Applied Language Learning

Action Research

An Artificial Intelligence-Enabled Tool for Real-Time Intercultural Competence Instruction

Tara L. Schendel
Defense Language Institute Foreign Language Center–Washington, D.C.

Inna Kerlin, PhD
Defense Language Institute Foreign Language Center–Washington, D.C.

This paper reports on a collaborative action research project between the Defense Advanced Research Projects Agency (DARPA) and the Defense Language Institute, Washington Office (DLI-W). The project examined the utility of using an existing interpretation system, not originally created for instructional purposes, for teaching and assessing intercultural competence. The system studied was the Cross-cultural Interaction Real-time Assistant for Negotiators and Operators (CIRANO), an artificial intelligence (AI)-enabled, consecutive interpretation and cultural norm violation detection system. CIRANO extends the ability of normal digital interpretation tools by adding capabilities to identify cultural norm violations and to detect emotional cues, and provides real-time rephrasing suggestions to ease interpersonal communication. CIRANO's utility for culturally-based foreign language instruction had not been explored. DARPA and DLI-W thus collaborated to pilot CIRANO with 19 U.S. Air Force students and four native-speaking instructors participating in a 64-week Mandarin Chinese language program. Students and instructors incorporated the tool into a scenario-based lesson. Analysis of student and instructor feedback and research assistants' observation notes indicated that CIRANO enhanced awareness of cultural differences, supported the practical application of intercultural competence principles, and generated valuable instructional interactions. However, serious limitations included the absence of monolingual support, the lack of context-sensitive feedback, and a heavy reliance on reading screen prompts. CIRANO shows potential for classroom integration, contingent upon further refinement to align with DLI-W's pedagogical environment, and to advance broader Department of War objectives in language and culture education.

Keywords: *Action Research, Artificial Intelligence (AI), Intercultural Competence, Foreign Language Instruction, Human Language Technology (HLT), Real-Time Interpretation*

BACKGROUND

This paper describes the findings of an action research project by the Defense Language Institute, Washington Office (DLI-W), which piloted a prototype of the language and cultural application CIRANO (Cross-cultural Interaction Real-time Assistant for Negotiators and Operators), developed by the Defense Advanced Research Projects Agency (DARPA) and its partners (Defense Advanced Research Projects Agency, n.d.; Defense Advanced Research Projects Agency, 2021). The pilot's main purpose was to explore CIRANO's potential to enhance and enrich DLI-W's training programs related to its main mission, which is to provide regional and culturally-based foreign language education globally in support of the Department of Defense.

DLI-W is seeking instructional tools to enhance its programs for intercultural competency, which applies to all students regardless of their military occupational specialty (per CJCSI 3126.01C Joint Chiefs of Staff, 2023 and DoDI 5160.70 (Office of the Under Secretary of Defense for Personnel and Readiness, 2016). Additionally, DLI-W is working with its vendor schools to implement its Final Learning Objectives for its Foreign Area Officers (FAO FLOs). FAO FLOs outline the minimum language skills, intercultural competence, and regional knowledge that FAOs must possess by the end of their full basic language course at DLI-Washington to effectively meet their key mission, which is to build relationships and alliances. These FLOs operationalize both CJCSI 3126.01C and DoDi 5160.70 at the ground level. However, selecting the most appropriate instructional approach and formative assessment of intercultural competence continues to be an ongoing challenge in the foreign language instructional environment (Al-Afifi et al., 2025; Jin et al., 2025). Additionally, the understanding of the concept of intercultural competence and the knowledge and skills associated with it continue to evolve (Byram et al., 2002; Deardorff, 2006; Luo & Chan, 2022). For example, Deardorff's (2004, p. 194) definition focuses on the "the ability to communicate effectively and appropriately in intercultural situations based on one's intercultural knowledge, skills, and attitudes" (as cited in Deardorff, 2006), while Luo and Chan (2022) recontextualize IC as an adaptive toolkit for a world where people simultaneously inhabit physical spaces and borderless digital landscapes. Preparing foreign language learners in the military for this range of intercultural interactions and competencies is an ongoing challenge. This action research project provided DLI-W with the opportunity to evaluate CIRANO and determine its potential for future adoption, adaptation, and maturation in the DLI-W foreign language program. It provided DARPA and its associates with real-world user input on the performance and utility of the application in a military foreign language training environment.

