Language Learning & Technology
ISSN 1094-3501

October 2017, Volume 21, Issue 3
pp. 76–103

ARTICLE

The effect of corpus-based instruction on pragmatic routines

Kathleen Bardovi-Harlig, *Indiana University*

Sabrina Mossman, *Indiana University*

Yunwen Su, *Indiana University*

Abstract

This study compares the effect of using corpus-based materials and activities for the instruction of pragmatic routines under two conditions: implementing direct corpus searches by learners during classroom instruction and working with teacher-developed corpus-based materials. The outcome is compared to a repeated-test control group. Pragmatic routines used for agreement, disagreement, and clarification in academic English discussion are targeted. 54 students in seven intact communication classes participated. 43 students received instruction in four 50-minute lessons across two to three weeks. Input came from MICASE (Simpson, Briggs, Ovens, & Swales, 2002) with noticing and production activities. The corpus-materials group (N = 26) received corpus excerpts and the corpus-search group (N = 17) conducted equivalent searches. The pre- and post-tests were administered through a computer-delivered oral-production task that simulated group discussion and included 30 items: 10 agreement, 10 disagreement, and 10 clarification scenarios. The results showed that both corpus searches and the use of corpus excerpts led to a significant increase in the oral production of pragmatic routines. The corpus-materials group additionally showed an increase in the clarity of speech acts. The corpus-search group reported engagement in self-directed searches outside the classroom, captured by a post-test questionnaire.

Keywords: *Instructional Pragmatics, Instructional Effects, Speech Acts, Corpus Searches*

Language(s) Learned in this Study: *English*

APA Citation: Bardovi-Harlig, K., Mossman, S., & Su, Y. (2017). The effect of corpus-based instruction on pragmatic routines. *Language Learning & Technology*, *21*(3), 76–103. Retrieved from http://llt.msu.edu/issues/october2017/bardovi-harligmossmansu.pdf

Introduction

One of the challenges of second language acquisition is learning the pragmatics of the target language. Pragmatics encompasses the knowledge of how to say what to whom in what contexts (Bardovi-Harlig, 2013). This includes both speech acts and pragmatic routines. Speech acts, as their name implies, are acts accomplished by words, such as apologies, requests, promises, and agreements and disagreements. Pragmatic routines, one type of formulaic language, allow speakers to indicate the intended illocutionary force of their utterance and help interlocutors interpret speaker intention by identifying the speech act. This is especially important in multiparty conversations such as those found in academic group work. This instructional effect study uses an academic corpus to provide authentic input for pragmatic routines found in academic group work—namely, agreement routines (*That's right*, *You're right*, and *That's true*), disagreement routines (*Yeah but*, and *I agree... but*), and clarification routines (*What do you mean* and *You're saying* for other-clarification, and *What I mean* for self-clarification).

Even learners with high grammatical competence show variable mastery of pragmatics. One interpretation of high grammatical competence and variable (low to high) pragmatic competence among advanced learners is suggested by the noticing hypothesis, paraphrased as "what learners notice in input is what

becomes intake for learning" (Schmidt, 1995, p. 20). In order for learners to notice a pragmatic feature, there must be target language input. Like foreign language learners, second language learners in host environments may also lack relevant input. 30 years of comparisons of language textbooks and conversations suggest that input for the acquisition of pragmatics is unlikely to come from standard second or foreign language textbooks. The portrayal of pragmatics is either absent or inaccurate—or, in the case of pragmatic routines, decontextualized (e.g., Bardovi-Harlig, Mossman, & Vellenga, 2015a; Cohen & Ishihara, 2013; Eisenchlas, 2011; Gilmore, 2011; Ishihara & Cohen, 2010; Jiang, 2006; Vellenga, 2004; Williams, 1988). Thus, the development of instructional materials that illustrate authentic interaction in the target language has been of paramount importance to advocates of instructional pragmatics, where *authentic* is understood to be "naturally occurring attested language" (Flowerdew, 2015, pp. 15–16). In an aptly named article, *I prefer not text*, Gilmore (2011) compared the efficacy of using authentic materials to standard textbooks in the development of communicative competence, including pragmalinguistics and sociopragmatics. The 10-month study documented superior outcomes for the group that used authentic materials.

The advent of free online corpora has the potential to significantly change materials development at the classroom and program levels for the teaching of pragmatics. Corpora can be matched with instructional objectives to provide resources for pragmatics instruction. In this study, we further investigate the use of a corpus in pragmatics instruction, moving from using the corpus solely as the source of authentic interactions in materials development (Bardovi-Harlig, Mossman, & Vellenga, 2015b) to hands-on learning through guided corpus searches undertaken by learners, thus implementing a discovery-based approach (Boulton, 2010a). The potential for discovery identified by researchers of data-driven learning (DDL) has been emphasized in instructional pragmatics as well. Tomlinson (1994) and Clennell (1999) highlight *discovery* as an important part of noticing and as integral to the resulting pragmatic awareness. Clennell (1999) observed that "learners need to feel that they have arrived at their discoveries through their own efforts" (p. 87). Discovery may be supported by *guidance* (Vyatkina, 2016a), which is often called *focused noticing* in pragmatics (e.g., Bardovi-Harlig et al., 2015b).

In addition to a principled use of authentic language and an interest in promoting discovery, instructional pragmatics and DDL further converge in their foundational use of Schmidt's (1995) noticing hypothesis (Vyatkina, 2016a; for two additional SLA approaches in DDL, see Flowerdew, 2015). As Vyatkina (2016a) observes, the format of the stacked concordance lines may promote noticing through input enrichment (the large number of examples in any concordance list) and input enhancement (the centering and highlighting of the search word, phrase, or expression). An additional advantage claimed by DDL proponents is learner autonomy in engaging with the corpus as a learning resource (e.g., Vyatkina 2016b, p. 207), and this would be highly valued in instructional pragmatics as learners have few reliable resources other than developing their own ability to notice pragmatic features in ambient speech.

Although Boulton (2010a) and Vyatkina (2016b), among others, compared the use of paper-based concordances and direct searches, finding that there is no essential difference between them, the use of corpus searches has not yet been tested in the instructional pragmatics literature. This study investigates the use of teacher-prepared corpus-based materials and guided, direct corpus searches by learners in ordinary classrooms taught by ordinary teachers—as has been advocated by Boulton (2010b, 2011) and Vyatkina (2016a). By creating lessons for an existing curriculum and engaging program-appointed instructors who taught their regular classes, we hope to demonstrate the viability of corpus-based approaches for the teaching of pragmatic routines.

Previous Studies

The instructional pragmatics literature has begun to discuss the use of expert-speaker corpora for pragmatics instruction (Bardovi-Harlig & Mossman, 2016; Ishihara & Cohen, 2010; Schauer & Adolphs, 2006); similarly, DDL has begun to discuss the teaching of pragmatics (Flowerdew, 2012). Very few studies have tested the instructional effects of utilizing a corpus for pragmatics instruction— in a recent review, only

one out of 81 instructional effect studies did so (Bardovi-Harlig, 2015). Like other multi-word expressions and formulaic sequences, pragmatic routines and conventional expressions are particularly amenable to searches in a corpus. Three studies of the effect of corpus-based instruction on the use of pragmatic routines have been conducted using a variety of corpora: the Michigan Corpus of Academic Spoken English (MICASE; Simpson, Briggs, Ovens, & Swales, 2002) for the use of pragmatic routines in academic group work for ESL students (Bardovi-Harlig et al., 2015b); the Russian National Corpus for the use of pragmatic routines in social conversation for learners of Russian as a foreign language (Furniss, 2016), and online fan-transcriptions of *Friends* for conventional expressions for social conversation (Bardovi-Harlig & Vellenga, 2012).

Using fan-transcriptions of *Friends*, Bardovi-Harlig and Vellenga (2012) provided instruction on 30 conventional expressions to six ESL classes (N = 36). Input consisted of written transcripts with noticing activities. Three classes (Group A) engaged in noticing activities on one-half of the expressions (Set A) and three classes (Group B) on the other half (Set B). Three 1-hour lessons were delivered over three weeks. The pre-test and post-test, given four weeks apart, tested oral production for the expressions in context. The design was intended to test the effectiveness of instruction (Group A, Set A; Group B, Set B) against exposure (Group A, Set B; Group B, Set A). Both groups improved significantly on expressions in Set B, and Group B also improved significantly on Set A. The results suggested that learners benefitted from instruction and also from exposure to expressions present in the ambient input, but that were not instructional targets for their group.

Furniss (2016) identified nine pragmatic routines in Russian for corpus-based instruction drawing on the Russian National Corpus. She addressed the potential for the type of input confound identified by Bardovi-Harlig and Vellenga (2012) by excluding candidate expressions if they had appeared in textbooks; thus, all expressions had the potential for being instructionally new (although some participants lived in Russia at the time of the study). 34 learners of Russian as a second or foreign language volunteered to participate via the web and were assigned to either the control group (N = 16) or the instructional group (N = 18). Participants had two weeks to complete the 4.5–5.0 hours of self-paced instruction.[1] Materials included written corpus excerpts, film clips, and audio. A pre-test, post-test, and delayed post-test were administered. The tests included two scenarios that elicited a written production response, six multiple-choice scenarios, and an 18-item recognition task (9 target expressions and 9 modified expressions, following Bardovi-Harlig & Vellenga, 2012). Significant improvement was found on the production items and multiple-choice scenario tests from pre-test to post-test in the instructed group. Although recognition of conventional expressions did not improve, the instructed group was significantly better at rejecting non-authentic expressions at the post-test. The control group did not improve.

Both Bardovi-Harlig and Vellenga (2012) and Furniss (2016) emphasized noticing and neither included oral production practice as part of instruction. Bardovi-Harlig et al. (2015b) added communicative oral practice to supplement the noticing of pragmatic routines in input provided by corpus-based materials (in both written transcripts and re-recorded listening activities). The communicative activities provided learners with opportunities to produce what they had noticed. 16 pragmatic routines for agreements, disagreements, and clarifications in academic group work were targeted for instruction. Aural and written input was based on excerpts from MICASE. Four classes (N = 37) participated in four hours of instruction over two weeks; two classes participated as a repeated-test control. Oral production by learners was elicited by means of computer-delivered simulation of group work in a pre-test and a post-test. Learners who received instruction showed significant improvement in the production of both speech acts and pragmatic routines, whereas the control group did not.

