3. Multilingual versus monolingual classroom practices in English for academic purposes: Learning outcomes, student attitudes, and instructor observations

Amanda Brown, Robert J. Lally & Laura Lisnyczyj

Introduction

There is no dispute that to learn English or any other new language effectively in both a second or foreign language context, one must receive input in, produce output in, and interact with others in the second language (L2). However, uncertainty exists around the amount of input, output, and interaction necessary and sufficient for successful second language acquisition (SLA). Perhaps to mitigate risk, most contemporary theories of language pedagogy advocate for maximal use of the L2 with minimal/no use of the first language (L1) (see discussion in e.g., Ellis & Shintani, 2014). Indeed, institutional policies and practices mandate L2-only use in both second and foreign language classrooms (Copland & Neokleous, 2011; Debreli, 2016; Jenkins, 2010; McMillan & Rivers, 2011; Rivers, 2011; Sampson, 2012; Shvidko, 2017). However, exclusive classroom use of the L2 has been labeled a "monolingual" view of pedagogy (Hall & Cook, 2012), and is drawing criticism. Recognizing this discord, Macaro (2014) has stated that the decision between using a fully 'monolingual', 'immersive', or 'English-only' pedagogy versus a 'multilingual', 'non-immersive', English plus L1/other language pedagogy, alternatively labeled 'translanguaging'

(e.g., Canagarajah, 2011; Creese & Blackledge, 2010; García, 2009; García & Wei, 2014; Hornberger & Link, 2012; Lewis et al., 2012a, b), 'translingualism' (Canagarajah, 2013; Horner et al., 2011), 'codeswitching', 'dynamic bilingualism' 'fluid languaging practices', and 'plurilingualism' (e.g., García & Kano, 2014; Piccardo, 2013), is "probably the most fundamental question facing second language acquisition (SLA) researchers, language teachers, and policymakers in this second decade of the 21st century" (p. 10). In this chapter, we address this fundamental question, contrasting a monolingual versus a multilingual pedagogy and empirically investigating the effects on learning outcomes, student attitudes, and instructor observations.

Background

The issue of language use in L2 classrooms has garnered considerable attention, especially in recent decades. Cook (2001) traces the belief in maximizing L2 use and minimizing or eliminating L1 use as far back as the 1880s. As a result, in contemporary research, language teachers have described a sense of guilt associated with L1 use (Macaro, 2009), especially when cautioned against overreliance on the L1 (e.g., Macaro, 2001; M. Turnbull, 2001). Learners may at least partially support an exclusive L2 policy (Shvidko, 2017), only tolerate a less than 10% use of L1 (Tang, 2002), and may even be reluctant to use the L1 at all (Storch & Wigglesworth, 2003). Indeed, use of the L1 for L2 learning may not be beneficial for learning outcomes, with negative correlations found between L1 use and text quality in L2 English writing (Weijen et al., 2009). Furthermore, institutional policies in both foreign and second language contexts may disfavor or even prohibit L1 use, often with severe penalties for non-compliance (Jenkins, 2010). Thus, teachers may fail to report or avoid L1 use in the L2 English classroom (Copland & Neokleous, 2011; Debreli, 2016; Sampson, 2012), and others may implement institutional policies inconsistently (Shvidko, 2017). As Sampson (2012) notes, professional practices around L1 use in the English language classroom have remained consistent, with change slow to nonexistent in recent decades. Precise descriptions of current institutional language policies are

hard to obtain, perhaps because administrators are aware of their inaction in this area, and empirical evidence demonstrating the prevalence of L2-only policies is therefore often anecdotal. Burton and Ranjendram (2019), for example, describe the absence of an English-only policy in the context of a Canadian ESL study, but the pervasiveness of "English-only discourses circulat[ing] among students, administrators, and instructors" (p. 28). And Shvidko's (2017) "informal online survey" revealed that 14 out of 23 administrators acknowledged an active prohibition on L1 use, and 23 out of 28 lamented student L1 use in their Intensive English Programs (IEPs) in English-speaking countries. Thus, negative attitudes towards and practices around L1 use in L2 classrooms are alive and well.

However, from a theoretical perspective, a growing number of researchers have called for a multilingual approach, including use of the L1, in L2 teaching and learning (e.g., Atkinson, 1987; Cook, 2001; Crump, 2013; Cummins, 2007; Hall & Cook, 2012; Littlewood & Yu, 2011; Piccardo, 2013; Sampson, 2012; Turnbull & Dailey-O'Cain, 2009). Empirical research in this area has primarily focused on describing English as a foreign language (EFL) contexts, where, as discussed by Shvidko (2017), learners typically share an L1. Such research has documented varied frequencies of L1 use for diverse classroom functions, for example, explanations of grammar and vocabulary, implementation of task, maintenance of classroom discipline, and development of interpersonal relationships (e.g., Franklin, 1990; Inbar-Lourie, 2010; Littlewood & Yu, 2011; Ma, 2019; Macaro, 1997; Polio & Duff, 1994; Sampson, 2012; Yu & Lee, 2014). The frequency of L1 use in the L2 classroom is argued to be moderated by the type of learning task, learner age, learner proficiency level, extent of learner engagement, and teacher language background (Azkarai & May, 2015; Burton & Rajendram, 2019; DiCamilla & Anton, 2012; Lee & Macaro, 2013; Moore, 2013; Storch & Aldosari, 2010; B. Turnbull, 2018). Evidence of positive attitudes towards L1 use in L2 classrooms has been found among teachers (e.g., Debreli, 2016; Kim, 2015; Kim & Petraki, 2009) and learners (Debreli & Oyman, 2015; Kim & Petraki, 2009; Leeming, 2011; Neokleous, 2016; Shvidko, 2017), though learner and teacher attitudes may be moderated by learners'

age (Macaro & Lee, 2013) and proficiency level (Burton & Rajendram, 2019; Debreli & Oyman, 2015).

In contrast to the vast majority of descriptive work, little experimental work has been conducted on use of the L1 in L2 English classrooms (Hall & Cook, 2012; Macaro, 2009), with recent calls for more intervention studies (Yu & Lee, 2014). Regarding effects on learning outcomes, in one study, use of a multilingual pedagogy improved learner perceptions of learning outcomes in EFL writing and listening; however, no control group was included for comparison (Adamson & Coulson, 2015). With inclusion of both an experimental and control group, Berning (2016) documented gains in EFL writing scores after a multilingual approach to teaching, although the treatment condition was limited to one 15-minute session. In a similar design, Arshad et al. (2015) found that use of the L1 facilitated the teaching of grammar for beginning level EFL students, but made no difference to those at higher levels. And in a series of studies, Macaro and colleagues have investigated the effects of a multilingual approach in the domain of EFL lexical development. In a study of L1 Chinese learners of EFL at various proficiency levels in their first year at university, Tian and Macaro (2012), for example, found that presentation of vocabulary using both L1 and L2 proved slightly more advantageous in tests of listening comprehension than presentation of vocabulary using L2 English only. Similar results were obtained by Zhao and Macaro (2016) in immediate and delayed post-tests of reading-based vocabulary knowledge (concrete and abstract words) among second year L1 Chinese learners of L2 English (non-majors). However, it was not clear to what extent these gains could be maintained in the long term, and Macaro et al. (2009) found no differences in learning outcomes from use of a multilingual versus monolingual approach.

While a little (quasi-)experimental research has been conducted in EFL contexts, where students generally share an L1 and exposure to the L2 outside the classroom may be limited, almost no research has been conducted in English as a Second Language (ESL) contexts in order to determine whether existing findings hold when students do not share an L1 and L2 exposure may be abundant outside the classroom. An earlier intervention study (Brown & Lally, 2019) preceding the current one involved one instructor and ESL learners at two proficiency levels, lower and

upper-intermediate, with one class at each level designated as the control groups, experiencing a fully immersive English-only teaching and learning environment, and the others as the treatment groups, experiencing a non-immersive environment with use of both English and other languages. The quantitative analysis of assignment scores found no statistical differences in learning outcomes as reflected in assignment grades between the control and treatment conditions. A further analysis of course evaluations found possible student preferences in some areas for the non-immersive classroom environment, though importantly course evaluations were standardized and did not include items specific to language policy and use.

The current study

A lack of consensus exists on whether and to what extent the L1 should be used in L2 English language classrooms. On the one hand, in professional practice, English-only policies and practices remain frequently mandated (Copland & Neokleous, 2011; Debreli, 2016; Jenkins, 2010; McMillan & Rivers, 2011; Rivers, 2011; Sampson, 2012; Shvidko, 2017), yielding potentially serious negative consequences for teachers around classroom L1 use (Jenkins, 2010). Support for an English-only pedagogy is derived from the clear need for L2 input, output and interaction in TESOL and a lack of research-supported guidelines indicating how much L2 is necessary and sufficient for optimal acquisition (Ellis & Shintani, 2014), along with findings showing some positive student attitudes towards L2-only use and the possibility of negative effects of L1 use on L2 learning outcomes (Shvidko, 2017; Tang, 2002; Weijen et al., 2009). On the other hand, recent theoretical work argues against a monolingual, English-only pedagogy, instead favoring a multilingual approach or English plus L1/other language pedagogy (e.g., Atkinson, 1987; Cook, 2001; Crump, 2013; Cummins, 2007; Hall & Cook, 2012; Littlewood & Yu, 2011; Piccardo, 2013; Sampson, 2012; Turnbull & Dailey-O'Cain, 2009). Empirical work supporting this second position is chiefly descriptive (Hall & Cook, 2012; Macaro, 2009), showing the existence and functions of L1 use and positive attitudes towards L1 use among students and teachers (e.g., Azkarai &

May, 2015; Burton & Rajendram, 2019; Debreli, 2016; Debreli & Oyman, 2015; DiCamilla & Anton, 2012; Kim, 2015; Kim & Petraki, 2009; Lee & Macaro, 2013; Leeming, 2011; Moore, 2013; Neokleous, 2016; Shvidko, 2017; Storch & Aldosari, 2010; B. Turnbull, 2018). Experimental work is severely limited overall (Yu & Lee, 2014) and especially so in ESL settings, generally with advantages or no effects found for a multilingual approach (Adamson & Coulson, 2015; Arshad et al., 2015; Berning, 2016; Macaro et al., 2009; Tian & Macaro, 2012; Zhao & Macaro, 2016). Thus, Macaro's (2014) question around the decision between a monolingual versus a multilingual pedagogy remains largely unresolved.

