THE MULTIMODAL PRODUCTION OF COMMON GROUND UNDERSTANDINGS IN INTERCULTURAL FLIGHT TRAINING

Saeko Nomura and Edwin Hutchins

Cognitive Science Department, UCSD
La Jolla, CA

In this paper we show how Japanese airline pilots and American flight instructors overcome pronounced differences in language and culture and achieve effective collaboration. They do this by drawing on a deep body of shared professional pilot culture and by exploiting richly multimodality situated communication practices to produce common ground understandings.

Introduction

Effective collaboration requires the creation and maintenance of common ground understandings (Clark, 1992). This is an especially interesting problem in the case of intercultural collaboration, where communicative conventions may not be shared. However, intercultural collaboration often takes place in professionally relevant material settings and among people who share professional competence. This paper discusses ethnographic field data which we collected at the Boeing/Alteon training center in Renton, WA, where Japanese airline pilots received flight training from American flight instructors. We attempt to clarify how the pilots and the instructors effectively use multimodal representations to achieve intercultural common ground understandings.

Research Methods

As part of a multi-year multicultural study of the roles of language and culture in flight deck operations (Nomura, et al., 2006; Hutchins, et al. 2006), we observed training events at Boeing training center in Renton , WA. in 2005[1]. A native Japanese Human Computer Interaction researcher and a flight deck human factors specialist observed and recorded training courses in which Japanese pilots received training from American flight instructors. The flight training was conducted in a B777 and a B767 full-flight simulator.

A total of 40 hours of video data were collected. In the intercultural training settings, nearly all conversations were in English. Most of the Japanese pilots had quite high English proficiency, but some had difficulty engaging in complex discussions in English. The American instructors could not speak Japanese at all.

As observers we were allowed to bring two digital video cameras and audio microphones into the high-fidelity simulator. The flight deck activities were videotaped from two perspectives. One was an "over-the shoulder angle," which allows observers to see the interactions of pilots with flight instruments and displays. The other was a wide-angle view looking aft from a location low on the First Officer's side of the instrument panel. This view captures the pilots' facial expressions, gestures and actions that would otherwise be hidden by their bodies, and their interactions with the instructor.

The Intercultural Multimodal Production of Common Ground Understandings

Our video records contain hundreds of instances of interaction between the Japanese pilots and their American instructors. In this brief paper we present a few representative examples to illustrate the processes by which common ground understandings are created and maintained in intercultural flight training.

Speaking Practices

Because the training took place in the US, the trainee pilots made efforts to use English as much as possible throughout the simulated flights. However, the way the pilots spoke English varied considerably as a function of speaking context. Much of the variability in English language production is accounted for by what is known in the field of discourse analysis as recipient design (Sacks, et al., 1974). Recipient design is a process by which speakers shape their productions to fit the needs of their listener.

When simulating a conversation with Air Traffic Control (the part of the controller was played by the instructor) one of the pilots who was highly proficient in English, spoke nearly unaccented English. When this same pilot spoke English to his

[1] We also made field observations in Japan and New Zealand in 2005 and 2006. In these countries, we observed 54 legs of revenue flight in a variety of Boeing aircraft including the B777. We also made 67 hours of video recordings of recurrent and transition training held in the B777 and the B767 full flight simulators.

Table 1. Conceptualizing a checklist item with multimodal representations.

	Original Speech	Translated Speech	Hand Gesture and Eye Gaze
PM:	'Initial climb: Set maximum climb thrust and 10 degrees pitch.'	'Initial climb: Set maximum climb thrust and 10 degrees pitch.'	He reads the checklist item.
PM:	OK. Initial climbの場合には、maximum climb thrust をset して、 pitch を 10度 にして、その後降りられるようにする。	OK. In case of initial climb, set maximum climb thrust, and make pitch as 10 degree, to be able to descend later.	- While he is interpreting, he is looking down at the written material (the checklist). - "Initial climb": Holding his right hand palm-down over the center console he tips his fingers. upward to indicate "climb". - "maximum climb thrust": With his right hand in space above the thrust levers, handshape as if pushing the thrust levers, he moves his hand forward. - " pitchを10度": Returning to the the down-turned palm making 10 degree up and then angle the hand downwards to show descent. - After saying this, he looks at the PF's face to verify that PF understood his remarks.
PF:	はい。	Yes.	