In all three studies, instruction resulted in positive outcomes to varying degrees, suggesting that corpus-based instruction holds promise for the teaching of pragmatic routines and conventional expressions both in widely taught languages such as English and in less-commonly taught languages such as Russian. In these three studies, the researchers developed materials that were used as the basis for learner noticing. In the present study, we compare the use of teacher-prepared corpus-based materials to direct corpus searches

and investigate the following research questions:

1. Is there a difference between the performance of a group that receives teacher-developed corpus-based materials and the performance of a group that performs teacher-guided corpus searches, as measured by the production of speech acts and pragmatic routines in an oral group-work simulation task?
2. Does the corpus search group engage in corpus searches independently, beyond instructional activities?

Method

In order to compare the efficacy of corpus-based materials and direct corpus searches in the instruction of pragmatic routines, we compared the production data from the corpus-materials group (from Bardovi-Harlig et al., 2015b) to production data from the corpus search group, and compared both to the repeated-test control group who took the pre- and post-tests, but did not receive instruction on pragmatic routines in Bardovi-Harlig et al. (2015b).

Participants

54 students comprising eight ESL classes at a large public US university participated in the study. They were divided into three groups: the corpus-materials (CM) group (N = 26), the corpus-search (CS) group (N = 17), and the control group (N = 11). The learners were enrolled in seven Level 5 communication classes in a 7-level intensive English program and one equivalent English support course for matriculated students (assigned to the CM group); they could be described as low-advanced learners.

Only students who completed both the pre-test and post-test and who attended at least three of the four hours of instruction were included in the instructed groups. The 26 students in the CM group represented six language backgrounds (Arabic, 8; Chinese, 10; Japanese, 1; Korean, 3; Spanish, 3; Portuguese, 1); 12 students were female and 14 were male. The 17 students in the CS group represented eight language backgrounds (Arabic, 5; Turkish, 4; Portuguese, 2; Spanish, 2; Chinese, 1; Japanese, 1; Korean, 1; Thai, 1); 6 students were female and 11 were male.

Two Level 5 communication classes were recruited for the control group. The students represented seven language backgrounds; 10 students were male and 1 was female. The control group was included to gauge the influence of taking the test twice.

Instructors

Six program-appointed instructors were assigned to the communication classes during the three terms in which the experimental classes met. All teachers had completed an MA in TESOL and Applied Linguistics or Second Language Studies at the time of the study and had a range of 2 to 30 years of experience teaching English as a second or foreign language (M = 11 years).

The researchers met with the teachers at the outset to discuss the general approach and subsequently to discuss each of the four 50-minute lessons that comprised the unit. Instructors were given a lesson plan for each lesson that included a suggested script, all necessary materials, and a checklist. Instructors used the checklists to make sure that they included all parts of the lesson, and they turned them in to the researchers as evidence of fidelity to the lesson plans provided.

Materials Development

The target of instruction, agreements and disagreements, was identified by Level 5 communication teachers who asked us to develop new materials for teaching academic discussion. In response, we developed a unit for the intensive English program to teach pragmatic routines used for agreements and disagreements as well as clarifications, which may precede or lead to avoidance of disagreements (Pomerantz, 1984). To identify pragmatic routines for instruction, we examined the chapters on academic discussion in textbooks

favored by teachers in the program at the time of the study (Porter & Grant, 1998; Skillman & McMahill, 1996) and subsequently verified our selection of pragmatic routines by reviewing an additional 24 textbooks from six current series (Bardovi-Harlig et al., 2015a). We compared textbook expressions for agreement, disagreement, and clarification with expressions occurring in MICASE and added expressions identified from previous research on disagreements (Bardovi-Harlig & Salsbury, 2004). Because we were working with classroom teachers, we used a procedure that they could easily replicate for other speech acts. We chose MICASE because its academic content was consistent with the English for academic purposes curriculum. MICASE consists of 1.8 million words transcribed from almost 200 hours of speech from expert English speakers at the University of Michigan, including both non-native and native speakers.

A total of 16 pragmatic routines were identified for instruction, seven of which occurred between 65 and 120 times per million words or pmw (exceeding the count for frequent multiword units of 40 pmw proposed by Biber, Conrad, & Cortes, 2004), five of which occurred between 30 and 35 pmw, and four of which occurred between 10 and 18 pmw (meeting or exceeding the lower threshold of 10 pmw suggested by Biber, Johansson, Leech, Conrad, & Finegan, 1999), resulting in six agreement routines, three disagreement routines, two self-clarifications, and five routines for other-clarification. The pragmatic routines are given with their frequency counts and categorized by the speech act they introduce in Table 1.

Table 1. *Lesson Outline*

Lesson	Focus	Expressions (Occurrences pmw in MICASE)
1	Agreement	*That's right* (90+), *You're right* (75+), *That's true* (65+), *Good point* (18+)
2	Agreement and Disagreement	*I agree* (35+), *I agree with* (10+) *Yeah but* (120+), *Okay but* (90+), *I agree but* (10)
3	Self- and Other-Clarification	(Self) *What I mean* (100+), *In other words* (10+), (Other) *Do you mean* (36+), *What do you mean (*27+), *I have a question* (35+)
4	Other-Clarification	*You're saying* (90+), *What you're saying* (35)

Both CM and CS had at least three noticing activities per pragmatic routine. The focused noticing activities were essentially the same for CM and CS, although for CS, one noticing activity per speech act included frequency information. All other activities were held constant.

Corpus Materials

For corpus materials development, examples from interactive texts that illustrate the use of the targeted pragmatic routines were excerpted for input. Non-essential or non-sequential turns were pruned (see also Furniss, 2016; Ishihara & Cohen, 2010) and common synonyms were substituted for technical terms. Turns that contained the pragmatic routines always appeared in their original form.

All input was accompanied by noticing activities. The examples in Figure 1 and Figure 2 illustrate the input for disagreement, and the guided noticing activity is presented in Figure 3. (Underlining did not appear in the student materials; for additional examples, see Bardovi-Harlig et al., 2015a, 2015b.)

Dialogue 1

A: Did you look at the entrance to the lobby? You've got this very large enclosure outside the building and then you've got a bigger area inside. Did you look at putting some of the staircases actually into your glass box? I'm just saying that Difanico would think this was quite wasteful of space.
B: Originally we had that in, we had the stair in the glass box.
C: **Yeah but** where we did it was here. And then we began to work with our circulation, and when you put your main stair here, then if it's here, then you always have this.
A: It's not across the building, **I agree, but** on the other hand, you have that huge area kind of pretty well empty…

Transcript ID: STP125JG050

Figure 1. Dialogue 1 (Lesson 2); from MICASE

Dialogue 2

A: I saw this being a poem very much between, the narrator and the eyes of the poor, and not the narrator and this woman. I just saw that Baudelaire needed a situation for the narrator to be in so he could have this experience. But I didn't really see it as a central part of the text.
B: So, on the one hand **I agree** that, it's the relationship between him and these people he sees that brings out the conflict in him, **but** you know the woman there is used to be kind of this other half of him.

Transcript ID: SEM545MG083

Figure 2. Dialogue 2 (Lesson 2); from MICASE

In the accompanying guided noticing exercise (Figure 3), learners noted that *yeah but* and *I agree but* are used in disagreements, and that *I agree but* may occur contiguously, or be separated by as many as 18 words (*I agree… but*).

B. *Teacher script (and handout)*
In both of these conversations, one of the speakers expressed disagreement. Notice that neither of them is a direct disagreement. Instead, people usually disagree by using the same expressions they use for agreements…except they add the word "but" at the end. This allows them to provide a reason for the disagreement. In other words, a polite way to disagree, is by agreeing first! Now, read the conversations again, and answer the following questions.

Read the conversations again, and answer the following questions.
What are 2 disagreement expressions used in dialogue 1?

1. ***Yeah, but*** 2. ***I agree, but***

There are 2 disagreement expressions in dialogue 1. Which one includes a phrase you already learned as an expression of agreement? ***I agree, but***

In both dialogues, the speaker uses "I agree but" to express disagreement. How are they used differently?
In the first one the expression is together. In the second there is agreement information between "I agree "and "but".
How many words are there between the agreement part (I agree) and the disagreement part (but)?

18

Figure 3. CM Noticing Activity (Lesson 2)

Corpus Search

For the CS group, all written input based on corpus extractions was replaced by planned corpus searches. These were designed so that students would be able to conduct the searches on their own with little additional teacher guidance. When the searches were completed, teachers compared answers, summarized student findings, and emphasized key points. The search activities were piloted with a group of students who did not participate in the study and an instructor who ultimately taught one class in the experimental group.

Figure 4 shows a guided search that followed an earlier searching exercise in which the agree-before-disagree strategy was introduced with the expression *I agree but.* Here, learners are directed to the relevant section of the transcript that is the CS equivalent of Figure 1.

C. Now let's look at an example carefully.

- Search for "I agree but" in MICASE. Click the transcript ID.

1 matches in 1 transcripts

View results statistics | Download results as XML | Download results in tab-delimited format

Sort results by: None | None | None | Sort

Transcript ID: (click to view)	Left context	Match	Right context	View context
STP125JG050	it's not across the building	i agree but	on the other hand to have that huge area kind of pretty well empty (xx)	view

Figure 4. CS Search and Results (Lesson 2)

The learners are then directed to *click the transcript ID*, which expands the entry, but when the students first see it, it appears as a wall of text (see Figure 5).

Title: Architecture Critiques
Transcript ID: STP125JG050
Academic Division: Humanities
Publisher: Michigan Corpus of Academic Spoken English, English Language Institute, University of Michigan
Interactivity Rating: Highly interactive
Number of Participants:
Participants: 25
Speakers: 12
Recording Date: December 11, 1998
Recording Duration: 123 minutes
Word Count: 22596
Download entire transcript in XML

S1: we, we have two sites you haven't seen the (xx) before have you?