The current study constitutes part of a larger group of studies attempting to shed light on this important decision in second and foreign language teaching, across L2 proficiencies, and within and across students and teachers. This study expands on earlier research in several critical ways. First, we partially replicate a prior study (Brown & Lally, 2019) by focusing on new groups of ESL learners, implementing an experimental contrast between a monolingual versus multilingual pedagogy. Second, we generalize across instructors, testing whether effects of the intervention vary by teacher either in terms of learning outcomes or in teacher observations. Third, we focus on a mid-intermediate level of proficiency, where critical thinking is developed, to test whether the effects of intervention vary by proficiency level. And fourth, the inclusion of within and between-participant intervention-control design and a survey on classroom language practices facilitates direct examination of learner views towards the two approaches.

Method

Participants

Two researcher-practitioners and a total of 50 international undergraduate students drawn from four English for Academic Purposes (EAP) courses at a large university in the northeastern United States participated in this study. The researcher-practitioners rated the English proficiency level of the learners as B2 on the CEFR Global Scale. One researcher-practitioner taught two classes

that were highly diverse linguistically and culturally with a total of 13 L1s represented, and the other researcher-practitioner taught two classes that were considerably less diverse, with a majority of students from one country. All classes contained students from backgrounds traditionally considered EFL, e.g., China, as well as backgrounds that could be considered 'post-colonial ESL' contexts, e.g., India.

The researcher-practitioners themselves were also multilingual; one (simultaneously) bilingual in English and Italian, with CEFR-C1 (North et al., 2018) proficiency in Spanish, and the other a native speaker of English with a CEFR-B2 level of proficiency in Spanish. Although the dual role of researcher-practitioner introduces the potential for research bias, inclusion of the instructors as researchers was essential to (1) establish classroom environments including or excluding the use of non-target languages based on instructor professional judgements (see Macaro, 2009), and (2) contribute qualitative observational data in the form of instructor journals reflecting on their experiences. Participant demographic information is summarized in Table 1 below.

Procedures for data collection

Following definitions laid out in Ortega and Iberri-Shea (2005), this study was considered longitudinal, with data collected throughout a four-month period. All student participants were enrolled in CEFR-B2 level EAP classes focused on academic writing and critical thinking. The courses were taught using Communicative Language Teaching methodology with Focus-on-Form and writing-based activities that included considerable oral interaction in class. The study employed a between-within participant design with each class experiencing both the treatment and control conditions in a counterbalanced fashion. Thus, one course from each of the two instructors began with the treatment condition, while the other courses began with the control condition, and the conditions were switched midway through the study.

For the treatment condition, instructors drew on resources such as Celic and Seltzer (2013) and García et al. (2016) for instructional activities that facilitate multiple language use, which largely comprised those around pre-writing and project preparation

Table 1. Participant demographics, data, and analysis

	Researcher-Practitioner (RP) 1	RP1 Class 1	RP1 Class 2	Researcher-Practitioner (RP) 2	RP2 Class 1	RP2 Class 2
% (#) participating in study	100% (1/1)	41% (7/17)	61% (11/18)	100% (1/1)	94% (16/17)	89% (16/18)
% (#) participating in survey	NA	41% (7/17)	44% (8/18)	NA	71% (12/17)	89% (16/18)
English level	NS	CEFR-B2	CEFR-B2	NS	CEFR-B2	CEFR-B2
L1s represented	Italian English	Mandarin (6) Hindi (1)	Mandarin (10) Nepalese (1)	English	Arabic (2) Dutch (1) Hindi (1) Korean (1) Mandarin (7) Portuguese (2) Russian/Ukrainian bilingual (2)	Arabic (1) French (1) Hindi (1) Hungarian (1) Indonesian (2) Japanese/Mandarin bilingual (1) Mandarin (7) Spanish (2)
Condition	NA	Monolingual first	Multilingual first	NA	Monolingual first	Multilingual first
Data contributed	Reflective journal	Stage 1: Learning outcomes Stage 2: Attitude Survey	Stage 1: Learning outcomes Stage 2: Attitude Survey	Reflective journal	Stage 1: Learning outcomes Stage 2: Attitude Survey	Stage 1: Learning outcomes Stage 2: Attitude Survey

NA = not applicable; NS = native speaker

such as discussion of topics, analysis of assigned readings, and background research. Students were encouraged, whenever possible, to communicate with peers in their language of choice (see Moore, 2013; Macaro, 2009), whether that was their native/other languages or English. Statements requiring "English-only" were removed from course syllabi and supplementary materials. However, institutional polices required that grading be based on products submitted in English; thus, all final assignments were submitted in English (see Lee, 2016, for a discussion of the assessment of translingual writing). For the control condition, the standard "English-only" policy was implemented in all classes. This policy was present on course syllabi for classes that began with the control condition and was included in supplementary materials. The policy was enforced in the classroom through regular reminders to use English, although with no penalties for non-English language use.

The switch in conditions midway through the study was explicitly marked for students. Students transitioning from treatment to control conditions discussed their use of multiple languages up to that point and were challenged to continue in only English. In the reverse case, students were praised for their 'exclusive' use of English during class and challenged to flexibly switch languages henceforth. From that point, all procedures described above for the reverse condition were implemented across courses.

The between-within participant design with condition switch is desirable on ethical grounds as it does not withhold treatment from a control group, but it also facilitated a student survey on practices, policies, and attitudes towards classroom language use after students had experienced both pedagogical approaches. The survey was administered on paper during class in order to maximize participation, and thus was kept brief, comprising the following four questions:

- *Which language(s) did you use during your course?*
- *Was an English-only policy in effect at any time during your course?*
- *How do you feel about an English-only policy in your course?*

- *How do you feel about using your native or other languages during your course?*

Consent for research participation was sought at two periods. Early in the study, volunteer students (71% of students (50/70) from all four classes) consented for work generated as part of the course to be secondarily analyzed for research purposes. Participation varied considerably among classes as shown in Table 1, with the highest participation in the two courses taught by one researcher-practitioner, which coincidentally were more culturally and linguistically diverse. Later, 61% of students (43/70) from all four classes volunteered to complete the survey on attitudes towards classroom language use. The survey was completed during class, with those choosing not to participate given an alternative activity in the same packet in order to ensure confidentiality. Levels of participation by class are indicated in Table 1, and none of the students who declined to participate in the first stage elected to participate in the survey. All consent interviews and data collection were conducted by the non-instructor researcher in order to minimize the possibility of coercion.

Finally, both researcher-practitioners completed reflective journals after each lesson throughout the four-month period. Journals documented the pedagogical content of each lesson, teacher instructions for activities as they related to language use, the languages observed in use by students for those activities, and any additional researcher-practitioner remarks, including their perceptions of student attitudes towards language use and any similarities or differences noted between classes in activity completion. These journals served as field notes demonstrating fidelity-to-condition in the procedure as well as enabling the inclusion of instructor perspectives on the intervention.

Analysis

Quantitative analyses

Quantitative analyses of five main student writing assignments representing a variety of rhetorical patterns were conducted. These comprised (1) *a response paper*, in which students had to demonstrate their comprehension of a video and their ability to

think critically and reflectively on the video's presentation (2) *a source-based paper*, in which students employed and appropriately documented credible sources to support the main thesis of a paper, (3) *a genre analysis*, in which students explained findings drawn from analysis of an unfamiliar genre, (4) *an argument synthesis*, in which students developed an argument about the effectiveness of a visual public service announcement that presented a relevant social issue in a provocative manner, and (5) *a narrative paper*, in which students told a story in first or third person.

Scoring of assignments was based on demonstrated mastery of rhetorical modes as well as grammatical and lexical accuracy. The research-practitioners had freedom in how the standard course syllabus was implemented, and research-practitioner 1 (RP1) chose to formalize the drafting and revision process in the first half of the study, scoring each stage for some of the assignments. Thus, for RP1 a total of 12 data points comprised of ten discrete assignment scores, an interim course grade, and a final course grade for each student were available for analysis, where the final course grade was assembled from grades in final versions of each writing assignment, along with completion of non-scored blogs and class participation. Research-practitioner 2 (RP2) did not formally score drafts, but instead implemented individual oral presentations of the source-based and argument synthesis papers and written self-reflections around those oral presentations, which were scored. Thus, for RP2, a total of nine data points comprising five assignment scores, two presentation and self-reflection scores, an interim course grade, and a final course grade for each student were available for analysis, with the final course grade an assembly of grades in each writing assignment, grades in presentations and self-reflections, and class participation.

Given that scores generally reflected discrete, progress-based assignments, with each targeting a different rhetorical pattern, between-group statistical analyses focused on each score independently without inclusion of a repeated-measures analysis over time. As a result, no data points were excluded by the SPSS software even if an individual student missed one assignment, which maximized statistical power. Importantly, given the variation in assignment scoring by instructor, analyses were not conducted across instructors. Thus, one set of t-tests compared

scores for the two classes across conditions taught by one researcher-practitioner, and a separate set of t-tests compared scores for the remaining two classes across conditions taught by the other researcher-practitioner.

