count of the topic:

> **Excerpt 2 (FOR, 3/28):** I understand about reacting to foreigners like that too, lol. Since down here we are all either Hispanics or mixed it's very rare that we see people of different cultures. However, in the past few years, we have been having many people from different cultures start moving down here. Every winter season we have what we call, "Winter Texans," come down from all over the upper United States. Winter Texans are older retired couples who escape the cold / snow from their homes and come down here to the Valley, where it's summer all year round.

In this passage, she explains the concept of "Winter Texans," using devices such as phrasal post-modifiers (e.g. "people of different cultures"), pre-modifying adjectives (e.g. "older retired couples"), relative clauses (e.g. "couples who escape the cold"), and nominal clauses (e.g. "we have what we call…").

Michelle's score for both DC/T and CN/T is also high for the topic of movies, with scores between 1 and 2.25 across the two emails in which movies are discussed (3/13 and 3/17). On 3/13, for example, Michelle explains her opinion of films: "I believe that having something unique to a film is great, but just because a film is unique doesn't mean that it's actually a good film." This sentence consists of two T-units, the first with one dependent clause ("having something unique to a film") and the second with two dependent clauses ("because a film is unique" and "that it's actually a good film"). The complexity of the thought Michelle is trying to express in this sentence and in the overall topic may contribute to higher DC/T and CN/T scores.

Other connection topics, however, seem to elicit particularly low DC/T. Within the topic of music (3/17 and 3/24), for instance, Michelle makes simple comments on artists she and Belinda both like, not expressing any complex ideas, merely a list of likes and dislikes:

> **Excerpt 3 (MUS, 3/24):** Super Junior is good too! Yeah I really like HanGeng, he was one of favorite members. I really don't care for SM, they are not my favorite company but I do enjoy some of their artists. Avril Lavigne is fun, I enjoy her songs, especially her latest one.

This topic is straightforward and does not warrant particularly complex structures. Similarly, Michelle's DC/T is low when she describes festivals:

> **Excerpt 4 (FES, 3/30):** But I mainly go for the rides, I'm a ride junkie!! Then there's the food! Yum!! Ton of food from different shops. Like normal pizza and burgers to funnel cakes, roasted corn in a cup, fajita (beef skirt) tacos, brisket tacos, raspas (flavored crushed ice) and burritos (I don't know how to describe this but it's really good!). But it's really expensive and it smells too, lol, all the livestock in one area is not a great idea! hehe
> For holidays, I usually spend them with my family. Hispanics are known for their big family get-togethers. My family is no exception, my cousins, aunts and uncles all gather at my grandparents house, where we eat lunch and then spend the day together. It's fun and we get to catch up with each other. […]

Topic complexity seems related to Michelle's syntactic complexity: when explaining complex ideas, she resorts to both subordination and complex nominalization, and when narrating or listing events or other items, she employs little subordination and moderate complex nominalization.

Despite these few observations, it is difficult to pinpoint a specific topic feature that leads to higher or lower complexity. Correlations between complexity measures and topic features substantiate this absence of clear pattern. As can be seen in Table 4, DC/T and CN/T correlate only slightly—and insignificantly—with topic word count and WTC scores, making it difficult to generalize from the data. This may relate to Michelle's NS status, which would logically depend less on topic variations, as her IL is more stable than a NNS's would be.

TABLE 4
Correlations among Michelle's topic measures and complexity measures

	DC/T	CN/T
Word count	.08	.13
WTC total	.07	.22

While Belinda's syntactic complexity data also shows evidence of variability and connected growth between complex nominalization and subordination, the results seem slightly more linear than in Michelle's case. Both Figure 14 (by email) and Figure 15 (by topic) show variability for both subsystems. However, Belinda's complexity scores display a less erratic pattern and clearer evidence of connected growth between CN/T and DC/T. Indeed, the correlation between the two for Belinda based on emails is .88 and based on topic is .87, much stronger coefficients than for Michelle. This result may indicate that complex nominalization and subordination are more connected during IL development than at end states, such as with NSs.

FIGURE 14
Belinda's variability in syntactic complexity organized chronologically by email

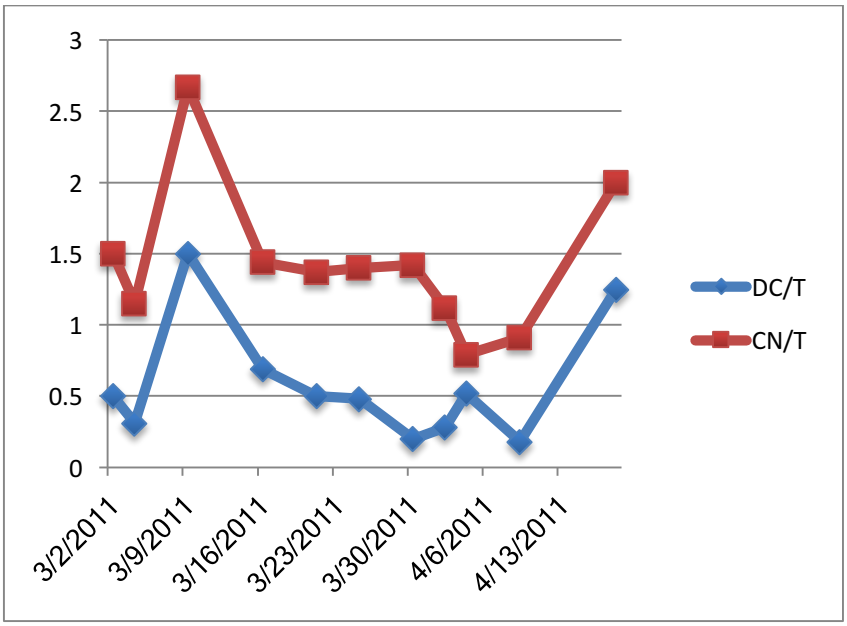
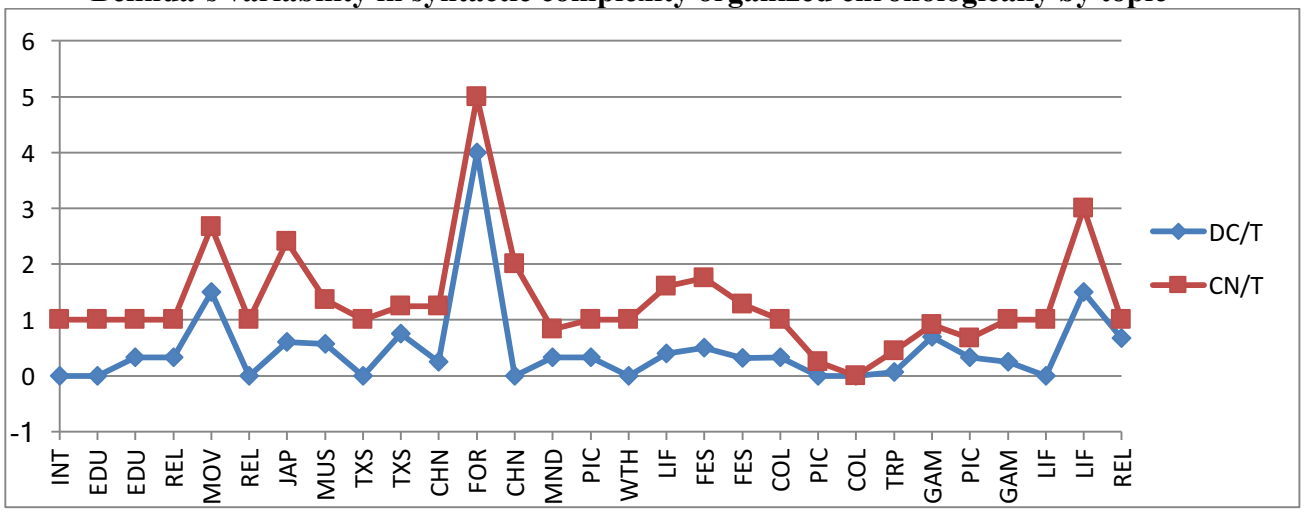


FIGURE 15
Belinda's variability in syntactic complexity organized chronologically by topic



The time lapse between Belinda's 3/16 and 3/30 emails seems particularly linear and may potentially indicates the presence of an attractor state. However, Belinda displays a wider range of complexity (CN/T between 0 and 5, DC/T between 0 and 4), while Michelle displays less variability (CN/T between 0 and 3, DC/T between 0 and 3). Belinda's CN/T and DC/T reach their full potential with the topic of foreigners, coinciding with Michelle's highest collective CN/T level. In this short segment, Belinda seems to experiment with constructions she does not normally use:

> **Excerpt 5 (FOR, 3/25):** It is still funny to remember how excited I was when I first saw a blonde here, for it was an alive one different from what I had seen in TV before. Hehe... A little bit stupid, right? Ok, back to the topic...

The first sentence alone contains four dependent clauses, one of which follows a low-frequency conjunction, *for*, meaning *because*. The use of this conjunction could suggest that Belinda is stretching the limits of her IL, and/or that she is unaware of the formality of *for*. The other topics in the same email present much more moderate complexity, which explains why this peak in complexity is not visible when the data is represented by email (Figure14).

Belinda's complexity also seems high while discussing the movies topic on 3/9. Like Michelle, Belinda expresses her opinion about movies in this segment: "I searched some information and found that though it won 0 awards in the Academy Awards, the film was highly praised by critics and audience. But according to the reactions of my classmates, the film is not as great as what the media has described…." Once again, the nature of the information Belinda communicates (an opinion rather than description) may lead to increased complexity.

As with Michelle, grouping CN/T and DC/T performance by topic collectively rather than chronologically (Figures 16 and 17) does not display particular trends. DC/T is again visibly lower than CN/T across all four categories. But unlike for Michelle, content categories display the highest overall complexity, followed then by connection topics and finally surface topics, a seemingly logical distribution given that content topics deal with more elaborate material. This may indicate that NNSs' syntactic complexity is more clearly affected by topic than NSs', although this cannot be confirmed from a single case.