Understanding CIRANO

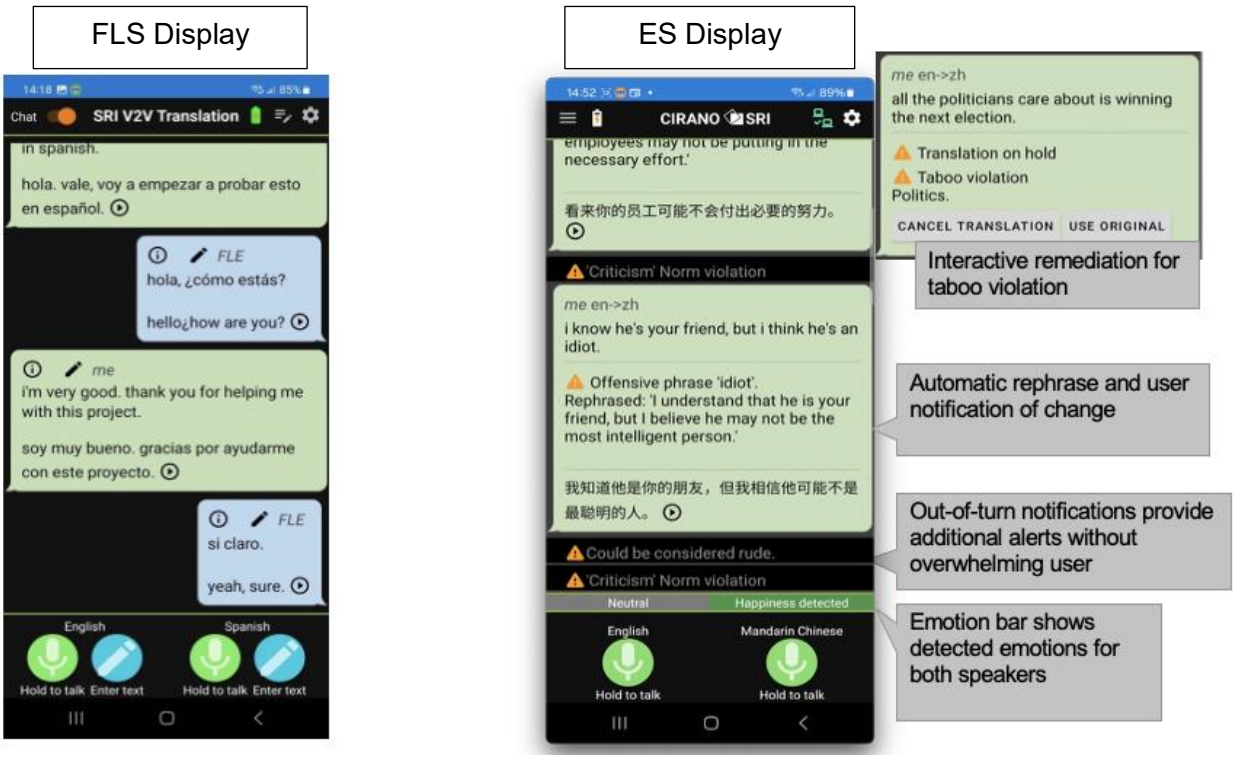
CIRANO is an updated version of previously developed an AI-assisted consecutive interpretation system currently used in the field by Special Operations Command forces when human interpreters are unavailable. These AI supported systems merely interpret, but produce no insight into non-verbal communications or possible cultural missteps that might be degrading communication and putting operators at risk.

DARPA took the existing interpretation system and added newly developed algorithms for emotion detection and detection of cultural norm violations in interactions between an English Speaker (ES) and a Foreign Language Speaker (FLS). DARPA funded and developed these algorithms as part of Cross-Cultural Understanding (CCU) Program. These algorithms use word choice, tone, and volume to detect violations of cultural norms and some speaker emotions (Defense Advanced Research Projects Agency, n.d.). Plutchik's (1980) general psychoevolutionary theory of emotion served as the foundational theoretical framework for the DARPA project's emotion detection. For the development of the CIRANO algorithms, Wang et al. (2023) note that a unique framework was created which drew from multiple other theories, including research on descriptive cultural norms (Cialdini et al., 1990) and situated reasoning (Emelin et al., 2021).

CIRANO was originally developed for use with an Oculus VR headset. For this pilot, DLI-Washington requested that it be re-designed as a cell phone application, which would be more appropriate for a training environment and more easily accessible to students and teachers. Each participant had a cellphone with CIRANO installed, one for ES and one for the FLS. The ES cellphone displays the most comprehensive information while the FLS cellphone has only limited information.

When either an ES or FLS speak, CIRANO determines the gender and approximate age of the speakers which impacts cultural interactions and language use in many cultures. When the ES or FLS speak, CIRANO transcribes what is said, but does not interpret it. Interpretation only occurs once the ES or FS signal CIRANO to do so. In addition to the transcription, the ES cellphone displays warnings about possible culture and taboo violations that could adversely impact communication (see Figure 1). In some cases, the system offers the ES alternative phrasing recommendations along with the cues to help mitigate potential rudeness. CIRANO also provides warnings to the ES about negative emotions such as anger or tiredness that either they or their interlocutors may be transmitting. Armed with this information, the ES can make adjustments by rephrasing a communication before allowing CIRANO to interpret or go forward without changes. The tool is designed to flag violations, but it is ultimately up to the user to determine the proper course of action and when an interpretation will occur.