Qualitative analyses

Qualitative analyses were conducted on the student survey responses and the researcher-practitioner journals. Handwritten data initially collected on paper to maximize participation were typed to facilitate coding. For all data, a thematic analysis was conducted following applicable procedures in Nowell et al. (2017). Given the relatively open nature of the survey and practitioner journal prompts, inductive, data-driven coding was applied in an iterative fashion (Fereday & Muir-Cochrane, 2006). For the student surveys, a team-based, consensus-driven inductive analysis was conducted on responses to the latter two questions eliciting student attitudes towards English-only versus multilingual classroom language practices. For the researcher-practitioner journals, the author of an individual journal was not involved in the inductive thematic analysis in order to minimize bias.

Results

Analysis of learning outcomes

Several analyses of learner outcomes were conducted. Because of assignment variation across instructors, separate analyses were conducted for each of the instructor's two classes to determine differences between the pedagogical treatment and control groups. The study design was considered within-participant to the extent that all students experienced both treatment and control conditions and could comment on their experiences of and attitudes towards each. However, given that the assignments during their experience of each condition were different, meaningful statistical comparisons between participants could only be made. In the following, monolingual refers to the control condition, where instructor and students used English almost exclusively, whereas multilingual refers to the treatment condition, where the students were
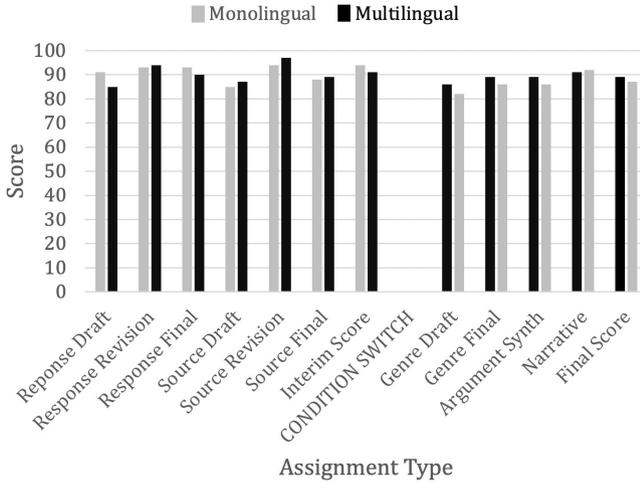
Figure 1. Mean student scores for research-practitioner 1 by course and condition

frequently and regularly encouraged to use their native or other languages as well as English.
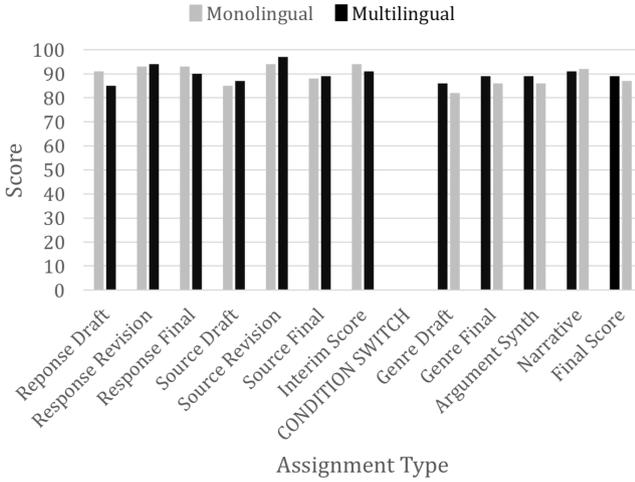
Figure 1 displays assignment scores for researcher-practitioner 1, who had the least cultural and linguistic diversity within courses. In this chart, the bars on the left for each assignment represent outcomes from one course, while the bars on the right represent outcomes from the second course. Note that since students in each class switched conditions midway through the study, the ordering of monolingual control and multilingual treatment changes after the interim score.

Overall, student scores were relatively high and at times close to ceiling levels. Throughout the study, the scores between classes were descriptively very comparable, with generally just a few percentage points between them. In terms of conditions, the students under the multilingual treatment condition descriptively outperformed those under the monolingual control condition at the majority of time points (8/12). However, twelve independent sample t-tests comparing scores under treatment and control conditions for each assignment revealed no statistically significant differences. The output of these analyses is shown in Table 2.

Table 2. Results of T-Test analysis by assignment for researcher-practitioner 1

Assignment	Monolingual First / Multilingual Second Mean (SD)	Multilingual First / Monolingual Second Mean (SD)	T-Test output
Response essay draft	91.00 (7.07)	84.58 (7.74)	t(16)= 6.42, p=.108
Response essay revision	93.40 (8.41)	94.42 (5.85)	t(5.1)= −16.58, p=.345
Response essay final	92.50 (6.72)	89.75 (6.06)	t(16)= 2.75, p=.394
Source-based essay draft	85.17 (11.41)	87.08 (9.06)	t(16)= −1.92, p=.702
Source-based essay revision	94.00 (7.70)	97.25 (1.71)	t(1.1)= −3.25, p=.632
Source-based essay final	88.00 (6.58)	88.63 (8.20)	t(16)= −.64, p=.865
Interim course grade	93.57 (5.79)	90.63 (4.85)	t(16)= 2.93, p=.273
MIDWAY CONDITION SWITCH			
Genre analysis draft	86.43 (6.29)	82.30 (8.68)	t(15)= 4.13, p=.301
Genre analysis final	89.14 (8.51)	85.80 (8.02)	t(15)= 3.34, p=.422
Argument synthesis final	89.00 (6.86)	85.60 (6.82)	t(15)= 3.40, p=.329
Narrative final	91.00 (3.46)	91.60 (4.45)	t(15)= −.60, p=.770
Final course grade	89.43 (7.18)	87.40 (10.07)	t(15)= 2.03, p=.655

* Note that not all assignments were completed by all students. Revisions of papers, in particular, were only completed by a few students. One student began but did not complete the course

Figure 2. Mean student scores for research-practitioner 2 by course and condition

Figure 2 displays assignment scores for researcher-practitioner 2, who had the most cultural and linguistic diversity across courses. As above, the bars on the left for each assignment represent one course, while the bars on the right represent the second course for this instructor, and the ordering of monolingual control and multilingual treatment switches after the interim score when the students in each class switched conditions.

Again, student scores were generally close to ceiling levels. The students under the multilingual treatment condition descriptively outperformed those under the monolingual control condition in three cases, while the reverse pattern was seen in three cases, and in three cases the mean scores were exactly the same. Nine independent sample t-tests comparing scores under treatment and control conditions revealed no statistically significant differences, with output provided in Table 3.

In summary, results from the quantitative analyses of assignment scores revealed that scores were high overall, at times close to ceiling. This is to be expected given that assignments were progress tests, measuring mastery of discrete rhetorical patterns.

Table 3. Results of T-Test analysis by assignment for researcher-practitioner 2

Assignment	Monolingual First / Multilingual Second Mean (SD)	Multilingual First / Monolingual Second Mean (SD)	T-Test output
Response essay final	85.93 (6.39)	87.00 (6.60)	t(27)= 2.60, p=.661
Source-based essay final	91.00 (6.84)	88.40 (5.18)	t(27)= −1.92, p=.257
Presentation 1	77.50 (4.83)	79.64 (9.09)	t(28)= −2.14, p=.418
Interim course grade	88.46 (5.13)	87.70 (5.36)	t(27)= .76, p=.698
MIDWAY CONDITION SWITCH			
Genre analysis final	90.56 (5.62)	87.44 (7.64)	t(30)= 3.13, p=.198
Argument synthesis final	89.25 (4.41)	90.63 (5.54)	t(30)= −1.38, p=.443
Narrative final	88.14 (5.63)	87.88 (5.52)	t(28)= .27, p=.896
Presentation 2	87.69 (9.27)	88.21 (4.64)	t(17.4)= −.52, p=.857
Final course grade	89.94 (5.03)	90.25 (5.50)	t(30)= −.31, p=.868

* Note that not all assignments were completed by all students

Descriptively, researcher-practitioner 1's students under the multilingual treatment condition generally outperformed those under the monolingual control condition, although results from researcher-practitioner 2's students were more variable. Crucially, these results cut across groups, with no individual class consistently surpassing another regardless of condition, suggesting a limited role for extraneous explanatory variables. Despite descriptive tendencies, no statistical differences between the conditions were found. Importantly, since repeated measures analyses were not conducted due to the differences in assignments throughout the

study, we cannot say at this point whether individual students performed better under one pedagogical approach versus the other.

Analysis of student attitudes

Quantitative and qualitative analyses of student attitudes towards classroom language practices elicited from a brief survey were conducted. A total of 61% of the total class population and 86% of students who volunteered consent during stage one to have their coursework secondarily analyzed for research purposes also volunteered consent to complete the survey (see Table 1 for a breakdown among classes).