FIGURE 16
Belinda's overall syntactic complexity organized collectively by topic

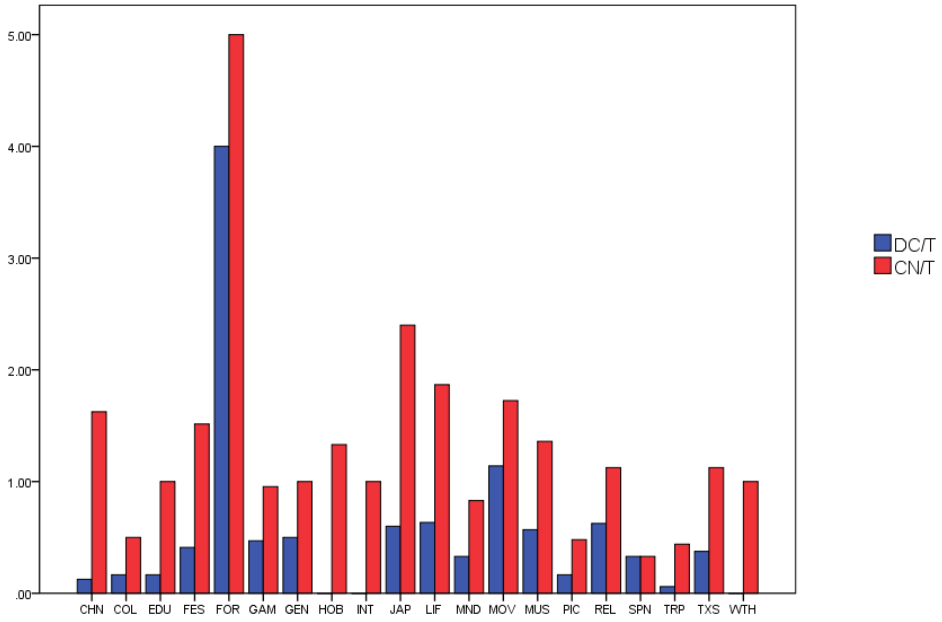
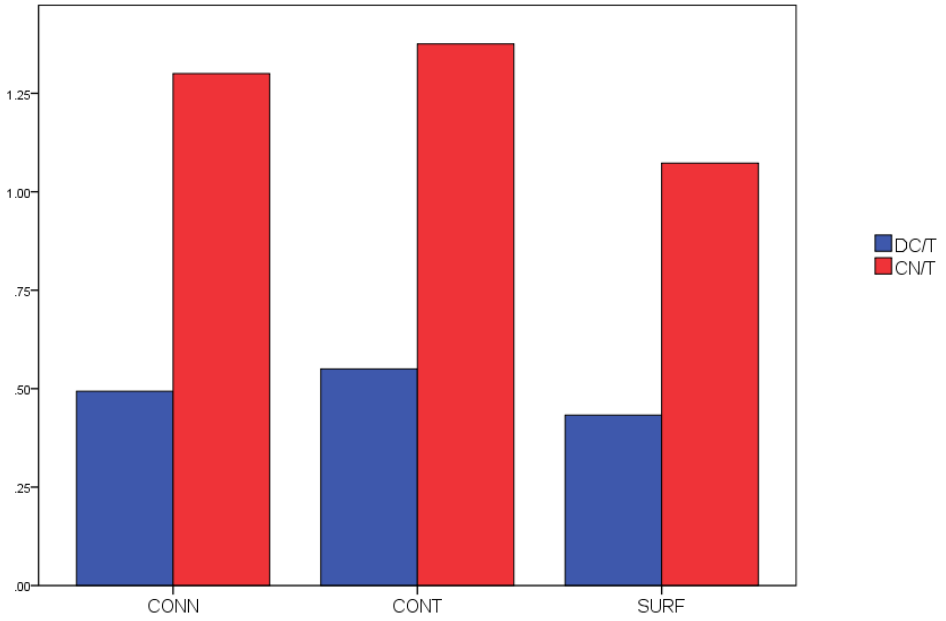


FIGURE 17
Belinda's overall syntactic complexity organized collectively by topic category



Still, no clear correlation exists between topic features and complexity measures. As can be seen in Table 5, DC/T and CN/T correlate only slightly with topic word count and WTC scores. In other words, topic features, at least as they are operationalized here, do not have any clear impact on complexity for the NNS.

TABLE 5
Correlations among Belinda's topic measures and total complexity and accuracy measures

	DC/T	CN/T
Word count	.08	.08
WTC total	.10	.06

Interestingly, Michelle and Belinda exhibit similar levels of complexity. Belinda's mean DC/T of 0.58 is only slightly lower than Michelle's mean of 0.59, and her CN/T of 1.43 actually surpasses Michelle's mean of 1.10. Both employ complex nominals more than subordination, possibly a feature of written discourse (Biber et al., 2008). However, neither CN/T nor DC/T is particularly high, which may reflect the conversational nature of email, which may "represent a convergence of both oral and written modalities" (Gonzalez-Bueno, 1998, p. 60).

From these analyses, it becomes apparent that certain topics, such as foreigners and movies, affect the participants' writing in similar ways. If each of Michelle's emails is paired with Belinda's subsequent email, the correlation between Michelle and Belinda's DC/T scores and their CN/T scores is weak and statistically insignificant (r=.12 and r=-.49, respectively). Still, Figures 18 and 19 show that these scores do coincide to some extent in certain segments of the email exchange.

FIGURE 18
Paired evolution of Michelle and Belinda's DC/T scores

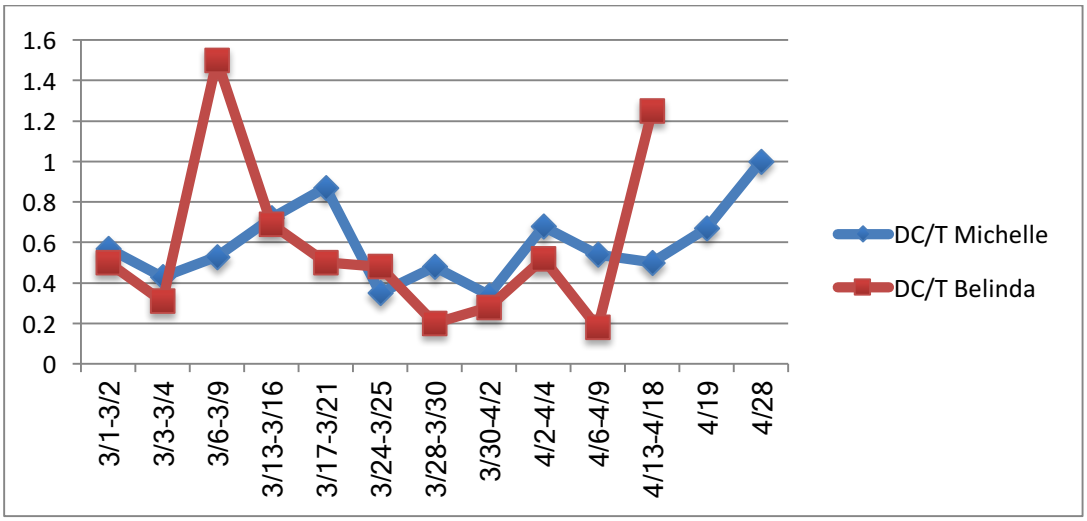
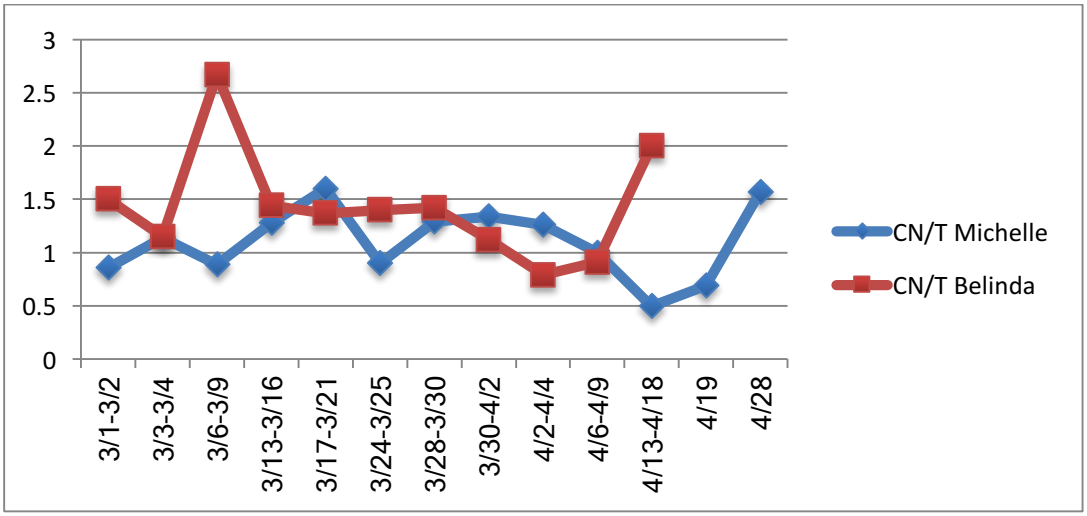


FIGURE 19
Paired evolution of Michelle and Belinda's CN/T scores



The participants' DC/T scores in Figure 18 start off with a similar complexity level and dip around the same time. Belinda's peak at 3/9, while abrupt, is followed by a slow increase in DC/T by Michelle. When Belinda's DC/T quickly decreases again, Michelle's eventually follows suit, and the two remain at a low point for several weeks about halfway through the exchange. Both participants' DC/T increases in the 4/2-4/4 turn, after which another dip ensues, but Belinda's second peak on 4/18 is followed once again by Michelle's steady increase in DC/T. Similarly, as seen in Figure 19, Michelle's CN/T seems to reflect Belinda's with a short delay. Belinda's peak on 3/9 is followed by a steady increase in Michelle's CN/T. Belinda's CN/T then remains stable from 3/16 to 3/30, as which point it slowly declines and peaks once again for her last email on 4/18. Michelle's CN/T eventually mirrors this pattern, stabilizing around 3/28, declining slowly from there, and ending in a peak in her last email on 4/28.

The common dip in DC/T and the stable period for CN/T occur at an intersection between connection topics and content topics. On 3/17, Belinda introduces the music topic, and the conversation then shifts towards a discussion of their countries and cultures. It therefore seems possible that complexity is lowest once an amicable relationship has been established and the participants feel at ease moving from surface and connection topics to content topics. DC/T picks up again when Michelle explains colors and Belinda brings up games, while CN/T declines for both at this point. Both DC/T and CN/T peak in each participant's final email, which are devoted entirely to surface topics.

