Designing interactive, automated dialogues for L2 pragmatics learning

Veronika Timpe-Laughlin[1], Keelan Evanini[1], Ashley Green[1], Ian Blood,[1] Judit Dombi[2] and **Vikram Ramanaranayan[3]**

[1]Educational Testing Service, Princeton, NJ, USA
[2]University of Pécs, Pécs, Hungary
[3]Educational Testing Service, San Francisco, CA, USA

vlaughlin@ets.org

Abstract

This paper describes an approach to designing interactive, automated dialogues for L2 pragmatics learning. It first outlines advantages and challenges of using automated multi-turn conversations to help learners practice pragmatic moves. In order to deal with a particular challenge–excessive variability in users' pragmatic performances–an interactive dialogue aimed at eliciting requests was deployed via a crowdsourcing platform. A total of 328 completed conversations, with both L1 and L2 English speakers, were collected and analyzed with regard to number of turns and requests as well as request strategies elicited in the conversations. Requests were coded based on head acts as direct (D), conventionally indirect (CI), and hint (H). The results revealed interesting patterns in both L1 and L2 speaker responses. For example, even though they were speaking to the same interlocutor, L1 speakers tended to use different request strategies for two distinct requests, dependent on the interaction sequence and prompts within the dialogue. Moreover, further culture-specific variability was identified. Finally, the implications of the findings for the design and use of systematic feedback on pragmatics in computer-assisted language learning applications is discussed.

1 Introduction

While the ability to communicate effectively and appropriately (i.e., pragmatic competence) is critical in general, it is particularly crucial in workplace contexts. For instance, pragmatic failure has been identified as a major cause of communication breakdown in workplace environments (Clyne, 1994). Moreover, pragmatic failure—unlike grammatical mistakes—has been shown to create negative impressions about the speaker (Thomas, 1983; Timpe, 2013; Washburn, 2001) insofar as many interlocutors do not recognize pragmatic infelicities as a language deficiency, but rather attribute pragmatic violations to the character of a speaker, perceiving them as impolite, crude, or direct. For example, (Holmes, 2000) interviewed employers about migrant workers in New Zealand. Although employers agreed that the workers had sufficient second/foreign (L2) abilities to perform their job, they highlighted that "they seem unfriendly or uncomfortable at work; they don't seem to fit in smoothly" (p. 9). Hence, pragmatic infelicities and the lack of pragmatic awareness are oftentimes major reasons for unsuccessful communication—especially when speakers involved in a communicative encounter do not share the same language and/or cultural background. However, despite potentially serious, high-stakes consequences, the inclusion of pragmatics in instructional materials, especially for Workplace English, is still very limited; this may leave English language learners either unaware of or ill-prepared for pragmatic challenges in the English-medium workplace.

In this study, we report on one aspect of a large-scale project that aims to design a self-access, interactive learning platform intended to help adult adult learners of English systematically raise awareness of pragmatic phenomena in the context of the English-medium workplace in the United States. Given the culture-dependency of pragmatics the tool focuses on one particular variety: American English pragmatics–a feature that may make the learning tool interesting and useful for for speakers of other varieties of English as well. The computer-delivered learning tool simulates the interrelated steps of a real-life career, starting with a *Job Hunt*, followed by a *Job In-*

terview, the first day on *The New Job*, up through the development of a regular *Job Routine*. Embedded in this scenario structure are nine learning modules, each of which focuses on a specific pragmatic phenomenon or speech act that is important for successful communication in the workplace such as requests, small talk, apologies, etc. A specific focus within the overall approach to designing this capability was the development of interactive speaking tasks for each learning module that deploy a spoken dialogue system (SDS) technology and allow L2 learners to engage in talk-in-interaction.