In response to Question 1 on language use in the classroom, 61% of students reported using languages other than English. These included Arabic, Hindi, Indonesian, Mandarin, Russian, Spanish, Ukrainian, a selection which overlapped with the known L1s of participants. However, some students did not report using languages that in particular researcher-practitioner 2 had witnessed in use during the highly linguistically diverse classes. These omitted languages included Korean and Portuguese. Furthermore, 39% of students reported using only English at some points in the course. Five of the respondents reporting English-only use were students of researcher-practitioner 1, representing 33% of the participating students from those courses. Four of the five were in the multilingual first condition such that they were experiencing the monolingual condition when the survey was administered, which may have affected their responses. Twelve of the respondents reporting English-only use came from the courses of researcher-practitioner 2, representing 43% of the participating students from those courses. However, they were generally distributed across conditions, suggesting little or no relationship between when the survey was administered (at the end of the study) and the condition applied (e.g., monolingual English-only) at that time. The difference between responses in this area from the students of researcher-practitioner 1 versus 2, albeit relatively small, might be explained by the fact that researcher-practitioner 1 saw lower participation in the study overall and thus those that did participate might have been more highly engaged with more accurate

self-perceptions of language use. Overall, the observations from the researcher-practitioner journals, which documented use of languages other than English including among those individuals who reported using only English, revealed that some respondents were not entirely accurate in their self-perceptions and reporting of classroom language use, which is a weakness of self-reported data in general.

In response to Question 2 on language policy, 65% of survey respondents accurately recognized that an "English-only" policy was in effect at some point in their courses. These responses were drawn from all courses of both instructors. 28% reported that such a policy was not in effect and 7% declined to answer the question. Respondents failing to recognize the existence of an "English-only" policy at some point were distributed across multilingual and monolingual first conditions, suggesting that the presence of that policy on syllabi at the beginning of the course did not impact some. Two of the respondents failing to recall an English-only policy were students of researcher-practitioner 1, representing 13% of the participating students from those courses. Ten of the respondents recalling no English-only policy were students of researcher-practitioner 2, representing 36% of the participating students from those courses. The disparity between instructors here is larger than that in responses to questions of language use, and it is not clear what underlies the difference. It is conceivable that researcher-practitioner 1 was more explicit about the distinction between treatment and control conditions as they related to classroom policies than researcher-practitioner 2, rendering heightened student awareness of the language policy in those classes.

The remaining two survey questions were analyzed through a team-based, inductive, data-driven thematic analysis applied iteratively to the data. In response to Question 3 eliciting attitudes towards an "English-only" policy, a total of six themes were identified. These concerned the extent to which the policy was considered expected, helpful especially for facilitating conversation or thinking in the target language, necessary, not necessary or even unhelpful. These themes are illustrated below with quotes from respondents.

(1) An "English-only" policy is expected in an English course.
   – *I do not oppose it, as this class is an English class. It makes sense to have an English only policy.*
(2) An "English-only" policy is helpful in general for learning/ improving English.
   – *I like English only policy, it helps me on learning better.*
   – *It is difficult sometimes but good for us.*
   – *This policy requires us to speak English at any time so it is an effective way to improve speaking for international students.*
(3) An "English-only" policy is helpful for facilitating cross-cultural conversation.
   – *It is good as it helps people of different countries to converse.*
   – *It helps understand each other better.*
(4) An "English-only" policy is helpful for thinking in English.
   – *It's helpful for students to cultivate an English-based mindset meaning for students to think in English.*
   – *I think it helps me change my channel to an English mind. I do feel that I behave differently when I am using English mind.*
   – *It's a way to make us think about issues in English, which will help us to organize words effectively.*
(5) An "English-only" policy "is necessary to improve English, at least in some cases.
   – *Yes. It makes me think that how it has to be.*
   – *Depends on the rigor of the [English] course. Lower [English] course could allow other languages to help expressing. Since [this class] is the highest [English] class for first year international students, English only should be in effect.*
(6) An "English-only" policy is not necessary and in some cases not helpful.
   – *It is good but not necessary.*
   – *Not really helpful because sometimes we can only understand things in our native language.*

– *It encourages you to adapt or become comfortable with the language but using sources in your own language and properly translating it helps you understand English better and develop a fluidity between your mother language and English.*
– *I can not use English only to finish all my tasks.*
– *Sometimes, the English conversations between 2 non-native speakers could be painful and frustrating.*

The most commonly expressed views were that an English-only policy would help English language development, followed by an equal number expressing the necessity of an English-only policy or expressing the lack of necessity and possible hindrance of an English-only policy. Varied other views were expressed by a minority.

In response to Question 4 eliciting attitudes towards the use of languages other than English in the English language classroom, a total of five themes were identified. These overlapped in part with those above, namely the extent to which use of languages other than the L1 was helpful for L2 development, made learners feel comfortable, was beneficial for the expression and development of ideas, was not necessary and was detrimental if the L1 of one student was not shared by other students in the class. These themes are illustrated below with quotes from respondents.

(1) Allowing languages other than the L2, e.g., the L1, in class is helpful for learning/improving the L2.
   – *It's still useful to construct meaningful conversations that help to improve my English writing.*
(2) Use of languages other than the L2, e.g., the L1, makes students feel comfortable.
   – *If I use my native languages, I will be relax.*
   – *It would be nice to talk in my native language as it can reduce my homesickness.*
   – *It is a convenient way to discuss some of the tough questions.*
(3) Use of languages other than the L2, e.g., the L1, is beneficial for expression and development of ideas.

- *I think I could explain more clearly with peer/group mate that also from my country.*
- *Sometimes it is more effective to talk with classmates who came from the same country with me.*
- *That will be easier to understand what they really means, because the level of speaking English isn't always the same in one class.*
- *I could understand the topic better by discussing in our native language with other students.*
- *It helps me to transfer and get better performance.*
- *I believe that while this is an English course it is beneficial get international students to use their native tongue as they can express themselves better.*

(4) Use of languages other than the L2, e.g., the L1, is not a necessity or it is a hindrance.
- *Don't like it. It confuses me.*
- *I still prefer that these parts could be completed in English.*
- *To learn more in English I need to talk more in English.*
- *When in class, students should be discouraged to speak in their own language because that divides groups into whatever language they speak. Instead, we should all be brought together during class time and encouraged to confidently interact with one another in English.*
- *It make me feel uncomfortable because it's* [an English class].
- *It could help explain some terms more easily, however, it did not impact majorly in learning for me since there is only one other* [specific language] *speaker in the class.*

(5) Use of languages other than the L2, e.g., the L1, can be unfair.
- *Maybe unfair to some guys.*
- *It is a good idea if multiple people in class speak that language.*
- *I never used my native language in the class, as no one would understand.*

Student views on the fourth question were that the use of languages other than English can be beneficial for the expression and development of ideas, but also that such language practices were not a necessity or were even a hindrance for English language development. However, collapsing responses in the first three categories – generally helpful, promotes comfort, beneficial for expression – indicated that the majority of views expressed by participants on the issue of use of languages other than the target language were positive.

Analysis of research-practitioner journals

Both researcher-practitioners kept reflective journals for the duration of the study (four months). The journals documented the pedagogical content of each lesson. Activities employed for the multilingual treatment condition included regular (daily or weekly) small-group and paired discussions as well as individual work, where students gathered and discussed information about various topics in languages other than English and subsequently produced associated writing assignments in English. The same activities were employed for the monolingual control condition, but students were asked to use English exclusively for pre-writing activities and discussion.

Given their involvement in the research, researcher-practitioner journals also focused on areas relevant to the study such as teacher instructions as they related to language use, the languages observed in use by students, and any additional remarks including instructor perceptions of student attitudes towards language use and similarities or differences noted between conditions in activity completion. The journals were examined using an inductive, data-driven, thematic analysis that was applied iteratively to the data, and the author of a given journal was not involved in the initial analysis, though was involved post hoc in the form of member checking (see e.g., Nowell et al, 2017).

Research-practitioner 1

The journal of researcher-practitioner 1 was relatively long, at 8,785 words. Excluding commentary simply recounting pedagogical content, eleven categories were initially identified in the remainder of the text. These were eventually reduced to the following five

main themes, confirmed by member checking, accompanied by representative illustrative quotes marked by the day of the observation. Note that the switch between conditions was made at Day 18.

(1) Encouragement and amount of L1 use
The L1 was specifically encouraged by the instructor in the multilingual condition, especially early in the condition.

– *I encouraged them to speak in their language if possible. (Mu1st Day 3)*
– *They were encouraged to speak in L1 if they wanted/ could. (Mu1st Day 8)*

The introduction of the multilingual condition generated some visible reactions:

– *I encouraged them to speak in their L1. Many were shocked and pleased to hear this. (Mu1st Day 1)*
– *I then reminded them that they can speak their language – the room exploded with discussion, mostly in Chinese. (Mu1st Day 2)*
– *I started class with the announcement that students can now speak in their L1 during group work. They actually CHEERED! (Mu2nd Day 18)*

As the study progressed, no explicit encouragement was used and students in the multilingual condition still recruited multiple languages, i.e., by weeks three and four:

– *They were not told they could use their L1, but many did. (Mu1st Day 12)*

However, there were also observations about the L1 not being used, particularly noted when the conditions were switched in a class:

– *When they spoke in their L1, they were whispering. It was as if they were nervous they were doing something*

    wrong. I told them it was okay to speak loudly in their
    L1. (Mu2nd Day 18)
– *Most of the pairs were not speaking in their L1 at the*
    start; I had to remind them that they could. (Mu2nd
    Day 26)

In terms of amount of L1 use, this was not recorded on a daily basis. There were a few reports of L1 use by a minority especially at the beginning of the multilingual condition:

– *Only one group spoke in their L1 at first. (Mu1st Day 4)*
– *Similar to Monday, there was not much in the L1 –*
    they seem to still be getting used to being allowed to
    speak in their L1. (Mu2nd Day 19 – after switch to
    multilingual condition)

However, there were more reports of L1 use by a majority:

– *Half the class spoke in their language. (Mu1st Day 1)*
– *They were not encouraged to speak in L1, but most*
    did anyway. (Mu1st Day 9)
– *Students were speaking to each other in L1 to get*
    further help on the self-review sheet. (Mu2nd Day 30)
– *They were mostly speaking in L1 and code switching.*
    (Mu2nd Day 33)

