Computer Assisted Language Instruction Consortium

CALICO JOURNAL (ONLINE) ISSN 2056–9017

Article

Mobile Application Use in Technology-Enhanced DCTs

Catherine Rockey[1], Jessica Tiegs[2], and Julieta Fernández[3]

Abstract

In response to calls for improving the quality of DCTs as data-gathering instruments, this study investigates the effectiveness of technology-enhanced discourse completion tasks (TE-DCTs) as a method for eliciting nonverbal speaker data. We used a mobile application to administer four TE-DCTs to native speakers (L1) and intermediate and advanced second-language (L2) speakers of Spanish. Each TE-DCT contained two scenarios with the goal of capturing nonverbal devices used in the speech act of attention-getting (i.e., devices used to draw the interlocutor's attention). The written description of each DCT scenario was supplemented with a short video clip to provide participants with nonverbal factors such as distance to interlocutor, bodily stance, and orientation of interlocutors. To capture nonverbal cues as part of the participant responses, the participant video recorded their oral responses to each scenario. The mobile application used was successful in capturing a variety of attention-getting elements, including nonverbal devices, in the majority of both L1 and L2 participant responses for all DCT scenarios. Drawing on this data, we argue for the use of mobile applications as an ecologically valid way to measure one type of pragmatic ability. In addition, we advocate their integration into L2 pedagogical practice.

KEYWORDS: DISCOURSE COMPLETION TASK; TE-DCTS; ATTENTION-GETTING DEVICES; GESTURES.

Affiliations

[1]The University of Arizona, Tucson, AZ.
email: crockey1007@email.arizona.edu

[2]The University of Arizona, Tucson, AZ.
email: jtiegs@email.arizona.edu

[3]The University of Arizona, Tucson, AZ.
email: julietafernandez@email.arizona.edu

CALICO JOURNAL VOL 37.1 2020 85–108
https://doi.org/10.1558/cj.38773

EQUINOX

equinoxonline

Introduction

One of the most researched areas in pragmatics is speech act performance (e.g., using language to apologize, promise, or make an offer). Based on their functional characteristics, different classifications of speech acts have been proposed in the pragmatics literature (e.g., Austin, 1962; Searle, 1976). A particular type of speech act that has received much attention is requests (e.g., Félix-Brasdefer, 2007; Pinto, 2005), including in technology-mediated contexts (e.g., Halenko, 2013; Schauer, 2009; Sykes, 2008).

The data elicitation methods used in the study of speech act performance in interlanguage and cross-cultural pragmatics vary, with a marked preference for the use of Discourse Completion Tasks or Tests (DCTs) (Asención Delaney & Fernández, 2016), whereby participants are asked to state how they would respond to a scenario developed to elicit a specific speech act (Golato, 2003). The development of new technologies has enabled advancements in the design and delivery of DCTs, which often involve the use of video prompts and/or a request for video responses. As seen in this study, these technology-mediated techniques, categorized as Technology-Enhanced DCTs (henceforth, TE-DCTs) (Culpeper, Mackey, & Taguchi, 2018), offer potential for the collection of a variety of types of pragmatic data.

This study investigates the use of TE-DCTs to elicit nonverbal attention getters—a communicative strategy that is a lesser-studied component of the speech act of requests in Spanish. To that end, we utilized the mobile phone application Flipgrid© to develop TE-DCTs that provided extralinguistic cues, such as physical context, as part of the scenario description. The TE-DCTs also gave participants the opportunity to use nonverbal cues in their video responses. As this is one of the first forays into the use of a TE-DCT to investigate use of nonverbal attention getters in the pragmatics literature, we address the following research questions:

1. To what extent do application based TE-DCTs capture nonverbal attention-getters in L1 and L2 participant responses?
2. What types of nonverbal devices do L1 and L2 participants use when performing the speech act of attention-getting in a TE-DCT?

Technology-Enhanced DCTs

The DCT has historically been the most used data elicitation method for evaluating pragmatic competence, due to its ease of administration, and more importantly the researchers' ability to manipulate contextual variables, such as formality, social distance, and power differences (Billmyer & Varghese, 2000). Despite is prolific use, the DCT has been the object of much methodological

review. As pointed out by Halenko (2016), criticism stems from the inverse relationship that tends to exist between control (over variables measured) and naturalness (of the data) in data collection methods. Indeed, researchers have found that there are differences between the results a DCT generates and patterns observed in naturally-occurring data, such as range of speech act strategies, length of turns, or number of turn construction units (e.g., Beebe & Cummins, 1996; Golato, 2003).

These differences could be attributed, in part, to the relatively few contextual details provided in DCT scenario descriptions (Félix-Brasdefer, 2010). Whereas traditional DCTs provide a written description of the setting, the relationship among the speakers, their social distance and the goals of the interaction, they often lack more extensive language cues, such as extralinguistic features (e.g., gesture or personal distance), which are involved in naturally-occurring speech.

A two-fold improvement on these issues is incorporating different DCT formats that allow for audio-visual supplementation in both the scenario description (i.e., multimedia elicitation tasks, METs) and the participant response (computer-delivered spoken DCTs) (Culpeper, Mackey, & Taguchi, 2018). Prior METs used have included pictures and audio (Schauer, 2009), video clips (Winke & Teng, 2010), computer-assisted interactive prompts (Yang & Zapata-Rivera, 2009) and Computer-Animated Production Tasks (CAPT) (Halenko, 2016). The methodological reasons behind the creation and use of METs vary. For example, Schauer (2009) developed a computer-based role-play MET to control for interlocutor effects (e.g., tone of voice or mood). According to Schauer, her MET ensured "equal conditions for every participant, while at the same time also providing rich audiovisual contextual information" (p. 79). Halenko (2013), on the other hand, compared a CAPT and a traditional written DCT (WDCT) and found that the CAPT, in addition to being perceived by participants as more engaging, provided the opportunity to include prosodic and paralinguistic features of language in the scenario prompt. These forms of TE-DCT prompts may, in turn, result in more detailed participant responses when compared to the "content poor" prompts used in many WDCTs (Billmyer & Varghese, 2000, p. 543).

Overall, using TE-DCT formats offers the potential to overcome some of the challenges of traditional DCTs, such as "lack of audio-visual input, limited authenticity, and neglected spoken features" (Culpeper, Mackey, & Taguchi, 2018, p. 65). An additional methodological advantage of using a TE-DCT (compared to administering an oral DCT in person) is the flexibility of administration: participants can complete the TE-DCT on their own time without the presence of a researcher, in an environment where they (potentially) feel more comfortable and less anxiety about performing.