SU-f: no

S1: (um. or Mark?) okay. we have two sites, one is on North Campus which uh, if uh this is the, Pierpont Commons the Media Union and the bell tower, this is Bonisteel here, and M current parking lot is, to the north of Pierpont Commons. or that slope. that whole area was open for, choice. and the other site is um, in town on Central Campus, uh State Street is h Thompson Street Thompson Street parking structure. this is currently used as grade level parking. alright um, and Blimpy Burger is here, (lying down,) okay? just to set the context. t sites through the afternoon. okay, some students have been working individually and some have been working in pairs. so um, this is obviously a pair. <LAUGH>

S2: um, in looking at the uh, North Campus, we found that, it's quite obvious um, that there's a suburban feel to, the environment and, what we've tried to do with our scheme, uh wa the Media Union has, um, on this side of the street to the, the front of the building and, as well as somehow try to, address, the um, quadrangle, uh which is, uh presently which pres building doesn't complete the quad and, uh we're trying to attempt to, somehow, um get a sense of a new, way of designing, uh buildings for, uh, North Campus. so what we've done building, interact with the street as well as with the quad. um in a larger uh... master plan, if you will.

S3: Speaker information restricted

S2: yeah Murfin, currently goes straight down, so what we've done is we curve it bring it over so that we could have, the road um, interact with the building as well as have, our build

SU-3: Speaker information restricted

S4: um, let me go through this uh, programmatically if uh

S1: George, i can't

S4: let me go through uh, programmatically. the uh, the main entrance uh, to our building, is where the roof, where there's an existing roof where this, this is the uh ground floor plan that exists there right now, we're planning on keeping that and using that as the main entrance into the building. whereby, you can sorta see it in this view right here, and um, so the into, a what you call, vestibule, because it's sorta separated from the rest of the building as a, sort of very public, uh primary ci- uh, circulation space. now this is one of, one of three allowed, publicly as long as this finger right here, um, when i describe this finger uh um, my uh concept was that, these fingers which is the circulation space, (were) sort of um uh co was very much, um, very much uh, the the sacred space of an architecture building so we wanted to, to really um, to uh enclose it and, sort of make this space uh, you know not nec public but, um something that you, you can go to but, can't really see from the exterior of the building. so, s- you enter through the building, and um, and um, these are the two finger public spaces as soon as you enter you get an exhibition space right here. um where we'd have a coffee shop, we have a lecture hall auditorium, a public I-T-D space which, not ne Architecture but, more for public use. we have a, computer (xx) upstairs i mentioned. and, around this other um, circulation space, on the ground floor we have a mechanical space right-hand side of it we get our, computer classes areas um, you know you get your seminars right here, you got your workshop right here. and on the other finger of the circulation s right here. now on the ground floor, what we did what we decided to do for the parking was that, we decided um, we're gonna use the existing parking but we're gonna cover it so th excavation we're building from the existing grade up. so, we're we decided that we're gonna have parking, within our building, coming off of a street, off of Murfin, entering into as a as our, as programmatically for the parking spaces that are required. um, ways of getting on- onto the second floor now there's a couple of ways of doing that. uh from the main entra or, you can go up um, a one-hundred-and-eighty-foot ramp, one-to-twelve ratio scale for the handicap accessibility up until the studio space. uh, you don't quite get right into studio runs back the other way, which, actually connects you to, this other piece which is another sort of face of our building, which is across the street, which um, is the admin, uh offices o the second floor, you bridge across, you get directly in- into faculty offices, which went two stories up from that, and one more top for G-S-Is, PhDs... so. um, getting back to over here two stories. so, you can actually access um, studio-level floor, from the studio-level floor down into the coffee shop into this two-story area space. so, basically, you get up to this poir center sorry, area, um sorta publicly um, publicly located, and then after we get an extension of our studio space which is on one floor. so an existing building on this side is only a h our cre- we get our c- uh computer (pool areas.) uh which is uh, interconnected within, within those studio spaces. now that's one part of the studio, but... um the crit spaces are loca view is actually sort of standing right here, and looking up that way. so i'm getting a view into studio, but i_ you get a you get a view also of the crit spaces that's located right here, or studio space, i bridge across, over these um, over these three bridges that exist one two and three, over into, our classroom area, um, which is another, which is sorta separated fro have our typical, all the, the fire exit uh requirements um, two hundred feet (there'll be a) fire exit (stair.) um... yeah, um let me just explain some of the views if you can't read it this is round going up... you get your exhibition space will be right here, and you get your auditorium, lecture hall, and your I-T-D in on the corner. so you can kinda... [SU: could you talk (x (hole) which is, this finger right here. there by, you know, go by, you tr- you're trying to introduce um, the public into the building, because, in a way this_ the_ there's there's a couple you get this, later, and after you get, you know, sort of this layer within a layer. so you get your studio space encapsulated with as i said with the, circulation finger space. whereby_ the architecture building a little bit better, which uh i don't think that this building provides, as much interaction, they seem sort of very closed off (lately...) these views right here, as i which exists, and that's our main entrance, with uh, with the (xx) center facing out. this is uh, looking south um... northwest, coming up... but you can

Figure 5. Expanded View (Lesson 2)

The subsequent instructions help them find *glass box* (see Figure 6). This directs them to the relevant section of the transcript that is then used for the accompanying noticing activity.

- Click "Ctrl + F" and you'll see a search box on the top right corner of the screen. Search "glass box".

S6: did you look at the um at the entrance to the lobby? you've got this very large enclosure outside the building and then inside you've got almost the well you've got a bigger area inside, for circulation and lobbies, at all levels. did you look at, trying to provide it by putting some of the staircases actually into your glass box? [S8: um] i'm just saying [S8: into here?] that (Difanico) would think this was quite wastefuls of space.

S9: originally we actually had a (xx)

S8: (originally) we had that in, we had the stair in the, in the glass box.

S9: yeah <*PAUSE:04*> but [S6: did you?] where we we did [S8: no] it was here

S8: it was further out

S9: it was somewhat round. and then we, we began to, work with our circulation and, um, when you put your main stair here, then if it's here then you always have this

S6: it's not across the building **i agree but** on the other hand [S8: but] to have that huge area kind of pretty well empty (xx)

S8: but i mean, as far as the auditorium we're th- i mean i guess one idea we were thinking that the auditorium isn't just, a normal architecture auditorium and that the whole campus could use it and so this would be a space where you could hold receptions or whatever afterwards and

Figure 6. Final instructions and results (Lesson 2)

In the CS noticing activity (see Figure 7), learners are directed to notice the use of *yeah but* as a disagreement expression.

- Look at the excerpt between the first "glass box" and "I agree but". What is another expression of disagreement that is not "I agree but"? Yeah but

Figure 7. Noticing Expressions for Disagreement (Lesson 2)—Shown With Answer

Noticing activities in both the CM and CS groups helped learners recognize the agree-before-disagree strategy common in American English (Bardovi-Harlig & Salsbury, 2004; Pomerantz, 1984). Following Reppen (2010), we tested all the searches, and because we could be sure what the learners would find when they conducted a search, we could direct their attention to a limited range of lines in the concordance produced by the search as shown in Figure 4.

The screenshots in Figures 8 to 10 illustrate three additional search activities carried out by the CS group. The first search activity focused on the frequency of four agreement expressions.

Open your browser and go to Google. Type in "MICASE corpus" and click on "the Michigan Corpus of Academic Spoken English".

Among all the possible agreement and disagreement expressions in English, some are more common than others. Now you're going to find out which expressions are really used by native speakers in real life. Let's start with just agreement expressions.

On the MICASE webpage, click on the "Find" box. Search for each AGREEMENT expression your instructor wrote in Table 1, and write down the number of times it appears in MICASE in the table below. Circle the three most frequent expressions. (**Teacher notes:** If the students search the full forms, tell them that they will have to search the contracted forms. You will also need to switch between the MICASE webpage and the word document that has Table 1 in it as needed.)

	EXPRESSION	FREQUENCY
1	*That's true.*	*118*
2	*Good point.*	*55*
3	*You're right.*	*139*
4	*That's right.*	*167*
5	*(I agree)*	*69*

Figure 8. Noticing Expressions for Agreement (Lesson 1)

Other searches were followed by noticing the different contexts in which expressions can be used (e.g., *what I mean* in Figure 9) and differences in form in related expressions (e.g., *you're saying* and *what you're saying* in Figure 10).

	2 words before	**what I mean**	**1 word/punctuation after**
1	*you know*	what I mean	*?*
2	*you know*	what I mean	*?*
3	*trying to*	what I mean	*is*
4	*you know*	what I mean	*?*
5	*you know*	what I mean	*?*
6	*you see*	what I mean	*?*
7	*you see*	what I mean	*?*
8	*community but*	what I mean	*is*

Figure 9. Noticing Expressions for Self-clarification (Lesson 3)

B. Search for "what you're saying" in MICASE again. Look at the first result in the list. Record the Transcript ID and up to 10 words that follow "what you're saying" in the table below. Then search for "you're saying". Look at the first result without a preceding "what". Also record up to 10 words that follow the expression in the table below.

Transcript ID	**Expression**	**10 words after**
LAB500SU089	what you're saying	*is that this independent variable is affecting psych people and*
LES280JG138	you're saying	*wo- does this bias the comparison between the probit and*

Teacher script: *What do people do after saying "what you're saying" or "you're saying"? We see that in both cases the speaker summarizes what the other person just said.*

Figure 10. Noticing Expressions for Other-clarification (Lesson 4)

Instruction

Speech acts and their corresponding pragmatic routines were the focus of each of the four lessons shown in Table 1. The lessons began with a warm-up activity and had three parts: noticing the routine (by the students), provision of metapragmatic information (by the instructor), and oral production practice (by the students). A total of 200 minutes of instruction was planned. Approximately 107–137 minutes were used for noticing, 42–48 minutes for metapragmatic information, and 51–61 minutes for oral production (see Table 2).