(2) Mandate for and amount of L2 English use
The L2 was at times encouraged by the instructor in the monolingual condition but more often mandated as indicated by verbs such as *asked, stressed strongly, forced.* A considerable portion of the journal was devoted to comments such as the following:

– *I encouraged them to only speak English. (Mo1st Day 3)*
– *I asked them to talk to partners in English. (Mo1st*
    Day 2)
– *I stressed strongly to speak in English and only English*
    once they cross the threshold of the classroom. (Mo1st
    Day 1)

- *I insisted that they only speak in English with their partner (M01st Day 5)*
- *I forced them to speak English. (M01st Day 2)*
- *I started class with the announcement that students will now be required to speak in English during group work. (M02nd Day 18)*

In terms of how much L2 was spoken, conspicuously few journal comments documented abundant language production during the English-only policy:

- *There was decent amount of conversation in the L2... (M01st Day 2)*
- * They were told to speak in English. They all did this very well. (M01st Day 12)*

Indeed, in a number of cases, the mandate to use the L2 was observed to result in markedly little oral language production:

- *Talking was limited. (M01st Day 1)*
- *Two groups (all Chinese speakers) were completely silent while one partner typed the answers. ... There was very limited if any English spoken in these two groups. (M01st Day 7)*
- *There was limited to no conversation within the groups. In fact, in one group, the only person who spoke was the student that was strongest in English. The rest of them remained silent. (M02nd Day 19)*
- *I reminded them to use English at this point. Most students were silent. (M02nd Day 30)*
- *When they got into pairs to do the exercise, there was not much talking and most were whispering. (M02nd Day 33)*

As indicated in the final quote above, in contrast to the stated policy, use of the L1 was noted in the monolingual, English-only condition, with the instruction to switch to the L2:

- *I had to regulate and stop many people from using their L1. (Mo1st Day 5)*
- *Some were speaking in L1 and I had to remind and almost scold them (playfully). (Mo1st Day 6)*
- *As soon as I walked over to them, they switched to speaking English. (Mo1st Day 7)*

While some L2 use was always observed, some use of L1 continued well into the monolingual condition:

- *I encouraged English, but the speech still continued in L1. (Mo1st Day 10)*
- *I had to remind several groups to speak English only. (Mo1st Day 14)*
- *Students were mostly whispering, and I could hear L1 as well as code switching when they were speaking in English. (Mo2nd Day 25)*
- *They were shouting out the vocab word in the L1 translation to find their partner. (Mo2nd Day 33)*
- *It was difficult to hear, but it seems most were trying to speak in their L1. (Mo2nd Day 33)*

(3) Presence of codeswitching / active translanguaging
Language switching, labeled by the research-practitioner as codeswitching or active translanguaging, was witnessed in the multilingual condition:

- *Then there was an explosion of code switching – students were reading statements from the reading to each other in English and then commenting orally in L1. (Mu1st Day 4)*
- *Lots of discussion and code switching. The students were pointing to words or sentences, reading them, and then making oral comments in L1. Then they would write the comments in the margin or on the worksheet in English. (Mu1st Day 5)*
- *Many spoke in L1 even though their notes were in English. Lots of code switching! (Mu1st Day 8)*

    – *Brainstorming in L1 with lots of code switching.*
      (Mu2nd Day 33)
    – *They were trying to mimic the emotion while speaking*
      in L1 and then writing in English on the worksheet.
      (Mu2nd Day 33)

A couple of these observations were around the influence of technology in this process:

    – *When they clicked on the Wikipedia website example*
      given to them, the language on some of their
      computers actually changed on their computers to
      their L1. (Mu1st Day 7)
    – *I noticed one pair where a female Chinese student*
      was speaking an answer in Chinese, and simul-
      taneously, her male Chinese partner was typing the
      answer in English while listening to her. (Mu1st
      Day 7)

(4) Assessment of students
A number of observations were made regarding student performance. Some assessments were made with distinctions more apparent by class than by condition. Thus, the group that experienced the multilingual condition first were observed to be a strong group, with little mention of difficulties even after the condition switch to monolingual:

    – *The answers came quickly and were detailed/thorough.*
      (Mu1st Day 2)
    – *There are several students in that class that are strong*
      and speak quickly and freely. (Mu1st Day 3)
    – *Their answers were still very good and quickly given.*
      (Mu1st Day 8)

In contrast, the group that experienced the monolingual condition first were observed to be a weaker group, including after the condition switch to multilingual:

    – *Answers were also limited. (Mo1st Day 1)*

- *There was decent amount of conversation in the L2, but comments seemed superficial. (M01st Day 2)*
- *This group does not seem as strong as the previous section. (M01st Day 3)*
- *This was a difficult task for them. They spoke very intently in their L1 to try to figure out the answers. They also asked me questions to verify their answers and that they were on the right track during this partner activity. (Mu2nd Day 25)*

Though some of the difficulty appeared specifically to be related to the monolingual condition:

- *I had to remind them often to speak in English. This was difficult for them. It was obvious it was easier for them to speak in their L1. (M01st Day 16)*
- *They had to speak in English, but this was challenging for them. (M01st Day 17)*

And there were some positive general assessments of performance of the weaker group, some of which related to comfort level after condition switch.

- *They did this well. It seemed it was an easy task for them and there was lots of discussion in English. (M01st Day 8)*
- *Again, discussion and output were fine. (M01st Day 8)*
- *The groups discussed the questions well in their L1 – they are getting used to speaking in their L1 now. (Mu2nd Day 21)*

One area where assessment was by condition to some extent was in observations of speed of activity completion. The following comments all described the same class of students, but with perceived differences in speed of performance under the different conditions:

- *The responses came quickly once prompted… (Mu1st Day 1)*
- *Their answers were still very good and quickly given. (Mu1st Day 8)*
- *It took time for the students to begin their discussions. (Mo2nd Day 18 – after condition switch)*
- *This took much longer than expected. (Mo2nd Day 23 – after condition switch)*

(5) Pedagogical commentary

Pedagogical commentary described activities that characterized each condition: mono- or multilingual:

- *They had to take notes … and could then speak about it to a partner in English. (Mo1st Day 8)*
- *Their job was to teach the other members the information they learned/collected/researched … They were instructed to speak in English only. (Mo1st Day 11)*
- *They were placed in pairs … and were told to only submit one answer per pair. They had to discuss the validity of the article together in their L1 and then come up with one response. (Mu1st Day 7)*

Also included were specific instances where the researcher-practitioner did not or could not implement a multilingual pedagogy

- *There was no translanguaging event connected to this rewrite. (Mu1st Day 13)*
- *There was no translanguaging event today. (Mu2nd Day 27)*

And finally, there were comments about potential off-task behaviors exhibited by the students, which could not be confirmed by the research-practitioner. These were entirely related to use of languages other than English, for the most part during the multilingual condition for both classes, but in one case during the monolingual condition when L1 was used:

  – *It seemed some were talking off topic about other
     things (perhaps weekend events), but I can't confirm
     this. (Mu1st Day 4)*
  – *At first, it was evident that several groups were not on
     task (one group was "chit-chatting" in Chinese all at
     once). I had to instruct them to complete the task and
     hovered there until I was convinced they were on task.
     (Mu1st Day 12)*
  – *In their groups, when speaking in L1, some were
     goofing off and making it a joke that they could speak
     in their L1. (Mu2nd Day 19)*
  – *However, at times, I noticed the conversation was not
     related to the topic. I had to monitor carefully and
     remind them to work on the assignment. I feel that
     since we made the L1 switch, the students have been
     side chatting more. (Mu2nd Day 21)*
  – *Students were speaking in L1 while doing the self
     review sheet; however, it was difficult to know if it
     was about the review sheet or not since they were
     not referring to the sheet or the essay. (Mo2nd
     Day 30)*

Research-practitioner 2. The journal of researcher-practitioner
2 was shorter, at 3,702 words. Excluding commentary simply
recounting pedagogical content, twelve categories were initially
identified in the remainder of the text. These were reduced to
the following four main themes, confirmed by member checking,
accompanied by representative illustrative quotes marked by the
day of the observation. Note that the switch between conditions
for this instructor was also made on Day 18.

  (1) Encouragement, purpose, and amount of L1 use
  The L1 was specifically encouraged by the instructor in the
  multilingual condition:

  – *L1 use was encouraged for those who could use it.
     (Mu1st Day 11)*

– *They were encouraged to use L1 in their groups and choose from movies in their L1. (Mu1st Day 16)*

The researcher-practitioner noted significant L1 use across groups during the multilingual condition for a variety of purposes:

– *I did hear various L1s being used to clarify my explanations (Mu1st Day 5)*
– *During the quiz, a lot of L1 use was taking place when deciding on the final answer. (Mu1st Day 7)*
– *Lots of L1 used in groups to discuss accuracy of English sentences. (Mu1st Day 14)*

Evaluative comments were also provided on the nature of the L1 interaction across groups during the multilingual condition:

– *This grouping produced, for the 1st time, very clear and loud use of students' L1s (Mu1st Day 4)*

Like researcher-practitioner 1, researcher-practitioner 2 also noted cases where the L1 was not being used or used minimally, which was observed throughout the multilingual condition:

– *A reminder that L1 use was OK. Did hear some L1 use in groups but not predominantly. (Mu1st Day 2)*
– *3 groups were using a fair amount of English. There were some brief interactions in L1s between Spanish, Indonesian, & Chinese pairs in bigger group, but English was dominant. (Mu1st Day 10)*
– *They were encouraged to use L1 … None did though. They chose English speaking movies. (Mu1st Day 16)*
– *Portuguese, Ukrainian, Mandarin, & Arab speakers seem hesitant to utilize their L1[s] together. (Mu2nd Day 21)*

In one case, upon investigation, this was because of a difference in L1 dialects:

– *I didn't hear any Arabic. .. I … learned that their*

dialects were very different. As a result, they preferred
to communicate in English. (Mu2nd Day 30)

In other cases, a lack of L1 use was due to L1 'singletons", i.e.,
those that did not share an L1 with anyone else in the class
(see Table 1):
.. the singleton pairs used English exclusively. (Mu1st Day 7)

(2) Mandate for and amount of L2 use
Researcher-practitioner 2 also encouraged English-only use
in the monolingual condition, even until late in the condition,
though was perhaps less forceful than researcher-practitioner
1, at least as indicated by the verbs employed.