Figure 1
CIRANO Cellphone Displays



The FLS cellphone displays only the transcript and translation without providing any cues or rephrasing options. This is by design as CIRANO is intended to aid the ES field operator. At the end of a dialog, CIRANO produces a full transcript of the exchange in .pdf format which includes translations, transcriptions, cues provided, remediation and rephrasing suggestions, and English-speaker choices.

The transcripts provided by CIRANO are a key feature to this tool. Teacher-student in-person interactions are undoubtedly crucial for learning both the language and intercultural communication, but they lack any physical artifact that can be used for future learning or reflection.

The decision was made to have students speak English rather than Chinese-Mandarin in this pilot for several reasons: 1.) CIRANO was not designed or optimized to support mono-lingual conversations, 2.) CIRANO was not optimized to process students' spoken target language, and 3.) the focus of the pilot was on intercultural competence vs. foreign language proficiency. Allowing students to be themselves in an interaction unhindered by language limitations would allow for more natural interactions and more norm violations.

ACTION PLAN

This section describes the project goal and objectives, participants, design, implementation, and data analysis for this pilot. The overarching goal was to explore the practical application of the CIRANO tool within a Department of War foreign language training environment.

> Research Question: What were the perceptions of the students, teachers, and research assistants with their interactions with CIRANO, including their perceptions of the tool's strengths and weaknesses for use as a training tool?

This project offered the opportunity to evaluate CIRANO and consider these new technologies for additional finding for future adoption, adaptation, and maturation; as well as giving DARPA the opportunity to receive feedback from DLI-W on how CIRANO might be improved to better serve DLI-W and DLIFLC's training needs.

The DLI-W team developed the pilot design and procedures in close collaboration with DARPA and their partners, as well as the selected DLI-W vendor school. Prior to conducting the pilot, DLI-W presented its pilot design and plan to DLIFLC's Human Research Protections Program (HRPP) for review. The project was approved by DLIFLC HRPP in January 2025 as it met the ethical review requirements for a project focusing on instructional technology applications.

Participants

DLI-W selected the participants using a purposeful and convenient sampling method based on student and teacher availability at the time of the pilot. The students were 19 U.S. military personnel training as cryptologic language analysts, four native Chinese-Mandarin teachers, and five research assistants (see Table 1). All participants had varying degrees of familiarity with instructional technology and AI-enabled tools based on their self-descriptions at the start of the pilot. The students and the teachers had no prior knowledge or experience utilizing CIRANO.

Table 1
Participants Overview

Category	Number	Details
Students	19	• 18 U.S. military personnel (cryptologic language analysts) • 1 military spouse • Enrolled in DLI-W full basic course for Chinese-Mandarin • First half of 64-week course • Estimated ILR proficiency: 1 to 1+ (speaking and listening)
Teachers/FLS	4	• Native Mandarin Chinese speakers • Assigned to teach the student group throughout course
Research assistants	5	• Representatives from DARPA, DLI-W, and vendor school

To further highlight the specifics relevant to the study's participants, Table 2 provides a description of the DLI-W instructional environment.

Table 2
DLI-W Learning Environment

Aspect	Description
Class Size	Maximum 5 students per section
Daily In-Person Instruction	6 hours
Daily Homework	Up to 2 hours
Instructional Format	Group instruction

Pilot Design and Implementation

The pilot design consisted of extensive development of the instructional scenarios and pre-briefing of the participants. The team designed the scenarios to align with the students' syllabus and learning objectives for their progress in their course. Scenarios were focused on situations that involved known differences between Chinese and American cultures to generate dialog and possible norm violations detectable by CIRANO The team consulted the Chinese instructors to devise the two final scenarios as represented in Table 3.

Table 3
Instructional Scenarios

Day 1: Formal Scenario	Day 2: Informal Scenario
You are unhappy with your teacher's feedback. Talk to your teacher. (Convince the teacher to change feedback).	You must decline an invitation from your Chinese friend. Find a way to do it without breaking the relationship.

Approximately two weeks before the pilot, vendor personnel, DLI-W, and DARPA representatives met with the students and teachers to introduce and discuss pilot goals and design, as well as to familiarize them with the CIRANO application for the first time. This was a hands-on exploration before the actual pilot. Teachers were given the role-plays a few days before the pilot to prepare their strategies and approaches in the interactions. Students received their role-play scenarios on pilot days in order to encourage more natural, impromptu interactions from students.

DLI-W and DARPA representatives conducted the pilot on March 4-6, 2025, at the DLI-W vendor facility. During the pilot, the students played the roles of English speakers while the teachers played the roles of Chinese speakers. This allowed participants to focus on intercultural competency and their personal identity without potential interference of the students' nascent foreign language ability (Byram et al., 2002).

The participants were broken into two groups:

- Group A: nine students and two teachers worked through a formal scenario. This scenario was used on the first day in both morning and afternoon sessions.
- Group B: ten students and two different teachers played through an informal scenario. This scenario was used on the second day in both morning and afternoon sessions.

