

---

## BRIEF REPORTS AND SUMMARIES

---

TESOL Quarterly invites readers to submit short reports and updates on their work. These summaries may address any areas of interest to Quarterly readers.

**Edited by LIA PLAKANS**

*University of Iowa*

**LAWRENCE JUN ZHANG**

*University of Auckland*

### *Creativity and EFL Students' Language Use During a Group Problem-Solving Task*

**KIM MCDONOUGH**

*Concordia University  
Montreal, Canada*

**WILLIAM J. CRAWFORD**

*Northern Arizona University  
Flagstaff, Arizona, United States*

**ALISON MACKEY**

*Georgetown University  
Washington, DC, United States,  
Lancaster University  
Lancaster, England*

*doi: 10.1002/tesq.211*

A long history of second language (L2) task research has explored how individual differences help account for variation in L2 learners' task performance. This research has investigated a wide range of cognitive and affective factors that may be implicated in task performance, such as working memory (Mackey, Adams, Stafford, & Winke, 2010), motivation (Dörnyei & Kormos, 2000), willingness to communicate (Cao & Philp, 2006), and anxiety (Sheen, 2008). These studies have provided evidence that individual factors affect diverse

aspects of L2 learners' task performance, such as quantity of speech, repetition of interlocutor's utterances, and repair and initiation behaviour. Situated within this framework, this exploratory study investigates whether creativity is another individual factor that provides insight into L2 learners' task performance. Adopting a corpus linguistics methodology, it examines whether creativity is related to variation in how L2 learners use language to fulfill different functions during task performance.

Creativity, which has its origins and a long history in cognitive psychology, is the ability to generate ideas, solutions, and plans that are new, novel, or unique when confronted with a specific problem (Feldhausen & Westby, 2003). Torrance (1969), a pioneer in creativity research and measurement, defined and consolidated the diverse processes associated with creativity into four components: fluency (production of ideas), flexibility (production of different ideational categories), originality (production of unusual ideas), and elaboration (persistence in providing details). This approach operationalizes creativity as a composite score that reflects quantitative measures of the number of ideas generated (fluency) and qualitative measures of the ideas' uniqueness and variability (flexibility, originality, and elaboration). Adapting and expanding on Guilford's (1968) tests of divergent thinking and problem solving, Torrance developed the verbal and figural versions of the Torrance Test of Creative Thinking (TTCT). These tests have been the subject of extensive psychometric research and are the most commonly used measures of creative ability (Althuizen, Wierenga, & Rossiter, 2010).

A potential role for creativity in L2 performance has been explored in two studies carried out in a Hungarian EFL context. Otto (1998) administered a Hungarian creativity test to secondary students taking a communicatively oriented English class. He hypothesized that creative students might be more successful in communicative language teaching (CLT) contexts due to CLT's emphasis on role plays and other simulation tasks. Although a positive correlation between creativity test scores and class grades was found, it is unclear how the students' class grades were calculated, making it difficult to determine the exact relationship between creativity and performance. Examining the relationship between creativity and students' narrative task performance more specifically, Albert and Kormos (2004) found moderate positive correlations between the fluency dimension of creativity and students' quantity of talk, as well as between originality and narrative clauses. Nonetheless, they reported a negative correlation between originality and quantity of talk, and no significant correlations between the creativity scores and linguistic measures of accuracy or complexity. They concluded that variation in creativity may help account for some

differences in students' task performance, although the relationships may be relatively weak.

Although these early studies have explored the possibility of a relationship between creativity and performance, further research is necessary in order to gain a clearer picture. The literature on creativity and L2 production is still nascent at this point, akin to how L2 researchers began to investigate working memory a decade or so ago, which raised questions about how to select and evaluate tests and how to best operationalize L2 use. With this in mind, the current study explored whether there was a relationship between EFL students' creativity and their performance on a group problem-solving task. Given the exploratory nature of the study and the early stage of L2 research into creativity, rather than select global linguistic measures *a priori*, we adopted a corpus linguistics approach in which the lexico-grammatical features of language that are relevant in a specific language use context emerge from the data (Biber, 1988; Biber & Conrad, 2009). This methodology was adopted in order to facilitate the identification of language features relevant for task accomplishment.