2 Background

2.1 Language learning using spoken dialogue systems

The multi-turn conversation items operationalized by means of an SDS offer a number of advantages for practicing and assessing L2 pragmatics in interaction. First, researchers in the field of L2 pragmatics have repeatedly highlighted the need for more use of pragmatics within discourse both for teaching and assessment (Kasper, 2006; Roever, 2011)—a capability provided by the SDS-based dialogues. Second, the automated SDS provides a low-stakes environment for practice. That is, learners can engage in the dialogues without running the risk of embarrassment when making mistakes. Third, they can practice anytime and anywhere they can access the internet. They do not need to find another human-being if they want to engage in a conversation and use English. Fourth, in contrast to L1 speaker interlocutors who tend to refrain from directly responding to pragmatic infelicities in a face-to-face conversation, SDSs provide the opportunity for systematic feedback implementation, thus making the learner aware of pragmatic violations. Finally, SDSs provide environments that allow for the operationalization of a number of principles that have been identified as key to effective L2 pragmatic pedagogy. They (a) allow for the design of dialogues that have a specific pragmatic focus or objective-orientation, (b) provide learners with enhanced, authentic, and relevant input, (c) promote their observational and reflective skills, (d) provide learner-oriented opportunities for interaction and practice, and (e) offer feedback and assessment (Limberg, 2016; Sykes and Cohen, 2008; Timpe-Laughlin, 2016)). Hence, SDSs constitute a beneficial environment that provide interactive activities, structured and scaffolded in ways that maximize noticing and awareness of the form-function-meaning relationship.

Due to the challenge of obtaining accurate automatic speech recognition (ASR) and semantic understanding results for open-ended spontaneous speech produced by L2 speakers, many interactive computer-assisted language learning applications have elicited restricted speech from the learners and have limited their feedback to pronunciation (Su et al., 2013, for example); however, some studies have attempted to automate the process of providing feedback to language learners about aspects of language proficiency that rely on accurate ASR, such as grammar (Morton and Jack, 2005; Lee et al., 2014; Baur, 2015) and even pragmatics (Bernstein et al., 1999; Johnson and Valente, 2009). This study extends on these previous efforts by investigating in detail how users respond to an interactive, dialogue-based language learning application that elicits a particular speech function (namely, requests) and what type of pragmatic strategies are employed.

2.2 Requests

A particular pragmatic phenomenon that tends to constitute a challenge for L2 learners due to its face-threatening potential are requests. Categorized as directives (Searle, 1969), requests are generally defined as "attempts by the speaker to get the hearer to do something" that benefits the speaker (Searle, 1979, p. 13). According to Leech (2014, p. 135), requests can be verbalized in a variety of ways along a "continuous scale of optionality" that ranges from direct requests that hardly leave the hearer a choice for non-compliance to very indirect requests (i.e., hints) that provide the hearer with an increasingly greater choice to refuse compliance. Along this continuum of optionality or indirectness, requests head acts have been classified according to three levels of directness (Blum-Kulka et al., 1989; Trosborg, 1995): (a) direct strategies (e.g., *Please clean up your room, Martin.*), (b) conventionally indirect strategies (e.g., *Could you clean up your room, Martin?*), and non-conventionally indirect requests or hints (e.g., *It looks like a bomb exploded in here, Martin.*). With regard to request use, Leech (2014, p. 134) noted that English "exhibits a tendency to favor indirectness of requests more than other

languages, indirectness being closely connected to politeness". Hence, Leech (2014) as well as others (Brown and Levinson, 1987, for example) have argued that higher levels of indirectness result in higher levels of politeness. However, Blum-Kulka (1987) mediated Leech's stance, arguing that in order to be polite every speaker has to strike a balance between pragmatic clarity and avoiding coerciveness. That is to say, while more direct strategies tip the balance toward being more coercive and thus impolite, hints may result in unclear messages which may also be perceived as impolite given that they violate the cooperative principle of clarity. In the following, we will describe the development of the dialogue that aims to elicit requests—the focus of this study.

3 The study

3.1 The dialogue task

The dialogue was couched in a task-based design. Accordingly, the learners received instructions that provided the needed contextualization for the task, featuring a clearly-defined interlocutor as well as goals that are to be achieved in the conversation. Given that the ultimate objective is to implement the dialogue task into *The New Job* unit of the pragmatics learning tool, the task features one of the interlocutors from the learning tool–the boss, Lisa Green. The following instructions were provided before learners engaged in the conversation.

> *Imagine that you are calling your boss, Lisa Green. Your goals are to (1) get her to agree to have a meeting with you and (2) ask her to review the presentation slides that you made so that you can discuss them at the meeting. Your schedule is free for the rest of the week so any time proposed by Lisa will work for you.*

Given that the SDS requires prompts that can be generated as responses to what users say during the conversation, any dialogue needs to be carefully conceptualized before it is implemented into a SDS.