During the pilot, morning and afternoon sessions were held with a lunch break between the two sessions. Four to five students used CIRANO during each session. Each session consisted of one student and teacher interacting with CIRANO for approximately 10 minutes. Two official research assistants monitored each exchange in order to fill out standardized observation checklists. Up to three additional research assistants were present at any given time but did not fill out the official observer survey. Research assistants changed at the end of each morning session.

After their individual sessions with CIRANO, students filled out a written feedback form in Microsoft Forms intended to capture their first impressions of the tool without the influence of others (see Appendix A). Four to five students then spoke with each teacher before the teacher also filled out a written feedback form to capture their first impressions (see Appendix B). The students received their own transcripts as well as the transcripts of their fellow students in the same session. After-action review (AAR) sessions were conducted at the end of each session with teachers, students, and research assistants using non-structured open-ended questions to facilitate group discussion. These sessions provided teachers and students with the opportunity to discuss their interactions while reviewing the transcripts, to ask questions, and to get more in-depth information about the various violations and cues CIRANO provided.

Several days after the pilot phase, the team conducted two separate focus group sessions (one with students, one with teachers) using semi-structured protocols. Each focus group session lasted about one hour. The purpose of these sessions was: (a) to capture any additional comments or ideas that arose after a period of reflection; and (b) to compare the focus group feedback to the individual feedback forms to identify any significant differences.

Data Collection and Analysis

The pilot design allowed for collection and analysis of both quantitative and qualitative feedback from the students, teachers, and research assistants. The quantitative and qualitative data were collected from the written feedback forms mentioned above. The forms contained a mix of Likert Scale questions (i.e., Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree) and open-ended prompts. The team gathered additional qualitative data from the focus groups and the research assistants' written notes. Overall, the quantitative and qualitative data supported each other. There was a high degree of agreement and consistency among participants across the various feedback instruments. Comments on the individual feedback forms aligned closely with those made during the focus groups and the AARs. Research assistants' comments also aligned well with those of students and teachers.

Our thematic analysis of the qualitative data produced 81 initial codes and 13 preliminary categories, which resulted in two final themes—Technology Performance and Learning and Instruction—and 8 final categories (see Table 4).

Table 4

Final Themes and Categories

Themes	Categories
Technology Performance	• Technical Features • Translation Quality • Emotion Detection, Culture Cues, Taboo Cues • Authenticity of Conversation
Learning and Instruction	• Learning • Instruction • Intercultural Competence • Tool Applicability

Limitations

Several limitations were observed during the pilot. The first was the student sample chosen for the project. Initially, DLI-W intended to recruit participants from among students who serve as Foreign Area Officers or Defense Attachés. Their jobs require them to develop strong speaking and participatory listening skills as well as intercultural competence. However, at the time of the pilot planning and implementation DLI-W did not have any FAOs enrolled in Mandarin Chinese language training. The 19 students who were ultimately selected to participate in the pilot were all Air Force Cryptologic Language Analysts. Cryptologic Language Analysts predominately work in secure, compartmentalized environments utilizing their reading and listening skills. Triaging material and completing tasks associated with transcription and translation work take up much of their time. They do not use their speaking skills for their work and are passive listeners. However, they learn speaking skills and intercultural competence as part of their language training with DLI-W. Therefore, the students were very receptive to the opportunity to interact with a native speaker, even in English, and to engage with the target language culture.

A second limitation was the potential for personal bias from both participants and research assistants. To address this, the team used several methods to collect qualitative and quantitative information, such as individual written surveys, focus groups discussions, and observer notes. All these steps helped strengthen the reliability and trustworthiness of the team's analysis and findings.

FINDINGS AND DISCUSSION

Overall, the results indicate that the students, teachers, and research assistants perceived the CIRANO tool as facilitating a successful instructional interaction that resulted in culturally-based language learning. It also provided insight into how such technology could be integrated within a foreign language training context related to the guidance in CJCSI 3126.01C (Joint Chiefs of Staff, 2023). However, the results clearly indicate that the tool requires further development to be applicable within the DLI-W instructional environment.

Intercultural Competence

Both teachers and students agreed that the tool helped facilitate intercultural communication, with 42% of students strongly agreeing with that statement. Multiple teachers and students reported that alternative phrasings and suggestions provided by CIRANO helped improve cultural appropriateness, which helped facilitate communication. The research assistants noted that the communication was smooth and without major interruptions. Additionally, a majority of students (80%) reported that they adjusted their communicative approach based on the suggestions provided by the tool. These results align with the qualitative feedback about the value of language phrasing options in choosing the most appropriate speaking strategy. This also aligns with the concept of the intercultural competence skills of *discovery* and *interaction* described by Byram et al. (2002).