Table 2. *Instruction in CM and CS Conditions*

	CM Group	**CS Group**
Lesson 1	Warm-up activity (5–10 minutes)	
	INPUT: Conversational excerpts from MICASE NOTICING ACTIVITIES: language use (~30 minutes)	INPUT: Concordance from MICASE NOTICING ACTIVITIES: Language use and frequency (~30 minutes)
	Production activity (6–9 minutes)	
Lesson 2	Warm-up activity (5–7 minutes)	
	INPUT: Conversational excerpts from MICASE NOTICING ACTIVITIES: Language use (~20 minutes)	INPUT: Concordance from MICASE NOTICING ACTIVITIES: Language use and frequency (~20 minutes)
	Aural input (10 minutes)	Aural input (10 minutes)
	Production activity (10–15 minutes)	
Lesson 3	Warm-up activity (3–5 minutes)	
	INPUT: Conversational excerpts from MICASE NOTICING ACTIVITIES: Language use (3–5 minutes)	INPUT: Concordance from MICASE NOTICING ACTIVITIES: Language use and frequency (10 minutes)
	Aural input (6–8 minutes)	Aural input (6–8 minutes)
	Production activity (10 minutes)	
	INPUT: Conversational excerpts from MICASE NOTICING ACTIVITIES: Language use (~10–15 minutes)	INPUT: Concordance from MICASE NOTICING ACTIVITIES: Language use and frequency (~10–15 minutes)
	Production activity (10 minutes)	
Lesson 4	INPUT: Aural & conversational excerpts from MICASE NOTICING ACTIVITIES: Language use (~30 minutes)	INPUT: Aural & concordance from MICASE NOTICING ACTIVITIES: Language use and frequency (~30 minutes)
	Production activity (15–20 minutes)	

For CM, input consisted primarily of the corpus excerpts with noticing activities. For CS, input comprised the corpus searches with instructions designed to focus student noticing on pragmatically relevant points. Focused-noticing activities included recording frequency counts for targeted expressions (for CS), and for both groups, activities included using corpus data to fill in tables to emphasize patterns and asymmetries,

paying attention to discontinuous elements (notably *I agree… but*), and writing expressions down. Metapragmatic information was provided by the instructor at the end of the noticing activities.

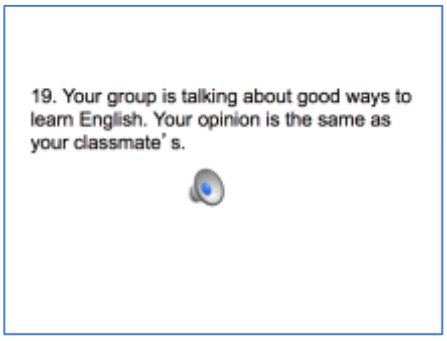
Learners in both groups participated in the games and tasks that comprised the oral production activities. These provided interactive oral practice that mimicked the unpredictability of turn-taking in conversation. Oral practice was cumulative so that the pragmatic routines from previous lessons were used in subsequent lessons. To practice agreements, groups of three students played a game in which one student read a non-controversial statement from a card, a second agreed with it using one of the expressions from class, and a third judged whether the second was successful. In the next lesson, disagreements were added in an activity where students tried to identify the subjects of ambiguous images and their partners used target expressions to agree or disagree with them as instructed on a card they drew. The final activity was a board game that provided students with opportunities to use all the expressions. Students moved along a path of squares, each containing a statement that they had to agree with, disagree with, self-clarify, or request clarification about as indicated on a card they drew. Successfully carrying out the required speech acts enabled students to advance in the game. Students rotated through the different roles during the activities, which ensured that they had equal opportunity to practice the speech acts and expressions.

Assessment Instruments

Because oral expressions for group work were the target of instruction, assessment was also oral. The task simulated academic group work through a computer-delivered spoken production task which was included in the instructional unit as an ungraded language-lab activity. The 30-item task included 10 agreement, 10 disagreement, and 10 clarification scenarios which were divided into two sections of 15 randomly arranged items (see Appendix A). The sections were reversed to create two equivalent tests. Learners were familiarized with the task through two examples followed by two practice items.

Each item briefly described the topic (e.g., your group is discussing transportation, newspapers, or learning English). Agreement and disagreement scenarios gave learners a position either by a relational statement (*your position is {the same as/different from} your classmates'*; see Figure 11) or a content statement (*you think that small cars save gas*; see Figure 12). Students saw the descriptions and their position on the screen. They next heard a classmate's turn to which they responded orally. Male and female voices alternated so that learners could easily distinguish the classmate's turn from the narrator's.

Screen 1.



Before Screen 2: Classmate's turn (audio only): Doing your homework is the best way to learn English.

Screen 2.

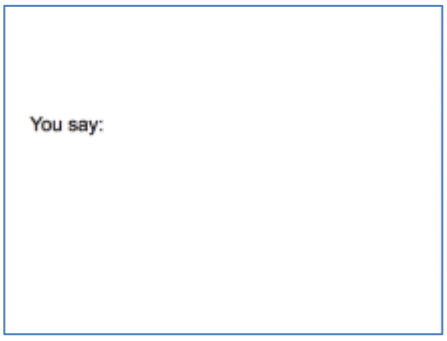
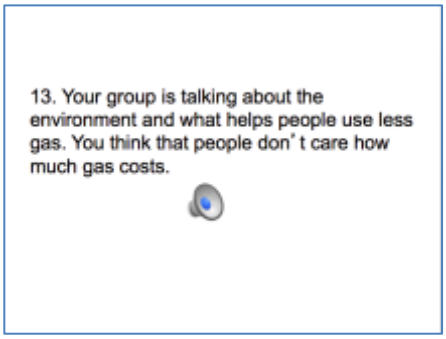


Figure 11. Agreement Item

Screen 1.



Before Screen 2: Classmate's turn (audio only): High gas prices help people use less fuel.

Screen 2.

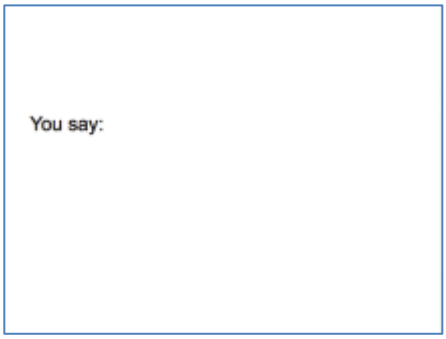
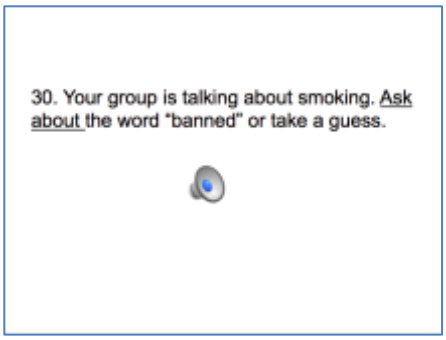


Figure 12. Disagreement Item

Clarifications included five other-clarifications and five self-clarifications. Other-clarification items instructed students to either request clarification or provide a comprehension check (see Figure 13). In self-clarification items learners had to rephrase a sentence given by the item when prompted by a screen stating *people look confused* (see Figure 14).

Screen 1.



Before Screen 2: Classmate's turn (audio only): Smoking should be banned in all public places.

Screen 2.

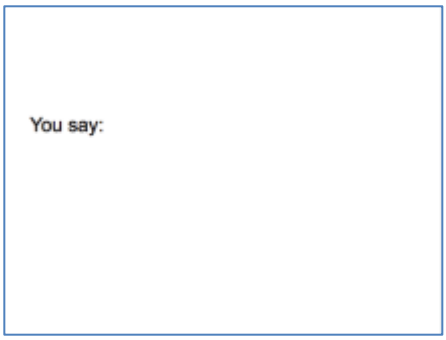
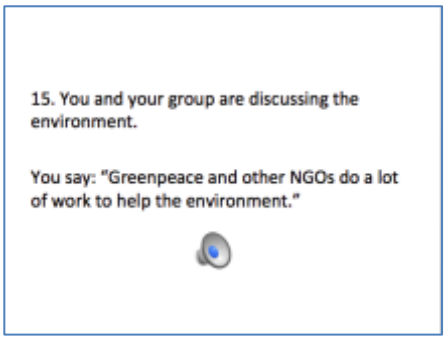


Figure 13. Other-clarification Item

Screen 1.



Before Screen 2: Classmate's turn (audio only): Doing your homework is the best way to learn English.

Screen 2.

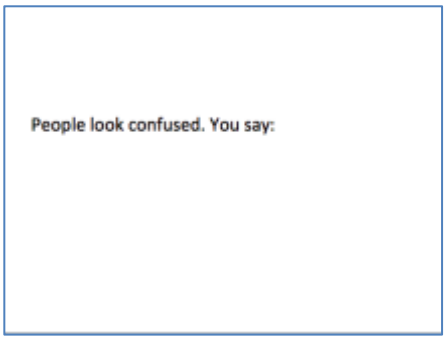


Figure 14. Self-clarification Item

Procedure

Prior to instruction, the CS students were introduced to direct corpus searches on MICASE by their instructor. Students completed two corpus searches related to the content of their class, but distinct from the pragmatic routines targeted in upcoming lessons. CS classes met in a language-dedicated computer lab to give students access to individual computers with internet connections for the corpus searches.

All students took the pre-test in the first week. Lessons 1 and 2 followed, allowing a day in between for activities that had not been covered the preceding day, if needed, to spread out the instruction. Lessons 3 and 4 were taught in the second and third weeks. The post-test was given the day after the instruction ended. The control group was tested at the same interval.

The pre-test and post-test were loaded on individual computers in the language lab. The scenarios were simultaneously presented visually on the screen and with sound through headsets. After each scenario, a classmate's turn was played in audio only. Students were then prompted to respond orally by a screen showing *You say*. 10 seconds were allotted per response, after which the next scenario started. Headset microphones recorded the learners' oral responses. Students who sat next to each other received alternate forms of the test. Following the post-test, a short questionnaire was given to the CS group to determine whether students had engaged in independent searches (see Appendix B).