## METHOD

### Participants

The participants were 55 first-year Thai university students (19 men, 36 women) who were taking a class in English as a foreign language (EFL) focused on oral communication. They were all native speakers of Thai with a mean age of 18.5 years ( $SD = 0.5$ ), and all had studied English in primary and secondary school ( $M = 13$  years,  $SD = 1.7$ ). Only 10 participants reported prior visits to an English-speaking country, with stays ranging from 5 to 30 days; three participants had lived in the United States for one year. An *a priori* power analysis using the software package R revealed that a sample size of 67 was necessary for a study with a medium effect size and sufficient power to detect relationships. To increase power, alpha was adjusted to .10, which resulted in a required minimum sample size of 49, and power for the obtained sample was .83.

### Materials and Procedure

In order to assess the students' creativity while reducing demand on their L2 verbal abilities, versions A and B of the Figural TTCT (Torrance, 2008), both of which have three tasks, were administered. The picture construction task provides a common shape (e.g., jelly

bean or teardrop) that the participant incorporates into a picture or object. The picture completion task consists of ten incomplete figures that are used to sketch objects or pictures. The third task provides lines (Version A) or circles (Version B), which are used to create objects or pictures by either using each line or circle independently or incorporating multiple lines or circles into a single image. For each task, a name or title is provided for each image or object. The Norms-Technical manual (Torrance, 2008) reports a KR21 reliability coefficient of .89 for both forms (age 16+).

To elicit task interaction, a popular ESL group problem-solving task, *Shipwreck*, was adapted from an ESL resource book (Maggs & Hird, 2002). This type of problem-solving task was familiar to the students because their oral communication EFL class used similar communicative activities. The scenario for this task is that four passengers become stranded on an island when their boat is shipwrecked. Although there is a lifeboat, it can support only three people. The goal of the task is for the group to solve the problem by deciding which person does not get a seat in the lifeboat. To ensure that each student had sufficient information to carry out the task, four passenger cards were created (veterinarian, science researcher, retired marine, sailor) with each card stating specific character attributes including one flaw or weakness. Framed in terms described by Pica and colleagues (Pica, Holliday, Lewis, & Morgenthaler, 1989), the *Shipwreck* activity was a problem-solving task in which all group members held, requested, and conveyed information with an optional two-way flow of information and one convergent task outcome.

The EFL students completed the TTCT (35 min) followed by the *Shipwreck* task (15 min) during a regularly scheduled English class. Per the instructions in the TTCT manual, the students were given 10 min to complete each of the three activities in the test booklet. The instructions for each task were provided aurally and in writing in English and given aurally in Thai by the first researcher and the instructor. Students were given the option to write picture captions in either English or Thai; all chose to write in English. After completing the TTCT, the students formed groups of four through self-selection and were given a digital audio recorder and the character cards. One group of three was formed due to an odd number of students in the class, and the sailor card was randomly selected for removal from the set. Any questions that arose before or while the students were carrying out the *Shipwreck* task were answered by the first researcher or the instructor.

## Analysis

The students' figural TTCT booklets were returned to Scholastic Testing Services to be scored professionally. Reflecting Torrance's operationalization of creativity described in the introduction, standardized age-based scores of fluency (40–149), originality (40–154), elaboration (40–160), abstractness of titles (40–160), and resistance to premature closure (0–160) were summed and averaged. Creative strengths scores (0–16) were awarded based on thirteen categories that reveal qualitative characteristics of creative performance, such as emotional expressiveness, movement or action, humor, fantasy, and unusual visualization. A review of studies involving college students (Torrance, 2008) reported interrater reliability ranging from .91 to .99 for each subscale with the exception of resistance to premature closure (.78). In light of the recommendations for appropriate usages of TTCT scores (Torrance, 2008) and previous studies showing high correlations among the subscores (Clapham, 1998; Kim, 2006), the creativity index was used for the statistical analysis rather than considering each subscore separately.