Students also reported that using the tool raised their overall self-awareness, making them more thoughtful about their posture and lack of eye contact with their dialog partner. This resulted in adjustment of their behavior, aligning with what Deardorff (2006) describes as the desired external outcome at the top of her pyramid model, i.e., effective communication and behavior in intercultural contexts. Additionally, students commented that CIRANO raised their awareness of their own communication patterns and linguistic choices in English, such as using "like" frequently as a filler word. By uncovering their automated linguistic habits, the students engaged in what Byram et al. (2002) define as *savoir s'engager* (critical cultural awareness)—the ability to critically evaluate features of one's own culture and communication practices rather than treating them as the universal standard. Furthermore, this sudden hyper-awareness of internal speech patterns aligns with Deardorff's (2006) process model, which asserts that true intercultural competence begins with an inward-facing baseline of cultural self-awareness. Students reported that the emotion cues raised awareness of their own voice, tone, lexicon, and intonation. Students reported changing their tone to achieve a more "neutral" emotion indication from CIRANO.

Teachers had mixed feelings about how CIRANO represented their speech. Some felt it made them sound either more abrupt or more polite than intended. During the focus group sessions, both students and teachers suggested that CIRANO's segmentation of Chinese sentences may have contributed to this discrepancy. This is consistent with the research literature on the subject of rendition of oral speech in written language. Unlike structured written language,

natural, spontaneous speech is inherently characterized by disfluencies, fragmented phrases, unfinished sentences, and continuous run-on sentences. This organic oral communication diverges sharply from the neatly bounded grammatical rules governing written discourse. As a result, the automated evaluation software struggled with the strict application of written language rules—such as rigid syntactic structures and punctuation mapping—when forced to process authentic spoken language. This fundamental structural tension is heavily documented in natural language processing research; for instance, Fraser et al. (2015) stress that the definitive concept of a "sentence" as structured in written typography does not naturally exist in spoken language. Because automated pipelines struggle to superimpose written syntactic boundaries onto the messy, continuous flow of oral delivery, segmentation errors frequently compound. When evaluating languages that lack explicit orthographic word boundaries like Chinese, these architectural constraints become even more acute. As noted by Tsang et al. (2025), Chinese characters are processed without standard visual spacing, making word segmentation and subsequent sentence boundary detection inherently ambiguous and highly prone to error transmission in automated software pipelines. By trying to fit spontaneous speech into pre-programmed written standards, CIRANO introduced punctuation and structural segmentation errors that artificially penalized the students' authentic oral productions.

Additionally, both students and teachers appeared to enjoy the interactions via CIRANO, as noted by the research assistants in their observation checklists. Students and teachers used idiomatic expressions to intentionally test the system's responses. In some instances, students and teachers attempted to come across as rude or forceful in order to trigger a warning from CIRANO. The research assistants interpreted this behavior as curiosity about the technology and as maintaining motivation and engagement, rather than as detracting from the main goal of the activity. Periodically, the system would reveal an awkward moment eliciting laughter from the group. In one scenario, to avoid attending a party, the student used his dog as an excuse. Immediately realizing the inappropriateness of this excuse (even without a cue from CIRANO), he and his teacher laughed, and he tried again. This instance came up again in a focus group with the student still embarrassed by his tactic. This finding aligns with the literature which suggests that attitude is a fundamental requirement for the successful development of intercultural competence (Byram et al., 2002; Deardorff, 2006).

In another scenario, one or two students planned ahead and took time to prepare to see if they could successfully navigate the interaction with a minimum of assistance from CIRANO. In the scenario, the student, originally from the Philippines, attempted to convince her teacher (the interlocutor) to call her father about her progress instead of her mother. When the teacher responded that it was only appropriate for her to contact the mother, the student called upon her knowledge of Filipino culture and invited the teacher to her home for lunch on the weekend. This behavior was not flagged by CIRANO because, it turned out, this was an appropriate response in Chinese culture as well.

Taboo and Culture Violations

Overall, students found the culture violation warnings and taboo indicators useful, even if sometimes inaccurate. Over half of the students (68%) reported that the cues provided by CIRANO helped them to understand the Chinese culture norms (Appendix A) However, students indicated that taboo and culture cues from CIRANO were rather formulaic and confusing at times, as they lacked clear labeling, explanation or guidance on how to repair the violation. They also noted that a better understanding of what triggered the taboo and cultural warnings would be helpful as would more on-screen guidance on how to repair a violation.

Non-verbal Cues

When asked whether using the tool during the interaction increased their awareness of non-verbal cues, only 37% of the students agreed with the statement, while more than half either provided a neutral rating (53%) or disagreed (11%). (Appendix A). The research assistants, students, and teachers noted that because there was a lot of reading involved while using the tool, they spent most of the time looking at the screen of the phone rather than in more natural conversation with eye contact.