– *I moved around to groups … encouraging the use of*
   English. (Mo1st Day 2)
– *English was encouraged and used in groups. (Mo1st*
   Day 16)

Researcher-practitioner 2 appeared to see some success with
the English-only policy:

– *They were reminded to use only English, and I didn't*
   detect any other language use in the groups. (Mo1st
   Day 6)
– *Didn't hear/notice any language other than English.*
   This has been the case for some time. (Mo2nd Day 30)

Success with the English-only policy appeared to some extent
to be related to groupings:

– *Pairings included some Mandarin L1s together, so*
   there was a need to remind them to use English L2 at
   times. (Mo2nd Day 30)

(3) Assessment of students
A number of observations were made regarding student perfor-
mance, to some extent on the basis of condition. Observations
of good performance were made for the multilingual condition,

and in one case in contrast to the monolingual condition, though these were generally about one specific group (i.e., that experiencing the multilingual condition first):

–   *Quite similar to* [the other class] *but I noticed less questioning about what was to be done … The answers from this class, in some groups, tended to be richer in detail. (Mu1st Day 3)*
–   *2 groups were able to choose the correct or best summary / unlike the monolinguals. Two groups produced very accurate summaries (Mu1st Day 7)*
–   *Open class discussion produced desired answers. (Mu2nd Day 22)*

The superior performance of the group noted above was supported in the above:

–   *I am able to get through work and can count on this group being prepared more so than the first group. (Mo2nd Day 20)*

Overall, fewer observations of good performance were made of the monolingual condition across groups. Included here is one positive comment about students in the multilingual condition, who were unable to use their L1 as they were 'singletons':

–   *The auction was lively and for the most part their decisions were accurate. (Mo1st Day 14)*
–   *Same class but majority had their books & were able to give more in-depth responses. (Mo2nd Day 31)*
–   *Noticeably, the one group of singletons finished first and were fairly accurate. (Mu1st Day 6)*

And a couple of observations were made about comparable performance across conditions:

–   *An interesting development was that 2 pairs were very quick compared to the others. Interestingly, one was an L1 pair, 1 Mandarin & 1 Mandarin/Japanese, and*

> *the other pair was a pair of singletons, Hungarian &*
> *Hindi. Mu1st Day 8*
> – *Action & accuracy mirrored monolingual class (Mu1st*
> *Day 14)*

In one collection of observations, researcher-practitioner 2 noted poorer performance specifically in one group, but also in the monolingual condition:

> – *The accuracy of their answers could have been sharper.*
> *Some groups struggled to finish in the allotted time.*
> *(Mo1st Day 6)*
> – *I was surprised that no group chose the correct answer.*
> *(Mo1st Day 7)*
> – *Results were mixed and answers were not sufficiently*
> *in-depth. (Mo2nd Day 21)*

In contrast, there was only one negative comment about the other group during their simultaneous multilingual condition:

> – *Two of the* [language] *groups were rather slow and*
> *barely finished on time. (Mu1st Day 6)*

(4) Pedagogical commentary

Much of the pedagogical commentary from researcher-practitioner 2 revolved around the efforts needed to make groupings of students that were appropriate for the condition. For the multilingual condition, the following was noted:

> – *I need to be a bit more systematic about grouping L1s*
> *when possible. (Mu1st Day 2)*
> – *The activity here mirrored the Mono 1ˢᵗ class, but the*
> *difference was the attention given to the grouping of*
> *L1s (Mu1st Day 4)*
> – *Wasn't able to make groups with uniform L1s, but L1*
> *use was present at times. (Mu1st Day 13)*

– *Groups of 3 & 4, so not of L1 use. Later, I mixed up groups for comparison. (Mu2nd Day 26)*

For the monolingual condition, the following was noted with comparable strategies used across classes and time and special care taken for the number of Mandarin speakers who were in the majority in both classes despite the linguistic diversity of RP 2's classes:

– *All the groups had some diversity of L1. (Mo1st Day 9)*
– *Pairs consisted of one Chinese L1 in every pair to ensure use of English. (Mo2nd Day 19)*

Finally, the researcher-practitioner noted a logistical challenge, tardiness, that affected his ability to implement effective groups, which appeared to emerge in the class under the multilingual condition towards the end of the study:

– *Lateness in class is creating a problem for in-class translanguaging activities. I'm unable to form the groups I'd like to due to late arrivals. Need to make adjustments. (Mu2nd Day 20)*

Discussion

In this study, two US-based university-level ESL courses at an intermediate CEFR-B2 level were instructed by two researcher-practitioners with the standard monolingual, English-only pedagogy. Two comparable courses were taught by the same instructors using a multilingual approach, where students were given agency in their choice of language – L1 or otherwise – for in-class discussion, project preparation, and pre-writing activities. Quantitative analyses focused on learning outcomes, specifically the development of writing (paragraph and genre) and presentational speaking as well as student perceptions of classroom language practices. Qualitative analyses were conducted of responses to a survey eliciting student attitudes towards classroom

language use and instructor views in the form of a researcher-practitioner's reflective journal.

Quantitative results showed generally high scores for all assignments, which was not unexpected given their nature, i.e., progress tests based on mastery of discrete aspects taught in prior lessons, e.g., rhetorical patterns. Quantitative analyses showed a few descriptive differences, to some extent favoring the multilingual treatment, but no statistically significant differences between conditions throughout the study. This is striking for a number of reasons. First, with conditions switched throughout the study on groups of different sizes, findings suggest a lack of effect by student group regardless of size. In addition, the study was conducted across research-practitioners, indicating a lack of effect by instructor. Finally, an additional unplanned variable emerged in this study of intact classes – that of cultural and linguistic diversity. Coincidentally, the classes of researcher-practitioner 1 were relatively homogenous, dominated by Chinese speakers, and similar to the classes examined in Brown and Lally (2019). However, both classes of researcher-practitioner 2 were highly diverse linguistically and culturally, which was unusual and offered the opportunity to examine the effects of pedagogical treatment by level of diversity in student demographics. Again, no statistical differences in learning outcomes were observed. Thus, like Brown and Lally (2019), this study failed to find statistically robust differences in learning outcomes associated with the form and environment of instruction as it related to student language use (target versus L1/other) in the classroom at an intermediate proficiency level. Moreover, comparable results across diverse classroom contexts (teacher, student demographics) diminished the possibility of non-relevant (to this study) differences between groups as explanatory variables.

We recognize that there was some student L1 use in the monolingual condition that was documented in both researcher-practitioner reflective journals. This may demonstrate the impracticality and even impossibility of asking students to 'switch off' their L1, especially in light of psycholinguistic evidence demonstrating the inevitability of simultaneous activation of all known languages in the mind (e.g., Runnqvist et al., 2012). However, the journals also documented strong efforts by the instructors

clearly to differentiate the conditions by maintaining an English-only policy during the monolingual condition and promoting student use of different languages in the multilingual condition; therefore, we have evidence of fidelity-to-condition at least as it related to instructor behavior. Overall, the findings presented here are in line with studies summarized in Macaro et al. (2009), and to some extent Zhao and Macaro (2016) and Tian and Macaro (2012), who found at least no negative outcomes and even some slight advantages of a multilingual pedagogical approach.

In terms of student perceptions of language use, notable in the current study was that some students did not accurately report their language usage and were not cognizant of the conditions employed in their classes, despite explicit instructor instructions during both conditions documented in both researcher-practitioner journals. More than one third of students reported using only English throughout, which was accurate for the few singletons in the study, but inaccurate for most of the respondents, who were specifically observed by the researcher-practitioners using languages other than English. In addition, only two thirds reported the existence of an English-only policy, which was in effect at some point for all courses. It is difficult to say what underlies these discrepancies, but their existence does have methodological implications, reflecting the weaknesses of self-report data and the need for a triangulation of data, such as the instructor journals.

Regarding student preferences for language use, some expressed positive attitudes towards an English-only policy and against multilingual classroom language practices, especially at their intermediate level of proficiency (cf. Burton & Rajendram, 2019; Debreli & Oyman, 2015; Shvidko, 2017; Tang, 2002), and raised some potentially serious issues of student equity and polarization of a multilingual policy. However, the majority expressed generally positive attitudes towards multilingual classroom language practices, emphasizing benefits such as the general helpfulness of the L1, promotion of student comfort, and assistance with development of ideas and expression. Such positive views are in line with a number of other studies (e.g., Debreli & Oyman, 2015; Kim & Petraki, 2009; Neokleous, 2016; Shvidko, 2017),

especially in claims that use of the L1 can lessen cognitive load for learners (Macaro, 2005).