Analysis

The task yielded 3,240 responses that were transcribed and coded for speech acts and targeted expressions. If a student produced the targeted speech act (e.g., an agreement after an agreement prompt), the response received a score of one point. Mismatched speech acts (e.g., an agreement after a disagreement prompt) earned no points. Percent agreement for inter-rater reliability for speech act identification for the entire data set by two author-coders was 91%. Disagreements were resolved by discussion.

The second coding recorded the targeted pragmatic routines for agreements, disagreements, and self- and other-clarifications. Most of the responses contained only one pragmatic routine. When a response contained two routines, only the first expression was scored, so that responses like *That's true. I agree* were logged as *That's true.* The maximum score for any response was one point. Grammatical routines were given one point; ungrammatical routines (e.g. *that true*) scored zero points. Agreements and disagreements each had a total possible score of 10 points, and self- and other-clarifications each had a total of five points. The scores of appropriate (and well-formed) pragmatic routines were calculated for each student, and the scores and ratios of appropriate and well-formed responses were aggregated by speech act for both the pre-test and the post-test (see Table 3).

Results

To ensure that the three groups were not different in their production of the target speech acts or pragmatic routines before instruction, one-way ANOVAs were performed on the pre-test scores. There were no significant between-group differences in any of the four speech-act types (agreement, disagreement, other-clarifications, and self-clarifications) on either measure (the scores for speech acts that convey the intended illocutionary force, henceforth *speech act scores*, or the scores for the pragmatic routines, henceforth *routine scores*). Two-way mixed-model ANOVAs with assessment (pre-test, post-test) as the within-subjects variable and treatment (CM, CS, control) as the between-subject variable were conducted to examine the effect of instruction. They were followed up with univariate ANOVAs on gain scores.

Both the number of appropriate speech acts and the number of targeted routines increased from the pre-test to the post-test in all three groups (see Figure 15). Results of mixed-model ANOVAs showed significant interactions between pre-test and post-test and treatment type on both speech act scores ($F_{(2, 51)} = 4.777$, $p = .013$, partial $\eta^2 = .158$) and pragmatic routine scores ($F_{(2, 51)} = 6.147$, $p = .004$, partial $\eta^2 = .194$) with large effect sizes (Cohen, 1988).[2] A post-hoc Tukey HSD test of speech act gain scores indicated a significant difference between CM and the control with a large effect size ($p = .023$, Cohen's $d = .954$; see Plonsky &

Oswald, 2014), but no significant difference between CS and the control (p = .723, Cohen's d = .291). A post-hoc Tukey test of gain scores for pragmatic routines indicated that both experimental groups significantly outperformed the control with large effect sizes (CM, p = .006, Cohen's d = 1.346; CS, p = .009, Cohen's d = 1.341), but the difference between the two experimental groups was not significant (p = .991, Cohen's d = .038).

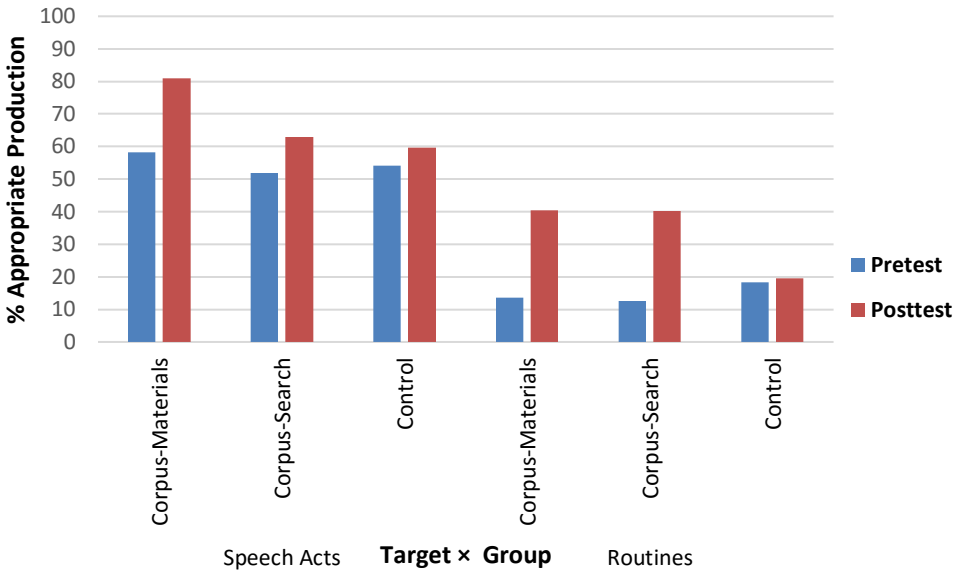


Figure 15. Comparison of appropriate speech act and routine production by the CM, CS, and control groups

Speech Acts

Speech acts were next considered by type. All three groups improved on speech act scores in all four speech-act types (see Table 3). CM showed greater improvement than CS (see Figure 16).

Table 3. *Production of Speech Acts Before and After Instruction by CM, CS, and Control Groups*

	Corpus-Materials (N = 26)				Corpus-Search (N = 17)				Control (N = 11)			
	Pre-test		Post-test		Pre-test		Post-test		Pre-test		Post-test	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Agreement (k=10)	.677	0.180	.862	0.142	.665	0.212	.800	0.170	.627	0.348	.691	0.311
Disagreement (k=10)	.681	0.240	.858	0.133	.565	0.259	.594	0.249	.582	0.322	.636	0.242
Other-Clarification (k=5)	.415	0.271	.831	0.202	.341	0.306	.506	0.317	.455	0.391	.545	0.336
Self-Clarification (k=5)	.362	0.330	.585	0.362	.318	0.325	.482	0.400	.364	0.398	.382	0.275
Total (k=30)	.582	0.166	.809	0.118	.519	0.164	.629	0.197	.541	0.302	.597	0.238

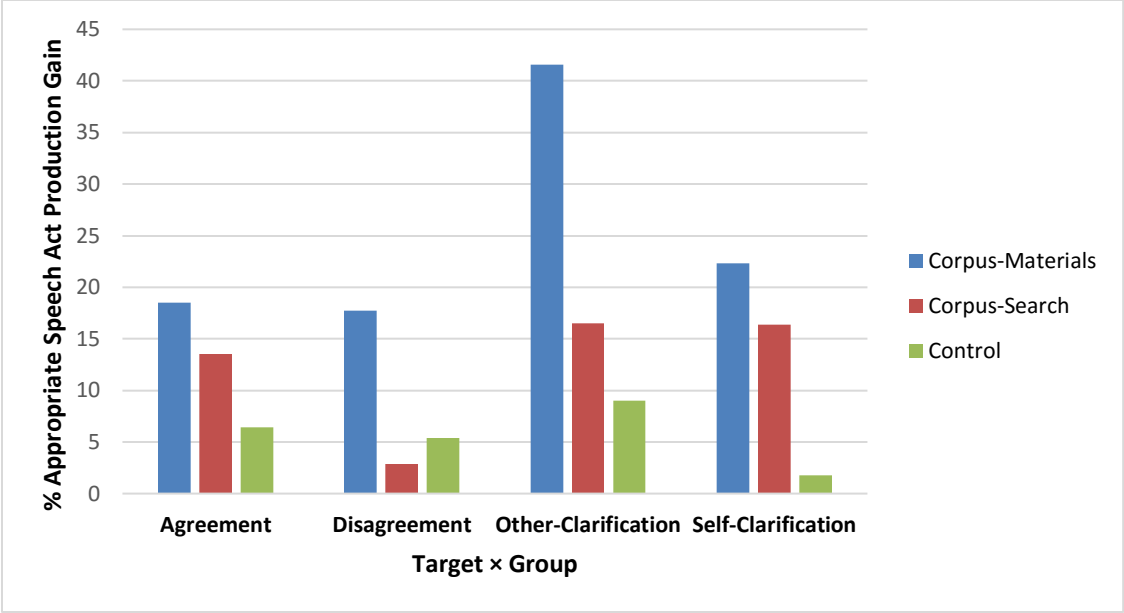


Figure 16. Gain scores in production of speech acts by CM, CS, and control groups

Mixed-model ANOVAs were used to determine whether the observed differences in gain scores of speech act production were statistically significant. There was a significant interaction between treatment and assessment (pre-test and post-test) for other-clarification ($F_{(2, 51)} = 8.798$, $p = .001$) with a large effect size (partial $\eta^2 = .257$), but not for the other three speech-act types. This indicates that the improvement in other-clarifications differed significantly between groups.

Significant main effects of assessment (pre-test and post-test) were found for the other three speech-act types with medium to large effect sizes (agreement, $F_{(1, 51)} = 18.821$, $p < .001$, partial $\eta^2 = .270$; disagreement, $F_{(1, 51)} = 5.642$, $p = .021$, partial $\eta^2 = .100$; self-clarification: $F_{(1, 51)} = 9.150$, $p = .004$, partial $\eta^2 = .152$). This means that all three groups improved from the pre-test to post-test on these speech act types, but the differences in gains between the groups were not significant.

Targeted Routines

Both experimental groups improved on routines in all four speech-act types, whereas the control showed either minimal or no increase—or, as in the case of disagreement routines, a decrease (see Table 4). CM showed greater improvement than CS on production of agreement routines, and CS showed greater improvement than CM on production of the other three types of routines.