Japanese co-pilot, however, he produced English words using Japanese phonology. In addition, he changed speech cadence and stress patterns to match those normally observed in Japanese.

Fig. 1. Entering the flight data into the FMC: the pilot reads aloud the information on the paper while inputting the data.

For example, while entering data into the flight management computer system (FMS) via the control and display unit (CDU), the pilot spoke aloud the numbers he was reading from flight papers and selecting on the CDU (See Figuer1). Integrating the visual and motor activities of reading and character selection with speech creates representations of the information in additional modalities (speech and audition) which makes the transfer of information more robust. This sort of activity is often observed even in mono-lingual English speakers. However, for non-native English speakers, producing accurate English pronunciation imposes additional cognitive loads. So when Japanese pilots want to concentrate on entering precise numbers into the flight management computer (FMC), they tend to pronounce English words using Japanese phonology. In addition, as CDU data entry is an individual task, pilots do not normally need to communicate with their partners while doing this. The speech tends to be near sub-vocal and appears as self-regulatory muttering.

When a pilot monitoring (PM) read aloud a checklist for his Japanese pilot flying (PF), he superimposed Japanese phonology on English vocabulary, and made other transformations as well. He raised his voice to a high pitch and put a distinctive rhythm on it. In the checklist procedure, for example, the pilot extended the pronunciation of words at the end of sentences. Cabin altitude was pronounced as "cabin arutitudooo", checklist as "checkriiistoooo", checklist is completed as "checkrist is compreteeedo", and so on. He also slowed down while reading the important portion of the sentences (e.g. numbers, alternative actions, etc.). Especially for non-normal checklists in emergency conditions, the PM needs more careful attention from the PF. He indicated this by making his voice tone different from that used in other settings. All of these practices are forms of recipient design that increase the congeniality of spoken English for a Japanese listener.

Bodily Practices

Conceptualization of English Japanese pilots use more gestures when speaking English than they use when speaking Japanese in the same setting.

For example, when they receive a clearance for a new heading in English, pilots sometimes draw the number in the air in front of them before reading it back. For the readback itself, pilots are required only to repeat the heading as assigned by ATC. However, before accepting any clearance a prudent pilot assesses the implications of the clearance. A pilot might ask himself, "Does that number represent an appropriate direction of flight for my airplane?" To answer that question, the pilot must understand the number as a spatial concept. Even the pilots whose English is excellent used alternative representations to conceptualize the meaning of numbers expressed

in English. Sometimes this was accomplished by translating the number into Japanese, and sometimes it was accomplished by producing visual representations with gestures (McNeill, 2005).
The pilots also use Japanese language augmented by hand gestures to interpret a complex section of the checklist written in English. Here is a example from training B767.

Fig. 2. Bodily practices facilitate pilots' conceptualization of maneuvers written in English.

Table 1 and Figure 2 show a conversation between pilots while doing the "Unreliable Airspeed Checklist" in the B767. This checklist is performed when the airspeed/mach indicator on either the PM's side or the PF's side is suspected to be unreliable. The checklist begins "Compare Captain and first officer airspeed indications to standby indicator. An Airspeed display that differs by more than 15 knots from the standby indicator should be considered unreliable." If the displays indicate more than 15 knots difference, the procedure continues 'Autopilot…off, Autothrottle …disconnect, Flight director…off, and Attitude and thrust… adjust.' One pilot (PM) reads the checklist aloud and manipulates the switches in the flight deck. The (PF) watches the PM's actions and gives him feedback, such as "Hai ('Yes' in Japanese)" and "Okay," to let the PM know that he confirms the actions. The simple actions in the first part of the procedure do not require the pilots to discuss the meanings of the actions.
Table 1 shows the last part of the checklist, which does require some interpretation. In this procedure, after reading the checklist aloud in English (the first line), the PM interprets the meaning of the procedure in Japanese, augmenting speech with iconic and environmentally coupled gestures (Goodwin, 2000; Goodwin, 2006) (the second line). For "Initial climb" and "pitch as 10 degree," the PM imitates the airplane nose up pitch attitude with his right hand (See Figure 2), and for "set maximum climb thrust," he positions his right fist above the thrust levers as if he is gripping them He then moves his fist forward, indicating increased thrust. His use of gesture here not only helps the PF understand the procedure, it also embodies his own conceptualizations of the airplane's movement and the movements to be made by the PF.