A significant amount of thematic overlap was observed in analyses of the researcher-practitioner journals, which signals the trustworthiness (Nowell et al., 2017) of the qualitative side of this mixed methods study. Both noted their efforts to encourage L1 use in the multilingual condition and to mandate L2 use in the monolingual condition, which as above, demonstrated fidelity-to-condition, as well as commenting on the amount of L1 and L2 use. They also provided some evaluative comments on student performance, which in some cases favored the multilingual condition. These comments revolved around the richness of discourse, accuracy and potentially speed of activity completion (see also Adamson & Coulson, 2015, for efficiency of task completion), though observations were somewhat mixed by language use, e.g., singletons using only English in the multilingual condition were also observed to perform quickly. The lack of clear distinctions in evaluations of performance by group or by condition observed by both research-practitioners, however, is supported by the lack of statistical differences found in the quantitative analysis of learning outcomes.

Although both researcher practitioners commented on various aspects of pedagogy, their focus was slightly different. Researcher-practitioner 1 described the types of activities that generated monolingual versus multilingual interactions and documented potential off-task behaviors. Researcher-practitioner 2 commented on the efforts taken to manage groupings to facilitate monolingual versus multilingual interactions (see Burton & Rajendram, 2019, for discussion of the impact on teacher attitudes of 'singletons' in the classroom) and to the logistical challenges, i.e., student tardiness, that complicated those efforts, which might have arisen due to the linguistic diversity of those groups.

Some limitations of this study offer additional opportunities for further research. Importantly, since repeated measures analyses could not be conducted, we cannot say whether individual students performed better under one pedagogical approach versus

the other or the effect that cultural and linguistic background has on this issue. Further, in line with CLT, the implementation of the control and treatment conditions was primarily through oral in-class tasks, while learning outcomes were operationalized primarily through written products. Thus, an open question regards whether the impact of condition would be more visible in oral products. All above areas warrant further investigation. In addition, both researcher-practitioners commented on the amount of L1 use, which, as noted above, may have compromised to some extent the monolingual control condition. Penalties for L1 use are employed in some contexts, but this would not have been in line with the pedagogical philosophy of either of the researcher-practitioners involved here and would likely not have changed the outcome given comments above on the impossibility of deactivating the L1. Instead, a detailed description of student language use based on recordings and associated transcripts could render language use a continuous rather than a nominal variable, facilitating correlations of amount of L1/L2 use by learning outcomes. Finally, the control and treatment conditions in this study revolved around student language use; there were no descriptions of non-English use by the teacher (see Burton & Rajendram, 2019, for a discussion of teacher translanguaging), and further research could examine this issue as it relates to learning outcomes especially in an ESL context with heterogenous learner L1s.

The multilingual pedagogy employed in this study is compatible with the theoretical framework in SLA of 'multi-competence' (Cook, 1992; Cook & Wei, 2016), which considers all the known languages within an individual mind as one system, as well as with the paradigm of 'translanguaging' (e.g., Canagarajah, 2011; Creese & Blackledge, 2010; García, 2009; García & Wei, 2014; Hornberger & Link, 2012; Lewis et al., 2012a, b), alternatively labeled as 'translingualism' (Canagarajah, 2013; Horner et al. 2011), 'codeswitching', 'dynamic bilingualism' 'fluid languaging practices', and 'plurilingualism' (e.g., García & Kano, 2014; Piccardo, 2013). The study has implications for language teaching and teacher training and is in line with several decades of

theoretical and descriptive work on L1 use in the L2 classroom and conclusions that an English-only mandate denies English "learners the opportunity of using an important tool" (Storch & Aldosari, 2010, p. 372).

Despite the burgeoning research literature, as Sampson (2012) has discussed, change in professional practices around L1 use in the English language classroom has been almost non-existent. In some professional contexts where institutional policies discourage or prohibit L1 use, particularly where English is taught as a foreign language, a multilingual pedagogy could be catastrophic for teachers, resulting in dismissal. In other professional contexts that have witnessed a growth in translingual practices, especially where English is taught as a second language, anecdotal and published sources still document the widespread perpetuation of English-only policies (García, 2009, Shvidko, 2017). Indeed, as seen here, at least some students still appear to prefer an English-only environment (see Storch & Wigglesworth, 2003, versus Berning, 2016, for differences in the use of L1 by learners and Shvidko, 2017, for multifarious student views).

To conclude, while understanding the potential consequences of a multilingual pedagogy in reductions of the amount of L2 input and interaction (Macaro, 2014), the results of this study support those from the larger group: four ESL classrooms examined in Brown and Lally (2019) and two French and two Arabic as a Foreign Language classrooms examined in Brown (2021). All together, these studies report either no difference or a facilitative effect of a multilingual pedagogy across 12 classrooms, in second and foreign languages, and across different learners, proficiencies, and instructors. Such parallel findings suggest at the very least no cost of a multilingual pedagogy in terms of learning outcomes as measured by assignment scores, a possible gain in terms of student attitudes, and a possible gain in terms of instructor perceptions, all findings which cast serious doubt on arguments for immersive L2 pedagogical practices. Further research is needed to determine whether and how language use in language classrooms may be adapted to optimize benefits, and all stakeholders in the learning process should feel reassured that such future research will likely not negatively impact learning.

References

Adamson, J., & Coulson, D. (2015). Translanguaging in English academic writing preparation. *International Journal of Pedagogies and Learning, 10*, 24–37. DOI: https://doi.org/10.1080/22040552.2015.1084674

Arshad, Z., Abdolrahimpour, M., & Najafi, M. R. (2015). The use of L1 as a consciousness-raising tool in teaching grammar to beginner and upper-intermediate EFL students. *Journal of Language Teaching and Research, 6*, 633–638. DOI: https://doi.org/10.17507/jltr.0603.21

Atkinson, D. (1987). The mother tongue in the classroom: A neglected resource?'. *ELT Journal, 41*, 241–247. DOI: https://doi.org/10.1093/elt/41.4.241

Azkarai, A., & May, M. P. G. (2015). Task-modality and L1 use in EFL oral interaction. *Language Teaching Research, 19*, 550–571. DOI: https://doi.org/10.1177/1362168814541717

Berning, B. S. (2016). *Can't I speak Japanese? The use of students' L1 as a cognitive tool for collaborative writing tasks in a Japanese EFL context*. Hamline University School of Education Student Capstones and Dissertations. Paper 4114.

Brown, A. (2021). Monolingual versus multilingual FLT: French and Arabic at beginning levels. *Language Teaching Research. Online FirstView:* DOI: https://doi.org/10.1177/1362168821990347

Brown, A., & Lally, R. (2019). Immersive versus non-immersive approaches to TESOL: A classroom-based intervention study. *TESOL Quarterly, 53*, 603–629. DOI: https://doi.org/10.1002/tesq.499

Burton, J., & Rajendram, S. (2019). Translanguaging-as-resource: University ESL instructors' language orientations and attitudes towards translanguaging. *TESL Canada Journal, 36*, 21–47. DOI: https://doi.org/10.18806/tesl.v36i1.1301

Canagarajah, A. S. (2011). Codemeshing in academic writing: Identifying teachable strategies of translanguaging. *Modern Language Journal, 95*, 401–417. DOI: https://doi.org/10.1111/j.1540-4781.2011.01207.x

Canagarajah, A. S. (2013). *Translingual practice: Global Englishes and cosmopolitan relations*. Routledge.

Celic, C., & Seltzer, K. (2013). *Translanguaging: A CUNY-NYSIEB Guide for Educators (print guide)*. CUNY-NYSIEB.

Cook, V. J. (1992). Evidence for multicompetence. *Language Learning, 42,* 557–591. DOI: https://doi.org/10.1111/j.1467-1770.1992.tb01044.x

Cook, V. (2001). Using the first language in the classroom. *Canadian Modern Language Review, 57,* 402–423. DOI: https://doi.org/10.3138/cmlr.57.3.402

Cook, V., & Wei, L. (2016). *Cambridge Handbook of Linguistic Multi-competence*. Cambridge University Press.

Copland, F., & Neokleous, G. (2011). L1 to teach L2: Complexities and contradictions. *ELT Journal, 65,* 270–280. DOI: https://doi.org/10.1093/elt/ccq047

Creese, A., & Blackledge, A. (2010). Translanguaging in the bilingual classroom: A pedagogy for learning and teaching? *The Modern Language Journal, 94,* 103–115. DOI: https://doi.org/10.1111/j.1540-4781.2009.00986.x

Crump, A. (2013). Fostering multilingual spaces in second and foreign language classes: Practical suggestions. *The Journal of Language Teaching and Learning, 3,* 65–71.

Cummins, J. (2007). Rethinking monolingual instructional strategies in multilingual classrooms. *Canadian Journal of Applied Linguistics, 10,* 221–240.

Debreli, E. (2016). Perceptions of non-native EFL teachers' on L1 use in L2 classrooms: Implications for language program development. *English Language Teaching, 9,* 24–32. DOI: https://doi.org/10.5539/elt.v9n3p24

Debreli, E., & Oyman, N. (2015). Students' preferences on the use of mother tongue in English as a foreign language classrooms: Is it the time to re-examine English-only policies? *English Language Teaching, 9,* 148–162. DOI: https://doi.org/10.5539/elt.v9n1p148

DiCamilla, F. J., & Anton, M. (2012). Functions of L1 in the collaborative interaction of beginning and advanced second language learners. *International Journal of Applied Linguistics, 22,* 166–188. DOI: https://doi.org/10.1111/j.1473-4192.2011.00302.x

Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. Routledge.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Research, 5*, 80–92.

Franklin, C. E. M. (1990). Teaching in the target language. *Language Learning Journal, 2*, 20–24. DOI: https://doi.org/10.1080/09571739085200371

García, O. (2009). *Bilingual Education in the 21st Century: A Global Perspective*. Wiley-Blackwell.