Attention-Getting Devices

Defined as "external elements which function to draw the interlocutor's attention to a request" (Félix-Brasdefer, 2005, p. 73), the communication strategies under study here have been referred to as "summons" (Schegloff, 1967), "alerters" (e.g., Lorenzo-Dus & Bou-Franch, 2003), "attention-getting devices" (e.g., McCollum, 1980), and "precursors" (e.g., Félix-Brasdefer, 2005). For the present purposes, the act of getting the attention of an interlocutor is considered a separate speech act in and of itself as it is a request for their time and attention. The few studies to examine attention-getting devices as a speech act in both the L1 and L2 pragmatics literature were conducted within the classroom setting (e.g., McCollum, 1980; Cekaite, 2008). These studies suggest that extralinguistic features such as prosody, body posture, and gestures are intrinsically linked to linguistic features in attention-getting scenarios (Cekaite, 2008). It has also been argued that contextual variables, such as the relationship between the interlocutors, have an important bearing on the type of attention-getting device used. McCollum (1980), for example, proposes that deference might play a role in how a speaker chooses to get an interlocutor's attention and suggests that "certain combinations of verbal and nonverbal alternatives which might be chosen to get one's attention are more deferential than others" (p. 16). Given that there are very few studies on attention-getting as a speech act, we also include studies involving precursors to requests, both in L1 and L2 contexts, in our review of attention-getting.

Among the research on precursors to requests, pragmatic variation has been studied in L1 contexts (e.g., Lorenzo-Dus & Bou-Franch, 2003 for Spanish vs. British English; Marquez Reiter, 2002 for Uruguayan vs. Peninsular Spanish) as well as in L2 contexts (e.g., Byon, 2004 for Korean; Félix-Brasdefer, 2007 for Spanish; Tateyama, 2008 for Japanese). The studies focused on L1 pragmatic variation examine verbal precursors with respect to regional/national variety of Spanish. The results of these studies suggest that variation exists in terms of directness, formality, deference, and overall range of available types of attention-getting devices. Relevant findings of L2 studies include that L2 learners differ from L1 speakers in request speech act performance, such as using fewer conventional forms (Byon, 2004) and precursors (Tateyama, 2008). Overall, studies of speech acts, whether in L1 or L2 contexts, have traditionally focused on spoken language more than on nonverbal features of communication.

The linguistic expression of speech acts in general can be highly routinized (Wolfson & Manes, 1980) and requires knowledge of conventional expressions, otherwise known as formulaic language or pragmatic routines (Bardovi-Harlig, 2019); likewise, nonverbal features also appear to be formulaic in the sense that they can be predictably matched to specific social and cultural contexts (Kendon, 1995; McNeill, 1992). While recognizing that a wide variety of strategies exists

for attention-getting and that they are likely to work in concert, the present study focuses solely on the nonverbal devices (e.g., body movements and gestures) used for attention-getting in order to assess the effectiveness of the technology used.

Gesture and Nonverbal Devices

Although some theories such as the "gesture-first" theory argue that gesture develops separately from speech, McNeill (2005, 2013) thoroughly investigates the relationship between gesture and language and regards gestures as "components of speech, not accompaniments but actually integral parts of it" (McNeill, 2013, p. 481). This conceptualization of gesture, also supported by another influential approach (Kendon, 1980), is the framework we draw upon for the present study. The Gesture Continuum (Kendon, 1980) describes types of gestures, ranging from least language-like to most (Figure 1).

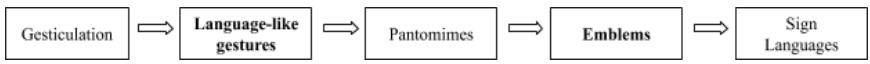

Figure 1. The Gesture Continuum (Kendon, 1980).

At the least language-like end of the continuum, gesticulation is defined as an extra-linguistic embellishment that accompanies language but does not replace it. Pantomimes, situated in the middle of the continuum, are more language-like than gesticulation and are not usually accompanied by speech; however, they do not replace language, but rather reenact narrative events holistically. Sign language, at the opposite end of the continuum, replaces language and does not accompany speech. The types of movements most relevant to this study are language-like gesture and emblems because they are situated in the middle of the continuum—but unlike pantomimes directly in the middle, they can accompany speech and also have context-dependent meaning. Emblems are culturally specific and capable of replacing speech entirely (however they still are able to accompany speech and often do); for example, raising one's hand in class is emblematic of saying, "I have a question", but it is common for a student to also say "Excuse me" or "Professor". On the other hand, language-like gestures are intricately tied to speech but are not independent of it. For example, in a restaurant, a person may turn their body and lean in the direction of a waiter to signal they need something, but this would usually not be sufficient to get their attention on its own—the gesture needs to accompany speech for the full meaning to be conveyed. To disambiguate the terminology, as other representations of gestures are exclusive to hand movements, we refer to bodily movements in general as *nonverbal devices* throughout the rest of our study, using the more specific terms *nonverbal*

attention-getting devices or *nonverbal attention-getters* for only those used for attention-getting.

Methodology

Participants

The participants included a convenience sample of eight L1 and five L2 users of Spanish. The L2 participants were American English L1 students at an estimated intermediate (two participants) and advanced (three participants) L2 Spanish proficiency who were taking Spanish courses at a large university in the United States. Course placement was used as the proficiency indicator, as determined by participants' results in the Spanish WebCAPE Computer-Adaptive Placement Exam, which students responding to the recruitment email took prior to being invited to complete the TE-DCT[1]. Intermediate and advanced students were selected because prior studies have shown that this population makes use of more overall gestures when speaking in their L2 than beginner learners (e.g., Gullberg, 2010; Gullberg, De Bot, & Volterra, 2008; Stam, 2006). The L1 participants were Mexican nationals. They were not studying or working in a language-related field (e.g., language teaching or applied linguistics). In order to control for gesturing that may be related to a specific languaculture (Agar, 1994), all L1 speakers were from the northern Mexican state of Sonora. Given the location of the university (close to the Mexican border), the L2 participants had had some exposure to Mexican culture and a Mexican variety of Spanish, either through formal education or more informal experiences.