Accuracy

As would be expected with any NS, Michelle's article usage accuracy performance was high (97.82%). However, as can be seen in Table 6 and Figure 20, some variability occurs. Michelle's use of the indefinite article *a* is generally 100% accurate, with the exception of her email on 4/19 (50% accuracy). As a single case, this outlier may be due to a typographical error (the omission of *a* in "just to give you _ idea how it looks") and may not be systematic.

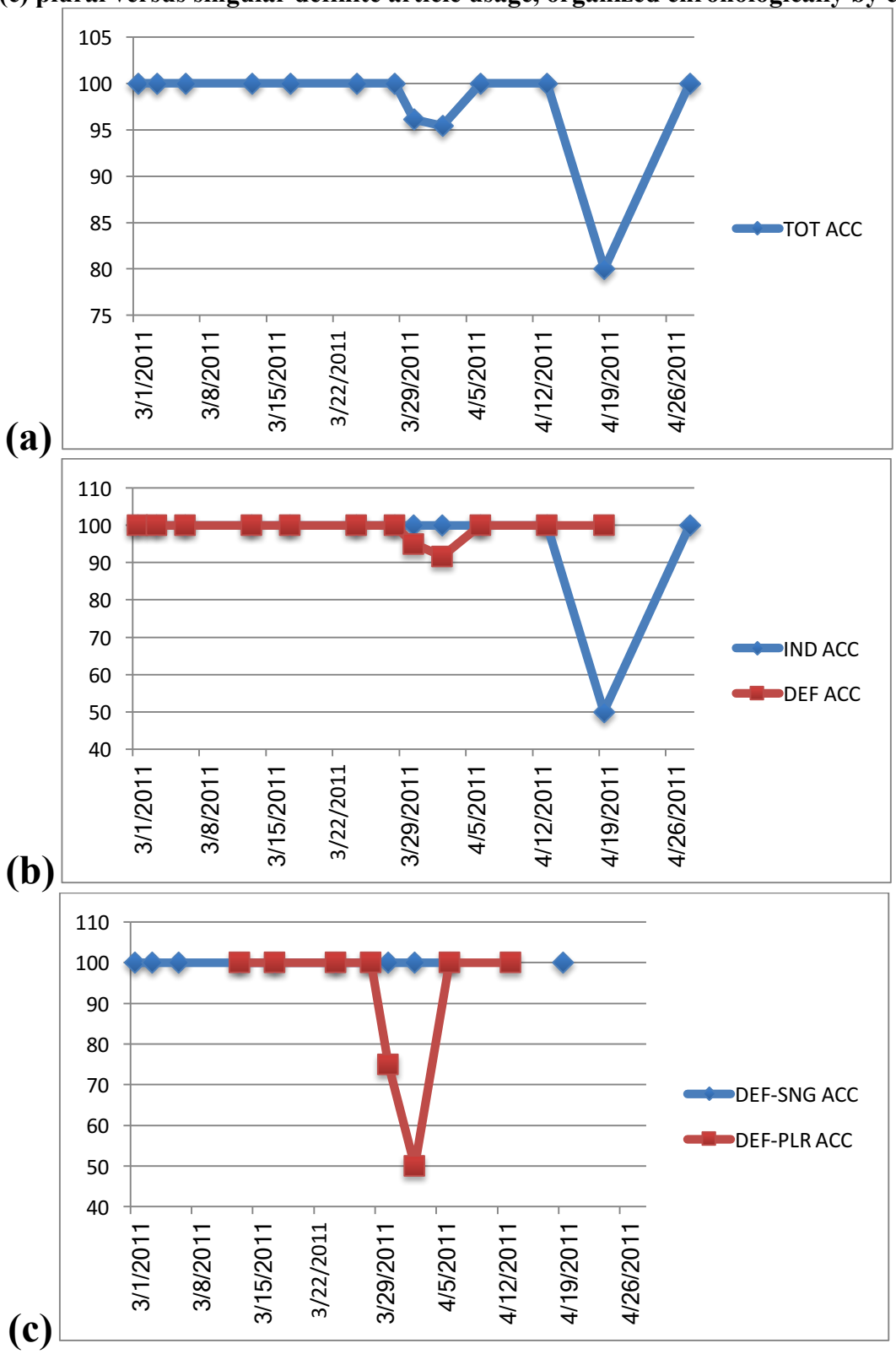
TABLE 6
TLU analysis of Michelle's article accuracy by email

Email	singular definite				plural definite				singular indefinite				total
date	OC	CS	OS	%	OC	CS	OS	%	OC	CS	OS	%	%
3/1/2011	2	2	0	100	0	0	0		1	1	0	100	100
3/3/2011	1	1	0	100	0	0	0		4	4	0	100	100
3/6/2011	2	2	0	100	0	0	0		9	9	0	100	100
3/13/2011	4	4	0	100	2	2	0	100	9	9	0	100	100
3/17/2011	2	2	0	100	1	1	0	100	6	6	0	100	100
3/24/2011	17	17	0	100	2	2	0	100	3	3	0	100	100
3/28/2011	8	8	0	100	3	3	0	100	5	5	0	100	100
3/30/2011	16	16	0	100	3	3	1	75	6	6	0	100	96.15
4/2/2011	10	10	0	100	1	1	1	50	10	10	0	100	95.45
4/6/2011	13	13	0	100	1	1	0	100	8	8	0	100	100
4/13/2011	0	0	0		1	1	0	100	1	1	0	100	100
4/19/2011	3	3	0	100	0	0	0		2	1	0	50	80
4/28/2011	0	0	0		0	0	0		2	2	0	100	100

OC = obligatory context CS = correct suppliance OS = oversuppliance

However, the variability in her use of the definite article *the* on 3/30 and 4/2 seems slightly more systematic. Indeed, the issue pertains only to her use of the plural definite article, which may confirm that the singular and plural definite article constitute two separate subsystems. The two instances of this error may result from overuse of the plural *the* in cases in which the subject is indefinite and requires the zero article ("In Mexico, the people go way out..." on 3/30 and "As to the colors, in our culture…" on 4/2). An analysis of Michelle's article accuracy as a function of topic yields near-identical results and therefore is not reported here. Raw numbers for Michelle's article use by topic can be found in Appendix B.

FIGURE 20
Michelle's variability in (a) overall article usage, (b) definite versus indefinite article usage, and (c) plural versus singular definite article usage, organized chronologically by email



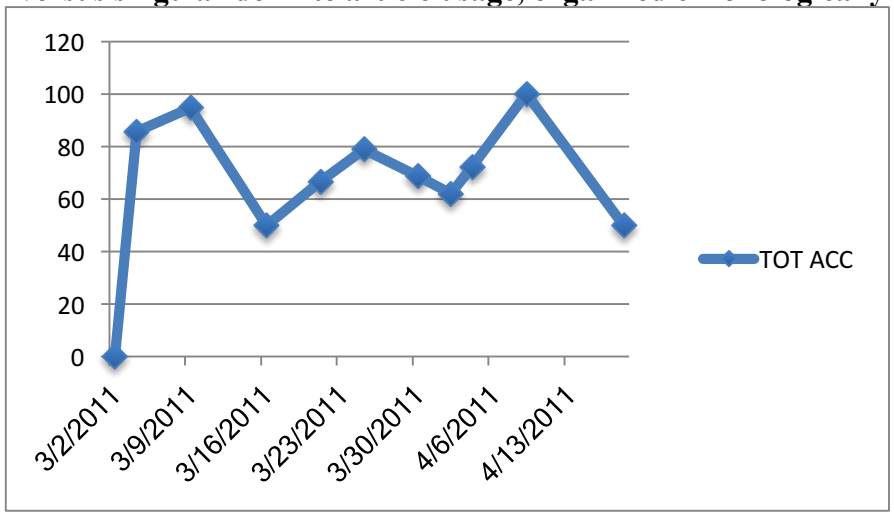
Belinda's accuracy percentages, on the contrary, reveal great variability in her article usage. While her mean of 66.28% indicates only partial acquisition of *the* and *a*, Table 7 and Figure 21 reveal fluctuation over the course of interaction.

TABLE 7
TLU analysis of Belinda's article accuracy by email

Email	singular definite				plural definite				singular indefinite				total
date	OC	CS	OS	%	OC	CS	OS	%	OC	CS	OS	%	%
3/2/2011	1	0	0	0	0	0	0		1	0	0	0	0
3/4/2011	1	1	0	100	0	0	0		6	5	0	83.33	85.71
3/9/2011	12	12	0	100	3	3	0	100	4	3	0	75	94.74
3/16/2011	4	2	0	50	1	1	0	100	1	0	0	0	50
3/21/2011	3	2	0	66.67	0	0	0		5	4	1	66.67	66.67
3/25/2011	2	2	0	100	1	1	0	100	16	12	0	75	78.95
3/30/2011	8	5	1	55.56	1	1	0	100	6	5	0	83.33	68.75
4/2/2011	19	12	0	63.16	2	1	0	50	8	5	0	62.5	62.07
4/4/2011	41	28	0	68.29	10	9	0	90	3	2	0	66.67	72.22
4/9/2011	1	1	0	100	1	1	0	100	4	4	0	100	100
4/18/2011	0	0	1	0	0	0	0		1	1	0	100	50

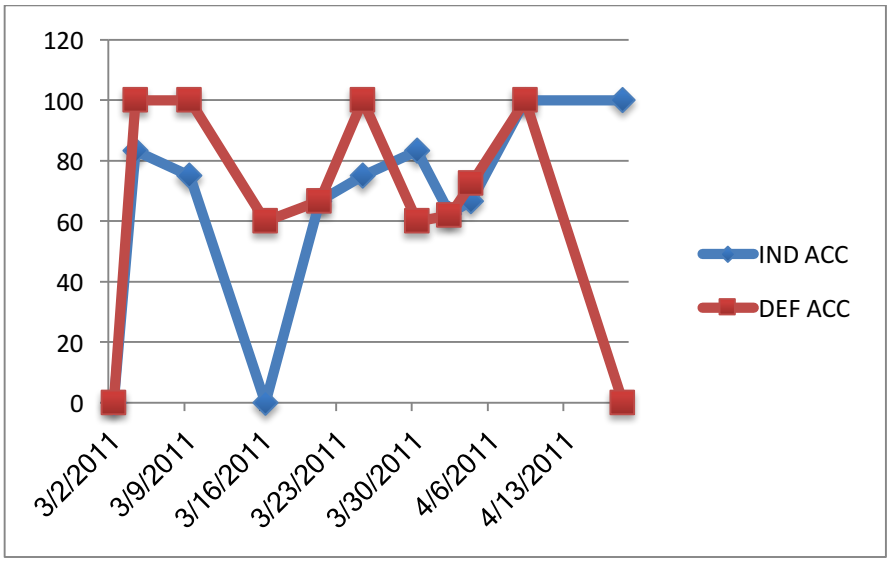
OC = obligatory context CS = correct suppliance OS = oversuppliance