CIRANO is designed to facilitate oral communication by removing the language barrier while supporting better cross-cultural awareness through written feedback on the phones. In a foreign language teaching environment, the amount of attention required for reading CIRANO prompts interferes with the authenticity of oral communication and distracts from focusing on speaking and possibly listening. This is an important observation, as it clearly indicates the tool's limitations.

Authenticity of Conversation

Teachers and students commented that the system speed was impressive and maintained a good conversational flow. While the system requires turn taking, the gaps between conversation turns were minimal, with the system able to transcribe, translate, and voice quickly.

Teachers reported that the translations produced by CIRANO were generally accurate but overly literal, especially with idiomatic expressions and colloquial language. This resulted in odd translations and misunderstandings between speakers (Appendix B).

In general, teachers commented that having students speak English allowed them to see them as their more natural selves. They noted that the low levels of speaking proficiency limit students' ability to authentically represent themselves. They often lack the ability to express emotions, humor, and use complex grammar and vocabulary to express themselves naturally in the foreign language at this early stage of their study. This aligns with the literature on multiple authentic identities that people who study and interact with other cultures develop while engaging in intercultural interactions (Byram et al., 2002).

Just over half (63%) of the students agreed with the statement that they would like to use the tool in their training, while 11% of the students disagreed with this statement (Appendix A). In their comments, these students in the minority indicated that the information received from the tool was readily available to them already from their teacher, and therefore, the tool was redundant. According to participants' comments, enabling the tool to perform in a monolingual setting would make it more relevant to their learning program.

Additionally. multiple participants and research assistants commented that receiving printed transcripts, including the transcripts of others, was extremely beneficial for their learning. It allowed them to see the different options for expressing similar ideas in Mandarin Chinese, as well as introduce them to the novel communication strategies employed by their fellow students as they negotiated a culturally sensitive situation. Reviewing other students' cultural violations and taboos was very beneficial as it facilitated interesting discussions about new concepts related to cultural norms. These results mirror the intercultural competence frameworks of Deardorff (2006) and Byram et al. (2002), specifically regarding openness, adaptability, and the ability to see things from others' perspectives.

Most of the teachers (75%) either agreed (50%) or strongly agreed (25%) that they would like to use the tool in the future to teach intercultural competence (Appendix B). The teachers mentioned a couple of teaching approaches for using this tool in the future, such as making it a game to recognize cultural violations. They acknowledged that the tool enhanced role-play, a common and widely used instructional activity, during which students and instructors play out a scenario to accomplish a specific language task. Many teachers might be uncomfortable reacting to a cultural taboo or explaining why a statement is rude, thereby raising the usefulness of using a tool such as CIRANO as an intermediary.

Lastly, CIRANO presents a novel and intriguing tool in our training environment. A 64-week intensive course in Mandarin Chinese is long and difficult. Maintaining motivation can be challenging for both students and teachers. Offering novel tools presents opportunities to exercise different modes of critical thinking, break up monotony, and maintain motivation and drive (Cabrera-Solano, 2022). However, all the participants and research assistants agreed that the tool needs further development to be useful within the DLI-W foreign language training environment. For instance, CIRANO needs to support monolingual interactions with the culture cues to promote not only the acquisition of cultural knowledge and skills, but also the development of strong communication skills in the target language.

Recommendations

The team provided several recommendations for using CIRANO and similar AI tools to build intercultural competence in foreign language training. These suggestions align with CJCSI 3126.01C (Joint Chiefs of Staff, 2023) guidance on expanding military language, regional

expertise, and culture (LREC) capabilities. Specifically, the team focused on identifying multi-language technologies that support these mission goals:

- **Cues for Both Users:** For instructional purposes, it is recommended that the users have equal access to all the prompts, violation cues, and other information to facilitate real-time corrective action and feedback and allow for more instructional flexibility.
- **Flagging a Wider Array of Cues and Emotions:** CIRANO is designed to flag a limited number of norm violations and is tuned to focus on emotions and violations that may negatively impact communication. Widening the variety of cues and including positive feedback would help students understand what works when interacting with native speakers.
- **Highlighting What Caused a Violation:** Providing any information on what triggered a cultural violation (e.g., a word, intonation, or topic) would significantly enhance acquiring the knowledge and skills associated with intercultural competence (Deardorff, 2006) and provide an understanding of why rephrasing was suggested.
- **Etiquette Rules or Guidelines for Using AI-enabled Communication Devices:** Like in the past, when people adjusted to the peculiarities of talking over the phone or through a video-conferencing platform, or when using a human interpreter, it would be valuable to develop etiquette rules for the users of devices and tools like CIRANO. These rules would address proper eye contact, how to disclose device use, where and when it is appropriate to use the device, or what to do when the device fails or makes a mistake. Use of such devices and the digital social norms associated with them are culturally laden in and of themselves (Heitmayer & Schimmelpfennig, 2023; Walton, 2024). In the DLI-W context, this would be taught as part of the initial orientation to learning tools.
- **Monolingual Input as an Option:** The goal of DLIFLC is to teach foreign languages; therefore, the support of monolingual dialogues conducted in the same language by both speakers, e.g. either all in English or all in the target language, would be most beneficial for teaching the principles of IC. Moreover, the resulting transcript would align with the widely used instructional approach of integrating a variety of modalities into a single learning activity (Dai et al., 2024; Joohoon Kang & Kim, 2024). Given the usefulness of the option of beginning students speaking English, this option should still be retained.
- **Natural Oral Speech Accuracy:** It is recommended that the tool developers continue to improve the rendering of natural oral speech that differs from written speech. For instance, it is normal and acceptable to have filler words, pauses, unfinished and "run-on" sentences (Siepmann, 2023). However, CIRANO currently adheres to written language norms, for example eliminating filler words and truncating long sentences. It also puts periods on unfinished sentences. Additionally, because CIRANO uses a generic mechanical voice for all interactions, it eliminates pauses and other aspects of spoken language that are key carriers of emotion. CIRANO needs to accommodate different voice options for its vocalizations that are more refined.