Technology-Enhanced DCT Design

In order to investigate the effectiveness of a TE-DCT for eliciting nonverbal devices, we chose to study the speech act of attention-getting, such as using one's hand to signal a waiter or ask a question in class. In calling someone's attention, the physical aspects of the scene, such as the interlocutor's distance from the speaker, bodily stance, and direction of gaze are crucial visual details. A method to provide sufficient detail to the participants and to prompt a richer response was the inclusion of a brief video clip along with the written explanation as part of the scenario description (or MET). This allowed for a visual interpretation of spatial elements and orientation to the interlocutor. To give participants more information as to where they would be at the moment of engaging in the speech act (in terms of distance to interlocutor, bodily stance, and direction of gaze), there were two actors. One actor represented the interlocutor, and one represented the participant (regardless of age or gender, because participants were asked to act as they would, not as the actor would).

The written descriptions were provided in the participants' L1 (i.e., English for Spanish L2 learners and Spanish for the Spanish L1 participants).

In addition to the video-enhanced descriptions, the participants video-recorded responses using their cellphones. This served to capture the non-verbal attention-getting strategies in an ecologically valid manner, given all our participants exhibited no difficulty in accessing and using phone video capabilities (e.g., Facetime; Snapchat).

In order to create and administer the TE-DCT, we chose to use a mobile application because we wanted to focus on the ease and the ubiquity of mobile phone use; however, we recognize that this type of task could also be performed on a laptop or other device. While there are many educational tools available with similar functions, we chose the mobile application FlipGrid© because it has a simple design and is compatible with different smart phone operating systems. FlipGrid© is streamlined and user-friendly for the purpose of uploading video prompts with the option to add text description and its ease for recording video responses. Figure 2 shows the video clip still and written description for DCT scenario 7 from the view of the creator (called "pilot" in FlipGrid©).

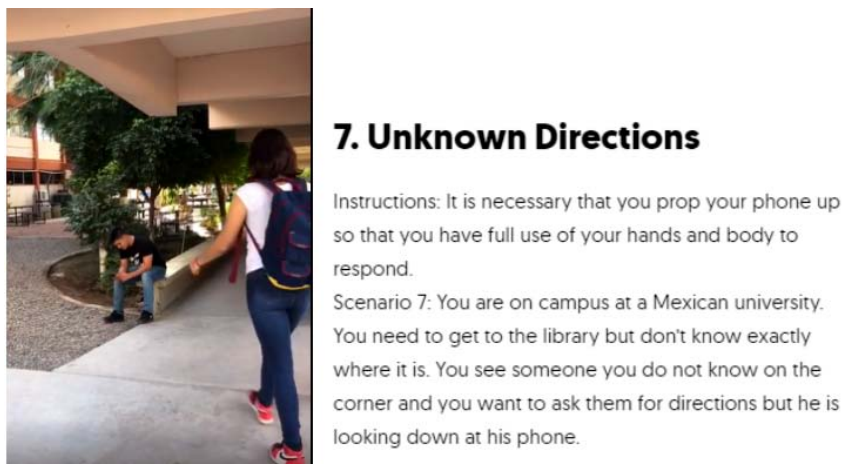

Figure 2. Written description and video clip still of DCT scenario 7, *Unknown Directions*.

Task

Four DCTs were created, with two scenario versions of each (Table 1). These scenarios controlled for social distance (SD)—whether the interlocutor is known or unknown— and power differential (PD) of the interlocutors, based on the interlocutors' relationship to the participant. In the *Classroom* and *Hallway* scenarios, the interlocutor is a professor, giving the scenario the trait +PD. The *Coffee Shop* and *Directions* scenarios are –PD of interlocutor, the

interlocutors being a server and a student in the street, respectively. The *Directions* scenario is depicted in Figure 2.

The clips were meant to set up the situation for the participant to initiate the conversation by getting the interlocutor's attention. Instead of centering the scenarios around the attention-getting task itself (i.e., asking directions from different people vs. asking to make an appointment with different people), the DCT design focused on the physical setting. The scenarios were set in two different contexts, one with more physical distance between interlocutors in a public setting and the other in a more intimate, academic setting (i.e., in a coffee shop vs. inside a small classroom). Although public settings are often considered "noisier", there were no purposeful changes in noise level, nor spoken words in the video clips. To increase the likelihood of use of an attention-getter, in all of the settings, the interlocutors were looking away.

Table 1
Scenarios by Social Distance and Power Differential

	+SD	**–SD**
+PD	• **Unknown professor classroom** (Student, sitting in class, wants to ask new professor, standing at the whiteboard, a question) • **Unknown professor hallway** (Student walking towards new professor in a hallway, wants to ask for an appointment)	• **Known professor classroom** (Student, sitting at desk, wants to ask familiar professor, standing at the whiteboard, a question) • **Known professor hallway** (Student walking towards familiar professor in a hallway, wants to ask for an appointment)
–PD	• **Unknown server coffee shop** (Person, seated in a coffee shop, wants to ask unfamiliar server walking by for a coffee) • **Unknown student public street** (Person walks up to a stranger sitting in a public place to ask for directions)	• **Known server coffee shop** (Person, seated in a coffee shop, wants to ask familiar server walking by for a coffee) • **Known student public street** (Person walks up to a friend sitting in a public place to ask for directions)

Procedure

The participants were instructed to download FlipGrid© on their cellphones and complete the study on their own time. As shown in Figure 3, within the application, participants were provided with a brief description and a 3–5 second video clip that accompanied each of the eight DCT scenarios. They were then instructed to press the green plus sign and record their video responses as naturally as possible. There was no time limit on responses; participants