Thus far, two iterations of the request dialogue have been developed and deployed in HALEF, an open-source, modular, web-based framework for designing and deploying SDS tasks (Ramanarayanan et al., 2017). As a first step, a team of pragmatics and natural language processing (NLP) experts conceptualized a short unbranched dialogue that was intended to elicit two different requests in line with the task instructions presented above. Table 1 below shows this initial version of the dialogue, featuring Lisa Green's turns (in italics). Lisa's turns, also referred to as "dialogue states", were unbranched and thus fixed in the initial version. Additionally, Lisa's turns were recorded by a voice actor in order to provide the intended intonation. In contrast, the user turns (T1-T5) are responses obtained from the study participants who called in and engaged in the conversation with Lisa. The notes featured in the brackets (see column labeled "Output" in Table 1) constitute the types of responses that we anticipated from the users when conceptualizing the dialogue. Request made by users, for instance, were anticipated in T2 (request for a meeting) and T4 (request to review the slides).

Once conceptualized, the dialogue was implemented using the OpenVXML design tool in HALEF (see Figure 1) and deployed via Amazon Mechanical Turk in order to obtain first insights into how users navigated the task. The responses collected in March 2016 were then used in order to refine the initial unbranched version of the dialogue, thus accounting for variability in user responses. For example, we observed that requests were not only made in T2 and T4 as anticipated, but that users made responses across all turns, oftentimes even combining both requests in one turn. As a result, the branching was implemented in an attempt to raise the authenticity of the system's responses, thus increasing the perceived naturalness of the interaction.

The system was set up in order to account for the variability as to where (i.e., in which turn) requests were made, responding accordingly based on semantic tokens identified in the respective utterance. If the ASR did not detect certain semantic tokens (e.g., *meeting*, *meet*, *slides* or a combination thereof), the system would repeat the prompt and thus expand the number of turns. While variability in turn count can be understood as a preliminary indicator of proper system performance, it also provides insights into user behavior—the focus of this study.

3.2 Research questions

The following research questions guided the analyses, aiming at investigating the dialogue re-

Dialog state (Turn)	Interlocutor	Response
Hello (T1)	Lisa Green	*Hello?*
	User	[greeting]
How (T2)	Lisa Green	*Hi, how's it going? What can I do for you?*
	User	[(positive sentiment) + request for meeting]
Friday (T3)	Lisa Green	*Yeah, sure I'm available on Friday at 12. Does that work for you?*
	User	[positive response]
Anything (T4)	Lisa Green	*Was there anything else you needed?*
	User	[request to review slides]
Sure (T5)	Lisa Green	*Sure, no problem. Send them over.*
	User	[expression of thanks]

Table 1: Request dialog template

sponses, particularly focusing on pragmatic phenomena elicited by the branched version of the task:

1. Based on the request moves, do the elicited dialogues differ in terms of length?

2. Where in the interactive dialogue do users tend to make the requests?

3. Which request strategies are being used in the interactive dialogue?

4. Do request moves differ between L1 and L2 speakers of English?

4 Methodology

4.1 Procedure

For the branched version, data were collected via the Amazon Mechanical Turk crowdsourcing platform from February to April 2017. Figure 2 below features a screenshot of the instructions and web-based video-telephony interface that participants saw during data collection.[1] A picture of Lisa Green was featured on the right-hand side of the screen and respondents would see themselves by means of their webcam on the left-hand side.

After completing the dialogue with Lisa Green, participants were asked to answer a background questionnaire, providing demographic information as well as feedback about their experience interacting with the SDS.

4.2 Participants

Out of a total 534 received calls[2], 328 calls were "complete calls," that is, calls that contain the *Sure* dialog state which is designed to prompt the final user response. Out of the 328 participants who completed the dialogues, 162 completed the background questionnaire, thus limiting the number of user responses for the analysis of L1 versus L2 speaker request behavior. Participants indicating English as their L1 were exclusively from the United States. Participants who completed the questionnaire reported a range of L1s, including U.S. English (n=108), Hindi (n=13), Malayalam (n=7), Tamil (n=7), Spanish (n=6), Urdu (n=5), Telugu (n=4), Bengali (n=2), Filipino (n=2), Arabic (n=1), Greek (n=1), Gujarati (n=1), Kannada (n=1), Korean (n=1), Portuguese (n=1), Russian (n=1), and Slovakian (n=1). Similarly, respondents varied in terms of age, ranging from 18 to approximately 60 years of age with the majority of callers indicating that they were between 22 and 29 years old.

4.3 Analyses

The data were transcribed verbatim. Each dialogue state was annotated for (a) request type (request-meeting versus request-slide review) and (b) request strategies which were coded based on the requests' head acts as direct (D), conventionally indirect (CI), and hint (H). The coding was conducted by two annotators. Intercoder reliability was calculated for 20 randomly selected calls (i.e., a total of 94 turns). The obtained simple agreement was 94.70 percent and the quadratic

[1] A prototype version of this SDS task is available for demonstration purposes at `http://englishtasks.org/`.