The audio recordings were transcribed and verified, and analyzed using a corpus-based approach to identify lexico-grammatical features associated with language functions relevant for task performance (e.g., Biber, Conrad, & Reppen, 1998; McEnery, Xiao, & Tono, 2006). A small corpus (11,321 words) was formed and a word list was generated to identify the most frequent words. Words that occurred at least one time across groups and served relevant functions for *Shipwreck* task accomplishment were identified and grouped into six categories along with their associated linguistic forms: Interactivity: questions; e.g., *anyone want to say anything else?* Interactivity: subject/object *you* pronouns; e.g., *well because you know technically they're pretty much identical animals and humans so I can fix you if there's something wrong with you.* Reasoning: subordinate clauses with *because* or *so*; e.g., *the sailor have to be on the boat because he can help us find the dry land.* Reasoning: coordinated phrases, clauses, or sentences with *and*; e.g., *you guys suffer from extreme sea sicknesses and you guys already have minor symptoms.* Stance: verbs and modals including *think*, *can*, *can't*; e.g., *I can sail and this ability I think I can sail for you.* Predicting: conditionals; e.g., *I am the professional sailor so if all of you let me go on this boat surely that I can lead this boat back to the other island.*

Individual counts for each feature were summed and divided by total words to account for variation in the quantity of speech each student produced. Alpha was set at .10 for all statistical tests (one-tailed).

## RESULTS

The students' interaction while carrying out the *Shipwreck* problem-solving task lasted between 6 and 16 min, with a mean length of 12 min ( $SD = 3$ ). Each student produced a mean of 171 words, ranging from 39 to 613 words ( $SD = 117$ ). The students' creativity index scores ranged from 83 to 142, with a mean of 109.5 points ( $SD = 13.8$ ). The language features that the students produced while carrying out the *Shipwreck* task are summarized in Table 1 as raw frequency counts and proportions (feature/total words).

The goal of this exploratory study was to investigate the relationship between creativity and L2 language use during a group problem-solving task. As shown in Table 2, creativity had a significant, positive correlation with questions ( $r_s = .254, p = .031$ ) and coordination ( $r_s = .273, p = .022$ ). The 90% confidence intervals were fairly wide and the effect sizes were small. Creativity obtained significant probability values for two additional linguistic features, subordinate clauses ( $r_s = -.225, p = .050$ ), and conditionals ( $r_s = .208, p = .064$ ), but the confidence intervals spanned zero, which suggests that the relationships were not stable.

## DISCUSSION

The present study revealed a positive relationship between the EFL students' creativity and their production of questions and coordination, but there were no relationships with the other language features identified as useful for task accomplishment, namely, *you* pronouns, subordinate reasoning clauses, conditionals, and stance. The findings confirm those of previous studies in the Hungarian EFL context reporting links between creativity and class grades and features of narrative task performance (Albert & Kormos, 2004; Otto, 1998). Our data-driven approach to identify language features relevant for task

**TABLE 1**  
Descriptive Statistics for Language Features

Language features	Frequency counts		Proportions	
	Sum	Low-High	<i>M</i>	<i>SD</i>
Interactivity: Questions	55	0–17	.004	.009
Interactivity: <i>You</i> pronouns	347	0–43	.027	.029
Reasoning: Subordination	223	0–20	.026	.025
Reasoning: Coordination	155	0–10	.012	.010
Stance: <i>Think, can, can't</i>	376	0–39	.042	.025
Predicting: Conditionals	43	0–4	.004	.006

**TABLE 2**  
**Spearman Correlation Coefficients**

	Creativity index $r_s$	90% CI	$p$	Effect size ( $r_s^2$ )
Interactivity: Questions	.254	.02, .45	.031	.07
Interactivity: <i>You</i>	.081	-.16, .30	.279	.01
Reasoning: Subordinate clauses	-.225	-.44, .002	.050	.05
Reasoning: Coordination	.273	.06, .46	.022	.07
Conditionals	.208	-.32, .11	.064	.04
Stance	-.127	-.35, .10	.177	.02

accomplishment, as opposed to using global measures of accuracy or complexity, may account for divergence from Albert and Kormos's finding (2004) that creativity had no relationship to linguistic measures. However, the task demands of the *Shipwreck* task appear to require only a limited range of linguistic forms that have a relationship with creativity, suggesting that research with more diverse tasks is needed to determine if tasks that demand other language functions could profitably be investigated for a link to creativity. In addition, noninteractive tasks that elicit speech from a single speaker may also be investigated in terms of whether creativity accounts for differences in individual language use.