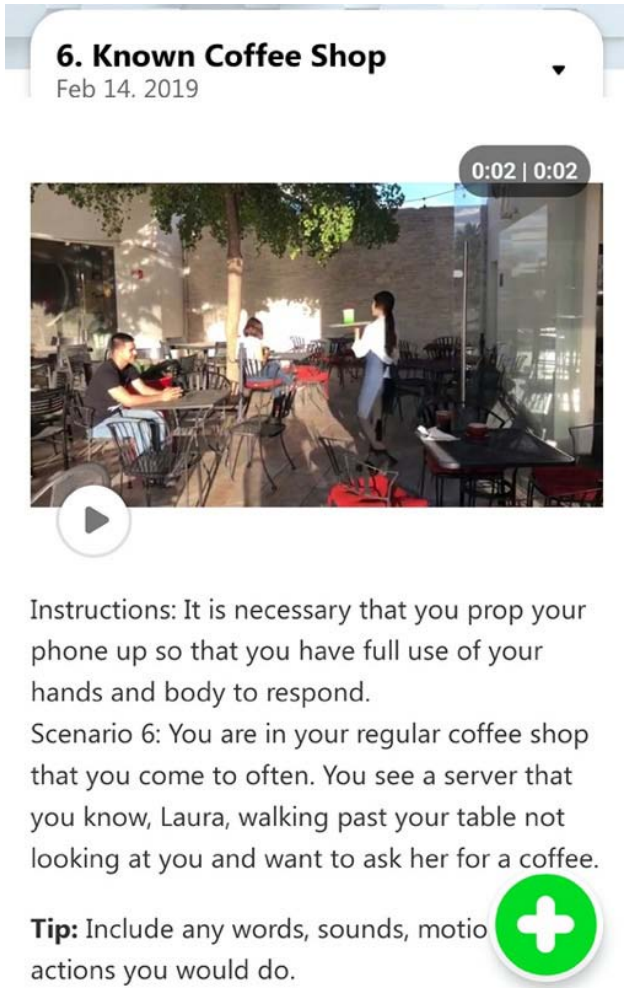

Figure 3. Participant view of DCT scenario 1.

created their video within the application but could restart the recording as many times as they needed to.[2]

Based on the results of a pilot study in which a number of participants held their phones, participants in this study were explicitly instructed to prop up their cellphones in order to have free use of their hands to complete the task. To make sure participants did not focus exclusively on spoken language, a "tip"[3] was included in the instructions of each scenario to "include any words, sounds, motions or actions you [the participant] would do." While this may have alerted some participants to the nature of the study, it was decided that

this was important to assist participants in being as natural as possible, without expressly telling them to use gestures.

Participants were also asked to record a spoken explanation of what they said or did once they had recorded each of their video responses to the DCT prompts. This was meant to act as an immediate recall procedure to collect participants' orientation to the variables considered (i.e., social distance and power differential) and their overall metapragmatic awareness (i.e., knowledge of the social meaning behind the use of their language choices) as related to the individual scenarios.

Coding

The data were imported into the qualitative data analysis software V-note® (Pro Version 2.5.4; Emig, 2019). The unit of analysis was taken to be the entire *gesture phrase* (a series of *gesture units, units* being each phase of one gesture: preparation, pre-stroke hold, stroke, post-stroke hold, retraction) as opposed to analyzing each gesture unit or step of the gesture phrase (McNeill, 1992).

Since, no standardized scheme exists for the coding of gestures (Gullberg, 2010), we decided to use Amory and Kisselev's (2016) categorization of features as the coding scheme that best suited our purposes due to its simplicity and readability. While acknowledging the value of previous coding conventions (i.e., McNeill,1992; Trippel et al., 2014), Amory and Kisselev's coding draws on Jeffersonian conversational analysis, and is thus feature-based, only describing what can be seen, not what a gesture may represent.

Our coding categories were:

- *general movement* (M): a generic hand movement in which the hand shape or wrist movement is unclear;
- *both hands* (2)*:* two-handed movement, one hand being considered default;
- *hand-shape*: all 5 digits extended (5), hand in fist with pointer finger extended (1), and hand closed in a fist (0); and
- *wrist movement* (W): hand movements involving wrist movement or rotation.

In addition to these categories (used to describe hand movements), we also coded *head* (H), *body* (B) and *facial* (F) movements (Table 2). Given that the participants were looking at themselves in their phone's forward-facing camera when recording their answers (rather than interacting with an interlocutor) and that we did not use eye tracking technology, we did not code for gaze even though we acknowledge its role in conversational engagement (e.g., Bednarik, Eivazi, & Hradis, 2012).

Table 2
Coding Categories

Code	Nonverbal attention-getter
M	all hand movements
2	both hands used
0/1/5	hand shape (fist, one finger, open hand)
W	wrist movement/rotation
H	head movement (e.g., tilt, turn, inclination)
B	bodily gesture (e.g., shoulders turning or torso moving)
F	facial movement (eyebrow raise/lower)
O	opt out

An example of the coding process is provided in Figure 4 below. In this case, both researchers coded that the participant used two wrist movements and four facial movements within the span of the approximately 1:15 minute video recording.

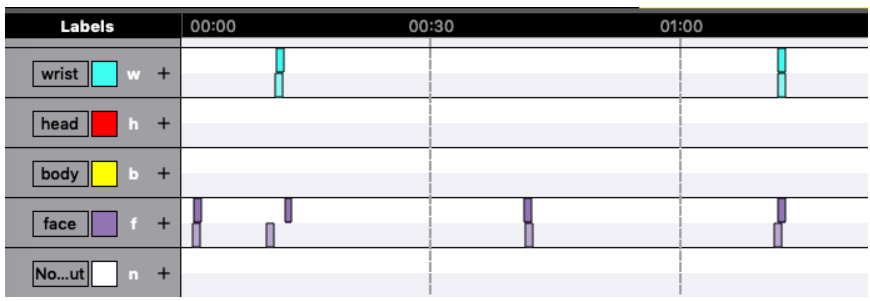

Figure 4. Coding system on V-Note® for Scenario 5.

Two of the authors coded all the data separately using V-note® (Figure 4). Interrater reliability was calculated by dividing the number of overlapping instances (data generated by V-note) by the total number of instances coded (overlap percentage). The original overlap percentage (for tokens of nonverbal devices, not duration) was 80.4%. For the 19.6% of codes that did not coincide, the two raters discussed each discrepancy until an agreement was reached; no data were discarded.

The data generated by the immediate recall task were compiled and analyzed qualitatively through a process of recursive thematic analysis to identify and reduce them to a number of connected themes or ideas.

Results

In response to our first research question—to what extent do mobile application based TE-DCTs capture nonverbal attention-getters in L1 and L2 participant responses? —the percentage of overall participants that used nonverbal devices in their video responses was between 83%–100%, depending on the scenario. Table 3 provides a summary of the percentage of participants who used at least one nonverbal attention-getter organized by scenario. In every scenario, between 11 (84%) and 13 (100%) of the thirteen participants used at least one nonverbal attention-getter and the majority of scenarios prompted 12 respondents to use a gesture (in 92% of the scenarios, 12 of 13 participants used a nonverbal attention-getter). The overall mean of participants that used at least one nonverbal attention-getter across all scenarios was 91.3% with a minimum of 10 of the 13 participants using nonverbal attention-getters in every scenario. One NS participant consistently did not use any nonverbal devices, deviating from this pattern in only one scenario (*Unknown Coffee Shop*, –PD, +SD). This finding is not altogether surprising given that individual and task variation are common in gesture research (Gullberg, 2010). This lack of gesture use is further considered in the discussion section.