1. Inst: Your airspeed was much better than two days ago, okay?
2. Pilot A: Uh, huh.
3. Pilot A: Final the **rudder is (.) .hhhh**


((the right hand palm down wagging hand at wrist from left to right to represent the yawing of the airplane as the pilot made rudder inputs.))

4. Inst: Yeah.
5. Pilot A: **Right?**

((looks up to the instructor's face))

6. Inst: Rudder was, rudder was lot of work.

Fig. 3. Combination of culturally meaningful speech and gesture completes the explanation of the context.

The Juxtaposition of Culturally Meaningful Bodily Techniques & Speech for Joint Reasoning Gestures frequently accompany speech in conversations between Japanese pilots and American instructors. Figure 3 depicts a conversation between a pilot and an instructor just after completing an ILS approach and landing in gusty cross-wind conditions. Although the pilot receives a compliment about landing airspeed from the instructor, the pilot is not satisfied with airplane's horizontal movement on final approach. In the third line of Figure 3, he begins to construct a statement, saying, "On final the rudder is…" At the same time, he places his right hand palm-down in the instructor's field of vision and wags it at the wrist from left to right to represent the yawing of the airplane as he made rudder inputs.

In this multimodal production, the pilot's gesture takes the place of words. The beginning of the utterance implies a completion that will somehow specify something about the rudder. Rather than providing that specification in speech the pilot offers a gesture for interpretation. The gesture is ambiguous in the sense that it is open to many possible interpretations. The hand might represent the airplane, it might represent the rudder itself, or perhaps it represents the rudder pedals. The pilot requests the instructor's verification of his observation by looking at the instructor and asking the tag question, "Right?" The instructor responds by providing a spoken completion for the sentence that was begun by the pilot. The instructor says, "Rudder was, rudder was a lot of work." Together, the pilot and the instructor have jointly produced a single declarative statement, "On final the rudder was a lot of work." In doing so, they have drawn on shared understandings drawn from professional pilot culture. It is interesting that the instructor's completion of the sentence does not resolve the ambiguity of the gesture. We suspect that this is not how the pilot would have completed the sentence if he had possessed the English skills needed to do so. It is probably also not a sentence that the instructor would have created spontaneously. Nevertheless, having produced this sentence together, the two accept it as a meaningful statement about the activity in which they are engaged (Goodwin, in press). This acceptance of the course of jointly produced meaning is a hallmark of common ground understanding.

Re-use of Previous Multimodal Actions by Later Speakers Japanese pilots also used gestures to complement their lack of English expressions.

The first line of Figure 4[2] provides an example. A pilot has made several hard landings. He asks the instructor to talk to him about the aiming point and the touchdown point. When introducing the "aiming point," he makes a small oval with the index fingers and thumbs of both hands and looks through the oval.[3] The oval made by the pilot's hands seems to represent the frame of reference with respect to which the apparent motion of the aiming point can be judged. Struggling with the articulation of the aiming point, the pilot shifts to a related concept, the touchdown point (or zone) which is where the aircraft's main wheels are expected to contact the runway. The pilot claps his hands together while searching for the English words "touchdown point." Once again, by using a combination of somewhat inarticulate speech with articulate gesture, the pilot has created an opening for the instructor to complete the discourse. Interestingly, when picking up the conversation, the instructor begins by re-using a gesture that was produced by the pilot. He refers to the aiming point by making the oval and looking through it just as the pilot did. But this gesture is then smoothly deformed into a very different representation. The two halves of the oval slide apart, the right half placed in front of the left. In this configuration and coordinated with the concurrent speech, the span between thumb and forefinger represent the distance, about 300 meters, on the runway between the threshold and the aiming point. This gestural configuration is transformed one more time, creating a new perspective. The right hand is brought close to the left, and the finger/thumb shapes are rotated to vertical. Now the right hand moves away from the left hand. The left hand indicates the runway threshold seen from the right side[4], and the right hand indicates the position of the PAPI (Precision Approach Path Indicator) along the runway.