The positive findings for questions and coordination suggest that creativity may help account for variation in how English L2 learners interact with each other during problem-solving tasks (e.g., asking for information, coordinating information across turns and speakers, clarifying their interlocutors' utterances). An example of an interaction between two high-creativity students involving questions can be seen in the following dialogue:

Science researcher: How long can your little medicine—

Veterinarian: Last? How long will it last? Well depends if you guys keep getting sick then it won't last long but ... if we use it wisely you guys can survive for a long time

Science researcher: Really?

Veterinarian: Yeah I'll just you know keep you guys healthy because I'm a vet

Science researcher: Vet? It's a medicine for an animals?

Veterinarian: Well technically our bodies work pretty much identical so—

Science researcher: Really?

Veterinarian: So what?

Their interaction illustrates the use of questions to anticipate and complete interlocutors' turns by framing and then answering their own questions (*Last? How long will it last?*), express doubt (*Really?*) and disagreement (*Vet? It's a medicine for animals?*), and oppositional stances (*So what?*). The students' use of questions may be illustrative of a creative individual's ability to explore and redefine problems, as has been identified in previous creativity research (Boden, 2004). To determine which specific dimensions of creativity may account for L2 students' higher use of questions and coordination, further research is needed. Because the TTCT creativity index is a composite of six measures that are highly correlated, it would be inappropriate to consider each subscore as reflecting an independent component of creativity. Carrying out a detailed qualitative analysis of the students' interaction and eliciting students' perceptions about their task performance could clarify whether these language features can be attributed to specific dimensions of creativity (e.g., a question that reflects fluency versus a question that illustrates elaboration).

In terms of potential implications for task research, the findings suggest that individual learner factors may interact with task design features to influence students' language use during task performance. For example, task complexity research (e.g., Robinson, 2007) has found that tasks with reasoning demands elicit complex language, but our findings suggest creative students produce higher rates of coordination, as opposed to subordination, when providing reasons. Task research has shown that tasks with an optional exchange of information tend to elicit fewer confirmation and clarification questions in negotiation of meaning sequences (e.g., Gass, Mackey, & Ross-Feldman, 2005; Newton, 2013) than required information-exchange tasks. For the *Shipwreck* task, creativity was associated with higher rates of question use, even though the task had an optional exchange of information. As highlighted by Dörnyei (2009), complex relationships exist among learner characteristics and features of the learning environment, with creativity potentially being another individual factor that impacts L2 task performance and helps account for variation across L2 speakers.

In terms of limitations and avenues for future research, our definition and measurement of creativity reflect the trait approach associated with cognitive psychology, which is not without criticisms due to its conceptualization of creativity as a stable attribute. Contrasting views of creativity, such as Sternberg's (2006) investment theory, argue



that people choose to engage in creative behaviour(s) based on their perceptions of risk and reward in specific situations. In other words, people make choices about whether to be creative based on what is potentially gained or lost by diverging from more conventional behaviour. In such a framework, assessing creativity through a standardized test such as the TTCT does not capture the situation-specific factors that potentially impact behaviour. This is akin to Dörnyei's recent reconceptualization of individual differences such as motivation as dynamic rather than static (e.g., Dörnyei & Ushioda, 2009). An alternate approach for assessing creativity includes the use of questionnaires that address interests, attitudes, and past behaviours (e.g., Raudsepp, 1981), and are based on the assumption that interests and attitudes indicate creative potential while past behaviours predict future behaviour (Clapham, 2004).

Whereas our approach focuses on creativity as a predictor variable (see also Mackey, Park, Akiyama, & Pipes, 2014), recent trends in assessment, such as the APT model (Baer & Kaufman, 2005; Kaufman & Baer, 2006), consider creativity as an outcome variable. In this approach, task performance within a specific domain is evaluated in terms of its creativity, after which the minimal requirements necessary for people to achieve creative task performance are identified. In terms of operationalizing and evaluating creative language use, recent English L1 studies on spoken interaction (Carter, 2004) and text messaging (Tagg, 2013) suggest that creative language use contains playful language, including unique discourse and morphological features and punning. The corpus methodology employed here offers a promising approach for the identification of creative language use. A corpus-driven *bottom up* approach aims to identify frequently produced lexical and grammatical forms and then link those forms to specific language functions. Unlike a top-down approach based on *a priori* coding categories, a corpus-driven approach reveals existing patterns of L2 language use that researchers can then associate with differences in task, genre, creativity, or other individual learner characteristics. Both investigating creativity as an outcome variable (rather than a predictor variable) and adopting corpus-driven analyses have potential to provide further insights into the complex relationships among L2 learner characteristics, learning environments, and task performance.