Table 3
Average Number of Participants Using Nonverbal Devices (N = 13)

Scenario	Percentage of participants that used nonverbal devices	N
Scenario 1	84%	n = 11
Scenario 2	92%	n = 12
Scenario 3	92%	n = 12
Scenario 4	92%	n = 12
Scenario 5	100%	n = 13
Scenario 6	84%	n = 11
Scenario 7	92%	n = 12
Scenario 8	92%	n = 12

To further elaborate on our response to the first question, we analyzed nonverbal attention-getters used by both L1 and L2 speakers of Spanish. Table 4 lists the number of each nonverbal device coded per language group. Given the sample size, we report descriptive statistics. As there was a different number of participants in each group, a rate of use (number of nonverbal devices/number of participants for each group) was calculated.

Table 4
Use of Nonverbal Devices by Proficiency

Proficiency	hand general	both hands	open hand	one finger	fist	wrist movement	head	whole body	face
L2, n=5	4	9	20	1	3	10	8	3	17
L1, n=8	4	3	19	12	0	5	22	11	38
L2 rate	0.8	1.8	4	0.2	0.6	2	1.6	0.6	3.4
L1 rate	0.5	0.4	2.4	1.5	0	0.6	2.75	1.4	4.75

The L1 speakers used slightly more one-finger gestures (L1 rate: 1.5; L2 rate: 0.2) as well as more facial expressions (L1: 4.75; L2: 3.4), head (L1: 2.75; L2: 1.6) and full body (L1: 1.4; L2: 0.6) movements than did the L2 group. The L2 group used more two- handed gestures (L2 rate: 1.8; L1 rate: 0.4) and wrist movements (such as waving; L2 rate: 2.0; L1 rate: 0.6)). The L2 group was the only group to use a closed-fist gesture (3 different participants, all on scenario 3, *Unknown Professor in the Classroom*).

Regarding our second research question (i.e., what types of nonverbal devices participants use for attention-getting), Table 5 provides a list of the number of each type of nonverbal attention-getting device used for the scenarios across all participants. While the type of scenario did not appear to affect the number of overall attention-getting devices, the types of devices varied. Figure 5 is a graphic representation of the same data, but slightly simplified for readability—all hand gestures are grouped into one category.

Facial expression was the most common nonverbal device used overall (57 tokens). Even participants who used very few hand or body movements tended to include facial movement, particularly eyebrow-raising. Hand movements

Table 5
Number of Total Nonverbal Devices Used by Scenario Type (Two Scenarios for Each Type)

Scenario type	hand general	both hands	open hand	one finger	fist	wrist movement	head	whole body	face	total
+PD, +SD	2	4	9	5	3	2	6	1	14	46
–PD, +SD	0	3	12	4	0	4	6	4	16	49
+PD, –SD	2	5	8	3	0	4	7	3	15	47
–PD, –SD	4	0	11	1	0	6	12	6	12	52
TOTALS	8	12	40	13	3	16	31	14	57	194

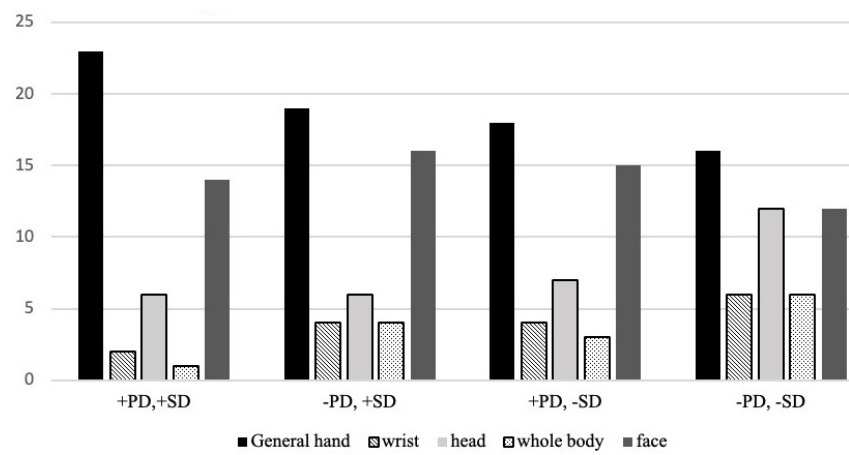

Figure 5. Number of nonverbal devices by scenario.

were also common, with the open hand being the most used for attention-getting (40 tokens). This nonverbal device was used more in scenarios of (–PD) than in the scenarios with a power differential between the interlocutor and the speaker. Head movements were most common in the (–PD, –SD) scenarios

(12 tokens), which were asking a friend for directions and ordering coffee from a known server. Some examples of the types of nonverbal device coded for are included in Figures 4–7, which depict stills from participant responses.

Figure 6. L2, Hand shape 5.

The examples included here are: hand shape 5 (open hand, Figure 6), hand shape 1 (one finger, Figure 7), head movement (Figure 8), and body movement (Figure 9).

To further understand participants' attention to the different contextual factors that influenced their choice, as well as their awareness of the presence and/or role of nonverbal elements in their response, we analyzed participants' qualitative immediate recall answers. Of the 13 participants, six completed this task: five L2 learners and one L1 speaker. Those that provided explanations did so for all of the scenarios. The participants' most frequently mentioned motivations for their speech were familiarity with the interlocutor (39 mentions), formality (22 explicit mentions), and politeness (11 mentions). Participants' linguistic examples of how they navigated the above-mentioned ideas included using the interlocutor's title (16 explicit mentions) (e.g., *profesora* "professor"),

Figure 7. NS, Hand shape 1

verbal attention-getters such as *disculpe* "excuse me" (second person singular formal) or perdón "pardon" (24 explicit mentions) and the *tú/usted* (informal/formal second person singular) distinction (15 explicit mentions).

Importantly, however, participants did not appear to show awareness of their own use of nonverbal devices, or the role of these in the speech act of attention-getting in the immediate recall. Explanations of nonverbal language used were only provided by one participant, who mentioned waving in four different scenarios. Overall, participants did not demonstrate awareness of nonverbal attention-getting devices having communicative weight.