[2] These include calls during which people did not say anything, hung up before completing the dialogue, or when the system encountered a technical difficulty.

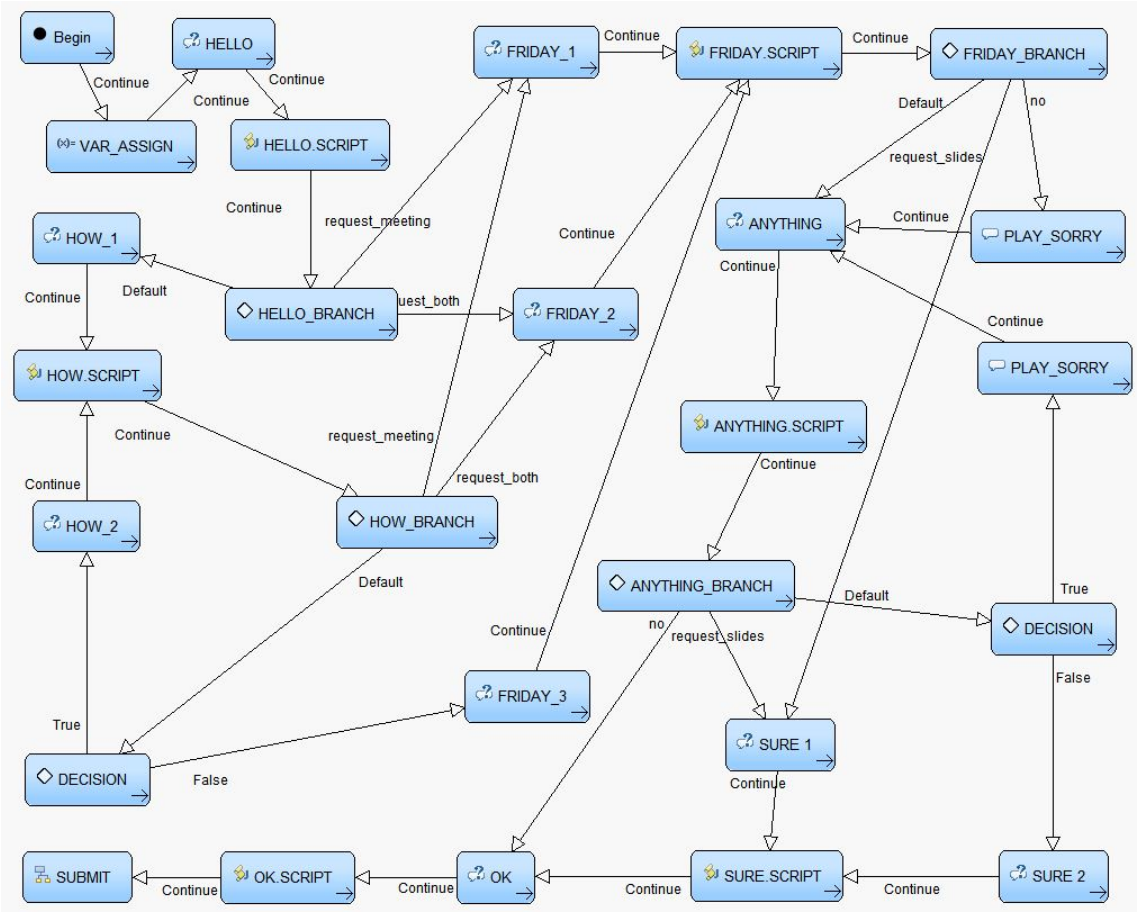

Figure 1: Flowchart of the branching version of the request dialogue in OpenVXML

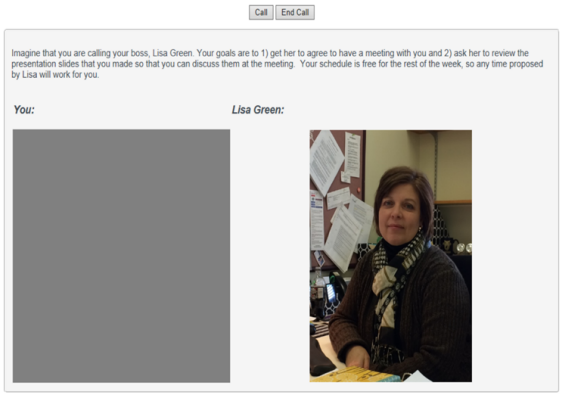

Figure 2: Screenshot of the web interface of the task

weighted kappa value was .86. Discrepancies in the codings were the result of slight misunderstandings with regard to the category "hint" which were resolved in a subsequent consensus coding. Based on the codings, frequency counts were tabulated to analyze the collected responses with regard to number of turns, requests, and request strategies elicited in the conversations.