Table 4. *Use of Pragmatic Routines Before and After Instruction by CM, CS, and Control Groups*

	Corpus-Materials (N = 26)				**Corpus-Search (N = 17)**				**Control (N = 11)**			
	Pre-test		**Post-test**		**Pre-test**		**Post-test**		**Pre-test**		**Post-test**	
	M	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Agreement (k=10)	.131	0.159	.500	0.291	.241	0.235	.471	0.329	.227	0.190	.273	0.195
Disagreement (k=10)	.212	0.188	.412	0.261	.129	0.165	.412	0.280	.282	0.199	.264	0.180
Other-Clarification (k=5)	.046	0.130	.308	0.345	.000	0.000	.400	0.374	.073	0.135	.073	0.185
Self-Clarification (k=5)	.092	0.228	.294	0.367	.024	0.066	.247	0.364	.000	0.000	.036	0.081
Total (k=30)	.137	0.106	.404	0.224	.126	0.111	.401	0.268	.183	0.125	.196	0.110

Mixed-model ANOVAs were used to examine if the observed differences in gain scores of routine production were statistically significant. There was a significant interaction between treatment and assessment (pre-test and post-test) on agreement ($F_{(2, 51)} = 4.916$, $p = .011$) with a large effect size (partial $\eta^2 = .162$), on disagreement ($F_{(2, 51)} = 3.695$, $p = .032$) with a medium effect size (partial $\eta^2 = .127$), and on other-clarification ($F_{(2, 51)} = 5.282$, $p = .008$) with a large effect size (partial $\eta^2 = .172$).

A significant main effect of assessment (pre-test and post-test) was found for self-clarification ($F_{(1, 51)} = 11.612$, $p = .001$) with a large effect size (partial $\eta^2 = .185$), though no significant interaction between assessment (pre-test and post-test) and treatment was found ($F_{(2, 51)} = 1.383$, $p = .260$, partial $\eta^2 = .051$). This means that all three groups improved from the pre-test to post-test on their routine scores for self-clarification, but the differences in gains between the groups were not significant (see Figure 17).

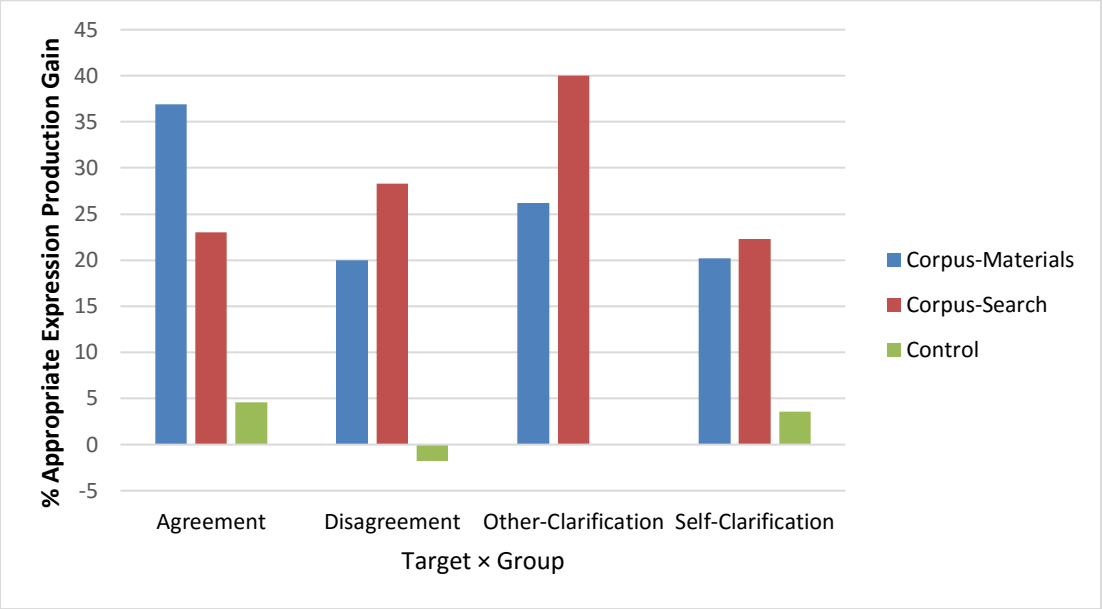


Figure 17. Gain scores in production of pragmatic routines by CM, CS, and control groups

The noticing events for the pragmatic routines were evenly balanced, but the range and number of routines that were actually used was up to the learners. On the post-test, the clear winners for CS were *That's right* and *That's true* for agreements; *Yeah but*, for disagreements; and *What do you mean?* for other-clarification (see Table 5). Similarly, *yeah but* and *What do you mean?* increased for CM. On the post-test, the three most frequent agreement routines in MICASE (*That's right, That's true*, and *You're right*) were used in 81% of the agreements by CS who had recorded frequency information as part of their noticing activities, and they used *I agree (with)* in only 20% of the pragmatic routines. CM, who received no frequency information, used the highly transparent illocutionary force indicating device *I agree (with)* in 54% of the routines and the three most frequent expressions only 44% of the time on the post-test.

Of special note for CS is the *creative agreement + but* category for disagreements, which showed 23 tokens. Learners were encouraged to notice that a variety of agreement expressions could be used with *but* to create a disagreement marker, and CS produced 23 of these on the post-test. Combinations included *That's true but* (8), *That's right but* (6), *That's a good point but* (3), and *You're right but* (1). Six pragmatic routines exhibited additional grammatical and content creativity. Three of those showed differing degrees of modality including *Maybe you are right but*, *You may be right but*, and *I think that is right but* and three showed particular sensitivity to the lexical content of the previous claim such as *Knowing a lot of words is important, but* in response to "knowing a lot of words is the best way to learn English," the classmate's turn in item I-12.

Both groups showed low scores for *in other words* for self-clarification, and *You're saying* and *What you are* saying for other-clarification. Tokens of *What do you mean* divide into *What do you mean?* as a stand-alone question, *What do you mean by* followed by a complement, and *What do you mean* followed by no complementizer (*What do you mean current events?*, said by learner CS14, in response to item I-6) or an interlanguage form (*What do you mean? [.] with this proverb*, said by learner CS9, to item I-8).

Table 5. *Pragmatic Routines Used Before and After Instruction (CM and CS Groups)*

	CM		**CS**	
Agreements	**Before N (%)**	**After N (%)**	**Before N (%)**	**After N (%)**
You're right	4 (12)	14 (11)	9 (21)	18 (23)
That's right	0 (0)	23 (18)	0 (0)	23 (29)
That's true	1 (3)	20 (15)	1 (0)	23 (29)
I agree	7 (21)	11 (9)	14 (33)	8 (10)
I agree with	22 (65)	59 (45)	18 (43)	8 (10)
Good point	0 (0)	3 (2)	0 (0)	0 [a](0)
Subtotal	34	130	42	80
Disagreements				
Yeah, but	0 (0)	28 (26)	3 (14)	24 (34)
Okay, but	0 (0)	9 (8)	0 (0)	16 (23)
I don't think so	55 (100)	52 (49)	16 (76)	6 (9)
I agree but	0 (0)	10 (9)	1 (5)	1 (1)
Creative agreement + but	0 (0)	8 (7)	1 (5)	23 (33)
Subtotal	55	107	21	70
Other-Clarification				
Do you mean	2 (33)	5 (13)	0 (0)	8 (24)
What do you mean	2 (33)	24 (60)	0 (0)	26 (76)
You're saying	1 (17)	7 (18)	0 (0)	0 (0)
Your point[b]	1 (17)	3 (8)	0 (0)	0 (0)
I have a question	0	0 (0)	0 (0)	0 (0)
What you're saying	0	1 (2)	0 (0)	0 (0)
Subtotal	6	40	0	34
Self-Clarification				
What I mean	1 (8)	16 (42)	0 (0)	8 (38)
I mean	11 (92)	16 (42)	2 (100)	12 (57)
What I'm saying	0 (0)	6 (16)	0 (0)	1(5)
In other words	0 (0)	0 (0)	0 (0)	0 (0)
Subtotal	12	38	2	21

Note. [a] *The three tokens of good point in the post-test (That's a good point, but) are included in the disagreement category creative agreement + but.* [b] *Your point was not taught in the CS group.*

CS received a questionnaire to gauge the effect of using corpus searches in class on their independent use of the corpus. When asked whether they carried out any additional searches on MICASE outside of class, 7 of the 17 students indicated that they did, and their examples showed that they searched for words and expressions from a range of registers.[3] Students reported searching for formulaic expressions not taught in the lessons and for words they thought were slang. For colloquial language, for example, in class they searched for expressions that contained *yeah* and *okay*, variants of *yes*. One student followed this up at home by searching for *yup*, *nope*, and *gosh*. Another student searched for *OMG*, *you're welcome*, and *are you sure*. A third searched for *got it!*, *got you!*, and *that's wassup.* Another student searched for *that is why* and *for this reason*, expressions that could be used to provide explanations. Thus, it appears that the training in using MICASE did not just teach students they can look words up, but also that they can use it to determine what expressions occur with the speech acts they want to carry out. We cannot compare the increase in autonomous searches by CM and CS directly, because we did not distribute a questionnaire to CM. However, we speculate that increased autonomy comes from being introduced to the corpus and being shown how to use it in class, whereas the citations to the corpus on the teacher-developed materials were not sufficient motivation for the students to seek it out. This interpretation can be tested in future research.

Discussion

This study compared the effect of using corpus-based materials and activities for the instruction of pragmatic routines under two conditions: implementing direct corpus searches by learners during classroom instruction and working with teacher-developed corpus-based materials. The instruction was consistent with the established curriculum of the intensive English program in which it was implemented, and was carried out in regular ESL classrooms and delivered by program-appointed ESL instructors. This study is a demonstration that both pragmatics instruction and DDL can be integrated successfully into established programs and ongoing language classrooms.

The use of MICASE, both in searches by learners and in teacher-developed materials, led to significantly improved use of pragmatic routines—a result that is consistent with the findings of Boulton (2012) and Vyatkina (2016b). In this study, in the CM format, the input in the paper-based materials provided three conversational excerpts per pragmatic routine rather than concordances. Crucially, they were accompanied by focused noticing activities, providing the guidance necessary for the students to engage actively with the input. The CS format provided the opportunity for discovery characteristic of DDL. Coupled with focused noticing activities, CS provided the learners with guidance that could compensate for any additional cognitive load associated with performing the searches themselves. Additional cognitive load may not be great, given that MICASE is relatively easy to use, but this could be an issue when corpora are less user-friendly.