## ACKNOWLEDGMENTS

This study was funded by an Insight Development grant from the Social Sciences and Humanities Research Council of Canada awarded to the first author. We would like to thank the research assistants who helped with data transcription and

coding: Sharon Crawford, Max Lavalley, Seonmin Park, Jieqiong Wang, Nicole Ziegler, and Laura Zurlo.

## THE AUTHORS

Kim McDonough is an associate professor and Canada Research Chair in applied linguistics in the Department of Education at Concordia University in Montreal, Canada. Her research interests include psycholinguistic approaches to second language acquisition, peer interaction, and task-based language teaching.

William J. Crawford is an associate professor in the Department of English at Northern Arizona University. His research interests are in corpus linguistics and second language acquisition with a specific interest in task-based language learning and second language writing.

Alison Mackey is a professor of linguistics at Georgetown University and Lancaster University. Her research interests are in second language acquisition, specifically the role of input, interaction, feedback, and cognitive factors in L2 learning. She also publishes in the area of second language research design and methodology.

## REFERENCES

- Albert, A., & Kormos, J. (2004). Creativity and narrative task performance: An exploratory study. *Language Learning, 54*, 277–310. doi:10.1111/j.1467-9922.2004.00256.x
- Althuizen, N., Wierenga, B., & Rossiter, J. (2010). The validity of two brief measures of creative ability. *Creativity Research Journal, 22*, 53–61. doi:10.1080/10400410903579577
- Baer, J., & Kaufman, J. C. (2005). Bridging generality and specificity: The amusement park theoretical (APT) model of creativity. *Roeper Review, 27*, 158–163. doi:10.1080/02783190509554310
- Biber, D. (1988). *Variation across speech and writing*. New York, NY: Cambridge University Press.
- Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge, England: Cambridge University Press.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge, England: Cambridge University Press.
- Boden, M. (2004). *The creative mind: Myths and mechanisms* (2nd ed.). London, England: Routledge.
- Cao, Y., & Philp, J. (2006). Interactional context and willingness to communicate: A comparison of behavior in whole class, group and dyadic interaction. *System, 34*, 480–493. doi:10.1016/j.system.2006.05.002
- Carter, R. (2004). *Language and creativity: The art of common talk*. London, England: Routledge.
- Clapham, M. (1998). Structure of figural Forms A and B of the Torrance Tests of Creative Thinking. *Educational and Psychological Measurement, 58*, 275–283. doi:10.1177/0013164498058002010
- Clapham, M. (2004). The convergent validity of the Torrance Tests of Creative Thinking and creativity interest inventories. *Educational and Psychological Measurement, 64*, 828–841. doi:10.1177/0013164404263883