5 Results

Taking a progressively fine-grained approach in the analysis, we first counted the number of participant turns per dialogue for all complete calls (n=328) as well as for the two subgroups of calls completed by L1 (n=108) and L2 (n=54) speakers of English respectively (Table 2 below). The number of different turns is a result of the branching within the dialogue, suggesting that the branching seemed to work when deployed operationally. Overall, the majority of dialogues featured between four to six turns. Moreover, L1 English speaker dialogues had on average fewer turns than dialogues completed by L2 English speakers. However, an independent sample t-test showed that this difference was not statistically significant ($t(10)$=1.481, p=.169).

The following examples show three distinct cases: a 5-turn, a 4-turn, and a 3-turn dialogue. All examples show the turns taken by male L1 English speakers (turns taken by the system omitted). The 5-turn dialogue shown in Table 3 features a

# Turns	Total (*n*=328)	L1 Eng. (*n*=108)	L2 Eng. (*n*=54)
2	1.8%	1.9%	2.0%
3	10.4%	15.7%	5.6%
4	26.5%	30.6%	16.7%
5	32.3%	29.6%	38.9%
6	19.5%	18.5%	16.7%
7	9.5%	3.7%	18.5%

Table 2: Turn counts for complete calls

participant response that is very much in line with the underlying, anticipated 5-turn schema. Participant ID176 makes the request for a meeting in T2, following the *How* dialogue state, then provides a positive response to the suggested meeting time, before making the request for a review of the slides in T4, immediately after the prompt embedded in the *Anything* dialogue state. By contrast, participant ID172 in T3 combines the acknowledgment of the suggested time with the request for the review of the slides (Table 4). In yet another pattern, participant ID166 (Table 5) makes the request for a meeting immediately in T1 and makes the request to review the slides in the *Friday* dialog state in T2, thus leading to a 3-turn dialogue, since the *How* and *Anything* dialogue states were bypassed.

Turn	Utterance
T1	Hello?
T2	Yeah hi, um I'd like to set up a meeting.
T3	Yeah that's fine. That sounds good for me.
T4	Yeah could you review the uh presentation slides that I made before we meet?
T5	All right, I'll send them in right away.

Table 3: Sample 5-turn dialogue (ID176, male, L1 English)

Hence, although the three examples all feature male speakers who identified English as their L1, there was quite a bit of variation noted in terms of how and where participants made the requests.

As a second step, we counted the number of requests per dialogue state (see Table 3 below). For instance, we observed 282 instances of requests for a meeting, 288 instances of requests to review the slides, and 21 instances of both requests made together. Overall, we found that requests were made across all dialogue states. For example, following the *Hello* dialogue state, we observed 30 instances of requests for a meeting (e.g., *Hello, I am calling to schedule a meeting.*, ID138), but no instance in which participants had asked only for the review of the slides. However, in 5 cases participants made both requests together (e.g., *Hi Lisa, it's Lina. Um I was wondering if you're available this week to have a meeting. I'd like you to uh review my presentation slides uh beforehand. If you have a chance, let me know if you're free this week. My schedule is pretty open. Uh so let me know if you'd be interested in doing that.*, ID305). While these two examples already provide preliminary insights into request variability per turn, the pattern shows that, as anticipated in the dialogue design, the majority of meeting requests were made in T2 after the *How* dialogue state (62.6%), whereas most requests for the slide review were made in T4 following the *Anything* dialogue state (56.5%).

Turn	Utterance
T1	Hello yes, um this is Lisa?
T2	Um yes I would like to set up um a meeting with you some time this week.
T3	Friday on twelve would work great. Um I also would like to ask you if you could review uh the presentation slides that I'm, I'll send you before the meeting.
T4	Okay great. Um let's see uh. That's, that's, that's everything I need then. Thank you.

Table 4: Sample 4-turn dialogue (ID172, male, L1 English)

Turn	Utterance
T1	Hi, can we have a meeting?
T2	Yeah that works for me. Do you mind reviewing my presentation slides first?
T3	All right I will send them to you.