FIGURE 21
Belinda's variability in (a) overall article usage, (b) definite versus indefinite article usage, and (c) plural versus singular definite article usage, organized chronologically by email

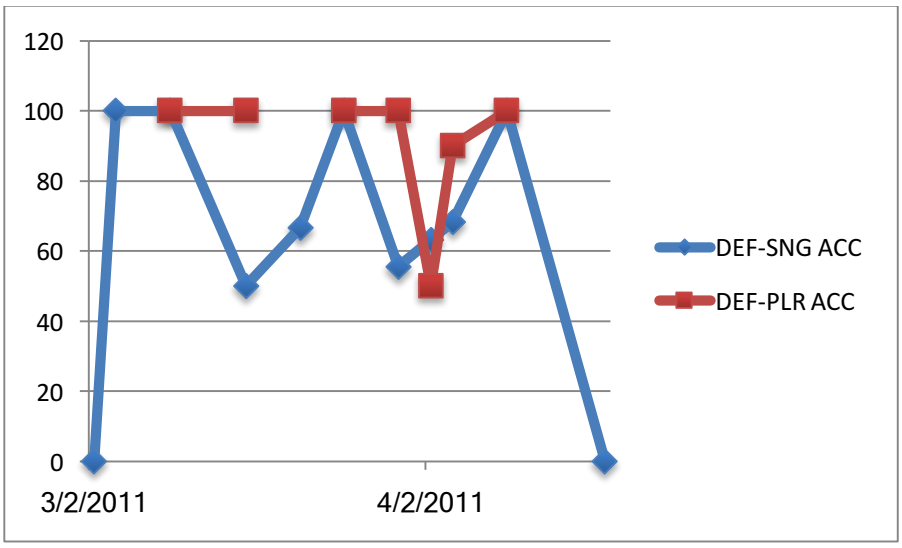


(a)

(b)



(c)



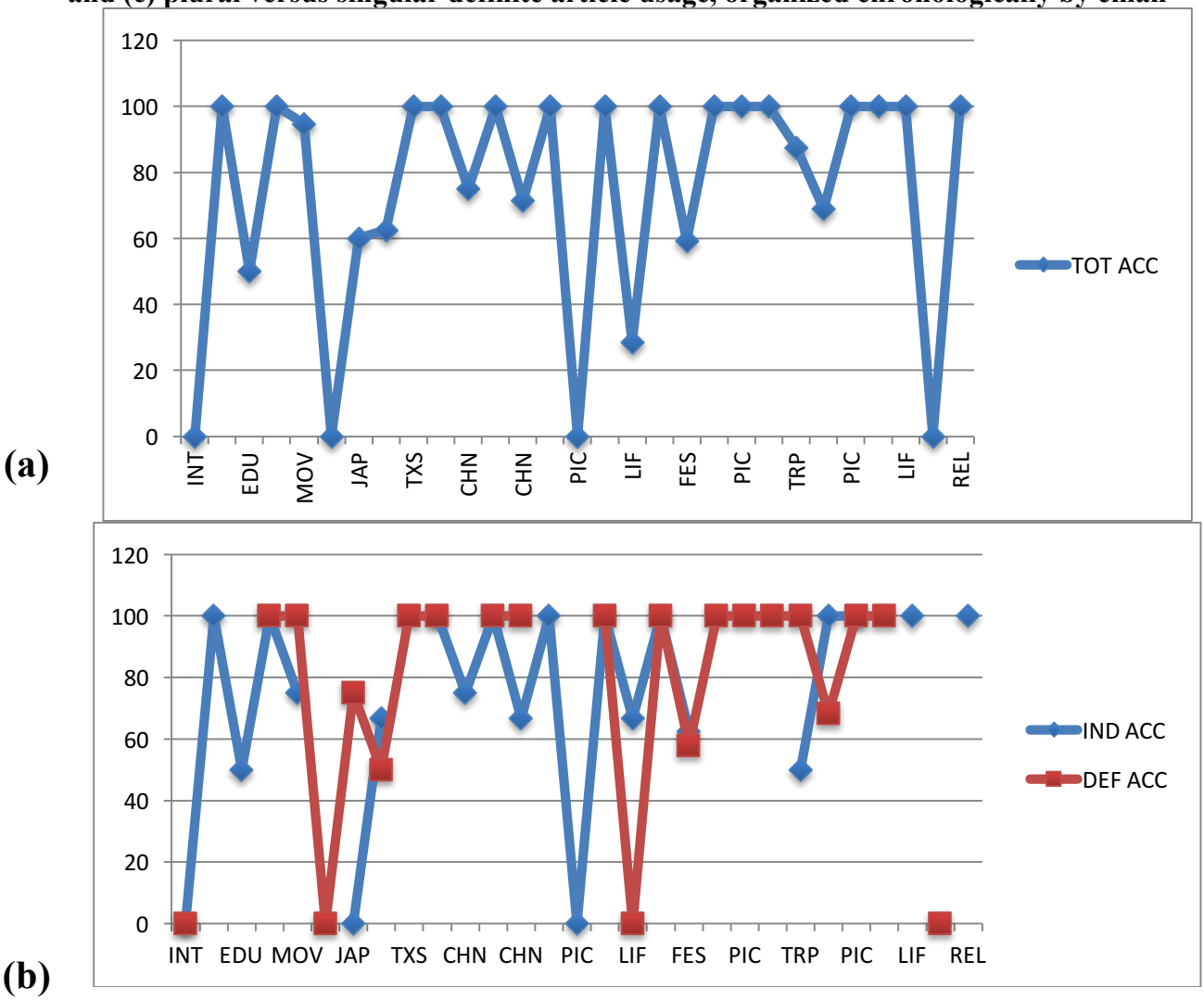
By topic, the fluctuations in Belinda's article usage are even greater. In Figure 22, article accuracy for all subsystems seems to fluctuate between 100% and 0% accuracy. Raw numbers for Belinda's article use by topic can be found in Appendix C. The correlations between subsystems are moderate and not statistically significant, with r=.55 between the indefinite and definite article, and r=.67 between the plural and singular definite article. Between the sparsity and variability of the data, little can be inferred as to connected or competitive growth. In cases of low accuracy, Belinda mainly underuses articles in obligatory contexts. For example, Belinda reaches approximately 69% article accuracy on 4/4, where her errors can be attributed to undersuppliance:

> **Excerpt 6 (GAM, 4/4):** First, the judge says: 'it's dark now; all close your eyes please.' All do what the judge says. Then __ judge say: 'Killer, open your eyes, and kill someone'. Next, the killer opens his/her eyes and motion the judge who he/she wants to kill and closes his/her eyes. Then, __ Judge says:'ok, police, open your eyes please.' __ Police opens his/her eyes and motion the judge who he/she

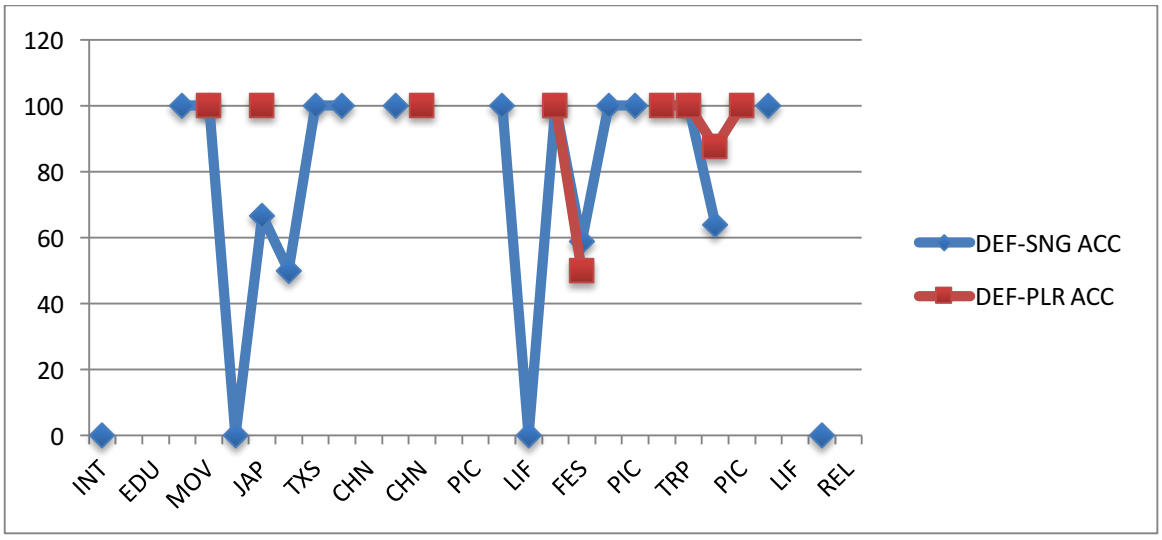
believes is the killer. Then __ Judge nods or shakes his/her head (just one chance). Then it is announced that the day breaks and all the people open their eyes. __ Judge tells the person who is killed and asks him/her to leave words. […] Being __ police, you have to find out __ killer and convince __ citizens to vote. Being __ murder, you have to pretend you are not the killer and use your logic to persuade others vote for another person. Being citizens, you should help ___ police (the one you believe he/she is) to catch __ murder.

In this excerpt, Belinda's use of *the* visibly fluctuates, particularly for the words "judge," "murder[er]," "killer," and "police." This may be a sign of emergent use, as she seems to understand the necessity of the definite article but employs it with little consistency.