- **Explore Options for Different Delivery Platforms:** For this pilot, CIRANO was delivered as a cellphone application. However, with the amount of information it provides to students, development of an alternative application that can be integrated into a learning management system and used on an interactive whiteboard is recommended. This would allow for both individual practice with an instructor as well as more group activity work. It would reduce the need for extra equipment and fit more naturally into the typical learning environment that enables intercultural competence development (Al-Afifi et al., 2025).

CONCLUSION

CIRANO shows strong potential to support the teaching and assessment of intercultural competence, especially when used as a springboard for deeper discussion and reflection, which are foundational in acquiring intercultural competence (Byram et al., 2002; Deardorff, 2006). Key recommendations for improving CIRANO's suitability for DLI-W's training environment include adding options monolingual interactions, improving context sensitivity and voice quality, integrating CCU features into a more versatile platform, providing information about cues and rephrasing prompts, improving natural oral speech rendering, and developing etiquette guidelines for using AI-enabled communication devices. Since this CIRANO pilot, there have been major advances in AI technology and more to come that could improve many aspects of CIRANO provided there is funding to continue development.

CIRANO is designed to facilitate oral communication through written feedback on the phones. In a foreign language teaching environment, the amount of attention required for reading CIRANO prompts interferes with the authenticity of oral communication and distracts from focusing on speaking and possibly listening. However, DLI-W's instructional approach aligns with educational research on multimodal learning, which supports integrating reading, listening, and speaking into training activities to maximize training results (Dai et al., 2024; Joohoon Kang & Kim, 2024). For educators, this highlights the importance of diverse learning models. For CIRANO's developers, it points to the need for further exploration of how the tool's current features support this approach, or what additional features might be necessary to support LREC requirements outlined in CJCSI 3126.01C (Joint Chiefs of Staff, 2023).

This pilot represents a successful model of collaboration and innovation between DARPA and DLI-W, advancing DoD training capabilities while providing DARPA with vital user-centered feedback to refine its tools for both operational and instructional environments.

Acknowledgement

This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under the Cross-Cultural Understanding (CCU) program. The views, opinions, and/or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Authors

Tara L. Schendel, M.A., is an Associate Professor serving as Academic Advisor for Educational and Information Technology at DLI Washington. Her current academic focus is on developing new technology tools to better gather and analyze student progress reporting data.

Inna Kerlin, Ph.D., is an Associate Professor and serves as an Academic Advisor at DLI-Washington. Her professional areas of expertise encompass academic leadership, language testing and assessment, and the application of adult educational principles in foreign language education.

REFERENCES

Al-Afifi, M., Nygren, T., Rasmusson, M., & Tväråna, M. (2025). Intercultural competence in the classroom: A systematic review of classroom-based research on teaching and learning practices for developing intercultural competence in different educational contexts. *Intercultural Education, 37*(2), 217-250. https://doi.org/10.1080/14675986.2025.2555782

Byram, M., Gribkova, B., & Starkey, H. (2002). *Developing the intercultural dimension in language teaching: A practical introduction for teachers*. Council of Europe. https://rm.coe.int/16802fc1c3

Cabrera-Solano, P. (2022). Game-based learning in higher education: The pedagogical effect of Genially games in English as a foreign language instruction. *International Journal of Educational Methodology, 8*(4), 719–729. https://doi.org/10.12973/ijem.8.4.719

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, *58*(6), 1015–1026.