The instructor thus demonstrates his shared understanding of the pilot's meaning by reusing the pilot's own gesture. The instructor's use of 'meters', rather than 'feet' is an example of recipient design which also aims at facilitating shared understandings. Since Japanese pilots use the metric system, the instructor converts length in feet into meters.

[2] Letters in a bold font show correspondence of speech with gestures.

[3] The aiming point is a location on the runway where the pilot focuses attention. If the aiming point appears to move up during the approach, the airplane will land short of the aiming point. If the aiming point appears to move down during the approach, the airplane will land beyond the aiming point.

[4] This detail is telling. Because pilot in command sits in the left seat in civil fixed-wing aircraft, the PAPI is located on the left side of the runway. The gesture here shows that the instructor is imagining the runway environment as seen from the right side of the runway looking across the runway to the PAPI lights.

1. Pilot A: .hhhh This, this. I (0.2) cannot understand it. (0.5) **Aiming point** is **a (0.5)** ah;;, (0.5) be a touch down point.


((Left: makes a oval with hands which indicates a region of runway as seen from the flight deck on final approach.
Right: claps his hands to represent touch down))

2. Inst: At the **aiming point** (0.1) should be::: **about 300 meters down the runway** ↑. Which i:s (0.1) **abeam the PAPI**.

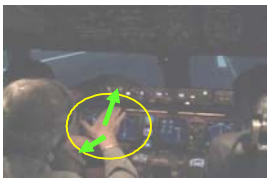


((Left: reuses the gesture of the pilot to represent aiming point as seen from flight deck.
Center: Then slides right hand to express '300 meters down the runway'
Right: slides right hand from left to right to represent the runway between the threshold and a point abeam the PAPI))

3. Pilot A: Ah:::
4. Inst: A:::nd, and **ILSes seem to be bringing us in a little bit below that aim point**.

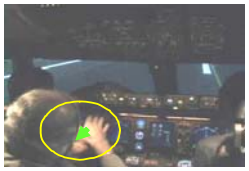
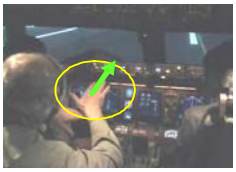
((This gesture enacts a view from the outside of the final approach looking down the ILS glide path to the touchdown zone. Makes the shape of glide path with right hand. Left hand roughly models the glide slope, and the right hand models the motion of the aircraft down the glide path to the runway.))

5.Pilot A: Ah::, ah::.
6. Inst: Yeah.
7. Inst: You're on the visual patterns you flew, (0.4) **your final and your aim point were good all the time**. (0.8) Ah::: couple of times you::: flared, (0.1) you **started the flare but didn't do:: near (0.1) quite enough**.

((Left: Right hand models the airplane on a final approach.
Right: Left hand models the touchdown point and the right hand models the airplane starting flare. Moves tips of fingers of right hand upwards to represent nose up rotation))

That's that's why you had the firmer touchdowns. But **(0.5) up (1)** most of 'em (0.1) a good flare height, and you were flying the **nose gear down**.


((Left: Hit back of left hand with heel of right hand three times.
Center: Right hand models flare of the airplane.
Right: Makes the motion of pilots manipulating the yoke to fly the nose gear down to the runway.))

8. Pilot A: Oh, I see.
9. Inst: Okay. So I think you're you're aim point, (0.5) and all control on final visual patterns were very good.
10. Pilot A: Hum.
11. Inst: Okay?
12. Pilot A: I see. Okay.

:

13. Inst: Do you wanna go back to three miles again? ↑

((Moves right hand palm facing to him toward him to express the airplane will be moved back to 3 miles from the runway.))