FIGURE 22
Belinda's variability in (a) overall article usage, (b) definite versus indefinite article usage, and (c) plural versus singular definite article usage, organized chronologically by email



(c)



When clustered by topic category, as seen in Figures 23 and 24, Belinda's article accuracy seems highest for content and connection topics and lowest for surface topics. This may be due to low word counts, as several categories with 0% accuracy offer only one obligatory occasion (see Appendix C), thus skewing the data. The results, therefore, are largely inconclusive.

FIGURE 23
Belinda's overall article accuracy organized collectively by topic

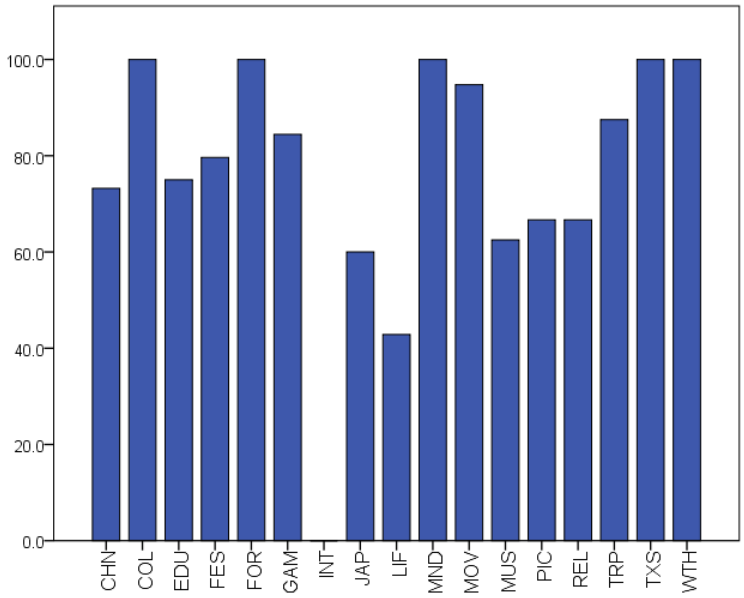
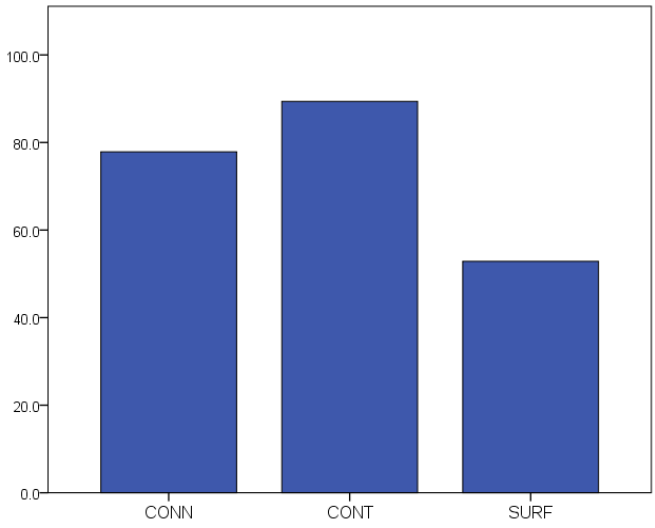
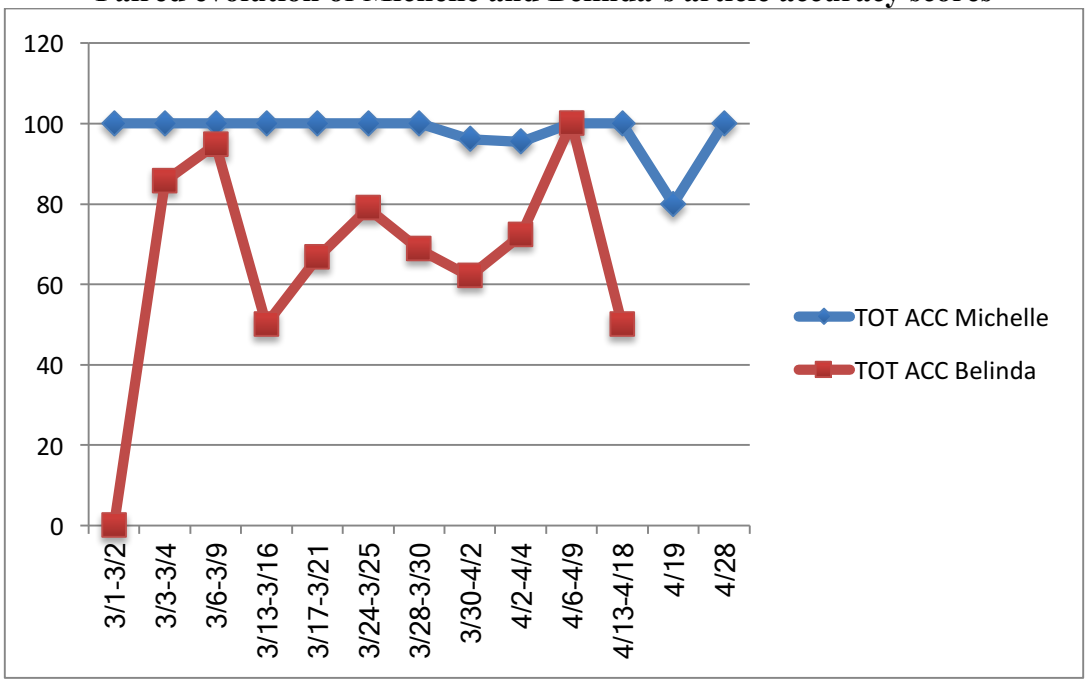


FIGURE 24
Belinda's overall article accuracy organized collectively by topic category



Overall, the correlation between Michelle and Belinda's article usage throughout the email exchange is insignificant and practically non-existent, with r=-.02. Figure 25 does not hint at any specific relationship, nor does a close, qualitative look at the data.

FIGURE 24
Paired evolution of Michelle and Belinda's article accuracy scores



Trade-offs between Complexity and Accuracy

Variability across topics in both accuracy and complexity makes it difficult to infer specific trade-offs between the two. Indeed, as can be seen in Tables 8 and 9, the correlations between accuracy and complexity for both participants hover around r=.00.

TABLE 8
Correlations among Michelle's syntactic complexity and article accuracy measures

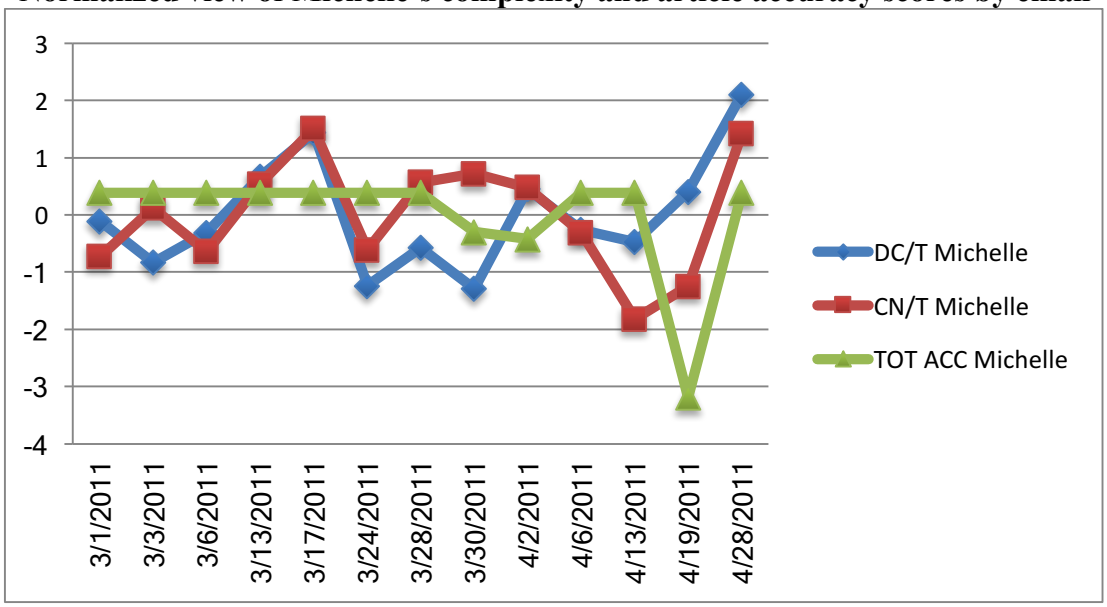
	TOT ACC	IND ACC	DEF ACC	DEF PLR ACC	DEF SNG ACC
DC/T	-.10	-.11	-.06	-.04	.[c]
CN/T	-.16	-.21	-.01	-.01	.[c]

TABLE 9
Correlations among Belinda's syntactic complexity and article accuracy measures

	TOT ACC	IND ACC	DEF ACC	DEF PLR ACC	DEF SNG ACC
DC/T	.05	.19	.05	.05	.10
CN/T	-.11	.02	-.09	.05	.07

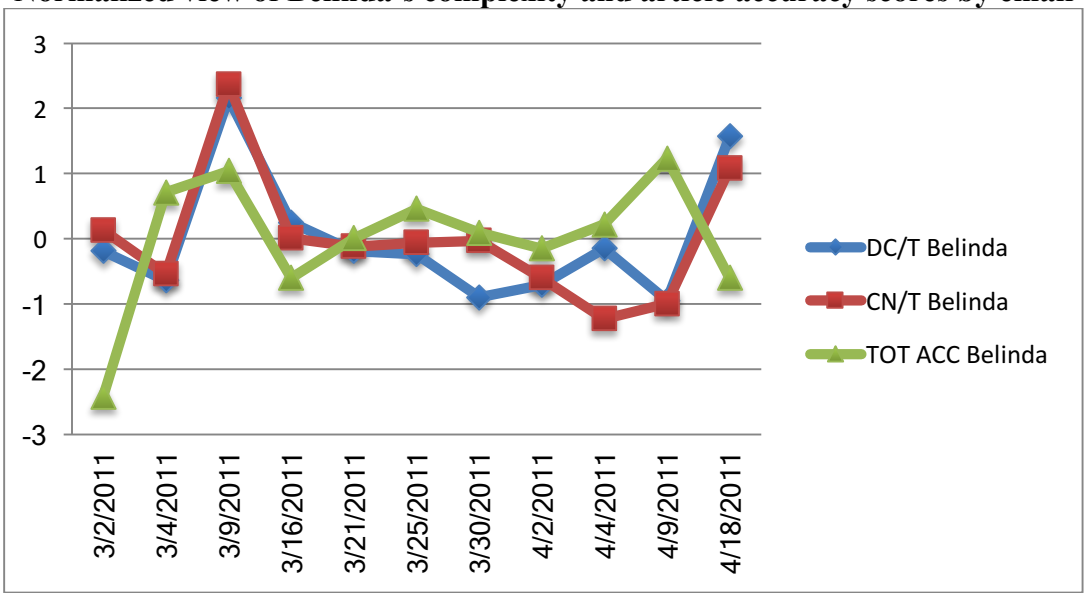
Although Michelle's complexity-accuracy correlations are all slightly negative, this result is mostly negligible due to the outlier identified as a typographical error. As shown in Figure 25, Michelle's article accuracy fluctuates much less than her complexity, a result compliant with her NS status. No particular relationship between the two subsystems can be determined.