However, the groups unexpectedly differed in speech act production. This may have less to do with hands-on or hands-off corpus use and more to do with format. The teacher-developed materials apparently gave students an advantage. The benefit may lie in the format of the materials contrasted with the format of the search results. The teacher-developed corpus-based examples highlighted speech acts that were presented as individual examples, as in Figure 18 (see also the examples in Figure 1 and Figure 2).

A: Oh it's definitely white on the tail feathers.
B: Okay well, *I agree with you*. Starling is totally wrong.
A: Thank you. <LAUGH> Finally. <LAUGH>
Transcript ID: LAB175SU026

Figure 18. Agreement Excerpt

In contrast, the concordance format in Figure 19 may focus the learners' attention primarily on the expression that they searched for, both by the number of tokens (input enrichment) and the highlighting with centering (input enhancement). However, the horizontal context is both less elaborated than in Figure

18 and less sufficient for pragmatics than for morphosyntax. A speech act takes more turns (and more lines) to complete than the context for a single word or phrase.

	Match	Right context	View context
feller feathers. okay well,	i agree	with you starling is totally wrong.	view
other improvement?	i agree	that there's, definitely a long way to go, but i definitely think there's been a lot of improvement	view
okay. and	i agree	with you Z greater than one does not exist well at equal one it exists, but if you leave it like thi	view
oh i remember this. yeah i	i agree	with that i mean i just feel that, it's not gonna, i can't imagine people even involved in the case	view
ame thing happens. with yeah.	i agree	. PAUSE duration :06	view
? i'm in the middle bec- like	i agree	, a lot with what (Julia) said like, the rights are there but, i've heard of problems with like, peo	view
hand up um, yeah	i agree	with, is it Debbie?	view
i wanna say which	i agree	it's it's it's a it's a uh i- it uh, uh it's conformal one might say or it's it's	view
ader, do with it what he will, that	i agree	with um... (you know) w	view
just said a minute ago i don't if	i agree	with that, i think that it's broader, and i think it has affected, i think it has affected, a lot of	view
and i and	i agree	. i i like the faculty lounge	view

Figure 19. Concordance lines for *I agree*

When the learners click on *view*, a longer excerpt appears, as in Figure 20. However, neither the concordance lines in Figure 19 nor the expanded view in Figure 20 mark the speaker or illustrate the turn.[4]

Utterance Type: U1

Speaker ID: SU-m
Native speaker status: Native speaker, American English
Speaker role: Senior Undergraduate
Speaker sex: Male
Speaker age: 17-23 years old

Transcript ID: LAB175SU026 (click to view)

... oh (so there's) (xx) oh there. oh it's definitely white on the tail feller feathers. okay well, **i agree** with you starling is totally wrong. thank you. *<LAUGH>* finally. *<LAUGH>* so, we'll call that a meadowlark. what do other people think? ...

Figure 20. Expanded text view in MICASE

Although Figure 18 and Figure 20 show the same sample, the excerpt in Figure 18 is clearer, providing speaker identification and turn breaks and presenting the speech act as a separable unit. The difference between the formats is even greater in longer exchanges where CM presentations pruned the extraneous turns, which CS expansions retained. Thus, the quality of the CM examples for speech act development may be superior.

A second issue of quality results not from the format, but from the selection of excerpts. The excerpts for CM input were individually selected for their clarity and comprehensibility. The first examples in the CS searches were also selected for the same features, but the subsequent examples were identified by the corpus search, and not by the materials developers. Thus, texts in the concordance may be less transparent to the learners.

Materials Development (Revisited)

At first blush, the difference between CM and CS may seem like a division of labor: Teachers work on the materials and learners work on the searches. However, the direct corpus searches that we designed provided substantial support for the learners through teacher-developed activities that directed learner attention to

pragmatic features. Both Flowerdew (2015) and Vyatkina (2016a) support guidance with the searches at early stages. Thus, designing the searches also requires materials development by teachers to support learners' initial searches so that learners can later search autonomously.

There are additional challenges to materials development, particularly for clarifications. Both CS teachers reported that learners had difficulty understanding parts of Lessons 3 and 4. For example, determining what should follow *Do you mean* and *What do you mean* is something that learners could not figure out for themselves and needed teachers' help with. Developing additional activities for self- and other-clarifications may be warranted. Testing self-clarifications was also challenging. We used a prompt that said *people look confused* to encourage a restatement of the original statement (see Figure 14). Creating alternative formats may be beneficial. Adding self-clarifications to the examples and the practice items at the beginning of the elicitation task might also help the learners.

Limitations

One challenge to our ability to compare corpus-based instruction across groups is the 4-year span between the treatments. In spite of the time span, the same test was used to determine the placement of new students in all the conditions, and the groups had similar scores on the pre-test. Nevertheless, we were aware of subtle differences. The lack of familiarity with some academic vocabulary may reflect changing program demographics: half of the students in CS were newcomers, whereas most of the students in CM continued from a previous term. This suggests the importance of comparing learners in the same or immediately contiguous terms if possible. The pre-test scores show that these differences were not a serious threat.

Conclusion

This study highlights the benefits of incorporating corpus-based instruction in the teaching of pragmatics. Instructionally supported searches of a corpus matched to the target register (in our case, a spoken academic corpus for academic group work) focus learners' attention on the pragmatic routines and encourage them to conduct independent searches. Teacher-prepared materials help learners improve the clarity of the illocutionary force of their turns in conversation. Although we contrasted the different approaches to corpus use in this study, the clear pedagogical implication is that a judicious combination of teacher-developed corpus-based materials for speech act clarity and supported corpus searches by learners for noticing pragmatic routines would be ideal. This study also shows that such instruction can be integrated into ordinary language classrooms, taught by classroom teachers, and designed for an established curriculum.

Acknowledgements

We thank the instructors, John Rothgerber and Kyle Swanson, who administered the lessons; Michael Frisby of the Indiana Statistical Consulting Center, Indiana University; and the Center for Language Technology (CeLT), Indiana University, and its staff for providing technical expertise in recording and formatting the computer tasks and assistance during the elicitation sessions. Audio-editing was done by Natasha Branch. We also thank Professors Yucel Yilmaz and Sun-Young Shin for insightful discussions of this paper.

Notes

1. Furniss (2016) reported 6.5 hours of participation for the experimental group and 1.5–2.0 hours for the control, which suggests that instruction was 4.5–5.0 hours.
2. Plonsky and Oswald (2014) suggested using field-specific benchmarks for effect sizes, but there are currently no available benchmarks for partial eta squared in second language research. Therefore, we follow the conventional benchmarks suggested by Cohen (1988; for further discussion of eta-squared and partial eta-squared, see Norouzian & Plonsky, 2017).

3. Three additional students in the class reported having searched MICASE independently. These students had missed the pre-test and thus were not included in the study, but as we worked with intact classes, they also received the instruction. One of those students reported looking for *dude* and *mate* (the latter not typical as an address term in American English, as he would discover).
4. The full conversation with turns is available on MICASE by clicking on the transcript number. The line that was being viewed is not immediately visible, but can be searched for in the longer transcript using the Ctrl-F function. We worked with the full transcripts, but learners may need significant support to do this.

References

Bardovi-Harlig, K. (2013). Developing L2 pragmatics. *Language Learning*, *63*(S1), 68–86.

Bardovi-Harlig, K. (2015). Operationalizing conversation in studies of instructional effects in L2 pragmatics. *System*, *48*, 21–34.

Bardovi-Harlig, K., & Mossman, S. (2016). Corpus-based materials development for teaching and learning pragmatic routines. In B. Tomlinson (Ed.), *SLA research and materials development for language learning* (pp. 250–267). New York, NY: Taylor and Francis.

Bardovi-Harlig, K., & Salsbury, T. (2004). The organization of turns in the disagreements of L2 learners: A longitudinal perspective. In D. Boxer & A. D. Cohen (Eds.), *Studying speaking to inform second language learning* (pp. 199–227). Clevedon, UK: Multilingual Matters.

Bardovi-Harlig, K., & Vellenga, H. E. (2012). The effect of instruction on conventional expressions in L2 pragmatics. *System*, *40*, 77–89.

Bardovi-Harlig, K., Mossman, S., & Vellenga, H. E. (2015a). Developing corpus-based materials to teach pragmatic routines. *TESOL Journal*, *6*, 499–526. doi: 10.1002/tesj.177

Bardovi-Harlig, K., Mossman, S., & Vellenga, H. E. (2015b). The effect of instruction on pragmatic routines in academic discussion. *Language Teaching Research*, *19*, 324–350.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at …: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, *25*, 371–405.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. London, UK: Longman.

Boulton, A. (2010a). Data-driven learning: Taking the computer out of the equation. *Language Learning*, *60*, 534–572.

Boulton, A. (2010b) Learning outcomes from corpus consultation. In M. Moreno Jaén, F. Serrano Valverde, & M. Calzada Pérez (Eds.), *Exploring new paths in language pedagogy: Lexis and corpus-based language teaching* (pp. 129–144). London, UK: Equinox.

Boulton, A. (2011). Data-driven learning: The perpetual enigma. In S. Goźdź-Roszkowski (Ed.), *Explorations across languages and corpora* (pp. 563–580). New York, NY: Peter Lang.

Boulton, A. (2012). Hands-on/hands-off: Varying approaches in data-driven learning. In J. Thomas & A. Boulton (Eds.), *Input, process, and product: Developments in teaching and language corpora* (pp. 152–168). Brno, Czech Republic: Masaryk University Press.

Clennell, C. (1999). Promoting pragmatic awareness and spoken discourse skills with EAP classes. *ELT Journal*, *53*, 83–91.

Cohen, A. D., & Ishihara, N. (2013). Pragmatics. In B. Tomlinson (Ed.), *Applied linguistics and materials development* (pp. 113–126). London, UK: Bloomsbury Academic.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Eisenchlas, S. A. (2011). On-line interactions as a resource to raise pragmatic awareness. *Journal of Pragmatics*, *43*, 51–61.

Flowerdew, L. (2012). Exploiting a corpus of business letters from a phraseological, functional perspective. *ReCALL*, *24*(2), 152–168.