Fig. 4. Re-use of gesture, iconic, and environmentally coupled gestures.

He also uses his right hand in one moment to model the movement of an airplane (finger tips as the nose of an airplane, and heel of the hand as landing gear), and just a moment later, he uses the same hand to model the movement of pilots' hand (easing the control yoke forward to fly the nose gear onto the runway). His hand gesture also shows that they are in a special setting; a flight training simulator. On the 13th line of Figure 4, the instructor moves his right hand palm facing to him toward him, saying "Do you wanna go back to three miles again?." It is impossible in real life to instantaneously reposition the airplane 3 miles from the runway threshold, but with the simulator, the instructor literally rewinds the approach and sets the airplane back to 3 miles from the runway threshold. This series of gestures and the words that are spoken with them mutually elaborate each other to produce a rich multimodal representation of pilot behavior, airplane behavior and the relations between those two.

Conclusion

This paper showed how Japanese airline pilots and American flight instructors overcome pronounced differences in language and culture and achieve effective collaboration.

First, we examined the speaking practices of the participants, showing how Japanese pilots reduce cognitive cost by pronouncing English in their native phonetic systems when they process flight data. We also showed how all parties engage in recipient design, adapting their language production to render their spoken utterances more interpretable to their interlocutors.

We then examined the bodily techniques of the pilots and instructors, showing how body orientation and hand gestures enact shared understandings of aircraft movement and pilot actions. The juxtaposition of bodily techniques with culturally meaningful objects in the flight deck was an especially important semiotic device the American instructor when offering explanations and advice to Japanese pilots with limited English skills. All participants in this study share technical knowledge as pilots, similar flight experience, and the social and physical context of the simulator. When they have problems producing meaningful utterances, they use other representations to facilitate mutual understanding. We also showed how elements of previous multimodal enactments are re-used by later speakers. Here, we discussed how the gesture and its lexical affiliate are not only produced by one person but are deeply intertwined in the development of a common structure of meaning (McNeill 2005). Shared projection of anticipated meanings, joint production of complex utterances, and the re-use of semiotic resources introduced by others are all indicators of the successful accomplishment of common ground understanding. Finally, we presented how some iconic and environmentally coupled gestures carry both a conceptual referent and a very specific perspective or point of view on that referent. We noted how fluidly communicators can switch among meanings and points of view.

The kinds of multimodal performances that we have described here are also observed in many other intercultural settings. The general argument here is that the richer the context, the easier it will be to produce intersubjective understanding in intercultural interactions.

Currently, we are continuing our analysis of the nature of multimodal communication among pilots and instructors who come from different linguistic and cultural backgrounds. We plan to conduct additional fieldwork in 2007 in Japan, the United States (at the training center in Renton, WA), and in Australia.

Acknowledgement

This research was funded by a grant from the Non-US operations and training program of the Boeing Flight Deck Concept Center.

References

Clark, H. H.: Arenas of Language Use. University of Chicago Press, Chicago, IL (1992)

Goodwin, C.: Action and Embodiment within Situated Human Interaction, Journal of Pragmatics, Vol. 32 (2000) 1489-1522

Goodwin, C.: Things and Their Embodied Environments (2006)

Goodwin, C.: Human Sociality as Mutual Orientation in a Rich Interactive Environment: Multimodal Utterances and Pointing in Aphasia. N. J. Enfield and S. C. Levinson (ed.) Roots of Human Society. Berg, New York, NY (in press) 97-125

Hutchins, E., Nomura, S., and Barbara Holder.: The Ecology of Language Practices in Worldwide Airline Flight Deck Operations. Proceedings of CogSci2006 (2006) 363-368

McNeill, D.: Gesture & Thought. The University of Chicago Press, Chicago (2005)

Nomura, S., Hutchins, E., and Holder, B.: The Uses of Paper in Commercial Airline Flight Operations, Proc. of CSCW2006 (2006) 249-258

Sacks, H., Schegloff, E. A., and Jefferson, G.: A Simplest Systematics for the Organization of Turn-Taking for Conversation. Language 50 (1974) 696-735