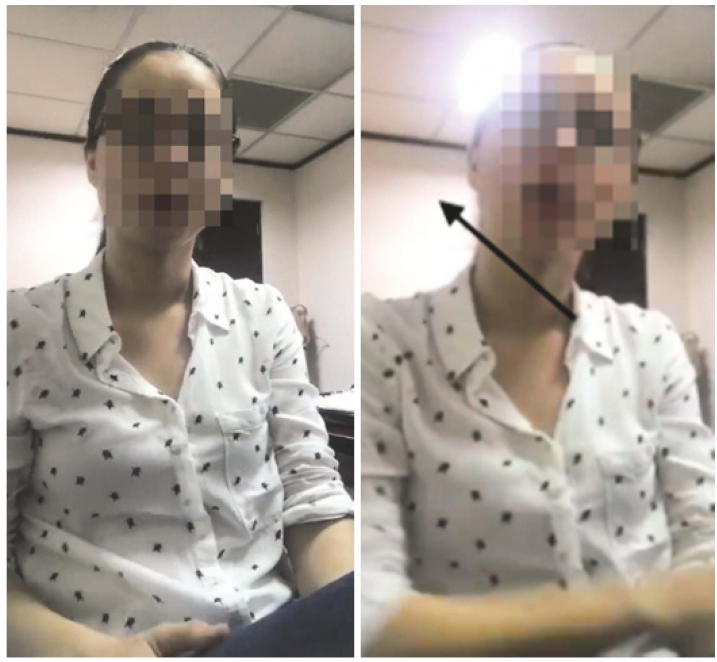

Figure 8. NS, Head movement.

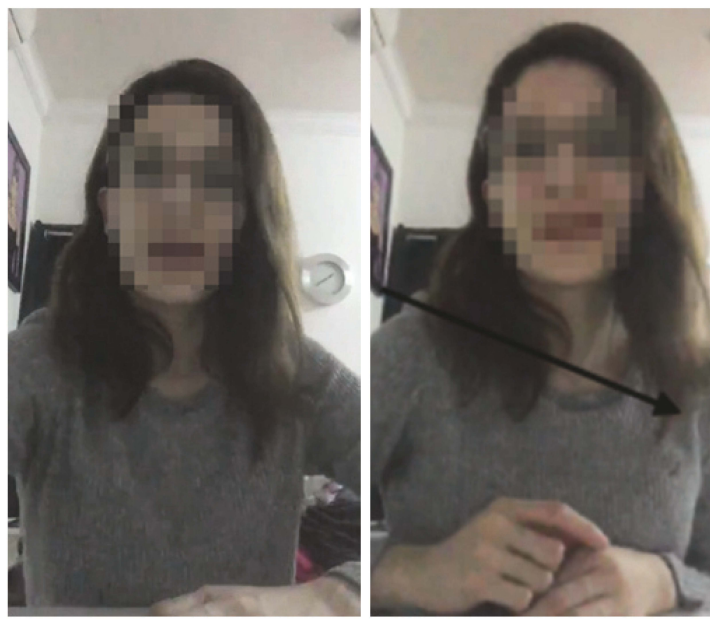

Figure 9. NS, Body movement.

Discussion

The TE-DCT developed for this study proved effective in capturing nonverbal attention-getters, which are an integral, yet often neglected, part of the performance of the speech act of requests. Overall, most participants included at least one nonverbal device with no significant differences between L1 and L2 speakers. All participants used a range of nonverbal devices in the speech act of attention-getting, with a marked preference for the use of facial expressions. This appears in line with prior literature which suggests that over half of distinguishable body movements are located in the region of the head and face (e.g., Burgoon, Buller, & Woodall, 1989). Another potential explanation, with direct connection to the particular technology that we used, is that the person's face is usually the focal point (i.e., where people point the camera to) in a self-video (or "selfie"), and therefore people are more likely to use facial expressions than other types of nonverbal devices, such as head or hand movements.

With respect to the variation found in participants' use of nonverbal attention-getters, the differences in types used according to PD and SD of scenario were minimal, suggesting that, unlike indicated in prior studies (McCollum, 1980), contextual variables such as deference did not play a role in our study. As one reviewer noted, it would be useful to further examine background noise as a factor in the scenarios, given that more nonverbal devices may be used in noisier environments. It was indeed found that a few more nonverbal attention-getters were used in the noisier scenarios (see Table 5). However, since the scenarios with more noise (the public, nonacademic settings) coincide with the –PD scenarios, this would be a confounding variable in our analysis. To confirm whether this is related to noise, a future study would need to isolate this variable to determine its effect on the use of nonverbal attention-getters. In sum, larger studies need to be conducted in order to find the significance of these factors in the act of attention-getting. Based on the results here, TE-DCTs offer one means for conducting this research on a larger scale.

Additionally, it must be acknowledged that other factors, such as individual variation or understanding of task instructions, may influence participants' use of nonverbal devices. While most participants used at least one nonverbal device per scenario, one Spanish L1 participant used only facial expressions and another Spanish L1 participant generally lacked any gestures, facial expressions or changes in intonation at all. The existence of individual variation is expected when analyzing differences in gesture use among participants (Gullberg, 2010). These individual differences in gesture use, both for L1 and L2 users, have been attributed to factors such as "intelligence, language aptitude, memory capacity, attitudes, motivation, personality traits, and cognitive style" (Gullberg et al., 2008, p. 164).

Variation in the amount of nonverbal devices used may also be attributed to participants' understanding and/or careful reading (or lack thereof) of the instructions; even as we attempted to include specific instructions to be hands-free and to use their body, participants did not always carry out the task as anticipated. For example, one participant focused on the instruction to "keep hands free in order to allow full use of your hands," and trained the camera in a way that cut the top of her head out of the frame. Difficulty in following instructions is not uncommon in human research, especially when the researcher is not present. An option would be to provide video instructions for the participant to view as a model, though this might have the unintended effect of priming them for the use of a particular pragmatic structure or nonverbal device.

Similar to what can happen during an in-classroom role-play, there will always be participants who are apt to perform the task more comfortably than others. Student or participant inclination to be a performer creates problems for eliciting ecologically valid responses, regardless if in a classroom or a privately-recorded TE-DCT. In a classroom setting, attention-getting is common behavior (i.e., when a student solicits the teacher's or another student's attention) that does not need to be elicited. Thus, a possible future direction may be to compare DCT responses with non-elicited data collected in a classroom setting to investigate how closely speech act performance in a DCT compares with natural data.