García, O., Johnson, S., & Seltzer, K. (2016). *The Translanguaging Classroom: Leveraging Student Bilingualism for Learning*. Caslon.

García, O., & Kano, N. (2014). Translanguaging as process and pedagogy: Developing the English writing of Japanese students in the US. In J. Conteh & G. Meier (Eds.), *The Multilingual turn in languages education: Opportunities and challenges* (pp. 258–277). Multilingual Matters.

García, O., & Wei, L. (2014). *Translanguaging: Language, bilingualism, and education*. Palgrave Macmillan.

Hall, G., & Cook, G. (2012). Own-language use in language teaching and learning. *Language Teaching and Learning, 45*, 271–308. DOI: https://doi.org/10.1017/S0261444812000067

Hornberger, N., & Link, H. (2012) Translanguaging and transnational literacies in multilingual classrooms: A bilingual lens. *International Journal of Bilingual Education and Bilingualism, 15*, 261–278. DOI: https://doi.org/10.1080/13670050.2012.658016

Horner, B., Lu, M., Royster, J. J., & Trimbur, J. (2011). Language difference in writing: Toward a translingual approach. *College English, 73*, 303–321.

Inbar-Lourie, O. (2010). English only? The linguistic choices of teachers of young EFL learners. *International Journal of Bilingualism, 14*, 351–367. DOI: https://doi.org/10.1177/1367006910367849

Jenkins, S. (2010). Monolingualism: An uncongenial policy for Saudi Arabia's low-level learners. *ELT Journal, 64*, 459–461. DOI: https://doi.org/10.1093/elt/ccq014

Kim, H. J. (2015). The use and perception of codeswitching among teachers and students. *The Journal of Language Teaching and Learning, 5*, 34–51.

Kim, Y., & Petraki, E. (2009). Students' and teachers' use of and attitudes to L1 in the EFL classroom. *Asian EFL Journal, 11*, 58–89.

Lee, J-H., & Macaro, E. (2013). Investigating age in the use of L1 or English-only instruction: Vocabulary acquisition by Korean ESL learners. *The Modern Language Journal, 97*, 887–901. DOI: https://doi.org/10.1111/j.1540-4781.2013.12044.x

Lee, J. W. (2016). Beyond translingual writing. *College English*, 79, 174–195.

Leeming, P. (2011). Japanese high school students' use of L1 during pair-work. *International Journal of Applied Linguistics, 21*, 360–382. DOI: https://doi.org/10.1111/j.1473-4192.2011.00284.x

Lewis, G., Jones, B., & Baker, C. (2012a). Translanguaging: Developing its conceptualisation and contextualisation. *Educational Research and Evaluation: An International Journal on Theory and Practice, 18*, 655–670. DOI: https://doi.org/10.1080/13803611.2012.718490

Lewis, G., Jones, B., & Baker, C. (2012b). Translanguaging: Origins and development from school to street and beyond. *Educational Research and Evaluation, 18*, 641–654. DOI: https://doi.org/10.1080/13803611.2012.718488

Littlewood, W., & Yu, B. (2011). First language and target language in the foreign language classroom. *Language Teaching, 44*, 64–77. DOI: https://doi.org/10.1017/S0261444809990310

Ma, L. P. F. (2019). Examining the functions of L1 use through teacher and student interactions in an adult migrant English classroom. *International Journal of bilingual Education and Bilingualism, 22*, 386–401. DOI: https://doi.org/10.1080/13670050.2016.1257562

Macaro, E. (1997). *Target language, collaborative learning and autonomy*. Multilingual Matters.

Macaro, E. (2001). Analysing student teachers' codeswitching in foreign language classrooms: Theories and decision making. *The Modern Language Journal, 85*, 531–548.

Macaro, E. (2005). Codeswitching in the L2 classroom: A communication and learning strategy. In E. Llurda (Ed.), *Non-native language teachers: Perceptions, challenges, and contributions to the profession (*pp. 63–84). Springer.

Macaro, E. (2009). Teacher use of codeswitching in the second language classroom. In M. Turnbull & J. Dailey-O'Cain (Eds.), *First language use in second and foreign language learning* (pp. 35–49). Multilingual Matters.

Macaro, E (2014). Overview: Where should we be going with classroom codeswitching research? In R. Barnard & J. McLellan (Eds.), *Codeswitching in university English-medium classes* (pp. 10–23). Multilingual Matters.

Macaro, E., Guo, T., Chen, H., & Tian, L. (2009). Can differential processing of L2 vocabulary inform the debate on teacher codeswitching behaviour: The case of Chinese learners of English. In B. Richards, M. Daller, D. Malvern, P. Meara, J. Milton & J. Treffers-Daller (Eds.), *Vocabulary studies in first and second language acquisition: The interface between theory and application* (pp. 125–146). Palgrave Macmillan.

Macaro, E., & Lee, J. H. (2013). Teacher language background, codeswitching, and English-only instruction: Does age make a difference to learners' attitudes? *TESOL Quarterly, 47*, 717–742. DOI: https://doi.org/10.1002/tesq.74

McMillan, B. A., & Rivers, D. J. (2011). The practice of policy: Teacher attitudes toward "English only." *System, 39*, 251–263. DOI: https://doi.org/10.1016/j.system.2011.04.011

Moore, P. J. (2013). An emergent perspective on the use of the first language in the English-as-a-foreign-language classroom. *The Modern Language Journal, 97*, 239–253. DOI: https://doi.org/10.1111/j.1540-4781.2013.01429.x

Neokleous, G. (2016). Closing the gap: Student attitudes toward first language use in monolingual EFL classrooms. *TESOL Journal, 8*, 314–341. DOI: https://doi.org/10.1002/tesj.272

North, B., Goodier, T., & Piccardo, E. (2018). *Common European Framework of Reference for Languages: Companion volume with new descriptors*. Council of Europe.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Research, 16*, 1–13. DOI: https://doi.org/10.1177/1609406917733847

Ortega, L., & Iberri-Shea, G. (2005). Longitudinal research in second language acquisition: Recent trends and future directions. *Annual Review of Applied Linguistics, 25*, 26–45. DOI: https://doi.org/10.1017/S0267190505000024

Piccardo, E. (2013). Plurilingualism and curriculum design: Toward a synergic vision. *TESOL Quarterly, 47*, 600–614. DOI: https://doi.org/10.1002/tesq.110

Polio, C. P., & Duff, P. A. (1994). Teachers' language use in university foreign language classrooms: A qualitative analysis of English and target language alternation. *The Modern Language Journal, 78*, 311–326. DOI: https://doi.org/10.1111/j.1540-4781.1994.tb02045.x

Rivers, D. J. (2011). Politics without pedagogy: Questioning linguistic exclusion. *ELT Journal, 65*, 103–113. DOI: https://doi.org/10.1093/elt/ccq044

Runnqvist, E., FitzPatrick, I., Strijkers, K., & Costa, A. (2012). An appraisal of the bilingual language production system: Quantitatively or qualitatively different from monolinguals? In T. Bhatia & W. C. Ritchie (Eds.), *The handbook of bilingualism and multilingualism* (pp. 244–266). Wiley-Blackwell.

Sampson, A. (2012). Learner code-switching versus English only. *ELT Journal, 3*, 1–11. DOI: https://doi.org/10.1093/elt/ccr067

Shvidko, E. (2017). Learner attitudes toward "English-only" institutional policies: Language use outside the classroom. *TESOL Canada Journal, 34*, 25–48. DOI: https://doi.org/1018806/tesl.v34i2.1261

Storch, N., & Aldosari, A. (2010). Learners' use of first language (Arabic) in pair work in an EFL class. *Language Teaching Research, 14*, 355–375. DOI: https://doi.org/10.1177/1362168810375362

Storch, N., & Wigglesworth, G. (2003). Is there a role for the use of L1 in an L2 setting? *TESOL Quarterly, 37*, 760–770. DOI: https://doi.org/10.2307/3588224

Tang, J. (2002). Using L1 in the English classroom. *English Teaching Forum, 40*, 36–44.

Tian, L., & Macaro, E. (2012). Comparing the effect of teacher codeswitching with English-only explanations on the vocabulary acquisition of Chinese university students: A Lexical Focus-on-Form study. *Language Teaching Research, 16*, 361–385. DOI: https://doi.org/10.1177/1362168812436909

Turnbull, B. (2018). The potential impact of cultural and educational background on foreign language teachers' use of the L1. *The Journal of Language Teaching and Learning, 8*, 53–70. DOI: https://doi.org/10.1111/0026-7902.00018

Turnbull, M. (2001). There is a role for the L1 in second and foreign language teaching, but... *The Canadian Modern Language Review, 57*, 531–540. DOI: https://doi.org/10.3138/cmlr.57.4.531

Turnbull, M., & Dailey-O'Cain, J. (2009). *First language use in second and foreign language learning* (Vol. 44). Multilingual Matters.

Weijen, D.V., Bergh, H.V., Rijlaarsdam, G.C., & Sanders, T. (2009). L1 use during L2 writing: an empirical study of a complex phenomenon. *Journal of Second Language Writing, 18*, 235–250. DOI: https://doi.org/10.1016/j.jslw.2009.06.003

Yu, S., & Lee, I. (2014). An analysis of Chinese EFL students' use of first and second language in peer feedback of L2 writing. *System, 47*, 28–38. DOI: https://doi.org/10.1016/j.system.2014.08.007

Zhao, T., & Macaro, E. (2016). What works better for the learning of concrete and abstract words: Teachers' L1 use or L2-only explanations? *International Journal of Applied Linguistics, 26*, 75–98. DOI: https://doi.org/10.1111/ijal.12080