- Dörnyei, Z. (2009). Individual differences: Interplay of learner characteristics and learning environment. *Language Learning*, 59(Suppl. 1), 230–248. doi:10.1111/j.1467-9922.2009.00542.x
- Dörnyei, Z., & Kormos, J. (2000). The role of individual and social variables in oral task performance. *Language Teaching Research*, 4, 275–300. doi:10.1191/136216800125096
- Dörnyei, Z., & Ushioda, E. (Eds.). (2009). *Motivation, language identity and the L2 self*. Bristol, England: Multilingual Matters.
- Feldhausen, J. F., & Westby, E. L. (2003). Creative and affective behavior: Cognition, personality, and motivation. In J. C. Houtz (Ed.), *The educational psychology of creativity* (pp. 95–105). Cresskill, NJ: Hampton Press.
- Gass, S., Mackey, A., & Ross-Feldman, L. (2005). Task-based interactions in classroom and laboratory settings. *Language Learning*, 55, 575–611. doi:10.1111/j.0023-8333.2005.00318.x
- Guilford, J. P. (1968). *Intelligence, creativity, and their educational implications*. San Diego, CA: Robert R. Knapp.
- Kaufman, J., & Baer, J. (2006). Intelligence testing with Torrance. *Creativity Research Journal*, 18, 99–102. doi:10.1207/s15326934crj1801\_11
- Kim, K. (2006). Is creativity unidimensional or multidimensional? Analyses of the Torrance Tests of Creative Thinking. *Creativity Research Journal*, 18, 251–259. doi:10.1207/s15326934crj1803\_2
- Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, 60, 501–533. doi:10.1111/j.1467-9922.2010.00565.x
- Mackey, A., Park, H., Akiyama, Y., & Pipes, A. (2014, March). *The role of cognitive creativity in L2 learning processes*. Paper presented at the Georgetown University Round Table, Washington, DC.
- Maggs, P., & Hird, J. (2002). *Timesaver speaking activities*. London, England: Mary Glasgow Magazines.
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. New York, NY: Routledge.
- Newton, J. (2013). Incidental vocabulary learning in classroom communication tasks. *Language Teaching Research*, 17, 164–187. doi:10.1177/1362168812460814
- Otto, I. (1998). The relationship between individual differences in learner creativity and language learning success. *TESOL Quarterly*, 32, 763–773. doi:10.2307/3588011
- Pica, T., Holliday, L., Lewis, N., & Morgenthaler, L. (1989). Comprehensible output as an outcome of the linguistic demands on the learner. *Studies in Second Language Acquisition*, 11, 63–90. doi:10.1017/S027226310000783X
- Raudsepp, E. (1981). *How creative are you?* New York, NY: Perigee.
- Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics*, 45, 193–213. doi:10.1515/IRAL.2007.009
- Sheen, Y. (2008). Recasts, language anxiety, modified output, and L2 learning. *Language Learning*, 58, 835–874. doi:10.1111/j.1467-9922.2008.00480.x
- Sternberg, R. J. (2006). The nature of creativity. *Creativity Research Journal*, 18, 87–98. doi:10.1207/s15326934crj1801\_10
- Tagg, C. (2013). Scraping the barrel with a shower of social misfits: Everyday creativity in text messaging. *Applied Linguistics*, 34, 480–500. doi:10.1093/applin/ams072

- Torrance, E. (1969). *Creativity: What research says to the teacher*. Washington, DC: National Education Association.
- Torrance, E. (2008). *Torrance Test of Creative Thinking: Norms-Technical manual*. Bensenville, IL: Scholastic Testing Services.

## ***Second Language Speakers at University: Longitudinal Development and Rater Behaviour***

**SARA KENNEDY, JENNIFER A. FOOTE, AND LARISSA KURTZ  
DOS SANTOS BUSS**

*Concordia University  
Montréal, Canada*

*doi: 10.1002/tesq.212*

University students studying in a second or additional language (L2) face many challenges. However, one potential benefit for these students could be the development of their L2 knowledge and skills, including the development of L2 speech. Researchers investigating effects of L2-medium instruction at university have mainly targeted study-abroad programs (short-term stays by international students) and content and language integrated learning (CLIL) programs for domestic and international students during all or most of the students' term of study. Recent findings on study-abroad students' L2 speech development have been mixed, with some students showing significant improvement in oral fluency measures and oral proficiency ratings (Segalowitz & Freed, 2004) but not in the pronunciation of individual sounds (Díaz-Campos, 2004). Only a few studies have explored L2 speech development in university-level CLIL contexts. In North America, undergraduate students who took L2 subject-matter courses together with adjunct language courses reported improved speaking ability over two semesters (Ready & Wesche, 1992) or showed significant improvement in at least one speaking task (Burger & Chrétien, 2001). In Spain, undergraduate translation students demonstrated significant L2 fluency gains after one term of study abroad, unlike in their two previous terms of formal instruction at home (Mora & Valls-Ferrer, 2012). Clearly, university students' longer term L2 speech development in L2 settings remains an underresearched area.

The L2 speech of university students in L2 settings is typically assessed before or upon admission, with formal assessments and trained raters (e.g., TOEFL). However, the L2 speech of university