Notwithstanding the difficulties that may arise in terms of participant performance (largely shared with other modalities of DCT elicitation), this method of administering a DCT has several advantages. The use of a forward-facing camera phone is a task already integrated into many of our participants' daily lives (e.g., Facetiming). The administration of a TE-DCT through a phone application also has the potential to reduce some of the demands on participants' language abilities. In face-to-face interactions, participants need to make online use of their oral communication skills. This can be particularly anxiety-inducing for some L2 students. Additionally, by including an audiovisual scenario, the participant does not have to entirely imagine[4] the scenario (as is the case in written or oral formats), thus potentially reducing their cognitive load. Finally, the absence of a physical researcher also reduces the observer's paradox (Labov, 1972), potentially allowing for more ecologically valid data than if the participant were performing in front of the researcher.

Based on our results, we argue that new technologies such as the one used in this study allow researchers to overcome some of the traditional limitations of DCT-elicited data, and capture the nonverbal features of a speech act that would be difficult to study otherwise. This, and other similar technologies, open a path for focusing on the rich extralinguistic features that are commonly understudied in linguistic and pragmatic research.

Conclusion

This study explored the efficacy of TE-DCTs administered through a phone application for eliciting multiple nonverbal devices in responses. Given that including nonverbal devices in a written DCT response is implausible (beyond a written description), we do not compare our results to those of a written DCT. A future study could explore the comparison of naturally occurring data with spoken DCT data elicited using a bot facilitated TE-DCT in which a "NPC" (nonplayer character) interacts with participants.

TE- DCTs have myriad applications for instruction, especially in areas of language use that have traditionally been neglected in the classroom, such as the teaching of L2 pragmatics in general, and the sociopragmatics of nonverbals in particular. Classroom applications of the TE-DCT have the potential to aid instructors in creating a comprehensive language-learning environment that accounts for both linguistic and extralinguistic features and results in a more accessible adaptation of attested language tasks. Integrating TE-DCTs into students' everyday use of phone apps allows for convenience and practicality: students could use a mobile application to complete a video response activity on their own time, in a comfortable setting. Additionally, the incorporation of mobile-application based tasks is beneficial for language learning because it "uses emerging media as (a) meaningful contexts for L2 language development and (b) a means for adding real world relevance to in-class uses of Internet communication tools" (Sykes & Thorne, 2008, p. 529). Mobile applications provide the opportunity for instructors to create tasks that allow students to increase their metapragmatic awareness of the context of particular languacultures while practicing their language use in response to verbal tasks.

While a TE-DCT is a methodological instrument increasingly used in the field of pragmatics, we have shown its applications for the investigation of lesser studied speech acts which involve the use of nonverbal devices, such as gestures. If the goal is to elicit rich data that includes extralinguistic features, TE-DCTs provide the conditions for participants to respond as naturally as possible through a medium (i.e., a cell phone) that they are accustomed to using on a daily basis. This application is an efficacious tool for potentially eliciting rich, comprehensive data.

Notes

1. We are aware of the fact that course level is not necessarily synonymous with proficiency, and we acknowledge this as a limitation of the study.
2. We acknowledge one does not get to perform a summons more than once (unless unsuccessful in prior attempts), but do not consider this to diminish the authenticity of the task, since it is not uncommon for users to use technology to record their "authentic" reactions to

different situations more than once in "real life." It is also consistent with DCT applications in that DCTs generally afford participants the opportunity to deliberate as much as necessary in their responses (as argued by Golato, 2003, they are an offline task that allows for introspection).

3. A "tip" is a brief message that pops up with every task prompt within FlipGrid© to remind students of some piece of information such as instructions.

4. Participants having to imagine themselves in a specific scenario with many factors (e.g., imagined relationship with the interlocutor) remains an issue for elicited pragmatic data which may result in a participant providing different responses on different occasions.

About the Authors

Catherine Rockey is a PhD student in Hispanic Linguistics and graduate associate in teaching and research in the Department of Spanish and Portuguese at the University of Arizona. Her research interests include second language acquisition and teaching and L2 pragmatics.

Jessica Tiegs is a PhD student in Hispanic Linguistics and graduate associate in teaching in the Department of Spanish and Portuguese at the University of Arizona. While her research is primarily in sociolinguistics and phonetics, she is interested in how technology is shaping second language teaching.

Julieta Fernández is Assistant Professor of Hispanic Linguistics in the Department of Spanish and Portuguese and an affiliated faculty member in Second Language Acquisition and Teaching at the University of Arizona. Her research centers primarily on pragmatic dimensions of language use, language learning and technology, and second language learning and pedagogy.

References

Agar, M. (1994). *Language shock. Understanding the culture of conversation*. New York, NY: William Morrow.

Amory, M. D., & Kisselev, O. (2016). Toward a systematic transcription of gesture in conversation analysis research. In J. Edlund, D. Heylen, & P. Paggio (Eds.), *Proceedings from multimodal corpora (MMC): Computer vision and language processing* (pp. 9–12). Portoroz, Slovenia.

Asención Delaney, Y., & Fernández, J. (2016). Spanish speech acts. In C. A. Chapelle (Ed.). *The encyclopedia of applied linguistics* (pp. 1–8). Malden, MA: Wiley-Blackwell. https://doi.org/10.1002/9781405198431.wbeal1483

Austin, J.L. (1962). *How to do things with words*. London, England: Oxford University Press.

Bardovi-Harlig, K. (2019). Formulaic language in second language pragmatics research. In A. Siyanova-Chanturia & A. Pellicer-Sanchez (Eds.). *Understanding formulaic language. A second language acquisition perspective* (pp. 97–114). London, England: Routledge. https://doi.org/10.4324/9781315206615-6

Beebe, L. M., & Cummings, M. C. (1996). Natural speech data versus written questionnaire data: How data collection method affects speech act performance. In S. M. Gass

& J. Neu (Eds.), *Speech acts across cultures* (pp. 65–86). Berlin, Germany: Mouton de Gruyter. https://doi.org/10.1515/9783110219289.1.65

Bednarik, R., Eivazi, S. & Hradis, M. (2012). Gaze and Conversational Engagement in Multiparty Video Conversation: An annotation scheme and classification of high and low levels of engagement. *Proceedings from GazeIn'12 international conference on multimodal interaction: 4th workshop on eye gaze in intelligent human machine interaction*. New York, NY: ACM. https://doi.org/10.1145/2401836.2401846

Billmyer, K., & Varghese, M. (2000). Investigating instrument-based pragmatic variability: Effects of enhancing discourse completion tests. *Applied Linguistics, 21*, 517–552. https://doi.org/10.1093/applin/21.4.517

Burgoon, J., Buller, D., & Woodall, G.W. (1989). *Nonverbal communication: The unspoken dialogue*. New York, NY: Harper & Row.