Table 5: Sample 3-turn dialogue (ID166, male, L1 English)

As a third step, we provided frequency counts of the request strategies (see Table 7), distinguishing requests in terms of their head acts according to direct requests (D), conventionally indirect requests (CI), and non-conventionally indirect requests/hints (H). Contrary to the expectation that requests to a person in a higher position of power

State	*n*	Mtg.	Slides	Both	None
Hello	285	10.5%	0%	1.8%	87.7%
How	396	62.6%	0.8%	3.3%	33.3%
Friday	310	0.3%	40.3%	0.7%	58.7%
Anythg.	283	0.7%	56.5%	0.4%	42.4%
Sure	328	0.3%	0%	0%	99.7%

Table 6: Frequency of requests per dialogue state

(+P) would be worded in a polite form, employing more (conventionally) indirect requests, it can be noted that most meeting requests were made in form of direct requests—a finding that may be due to the direct nature of the prompt embedded in the *Hello* dialogue state (*What can I do for you?*). Additionally, the request to review the slides is not only the second request users are to make, but it also has a potentially higher imposition associated with it because reviewing slides may take up more of someone's time than scheduling a meeting.

Finally, in order to explore potential differences between L1 English speakers and L2 speakers, we investigated request strategies for each group of participants. As shown in Table 8 below, there was a trend among L1 speakers of English to use direct request strategies for the meeting request and conventionally indirect request strategies for the request to review the slides—again, the latter may have more of an imposition associated with the request which would require a more indirect and polite wording. By contrast, L2 English speakers were found to primarily use direct request strategies for both requests. A chi square test of independence was performed to further examine the relation between native English background and directness of request strategy used. The relation between these variables was significant, $(2, N = 311) = 20.65$, $p = .000$. We speculate that this trend in the L2 English sample may be explained by the greater use of direct strategies by less proficient language users, whereas indirect ones were used by more advanced speakers who (a) have the linguistic repertoire to express CI requests and (b) are familiar with conversational conventions in English.

Overall, the dialogue elicited requests in a number of different turns with a variety of request strategies employed by participants. Despite the variability in user performances, distinct patterns and trends emerged in the data. For example, NS dialogues were on average slightly shorter than NNS ones. Moreover, NS seemed to prefer direct requests when responding to a direct prompt. That is to say, NS used more direct requests for making meetings than for requesting the review of the PPT slides. Hence, despite variability these observed trends can inform further development of the automated dialogues, including the implementation of feedback with regard to the use of certain pragmatic moves at particular turns.

6 Discussion and conclusion

Taking everything into account, the conversations elicited by the branched dialogue structure were found to elicit the intended speech act in various ways—a variability that is typical of real-life talk-in-interaction. While this variability is common in real-life, human-to-human communication, it poses a number of challenges for SDS technology which can be addressed by means of an empirically-driven development approach that analyses linguistic and in this case pragmatic phenomena in order to reveal trends in speaker behavior. These patterns and trends can then be used to inform the next steps in the development process. Further advancing the capability, we aim to account for the variation in SDS system responses in order to provide a more authentic user experience while also implementing feedback for users.

In this study, we looked in more detail at the number of turns, requests, and request strategies—both for L1 and L2 English speakers—that were elicited by the first iteration of the branched dialogue. As shown in Table 2, we found considerable variability in dialogue length based on variability of request moves—both within and across L1 groups. Although this indicates successful branching and system performance based on where users made the requests, we also found L1 dialogues to be slightly shorter, containing on average fewer turns than L2 dialogues. A closer look at the elicited responses showed that this finding might in part be due to challenges with the ASR and/or semantic understanding of the SDS as exemplified in the example in Table 9, where appropriate requests were identified as off-topic, leading to a re-prompt by the system.

Although this type of re-prompt was found in only very few cases, it constitutes an issue for further investigation in order to rule out bias that may consist in terms of semantic understanding relative to different accents.

Dialog state	*n*	Meeting			Slides		
		D	CI	H	D	CI	H
Hello	40	35%	50%	2.5%	12.5%	0%	0%
How	277	58.8%	31.4%	4.0%	3.6%	2.2%	0%
Friday	130	0%	2.3%	0%	26.2%	63.1%	8.5%
Anything	164	0.6%	1.2%	0%	45.1%	50.6%	2.4%
Sure	1	1%	0%	0%	0%	0%	0%

Table 7: Frequency of request strategies per dialogue state

Strategy	L1 English			L2 English		
	Meeting (*n*=107)	Slides (*n*=108)	Total (*n*=215)	Meeting (*n*=47)	Slides (*n*=49)	Total (*n*=96)
D	52.3%	21.3%	36.7%	74.5%	53.1%	63.5%
CI	43.9%	75.9%	60.0%	23.4%	40.8%	32.3%
H	3.7%	2.8%	3.3%	2.1%	6.1%	4.2%

Table 8: Request strategy type per L1 group

Dialog state	Utterance
Anything	Yes, uh so I want you to review my presentation slides before the meeting. Can you do that? I want you to review my presentation slides before the meeting.
Anything	I want you to review presentation slides before the meeting.