Dai, Z., Thipatdee, G., & Metjiranont, M. (2024). The development of a task-based Chinese speaking instructional model for Chinese as a foreign language learners in Thailand. *Journal of Education and Learning, 13*(1), 162. https://doi.org/10.5539/jel.v13n1p162

Deardorff, D. K. (2006). Identification and assessment of intercultural competence as a student outcome of internationalization. *Journal of Studies in International Education*, *10*(3), 241-266. https://doi.org/10.1177/1028315306287002

Defense Advanced Research Projects Agency. (n.d.). *Computational cultural understanding*. Retrieved April 25, 2025, from https://www.darpa.mil/research/programs/computational-cultural-understanding

Defense Advanced Research Projects Agency. (2021, June 3). *Broad agency announcement computational cultural understanding (CCU) information innovation office HR001121S0024 amendment 2*. https://www.darpa.mil/sites/default/files/HR001121S0024-Amendment02.pdf

Emelin, D., Le Bras, R., Hwang, J. D., Forbes, M., & Choi, Y. (2021). Moral stories: Situated reasoning about norms, intents, actions, and their consequences. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing* (pp. 698–718).

Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.emnlp-main.55

Fraser, K. C., Ben-David, N., Hirst, G., Graham, N. L., & Rochon, E. (2015). Sentence segmentation of aphasic speech. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 862–871). Association for Computational Linguistics. https://aclanthology.org/N15-1087/

Heitmayer, M., & Schimmelpfennig, R. (2023). Netiquette as digital social norms. *International Journal of Human–Computer Interaction*, 40(13), 3334–3354. https://doi.org/10.1080/10447318.2023.2188534

Jin, Y., Sercu, L., Zhou, J., & Chen, Y. (2025). Assessing intercultural competence in the digital age: A Chinese scale development and validation. *Journal of Studies in International Education*, *0*(0). https://doi.org/10.1177/10283153251371030

Joint Chiefs of Staff. (2023, March 8). *Language, regional expertise, and culture capability identification, planning, and sourcing* (CJCSI 3126.01C). https://www.jcs.mil/Portals/36/Documents/Library/Instructions/CJCSI%20312601C.pdf

Joohoon Kang, & Kim, H. G. (2024). Multimodality in English language teaching: A review of L2 multimodal composition. *Multimedia-Assisted Language Learning,* 27(4), 160–178. https://doi.org/10.15702/mall.2024.27.4.160

Luo, J., & Chan, C. K. Y. (2022). Qualitative methods to assess intercultural competence in higher education research: A systematic review with practical implications. *Educational Research Review, 37*. 100476. https://doi.org/10.1016/j.edurev.2022.100476

Office of the Under Secretary of Defense for Personnel and Readiness. (2016, December 30). *DoD instruction 5160.70: Management of the defense language, regional expertise, and culture (LREC) program*. Retrieved April 25, 2025, from https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/516070_dodi_2016.pdf

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In *R. Plutchik & H. Kellerman (Eds.), Theories of Emotion* (pp. 3–33). Academic Press. https://doi.org/10.1016/B978-0-12-558701-3.50007-7

Siepmann, D. (2023). This deserves a brief mention: A multi-corpus comparison of written and spoken academic discourse in English and French with implications for pedagogy and lexicography. *Languages in Contrast*, 23(1), 60–92. https://doi.org/10.1075/lic.21004.sie

Tsang, Y. K., Chen, X., & Jin, Z. (2025). *Parsing through boundaries in Chinese word segmentation*. arXiv. https://doi.org/10.48550/arXiv.2503.23091

Walton, D. J. (2024). Culturally responsive computing: An introduction into computer science, security, and technology. *Rotel*. Retrieved June 6, 2025, from https://rotel.pressbooks.pub/culturally-responsive-computing/

Wang, S. C.-H., Saakyan, A., Li, O., Yu, Z., & Muresan, S. (2023). Sociocultural norm similarities and differences via situational alignment and explainable textual entailment. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing* (pp. 3548–3564). Association for Computational Linguistics. https://aclanthology.org/2023.emnlp-main.215.pdf

APPENDIX A

CIRANO CCU Student Feedback Form

1. Please indicate your level of agreement with each statement about using the CCU CIRANO Application, choosing one of the following for each: *Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree*
    a. The application helped me communicate more effectively with my instructor.
    b. The cues provided helped me better understand cultural norms.
    c. The application increased my awareness of non-verbal communication cues.
    d. I am comfortable using AI powered tools like this application.
    e. I would like to use this application more as my training progresses.
2. Describe any specific instances where the application was particularly helpful or challenging.
3. I adjusted my approach to the conversation based on the suggestions provided by the application: *Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree*
4. What did you like most about using the application?
5. What did you like least about using this application?
6. If you have any concerns about using this application, please share them with us.

APPENDIX B

CIRANO CCU Instructor Feedback Form

1. Please indicate your level of agreement or disagreement with each statement: *Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree*.
    a. The tool's Chinese translations were accurate.
    b. The tool helped facilitate culturally appropriate communication.
    c. I would like to use this tool in the future for teaching intercultural competence.
    d. The tool was easy to use and navigate.
2. What did you like best about this application?
3. What did you like least about using this application?
4. Please share any suggestions you have for improving the tool for possible future use in the DLI-W foreign language program.