FIGURE 25
Normalized view of Michelle's complexity and article accuracy scores by email



Instead, for Belinda, while the correlation coefficients in Table 9 indicate no clear relationship between complexity and accuracy, Figure 26 does reveal points at which the two variables may have a positive relationship. Though the relationship starts and ends negatively, the two variables seem to follow a similar trajectory between 3/9 and 4/4. In other words, the bookends of the email exchange show little evidence of connected use, but the core of the exchange, in which connection and content topics predominate, could possibly host connected use. This would suggest that topic indeed affects how complexity and article accuracy interact, with certain topics inducing trade-offs and others not, though the evidence here is not strong enough to confirm this finding.

FIGURE 26
Normalized view of Belinda's complexity and article accuracy scores by email



For both participants, the correlations of article accuracy with topic word count and with WTC ratings were slightly negative, while the correlations of syntactic complexity measures with topic word count and with WTC were slightly positive, as can be seen in Tables 10 and 11. Though statistically insignificant and weak, these correlations may indicate that WTC and word count, as elements of topic, may have a negative relationship with accuracy and a positive relationship with syntactic complexity. The consistency across participants could imply that topic has the same general effect on both NSs and NNSs.

TABLE 10
Correlations among Michelle's topic measures and total complexity and accuracy measures

	DC/T	CN/T	TOT ACC
Word count	.08	.13	-.051
WTC total	.07	.22	-.077

TABLE 11
Correlations among Belinda's topic measures and total complexity and accuracy measures

	DC/T	CN/T	TOT ACC
Word count	.08	.08	-.03
WTC total	.10	.06	-.22

DISCUSSION AND CONCLUSION

Through a CDST lens, this study has attempted to examine whether situational factors such as topic constitute an integral part of an evolving, dynamic IL. As seen in the results, topic can be considered its own dynamic system, constantly in flux within and across emails, with one topic triggering another in a chain reaction. Several different patterns of topic evolution can be observed based on the number of words dispensed for each topic and the topic's category. This evolution seems fueled in part by the participants' WTC for each topic occurrence. In observing the interweaving of topics within and across emails, it is difficult not to draw a parallel with spoken conversation. It may be possible to identify what Jefferson (1984) calls stepwise transitions—that is, gradual disengagements from one topic leading into the next topic—as well as more abrupt interjections. Furthermore, it may be of interest to investigate adjacency pairs occurring through email, as its asynchronous nature dictates different conversational conventions. While outside the scope of this study, a survey of topic in computer-mediated communication merits further research, as it tends to display a convergence of both spoken and written conventions (Gonzalez-Bueno, 1998).

Within syntactic complexity, both complex nominalization and subordination also evolve dynamically by topic. In fact, much more variability is seen from topic to topic than from email to email, suggesting that fluctuations in dynamic systems occur at the micro-situational level. CN/T and DC/T both are more variable for the NS than for the NNS, which may suggest that while repeller states are typical of IL development in NNSs, they may also be characteristic of NS language use. Topic variation may dictate Michelle's syntactic complexity, which seems natural given that different topics elicit certain structures. However, this speculation is statistically unsubstantiated, as no general trends were apparent by topic category. Michelle's complexity seems more sporadically impacted by topic than Belinda's, with surface, connection, and content topics presenting similar complexity levels. Belinda's complexity scores, on the contrary, behave less erratically, reaching their highest for content topics, as might be expected. It may be possible, then, that NNSs are more clearly impacted by topic than NSs, whose ILs have reached a more stable state. While more linear, though, Belinda's range of complexity is also wider, achieving higher levels of complexity than Michelle. One would think the NS would display higher levels of complexity, as she has more linguistic resources to draw from, but instead the NNS seems to push the limits of her IL, hinting that at times she may have experimented with diverse syntactic constructions, which could be interpreted as a sign of development and of a slightly more formal writing style. The fact that Michelle stayed within a smaller range of complexity may be due to her more informal writing style, which could reflect the nature of the email genre (Gonzalez-Bueno, 1998). Belinda's tendency towards slightly more formal and complex syntax may have been a repercussion of her use of English almost

exclusively in academic contexts. In any case, CN/T remained on average higher than DC/T for both participants throughout the email exchange, reflecting, perhaps, the tendency in written discourse to rely on "compressed" language (Biber et al., 2008, p. 184).

The participants' performance on accuracy reveals that while NNS article usage displays the type of instability typical of a dynamic system, NS article usage generally remains constant. This conforms to CDST in that Michelle seems to have reached a permanent attractor state, while Belinda's IL is still in flux. Their performance also corroborates research on cross-linguistic influence. Indeed, Michelle's overuse of the definite article in cases of plural generix reference complies with prior studies suggesting that issues arise due to Spanish's requirement of an article for generic reference when English employs the zero article for generic reference (Diez-Bedmar & Papp, 2008). Similarly, Belinda's underuse of articles is likely due to the lack of an article system in her native language, Mandarin (Diez-Bedmar & Papp, 2008). Also, the low accuracy percentage for topics with higher word counts may validate prior findings that topic investment negatively impacts accuracy due to reduced monitoring when discussing an invested topic (Eisenstein & Starbuck, 1989), and that "email" as a discourse genre may entail less attention to grammatical form due to the "urgency of communicative flow" characteristic of peer interactions in this medium (Gonzalez-Bueno, 1998, p. 59).

This study's findings in terms of trade-offs between complexity and accuracy as a function of topic are limited, supporting neither Robinson's 1995, 2001, 2003) Cognition Hypothesis and Foster and Skehan's (1996; Skehan, 1998; Skehan & Foster, 1997, 1999) Limited Attentional Capacity Model. However, it is possible that certain topic features, identified as influences on WTC, positively correlate with syntactic complexity and negatively correlate with article accuracy. Such a finding would imply that certain situational WTC and the topic features associated with them lead to trade-offs between complexity and accuracy, which reinforces Foster and Skehan's argument. Although the statistics for this claim are weak and insignificant, WTC's slight negative correlation with accuracy and slight positive correlation with complexity begs further investigation. The data also intimate a difference between NSs and NNSs. While NSs' language use—having reached what could be called, in L2 development terms, an end state—is marginally influenced by situational variables, for NNSs, situational variables like topic may differentially impact whether complexity and accuracy compete for resources—supporting Skehan's trade-off hypothesis—or whether they are connected growers—supporting Robinson's hypothesis.

None of these findings, of course, can be completely confirmed in the present study due to its methodological limitations. In terms of topic, the boundaries between themes were difficult to determine regardless of iterative coding processes. Plus, the application of topic feature codes may not have been reliable, as the analyst had to rely on etic interpretations of naturalistic data alone, without the help of verbal report data as has been used in studies of situational WTC (Kang, 2004; MacIntyre & Legatto, 2010). A true evaluation of WTC for each topic occurrence would require detailed stimulated-recall data to triangulate the researcher's interpretations of WTC. Finally, simplifying topic feature variables into a single WTC score may have masked more fine-grained interactions of each sub-feature of WTC with L2 performance. The codes themselves may have diluted some of the subtle characteristics of each topic feature, masking other internal aspects of topic that may have catalyzed reactions in complexity and accuracy. For example, the illocutionary function—such as stating an opinion versus giving a list of events or other items—associated with each topic may have a greater impact on complexity than the topic itself. Besides these internal variables, external variables, such as the evolving relationship

between the two participants, may also have confounded the results. As is characteristic of dynamic systems, multiple variables, both overt and covert, may have influenced the participants' performance, making it difficult to isolate topic from other environmental and socio-cognitive factors and thus confounding the results.

In terms of complexity and accuracy, the number of words for each topic varied so greatly that some of the sections with smaller word counts may not have been substantial enough to correctly represent L2 performance. Although needed to identify variability within each email, operationalizing data points by topic boundaries rather than by temporal units may not be reliable due to this word-count imbalance. Some of the linguistic data may also have been misrepresented by the fact that, in such casual written interaction, not all idea units are necessarily clausal, consisting of only noun phrases at times. AS-units may have been a more appropriate measure than T-units, as the former also account for other phrases conveying pragmatic meaning (Ellis & Barkhuizen, 2005), although this conjecture cannot be confirmed without further empirical support. This issue, as well as the participants' overuse of commas to separate independent clauses, could have impacted Web-based L2 Syntactic Complexity Analyzer's ability to discern limits between T-units, which skew the complexity data. This automatic tool may have been more unreliable, at least for this dataset, than Lu's (2010) study suggests.

While positive correlations between the two complexity subsystems for both participants may constitute evidence of connected use, it is important to note that this finding may also mean that CN/T and DC/T are not in fact separate subsystems, but rather a single system, especially because some of their features overlap. Concurrently, the lack of sustained correspondence between the two complexity measures, which for Michelle, in particular tended to diverge at some points and converge at others, may expose the measures' deficiency in capturing syntactic complexity. Similarly, the selection of articles as a measure of accuracy may not have captured the participants' overall performance and may not have been appropriate to compare with the global complexity. Rather, an investigation of global accuracy through error analysis might have proven to be more appropriate than a TLU analysis of articles.

More glaringly, the definition of linguistic performance variables as discrete subsystems is itself questionable. Splitting the English article system by definite and indefinite articles does not recognize functional differences in article usage, such as the difference between anaphoric and cataphoric reference, which may represent a more valid distinction than mere formal differences between *a* and *the*. Also, as mentioned above, it is unclear whether complex nominalization and subordination truly constitute separate subsystems.