Table 9: Example of potential ASR challenge (ID233, L2 English)

With regard to feedback implementation, post-hoc feedback could be provided for the average number of turns based on the analysis of a large corpus of callers (one that would need to be larger than the one used in this study)–a step planned for future iterations in the development cycle. The trends revealed in such a large corpus could be used to establish a benchmark and provide formative feedback, making users aware how many turns speakers take on average in order to complete the task.

Somewhat interconnected to the variability in dialogue length, we found that users made requests across all dialogue states; however, these requests differed in terms of request strategies employed. As shown in Table 6, there was a clear trend supporting the original dialogue structure insofar as meeting requests were mainly elicited in the early turns (T1 and T2) and requests to review the slides later on in the conversation (T4 and T5). Within this larger pattern, L1 user responses showed a preference for direct request strategies to make the request for a meeting, while strongly favoring conventionally indirect request strategies when making the second request in terms of asking for a review of PPT slides. By contrast, L2 callers heavily relied on direct strategy use throughout.

These patterns highlight two interesting aspects in terms of feedback implementation and L2 speaker responses. With regard to feedback, it highlights the need to investigate responses from representatives of the target language and culture before making any decision regading appropriateness. That is to say, most native speaker judgments as well as textbooks (*if* they deal with pragmatic phenomena) recommend the use of conventionally indirect strategies as a polite means of making requests to a superior in the workplace. However, the data clearly show that such a blanket recommendation may not always be applicable. To make the meeting request in the dialogue L1 speakers used direct strategies even more frequently than CI request strategies. This finding may be explained by the direct question in Lisa Green's prompt (*What can I do for you?*), a direct question which seems to require a clear, concise, and direct response. Overall, this finding emphasizes the importance of the interaction sequence within the dialogue. That is, adjacent turns also need to be taken into consideration when determining appropriateness of a given pragmatic move. Hence,

pragmatics feedback cannot be provided a priori without consideration of the local context.

In addition to the L1 responses which may be used to systematically design and implement feedback, it is also important to consider the patterns in L2 speaker responses. The reason that L2 English speakers used primarily direct request strategies could be due to their lower English language proficiency. For example, direct requests in form of a command that uses an imperative is grammatically less challenging than a complex question format that uses modal auxiliaries (e.g., *I was wondering if you* + past tense). In addition to the more general issue of lower L2 proficiency, L1 culture-specific transfer could also play a role in L2 speaker's pragmatic moves–a critical issue especially with regard to providing learner-specific feedback which could be implemented into the system's dialogue state. Feedback could be implemented into Lisa Green's responses to provide input to users with regard to appropriate pragmalinguistic realizations. Thus, Lisa Green could be offended by the direct request for the slide review or confused if a hint is used which violated the principle of pragmatic clarity (Blum-Kulka, 1987).

However, further analyses, especially with regard to (culture-specific) variation within request strategies, still need to be conducted in more detail in order to allow for an even more fine-grained feedback adaptation and implementation. That is to say, request strategies were only categorized based on their head acts according to the three broad categories of direct, conventionally indirect, and hinted requests. However, a first glance at the data revealed considerable variation with regard to internal and external modification devices such as syntactic and lexical downgraders as well as supportive moves like grounders and disarmers. For example, a trend we noticed in the sample—primarily among speakers from India—was the use of a direct strategy in combination with mostly lexical and phrasal downgraders as shown in the following examples: *I want to, I want to meet you madam.* (ID87), *Uh please tell me which time you are available for me madam.* (ID201) or *Good morning madam. Uh I want to meet you madam.* (ID282). In addition to the term of address (madam), internal modifiers such as please are used to mitigate the force of the request. Hence, the range of internal and external modification devices will need to be analyzed from a qualitative perspective in order to provide further insights that can inform developments. Additional analyses should improve the dialogue and increase the user experience by gradually approximating real-life conversations. Future work will also focus on examining a larger data sample, with a wider range (and sufficient number) of non-native English speakers from different L1 backgrounds. The insights gained during these iterations, such as the one presented here, will be used to further advance the language model underlying the SDS and develop a branching structure that includes feedback to students regarding the linguistic realizations of requests, thus providing a more complete low-stakes environment for practicing pragmatic moves.