Flowerdew, L. (2015). Data-driven learning and language learning theories: Whither the twain shall meet. In A. Leńko-Szymańska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 15–36). Amsterdam, Netherlands: John Benjamins.

Furniss, E. A. (2016). Teaching the pragmatics of Russian conversation using a corpus-referred website. *Language Learning & Technology*, *20*, 38–60. Retrieved from http://llt.msu.edu/issues/june2016/furniss.pdf

Gilmore, A. (2011). "I prefer not text": Developing Japanese learners' communicative competence with authentic materials. *Language Learning*, *61*, 786–819.

Ishihara, N., & Cohen, A. D. (2010). *Teaching and learning pragmatics: Where language and culture meet*. London, UK: Longman.

Jiang, X. (2006). Suggestions: What should ESL students know? *System*, *34*, 36–54.

Norouzian, R., & Plonsky, L. (2017). Eta- and partial eta-squared in L2 research: A cautionary review and guide to more appropriate usage. *Second Language Research*. Advance online publication. doi: 10.1177/0267658316684904

Plonsky, L., & Oswald, F. (2014). How big is 'big'? Interpreting effect sizes in L2 research. *Language Learning*, *64*, 878–912.

Pomerantz, A. (1984). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 57–101). Cambridge, UK: Cambridge University Press.

Porter, P. A., & Grant, M. (1998). *Communicating effectively in English: Oral communication for non-native speakers* (2nd ed.). Boston, MA: Heinle & Heinle.

Reppen, R. (2010). *Using corpora in the language classroom*. Cambridge, UK: Cambridge University Press.

Schauer, G. A., & Adolphs, S. (2006). Expressions of gratitude in corpus and DCT data: Vocabulary, formulaic sequences, and pedagogy. *System*, *34*, 119–134.

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1–63). Honolulu: University of Hawai'i Second Language Teaching and Curriculum Center.

Simpson, R. C., Briggs, S. L., Ovens, J., & Swales, J. M. (2002). *The Michigan corpus of academic spoken English*. Ann Arbor, MI: The Regents of the University of Michigan. Retrieved from http://quod.lib.umich.edu/m/micase/

Skillman, P., & McMahill, C. (1996). *Springboard to success: Communication strategies for the classroom and beyond*. Upper Saddle River, NJ: Prentice Hall Regents.

Tomlinson, B. (1994). Pragmatic awareness activities. *Language Awareness*, *3*, 119–129.

Vellenga, H. E. (2004). Learning pragmatics from ESL and EFL Textbooks: How likely? *TESL-EJ*, *8*(2), n.p.

Vyatkina, N. (2016a). Data-driven learning for beginners: The case of German verb-preposition collocations. *ReCALL*, *28*, 207–226.

Vyatkina, N. (2016b). Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning & Technology*, *20*(3), 159–179. Retrieved from http://llt.msu.edu/issues/october2016/vyatkina.pdf

Williams, M. (1988). Language taught for meetings and language used for meetings: Is there anything in common? *Applied Linguistics*, *9*, 45–58.

Appendix A. Conversation Simulation

(Slideshow available from IRIS)

Practicing for Academic Discussion

This exercise is a role-play. In this exercise, you pretend that you are a student who is participating in a group discussion about controversial topics. You will hear statements made by your classmates and you will respond orally with the information in the role play instructions. Please notice that your opinion in the role play is given in the instructions. (We are not asking for your own opinion.) Sometimes you will need to ask for information, check your understanding, or help someone else understand what you have said or what a classmate has said.

Please speak clearly. You will be given 10 seconds to respond to each situation. You will answer when you see the slide that says "You say:"

Let's do two examples.

Example A

Moderator 1: You are a first year college student who is taking a German class. You think that your teacher speaks too fast in class. You talk to your teacher, a native speaker of German, after class.

Student 1: I was wondering if you could slow down a little. I have a hard time following you.

Example B

Moderator 2: You are working as a computer assistant in a computer lab. Your job is to make sure students are doing school-related work. While helping one student, you see another student playing a computer game.

Student 2: Excuse me, but playing games is not academic. Please go play the game on your home computer.

Now you try it. Speak clearly. You have 10 seconds to respond.

Moderator 1: Number 1. Your group is talking about what is important in life. You think something different from your classmate.

Student 1: It is better to be very beautiful than very smart.

Moderator 2: Number 2. Your group is talking about the news and media. You do not have the same opinion as your classmate.

Student 2: Blogs are a reliable source of news.

Set 1

1. Your group is talking about good ways to learn English. Your opinion is the same as your classmate's.

Classmate: Doing your homework is the best way to learn English.

2. Your group is discussing advantages and disadvantages of using the internet to study. You think that looking up information on the web is not always a good use of time.

Classmate: When you go online to look for information while studying, you can end up wasting time looking at websites.

3. Your group is discussing the way that people communicate. You have the same opinion as your classmate.

Classmate: People spend too much time talking on the phone these days.

4. Your group is discussing public health. You know that second-hand smoke (when *nonsmokers breathe the smoke)* is a problem.

Classmate: People should not smoke in public places, including universities.

5. Your group is discussing ways that students can improve their English. You like the many ways to use English on the internet.

Classmate: Using the internet is a good way for students like us to improve our English.

6. Your group is talking about the news and media. You want to know what your classmate thinks "current events" are.

Classmate: Facebook is a good place to learn about current events.

7. Your class is discussing how to improve English language skills. Your opinion is different from your classmate's.

Classmate: Studying grammar is more important than practicing conversation skills.

8. Your group is talking about motivating other people. Your classmate uses a proverb. Check your understanding of the saying.

Classmate: You can lead a horse to water, but you cannot make it drink.

9. You and your partner are talking about fashion. You have a different opinion.

Classmate: If a fashion is popular in Bloomington, then it will be popular in New York.

10. You are talking about the number of days in a year. In the western calendar there are generally 365 days in a year. In 2012 there are 366.

You say: "2012 is a leap year."

[Screen only: People look confused]

11. Your group is discussing good ways to work. You think the same thing as your classmate.

Classmate: Group work is good for some projects, and working alone is better for other projects.

12. Your class is discussing how to improve English language skills. You have a different opinion from your classmate's.

Classmate: Knowing a lot of words is important to speaking English well.

13. Your group is talking about the news and media. You think that newspapers like *The New York Times* and *The London Times* are still very important.

Classmate: Nobody reads newspapers these days.

14. Your group is talking about the environment and what helps people use less gas. You think that people don't care how much gas costs.

Classmate: High gas prices help people use less fuel.

15. You and your group are discussing the best way to learn math.

You say: Daily practice with problems is important for learning math.

[Screen only: People look confused]

Set 2

1. Your group is discussing healthy eating. You do not think the same thing as your classmate.

Classmate: Fast food restaurants serve healthy food.

2. Your group is discussing transportation and cars. You think that small cars save more gas.

Classmate: Big cars are better than small cars.

3. Your group is talking about smoking. Ask about the word "banned" or take a guess.

Classmate: Smoking should be banned in all public places.

4. Your group is discussing transportation and cars. You have the same opinion as your classmate.

Classmate: People who take the bus are more responsible environmentally than people who drive cars.

5. Your group is discussing calendars, but your classmate uses a term that you are unsure of. You think you have an idea of what it means. The term is "leap year." Verify your understanding.

Classmate: In the western calendar, every fourth year is a leap year.

6. Your group is discussing whether governments should influence the size of families. You think it is a family decision.

Classmate: Governments should not tell families how many children they should have.

7. Your group is discussing television and other media. Your opinion is the same as your classmate's.

Classmate: Television has a bad influence on society.

8. You and your group are discussing computers.

You say: Faster processers result in increased computer speed.

[Screen only: People look confused]

9. Your group is discussing transportation and cars. You think that big cars are more comfortable.

Classmate: Big cars are better than small cars.

10. You are talking about the government's influence on how many children families have. You do not understand the importance of your classmate's contribution.

Classmate: In some countries the government gives you money for children.

11. Your group is discussing transportation and cars. You think that having a car is very convenient.

Classmate: Owning a car has a lot of disadvantages including insurance and other expenses.

12. Your group is talking about what is important in life. You think something different from your classmate.

Classmate: Money is more important for a good life than health or happiness.

13. Your class is talking about the environment. You have heard that sea levels are rising and that average temperatures are rising.

Classmate: Global warming is a myth.

14. Your classmate has asked you to finish the project for the group.

You say: I'm not saying that I won't do it, just that other people should contribute.

[Screen only: People look confused]

15. You and your group are discussing the environment.

You say: Greenpeace and other NGOs do a lot of work to help the environment.

[Screen only: People look confused]

End of Test

Appendix B. Background Questionnaire

1. Native Language _____
2. Gender (Circle one) Male Female
3. Years studying English (Circle one) 1 2 3 4 5 other _____
4. Did you do all of the MICASE searches in the lessons yourself? (Circle one)

YES NO

If you circled NO, please explain.

5. Did you do additional MICASE searches on your own (not part of the assignments)? (Circle one)

YES NO

If you circled YES, what did you search for?

About the Authors

Kathleen Bardovi-Harlig is Professor of Second Language Studies at Indiana University where she teaches and conducts research on second language acquisition, L2 pragmatics, and tense-aspect systems. Her work on pragmatics has appeared in Language Learning, SSLA, and Intercultural Pragmatics. She is co-editor of *Interlanguage Pragmatics* (Erlbaum) and *Teaching Pragmatics*.

E-mail: bardovi@indiana.edu

Sabrina Mossman is a PhD candidate in Second Language Studies at Indiana University. She has 25 years' experience in ESL teaching and curriculum development in the United States and Mexico. Her work on using corpus materials to teach pragmatics has appeared in *Language Teaching Research*, *TESOL Journal*, and edited volumes.

E-mail: mossmans@indiana.edu

Yunwen Su is a PhD candidate in the Department of Second Language Studies at Indiana University. She has taught EFL in China, and ESL in the US. Her research interests include L2 pragmatics in English and Chinese. She has published in *Foreign Language Annals* and edited volumes.

E-mail: yunwsu@umail.iu.edu