In spite of these barriers, the necessity of unifying both linguistic and non-linguistic variables is quintessential for a comprehensive understanding of dynamic systems. Topic is one of many other factors that can be married to linguistic performance, and an idiodynamic study of changes moment by moment provides promising ground for future research in the field.

REFERENCES

Biber, D., Grieve, J, & Iberri-Shea, G. (2008). Noun phrase modification. In G. Rohdenburg & J. Schlüter (Eds.), *One Language, Two Grammars? Differences between British and American English* (182-193). Cambridge: Cambridge University Press.

Baba, K., & Nitta, R. (2014). Phase transitions in development of writing fluency from a complex dynamic systems perspective. *Language Learning*, *64*(1), 1-35.

Cherciov, M. (2013). Investigating the impact of attitude on first language attrition and second language acquisition from a Dynamic Systems Theory perspective. *International Journal of Bilingualism, 17*(6), 716-733.

Churchill, E. (2008). A dynamic systems account of learning a word: From ecology to form relations. *Applied Linguistics, 29*(3), 339-358.

Cornu, A.M., & Delahaye, M. (1987). Variability in interlanguage reconsidered: LSP vs. non-LSP talk. *English for Specific Purposes, 6*(2), 145-152.

Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, *26*, 42-65.

De Groot, F.O. (2012). Intra-Individual Learner Variation in English L2 Writing: A Case Study of Complexity, and Accuracy Development in Intensive English Language Programs. *LEARN Journal: Language Education and Acquisition Research Network, 1*(1), 49-69.

Díez-Bedmar, M. B., & Papp, S. (2008). The use of the English article system by Chinese and Spanish learners. *Language and Computers*, *66*(1), 147-175.

Dowd, J. L. (1984). *Phonological variation in L2 speech: the effects of emotional questions and field-dependence/field-independence on second language performance*. (Unpublished doctoral dissertation). Teachers College, Columbia University.

Eisenstein, M. R., & Starbuck, R. (1989). Investment in topic and verb system accuracy of advanced L2 learners. In S. Gass, C. Madden, D. Preston, & L. Selinker (Eds.), *Variation in Second Language Acquisition (Vol. 2): Psycholinguistic Issues* (125-137). Clevedon: Multilingual Matters.

Ellis, R., & Barkhuizen, G. P. (2005). *Analysing learner language*. Oxford: Oxford University Press.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*(3), 299-323.

Giles, H. (Ed.). (1984). *The dynamics of speech accommodation*. Amsterdam: Mouton.

González-Bueno, M. (1998). The effects of electronic mail on Spanish L2 discourse. *Language Learning & Technology*, *1*(2), 55-70.

Kang, S. J. (2005). Dynamic emergence of situational willingness to communicate in a second language. *System*, *33*(2), 277-292.

King, J. (2013). Silence in the second language classrooms of Japanese universities. *Applied Linguistics*, *34*(3), 325-343.

Larsen-Freeman, D. (2006). The Emergence of Complexity, Fluency, and Accuracy in the Oral and Written Production of Five Chinese Learners of English. *Applied Linguistics*, *27*(4), 590-619.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, *15*(4), 474-496.

Macintyre, P. D., & Legatto, J. J. (2011). A dynamic system approach to willingness to communicate: Developing an idiodynamic method to capture rapidly changing affect. *Applied Linguistics*, *32*(2), 149-171.

Master, P. (1997). The English article system: Acquisition, function, and pedagogy. *System*, *25*(2), 215-232.

Opitz, C. (2012). A dynamic perspective on late bilinguals' linguistic development in an L2 environment. *International Journal of Bilingualism, 17*(6), 701-715

Pica, T. (1984). Methods of morpheme quantification: their effect on the interpretation of second language data. *Studies in Second Language Acquisition, 6*(1), 69-78.

Polat, B., & Kim, Y. (2014). Dynamics of complexity and accuracy: A longitudinal case study of advanced untutored development. *Applied Linguistics*, *35*(2), 184-207.

Reigel, D. (2008). Positive feedback in pairwork and its association with ESL course level promotion. *TESOL quarterly*, *42*(1), 79-98.

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning, 45*(1), 99-140.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*(1), 27-57.

Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies, 21*(2), 45-105.

Rosmawati (2014). Dynamic development of complexity and accuracy: A case study in second language academic writing . *Australian Review of Applied Linguistics, 37*(2), 75–100.

Selinker, L., & Douglas, D. (1985). Wrestling with "context" in interlanguage theory. *Applied Linguistics, 6*(2), 190-204.

Skehan, P. (1998). *A Cognitive Approach to Language Learning.* Oxford: Oxford University Press.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research, 1*(3), 185-211.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, *49*(1), 93-120.

Spoelman, M., & Verspoor, M. (2010). Dynamic Patterns in Development of Accuracy and Complexity: A Longitudinal Case Study in the Acquisition of Finnish. *Applied Linguistics, 31*(4), 532-553.

Street, R. L., & Giles, H. (1982). Speech accommodation theory: A social cognitive approach to language and speech behavior. In Roloff, M.E., & Berger, C.R. (Ed.), *Social Cognition and Communication* (193-226). Los Angeles: SAGE Publications.

Verspoor, M., Schmid, M. S., & Xu, X. (2012). A dynamic usage based perspective on L2 writing. *Journal of Second Language Writing*, *21*(3), 239-263.

Yang, W., & Sun, Y. (2015). Dynamic Development of Complexity, Accuracy and Fluency in Multilingual Learners' L1, L2 and L3 Writing. *Theory and Practice in Language Studies*, *5*(2), 298.

Woken, M. D., & Swales, J. (1989). Expertise and authority in native-non-native conversations: The need for a variable account. In S. Gass, C. Madden, D. Preston & L. Selinker (Eds.), *Variation in second language acquisition (Vol. 1): Discourse and Pragmatics* (211-227). Clevedon and Philadelphia: Multilingual Matters.

Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity*. Honolulu: University of Hawai'i Press.

Zuengler, J. (1982). Applying accommodation theory to variable performance data in L2. *Studies in Second Language Acquisition, 4*(2), 181-192.

Zuengler, J. (1989). Assessing an interaction-based paradigm: How accommodative should we be? In M. R. Eisenstein (Ed.), *The dynamic interlanguage: Empirical studies in second language acquisition* (49-67). New York: Plenum.

Zuengler, J. (1993). Explaining NNS interactional behavior: The effect of conversational topic. *Interlanguage Pragmatics*, 184-195.

Zuengler, J., & Bent, B. (1991). Relative knowledge of content domain: An influence on native-non-native conversations. *Applied Linguistics, 12*(4)*,* 397-415.

APPENDIX A
Sequence of topics over time and their associated features

Participant	Date	Category	Topic	Word count	INV	PRS	SNS	EXP	GRP	Total WTC
Michelle	3/1/2011	SURF	INT	16	0	1	0	0	0	1
		SURF	REL	21	0	0	0	0	0	0
		SURF	INT	28	0	1	0	0	0	1
		SURF	REL	15	0	0	0	0	0	0
Belinda	3/2/2011	SURF	INT	32	0	1	0	0	0	1
		SURF	REL	50	0	0	0	0	0	0
Michelle	3/3/2011	SURF	REL	20	1	1	0	0	0	2
		CONN	EDU	20	0	1	0	0	0	1
		CONN	HOB	12	0	1	0	0	0	1
		CONN	EDU	26	0	1	0	0	0	1
		CONN	HOB	8	0	0	0	0	0	0
		SURF	GEN	3	0	0	0	0	0	0
Belinda	3/4/2011	SURF	GEN	5	0	0	0	0	0	0
		CONN	EDU	13	0	1	0	0	0	1
		CONN	SPN	19	1	0	0	0	0	1
		CONN	HOB	24	0	1	0	0	0	1
		CONN	EDU	35	1	0	0	0	0	1
		SURF	REL	28	0	0	0	0	0	0
Michelle	3/6/2011	SURF	GEN	23	0	1	0	0	0	1
		CONN	EDU	28	1	1	0	0	0	2
		CONN	SPN	52	0	1	0	1	1	3
		CONN	HOB	30	1	1	0	0	0	2
		CONN	EDU	63	0	0	0	1	0	1
		SURF	REL	29	0	1	0	0	0	1
		SURF	GEN	14	0	0	0	0	0	0
Belinda	3/9/2011	CONN	MOV	137	1	1	0	0	0	2
Michelle	3/13/2011	SURF	GEN	16	0	0	0	0	0	0
		CONN	JAP	30	1	0	1	0	0	0
		CONN	MOV	190	1	1	0	0	0	2
		SURF	GEN	4	0	0	0	0	0	0
Belinda	3/16/2011	SURF	GEN	11	0	0	0	0	0	0
		SURF	REL	20	0	0	0	0	0	0
		CONN	JAP	80	0	1	1	1	1	2
		CONN	MOV	86	1	1	0	0	0	2
Michelle	3/17/2011	CONN	JAP	15	0	0	1	0	0	-1
		CONN	MOV	131	1	1	0	0	0	2
		CONN	MUS	32	1	1	0	0	0	2