Byon, A. S. (2004). Sociopragmatic analysis of Korean requests: Pedagogical settings. *Journal of Pragmatics, 36*, 1673–1704. https://doi.org/10.1016/j.pragma.2004.05.003

Cekaite, A. (2008). Soliciting teacher attention in an L2 classroom: Affect displays, classroom artefacts, and embodied action. *Applied Linguistics, 30*, 26–48. https://doi.org/10.1093/applin/amm057

Culpeper, J., Mackey, A., & Taguchi, N. (2018). *Second language pragmatics: From theory to research*. New York, NY: Routledge. https://doi.org/10.4324/9781315692388

Emig, B. (2019). V-Note Video Analysis Software [Computer Software]. Retrieved from http://www.V-Note.org.

Félix-Brasdefer, J. C. (2005). Indirectness and politeness in Mexican requests. In D. Eddington (Ed.), *Selected proceedings of the 7th Hispanic linguistics symposium* (pp. 66–78). Somerville, MA: Cascadilla Press.

Félix-Brasdefer, J. C. (2007). Pragmatic development in the Spanish as a FL classroom: Cross-sectional study of learner requests. *Intercultural Pragmatics, 4*, 253–286. https://doi.org/10.1515/IP.2007.013

Félix-Brasdefer, J. C. (2010). Data collection methods in speech act performance: DCTs, role plays, and verbal reports. In A. Martínez-Flor & E. Usó-Juan (Eds.), *Speech act performance. Theoretical, empirical and methodological issues* (pp. 41–56). Amsterdam, The Netherlands: John Benjamins. https://doi.org/10.1075/lllt.26.03fel

Golato, A. (2003). Studying compliment responses: A comparison of DCTs and recordings of naturally occurring talk. *Applied Linguistics, 24*, 90–121. https://doi.org/10.1093/applin/24.1.90

Gullberg, M. (2010). Methodological reflections on gesture analysis in second language acquisition and bilingualism research. *Second Language Research*, *26*, 75–102. https://doi.org/10.1177/0267658309337639

Gullberg, M., De Bot, K., & Volterra, V. (2008). Gestures and some key issues in the study of language development. *Gesture*, *8*, 149–179. https://doi.org/10.1075/gest.8.2.03gul

Halenko, N. (2013). Using computer animation to assess and improve spoken language skills. *Proceedings of ICT for language learning.* Florence, Italy.

Halenko, N. (2016). *Evaluating the explicit pragmatic instruction of requests and apologies in a study abroad setting: The case of Chinese ESL learners at a UK higher education institution* (Unpublished doctoral dissertation). Lancaster University.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *The relationship of verbal and nonverbal communication* (pp. 207–227). The Hague, The Netherlands: Mouton Publishers.

Kendon, A. (1995). Gestures as illocutionary and discourse structure markers in Southern

Italian conversation. *Journal of Pragmatics, 23*, 247–279. https://doi.org/10.1016/0378-2166(94)00037-F

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Lorenzo-Dus, N., & Bou-Franch, P. (2003). Gender and politeness: Spanish and British undergraduates' perceptions of appropriate requests. In J. Santaemilia (Ed.), *Género, lenguaje y traducción* (pp. 187–199). Valencia: Universitat de Valencia/ Dirección General de la Mujer.

Marquez Reiter, R. (2002). A contrastive study of conventional indirectness in Spanish: Evidence from Peninsular and Uruguayan Spanish. *Pragmatics, 12*, 135–151. https://doi.org/10.1075/prag.12.2.02mar

McCollum, P. (1980). *Attention-getting strategies of Anglo-American and Puerto Rican students: A microethnographic analysis* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (8108603)

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

McNeill, D. (2005). *Gesture and thought*. Chicago, IL: University of Chicago Press. https://doi.org/10.7208/chicago/9780226514642.001.0001

McNeill, D. (2013). The co-evolution of gesture and speech, and downstream consequences. In C. Müller, A. J. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, & S. Tessendorf (Eds.), *Body - Language - Communication: An international handbook on multimodality in human interaction* (pp. 480–512). Berlin, Germany: De Gruyter Mouton.

Pinto, D. (2005). The acquisition of requests by second language learners of Spanish. *Spanish in Context, 2*, 1–27. https://doi.org/10.1075/sic.2.1.01pin

Schauer, G. (2009). *Interlanguage pragmatic development: The study abroad context*. London, England: Continuum.

Schegloff, E. (1967). *The first five seconds: The order of conversational opening* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (302223054)

Searle, J. (1976). The classification of illocutionary acts. *Language in Society, 5*(1) 1–23. https://doi.org/10.1017/S0047404500006837

Stam, G. (2006). Thinking for speaking about motion: L1 and L2 speech and gesture. *International Review of Applied Linguistics, 44*, 143–169. https://doi.org/10.1515/IRAL.2006.006

Sykes, J. (2008). *A dynamic approach to social interaction: Synthetic immersive environments and Spanish pragmatics* (Unpublished doctoral dissertation). University of Minnesota, Minneapolis.

Sykes, J. M., & Thorne, S. L. (2008). Web 2.0, synthetic immersive environments, and mobile resources for language education. *CALICO Journal, 25*, 528. https://doi.org/10.1558/cj.v25i3.528-546

Tateyama, Y. (2008). *Learning to request in Japanese through foreign language classroom instruction* (Doctoral Dissertation). Retrieved from ProQuest Dissertations and Theses. (3347676).

Trippel, T., Thies, A., Milde, J.-T., Looks, K., Gut, U., &, Gibbon, D. (2014). CoGesT: A formal transcription system for conversational gesture. In Proceedings of LREC 2004. Lisbon, Portugal: ELRA.

Winke, P., & Teng, C. (2010). Using task-based pragmatics tutorials while studying abroad in China. *Intercultural Pragmatics, 7*, 363–399. https://doi.org/10.1515/iprg.2010.016

Wolfson, N. & J. Manes (1980). The compliment as a social strategy. *Papers in Linguistics, 13*, 391–410. https://doi.org/10.1080/08351818009370503

Yang, H. C., & Zapata-Rivera, D. (2009). An exploratory study into interlanguage pragmatics of requests: A game of persuasion. *ETS Research Report Series, 2009*(1), 1–28. https://doi.org/10.1002/j.2333-8504.2009.tb02170.x