References

Claudia Baur. 2015. *The Potential of Interactive Speech-Enabled CALL in the Swiss Education System: A Large-Scale Experiment on the Basis of English CALL-SLT*. Ph.D. thesis, University of Geneva.

Jared Bernstein, Amir Najmi, and Farzad Ehsani. 1999. Subarashii: Encounters in Japanese spoken language education. *CALICO Journal*, 16(3):361–384.

Shoshana Blum-Kulka, Juliane House, and Gabriele Kasper. 1989. *Cross-cultural pragmatics: Requests and apologies*. Ablex Publishing Company, Norwood, NJ.

Shoshana Blum-Kulka. 1987. Indirectness and politeness in requests: Same or different? *Journal of Pragmatics*, 11(2):131–146.

Penelope Brown and Stephen C. Levinson. 1987. *Politeness: Some universals in language usage*. Cambridge University Press, Cambridge, England.

Michael Clyne. 1994. *Inter-cultural communication at work: Cultural values in discourse*. Cambridge University Press, Cambridge, English.

Janet Holmes. 2000. Talking English from 9 to 5: Challenges for ESL learners at work. *International Journal of Applied Linguistics*, 10:125–140.

W. Lewis Johnson and Andre Valente. 2009. Tactical language and culture training systems: using AI to teach foreign languages and cultures. *AI Magazine*, 30(2):72–83.

Gabriele Kasper. 2006. Beyond repair: Conversation analysis as an approach to SLA. *AILA Review*, 19:83–99.

Kyusong Lee, Soo-Ok Kweon, Sungjin Lee, Hungjong Noh, and Gary Geunbae Lee. 2014. POSTECH immersive English study (POMY): Dialog-based language learning game. *IEICE Transactions on Information and Systems*, 97(7):1830–1841.

Geoffrey Leech. 2014. *The Pragmatics of Politeness*. Oxford University Press, Oxford, England.

Holger Limberg. 2016. Teaching how to apologize: EFL textbooks and pragmatic input. *Language Teaching Research*, 20(6):700–718.

Hazel Morton and Mervyn A Jack. 2005. Scenario-based spoken interaction with virtual agents. *Computer Assisted Language Learning*, 18(3):171–191.

Vikram Ramanarayanan, David Suendermann-Oeft, Patrick Lange, Robert Mundkowsky, Alexei V. Ivanov, Zhou Yu, Yao Qian, and Keelan Evanini. 2017. Assembling the jigsaw: How multiple open standards are synergistically combined in the HALEF multimodal dialog system. In D. A. Dahl, editor, *Multimodal Interaction with W3C Standards: Towards Natural User Interfaces to Everything*, pages 295–310. Springer.

Carsten Roever. 2011. Testing of second language pragmatics: Past and future. *Language Testing*, 28(4):463–481.

John R. Searle. 1969. *Speech acts: An essay in the philosophy of language*. Cambridge University Press, Cambridge, England.

John R. Searle. 1979. *Expression and meaning*. Cambridge University Press, Cambridge, England.

Pei-Hao Su, Tien-Han Yu, Ya-Yunn Su, and Lin-Shan Lee. 2013. A cloud-based personalized recursive dialogue game system for computer-assisted language learning. In *Proceedings of SLaTE 2013*, pages 37–42.

Julie M. Sykes and Andrew D. Cohen. 2008. Observed learner behavior, reported use, and evaluation of a website for learning Spanish pragmatics. In *Selected Proceedings of the 2007 Second Language Research Forum*, pages 144–157.

Jenny Thomas. 1983. Cross-cultural pragmatic failure. *Applied Linguistics*, 4(2):91–112.

Veronika Timpe-Laughlin. 2016. Learning and development of L2 pragmatics as a higher-order language skill: A brief overview of relevant theories. Research Report No. RR-16-35.

Veronika Timpe. 2013. *Assessing intercultural language learning: The dependence of receptive sociopragmatic competence and discourse competence on learning opportunities and input*. Peter Lang, Frankfurt am Main, Germany.

Anna Trosborg. 1995. *Interlanguage Pragmatics: Requests, Complaints, and Apologies*. Mouton de Gruyter, Berlin.

Gay N. Washburn. 2001. Using situation comedies for pragmatic language teaching and learning. *TESOL Journal*, 10(4):21–26.