Participant	Date	Category	Topic	Word count	INV	PRS	SNS	EXP	GRP	Total WTC
		SURF	GEN	8	0	0	0	0	0	0
Belinda	3/21/2011	CONN	MUS	193	1	1	0	0	0	2
		CONT	TXS	34	1	1	0	0	0	2
Michelle	3/24/2011	CONN	MUS	85	1	1	0	0	0	2
		CONT	TXS	238	1	1	0	1	1	4
		CONT	CHN	20	1	0	0	0	1	2
		CONN	MND	16	1	1	0	0	0	2
Belinda	3/25/2011	CONT	TXS	47	1	1	0	0	0	2
		CONT	CHN	94	1	1	0	1	1	4
		CONT	FOR	43	1	1	0	0	0	2
		CONT	CHN	74	1	1	0	1	1	4
		CONN	MND	70	1	1	0	1	1	4
		SURF	PIC	27	0	0	0	0	0	0
Michelle	3/28/2011	CONT	CHN	37	1	0	0	0	0	1
		SURF	PIC	47	0	0	0	0	0	0
		CONT	FOR	93	1	1	0	1	1	4
		CONT	WTH	83	1	1	0	1	1	4
		CONN	MND	25	0	1	0	0	0	1
Belinda	3/30/2011	CONT	WTH	87	1	0	0	1	1	3
		SURF	LIF	49	1	1	0	0	0	2
		SURF	PIC	7	0	0	0	0	0	0
		CONT	FES	52	1	0	0	0	1	2
Michelle	3/30/2011	CONT	FES	504	1	1	0	1	1	4
		CONT	WTH	37	1	0	0	0	0	1
		CONT	FES	15	1	0	0	0	1	2
		SURF	GEN	32	0	0	0	1	1	2
Belinda	4/2/2011	CONT	FES	300	1	1	0	1	1	4
		CONT	COL	34	1	0	0	0	1	2
		SURF	PIC	18	0	1	0	0	0	1
Michelle	4/2/2011	CONT	FES	100	1	1	0	1	1	4
		CONT	COL	316	1	0	0	1	1	3
		SURF	PIC	55	1	1	0	0	0	2
Belinda	4/4/2011	CONT	COL	7	0	0	0	1	1	2
		CONT	TRP	118	1	1	0	0	0	2
		CONT	GAM	357	1	1	0	1	0	3
Michelle	4/6/2011	CONT	GAM	204	1	1	0	1	1	4
		CONT	TRP	42	1	0	0	0	0	1
		SURF	PIC	38	1	1	0	0	0	2
		SURF	LIF	27	1	1	0	0	0	2

Participant	Date	Category	Topic	Word count	INV	PRS	SNS	EXP	GRP	Total WTC
Belinda	4/9/2011	SURF	PIC	28	0	1	0	0	0	1
		CONT	GAM	46	0	1	0	0	0	1
		SURF	LIF	39	0	0	0	0	0	0
Michelle	4/13/2011	SURF	GEN	26	0	0	0	0	0	0
		SURF	PIC	15	0	0	0	0	0	0
		SURF	GEN	6	0	0	0	0	0	0
Belinda	4/18/2011	SURF	LIF	32	0	0	0	0	0	0
		SURF	REL	39	1	1	0	0	0	2
Michelle	4/19/2011	SURF	GEN	65	1	1	0	0	0	2
		SURF	LIF	51	1	1	0	0	0	2
		SURF	PIC	36	1	1	0	0	0	2
		SURF	LIF	26	1	1	0	0	0	2
Michelle	4/28/2011	SURF	GEN	76	0	0	0	0	0	0
		SURF	REL	32	1	1	0	0	0	2

APPENDIX B
Raw TLU analysis data for Michelle's article accuracy by topic

Email	Topic	singular definite				plural definite				singular indefinite				total
date	code	OC	CS	OS	%	OC	CS	OS	%	OC	CS	OS	%	%
3/1/2011	INT	1	1	0	100	0	0	0		0	0	0		100
3/1/2011	REL	0	0	0		0	0	0		0	0	0		
3/1/2011	INT	1	1	0	100	0	0	0		1	1	0	100	100
3/1/2011	REL	0	0	0		0	0	0		0	0	0		
3/3/2011	REL	1	1	0	100	0	0	0		0	0	0		100
3/3/2011	EDU	0	0	0		0	0	0		2	2	0	100	100
3/3/2011	HOB	0	0	0		0	0	0		0	0	0		
3/3/2011	EDU	0	0	0		0	0	0		2	2	0	100	100
3/3/2011	HOB	0	0	0		0	0	0		0	0	0		
3/3/2011	GEN	0	0	0		0	0	0		0	0	0		
3/6/2011	GEN	0	0	0		0	0	0		0	0	0		
3/6/2011	EDU	0	0	0		0	0	0		2	2	0	100	100
3/6/2011	SPN	1	1	0	100	0	0	0		1	1	0	100	100
3/6/2011	HOB	0	0	0		0	0	0		1	1	0	100	100
3/6/2011	EDU	1	1	0	100	0	0	0		2	2	0	100	100
3/6/2011	REL	0	0	0		0	0	0		2	2	0	100	100
3/6/2011	GEN	0	0	0		0	0	0		1	1	0	100	100
3/13/2011	GEN	0	0	0		0	0	0		0	0	0		
3/13/2011	JAP	1	1	0	100	0	0	0		0	0	0		100
3/13/2011	MOV	3	3	0	100	2	2	0	100	8	8	0	100	100
3/13/2011	GEN	0	0	0		0	0	0		1	1	0	100	100
3/17/2011	JAP	0	0	0		0	0	0		0	0	0		
3/17/2011	MOV	2	2	0	100	1	1	0	100	5	5	0	100	100
3/17/2011	MUS	0	0	0		0	0	0		0	0	0		
3/17/2011	GEN	0	0	0		0	0	0		1	1	0	100	100

3/24/2011	MUS	2	2	0	100	0	0	0		1	1	0	100	100
3/24/2011	TXS	13	13	0	100	1	1	0	100	0	0	0		100
3/24/2011	CHN	2	2	0	100	1	1	0	100	1	1	0	100	100
3/24/2011	MND	0	0	0		0	0	0		1	1	0	100	100
3/28/2011	CHN	1	1	0	100	0	0	0		1	1	0	100	100
3/28/2011	PIC	2	2	0	100	0	0	0		2	2	0	100	100
3/28/2011	FOR	2	2	0	100	2	2	0	100	0	0	0		100
3/28/2011	WTH	3	3	0	100	1	1	0	100	0	0	0		100
3/28/2011	MND	0	0	0		0	0	0		2	2	0	100	100
3/30/2011	FES	14	14	0	100	3	3	1	75	5	5	0	100	95.65
3/30/2011	WTH	1	1	0	100	0	0	0		1	1	0	100	100
3/30/2011	FES	0	0	0		0	0	0		0	0	0		
3/30/2011	GEN	1	1	0	100	0	0	0		0	0	0		100
4/2/2011	FES	1	1	0	100	0	0	0		2	2	0	100	100
4/2/2011	COL	7	7	0	100	1	1	1	50	7	7	0	100	93.75
4/2/2011	PIC	2	2	0	100	0	0	0		1	1	0	100	100
4/6/2011	GAM	10	10	0	100	0	0	0		5	5	0	100	100
4/6/2011	TRP	0	0	0		1	1	0	100	3	3	0	100	100
4/6/2011	PIC	2	2	0	100	0	0	0		0	0	0		100
4/6/2011	LIF	1	1	0	100	0	0	0		0	0	0		100
4/13/2011	GEN	0	0	0		0	0	0		0	0	0		
4/13/2011	PIC	0	0	0		1	1	0	100	1	1	0	100	100
4/13/2011	GEN	0	0	0		0	0	0		0	0	0		
4/19/2011	GEN	0	0	0		0	0	0		1	1	0	100	100
4/19/2011	LIF	2	2	0	100	0	0	0		0	0	0		100
4/19/2011	PIC	1	1	0	100	0	0	0		1	0	0	0	50
4/19/2011	LIF	0	0	0		0	0	0		0	0	0		
4/28/2011	GEN	0	0	0		0	0	0		2	2	0	100	100
4/28/2011	REL	0	0	0		0	0	0		0	0	0		

APPENDIX C:
Raw TLU analysis data for Belinda's article accuracy by topic

Email	Topic	singular definite				plural definite				singular indefinite				total
date	code	OC	CS	OS	%	OC	CS	OS	%	OC	CS	OS	%	%
3/2/2011	INT	1	0	0	0	0	0	0		1	0	0	0	0
3/2/2011	REL	0	0	0		0	0	0		0	0	0		
3/4/2011	GEN	0	0	0		0	0	0		0	0	0		
3/4/2011	EDU	0	0	0		0	0	0		2	2	0	100	100
3/4/2011	SPN	0	0	0		0	0	0		0	0	0		
3/4/2011	HOB	0	0	0		0	0	0		0	0	0		
3/4/2011	EDU	0	0	0		0	0	0		2	1	0	50	50
3/4/2011	REL	1	1	0	100	0	0	0		2	2	0	100	100
3/9/2011	MOV	12	12	0	100	3	3	0	100	4	3	0	75	94.74
3/16/2011	GEN	0	0	0		0	0	0		0	0	0		
3/16/2011	REL	1	0	0	0	0	0	0		0	0	0		0
3/16/2011	JAP	3	2	0	66.67	1	1	0	100	1	0	0	0	60
3/16/2011	MOV	0	0	0		0	0	0		0	0	0		
3/21/2011	MUS	2	1	0	50	0	0	0		5	4	1	66.67	62.50
3/21/2011	TXS	1	1	0	100	0	0	0		0	0	0		100
3/25/2011	TXS	1	1	0	100	0	0	0		2	2	0	100	100
3/25/2011	CHN	0	0	0		0	0	0		4	3	0	75	75
3/25/2011	FOR	1	1	0	100	0	0	0		2	2	0	100	100
3/25/2011	CHN	0	0	0		1	1	0	100	6	4	0	66.67	71.43
3/25/2011	MND	0	0	0		0	0	0		1	1	0	100	100
3/25/2011	PIC	0	0	0		0	0	0		1	0	0	0	0
3/30/2011	WTH	4	4	0	100	0	0	0		2	2	0	100	100
3/30/2011	LIF	3	0	1	0	0	0	0		3	2	0	66.67	28.57

3/30/2011	PIC	0	0	0		0	0	0		0	0	0		
3/30/2011	FES	1	1	0	100	1	1	0	100	1	1	0	100	100
4/2/2011	FES	17	10	0	58.82	2	1	0	50	8	5	0	62.50	59.26
4/2/2011	COL	1	1	0	100	0	0	0		0	0	0		100
4/2/2011	PIC	1	1	0	100	0	0	0		0	0	0		100
4/4/2011	COL	0	0	0		1	1	0	100	0	0	0		100
4/4/2011	TRP	5	5	0	100	1	1	0	100	2	1	0	50	87.50
4/4/2011	GAM	36	23	0	63.89	8	7	0	87.50	1	1	0	100	68.89
4/9/2011	PIC	0	0	0		1	1	0	100	2	2	0	100	100
4/9/2011	GAM	1	1	0	100	0	0	0		0	0	0		100
4/9/2011	LIF	0	0	0		0	0	0		2	2	0	100	100
4/18/2011	LIF	0	0	1	0	0	0	0		0	0	0		0
4/18/2011	REL	0	0	0		0	0	0		1	1	0	100	100