01/05/2024 08:14

(Article begins on next page)

Routledge
Taylor & Francis Group

Judgments Concerning the Valence of Inter-Turn Silence Across Speakers of American English, Italian, and Japanese

Felicia Roberts
Department of Communication
Purdue University, West Lafayette, IN

Piera Margutti
Department of Language Sciences
Università per Stanieri di Perugia, Perugia, Italy

Shoji Takano
Department of English
Hokusei Gakuen University, Sapporo, Japan

The fact that people with minimal linguistic skill can manage in unfamiliar or reduced linguistic environments suggests that there are universal mechanisms of meaning construction that operate at a level well beyond the particular structure or semantics of any one language. The authors examine this possibility in the domain of discourse by focusing on how gaps arising at the juncture between 2 persons' turns-at-talk (inter-turn silences) are evaluated by speakers of typologically distinct languages: English, Italian, and Japanese. This cross-linguistic design allows the testing of both universal and relative aspects in orientation to silence. For this study, the effects of inter-turn silence are tested using study participants' ratings of speakers' willingness to comply with requests or agree with assessments that were embedded in conversations. In a $3 \times 2 \times 3$ between-groups design, 3 silence lengths (0 ms, 600 ms, or 1,200 ms) were crossed with 2 speech act types (requests

331

and assessments) in manipulations of telephone conversations that were modeled on an actual telephone call between friends. Native-speaking study participants, in their home countries, provided ratings on Likert-type scales. Ratings significantly decreased within each language group at longer inter-turn silences, indicating a generalized response to the gaps; however, means were also significantly different between groups, indicating different expectations for agreement.

Gathering meaning from others' talk and embodied action is the ongoing business of daily life. The fact that people with minimal linguistic skill can manage even in very unfamiliar or reduced linguistic environments attests to the seamless integration of social and cognitive skills for human meaning-making (Goodwin, 2002; Levinson, 2006). It also suggests that there are mechanisms for meaning construction that are shared across languages and that operate at a level beyond the particular structure or semantics of any one linguistic system. In this article, we examine one such mechanism in the domain of discourse processing by focusing on how gaps arising at the juncture of turns-at-talk (inter-turn silences) are evaluated both similarly and differently by speakers of typologically distinct languages. This approach allows us to engage both sociocultural and cognitive perspectives, providing a necessary, if as yet preliminary, step toward fuller studies of the relation between interactional norms and cognitive processing of speech.

This project extends findings from a series of experiments on inter-turn silence and prosody in American English (Roberts, Francis, & Morgan, 2006) by replicating, for native speakers of Italian and Japanese, the design of the final experiment in the English language project. Although additional prosodic variables were of interest and were collected in the full replication, those findings are reported elsewhere (Roberts & Margutti, 2008; Roberts & Takano, 2007). Our main concern in this analysis is with perceptions of inter-turn silence, as that is a robustly meaningful cue across the languages studied. Thus, this article draws on subsets of the original (American English) data and subsequent replication efforts in Italian and Japanese. This allows us to make cross-linguistic comparisons that focus on inter-turn silence.

The reason for this focus is that inter-turn silence is a key component in the construction and display of intersubjective understanding. This is particularly salient in the context of "conditional relevance" (Schegloff, 1972), those points in conversation where a first action (e.g. an invitation) establishes the expectation of some next action or response (e.g., acceptance, rejection). When seamlessly produced this machinery of mutual understanding goes unnoticed. From an ethnomethodological standpoint, the accomplishment of intersubjectivity is, thus, an "operation" (Garfinkel, 1967, p. 30); when there is even a brief gap between two mutually relevant turns, intersubjectivity is no longer invisible. As demonstrated in qualitative analyses of talk-in-interaction, participants orient to these

glitches as some form of trouble, as displayed in their pursuit of responses or reformulations of prior talk (Davidson, 1984; Pomerantz, 1984).

This project is grounded in these descriptive findings and builds on them to develop an experimental approach that is based on actual interaction. The aim is to capture a possible overhearer's perception of others' conduct with regard to inter-turn silence (cf., Fox Tree, 2002) and to directly compare these judgments across language groups. This can help inform us as to the generality (or not) of overhearers' responses to disruptions in the forward motion of talk.[1] Admittedly, this is an "external view" of the phenomenon (Heritage, 1995, p. 406), not an analysis of members' methods for organizing the "occasions and resources of understanding" (Schegloff, 1992, p. 1299). Nonetheless, the cross-linguistic approach to the study of inter-turn silence affords a glimpse at one possible entry point for understanding the seamless interrelation of the social/relative and the cognitive/universal in language use (Levinson & Enfield, 2006). We begin by reviewing these two approaches to silence, then provide a rationale for the choice of the additional languages under study (Italian and Japanese), and finally turn to the experimental design and study results.

CULTURAL, COGNITIVE, AND CONVERSATION ANALYTIC CONTRIBUTIONS

Anthropologists and sociolinguists have described cultural differences in the tolerance for silence in conversation, demonstrating and theorizing how those differences can affect intercultural interactions (e.g., Basso, 1970; Clancy, 1986; Lehtonen & Sajavaara, 1984; Nakane, 2006; Nwoye, 1985; Scollon & Scollon, 1979; Tryggvason, 2006; for edited collections about silence in conversation, see Jaworski, 1997, and Tannen & Saville-Troike, 1985). Viewing orientations to silence from the perspective of norms that are culturally relative and distinct, these studies provide grounds for an assumption that members of different language groups (as a surrogate for cultural groups) will differentially perceive the valence of inter-turn silence (i.e., the degree of negative or positive attribution to gaps in conversation).

Conversely, psychologists and psycholinguists have studied silence in terms of response latency, primarily in the context of factual questions. In this processing paradigm, latency is considered a cue to uncertainty or deception (for an overview, see Andersen, 1999). This cognitive approach, built on the conceptualization of latency to respond as a symptom or signal of processing effort

---

[1]Our aim should not be confused with an analysis of the "meaning" of silence. Silence itself is featureless (Levinson, 1983), taking on shades of meaning only in the context of social interaction.

in finding an answer (Glucksberg & McCloskey, 1981; Smith & Clark, 1993), assumes that answerers, in the moments of hesitancy, are both doing a search and monitoring the search process (Nelson, 1993). The outward appearance of this searching/monitoring process can be characterized as uncertainty (Brennan & Williams, 1995) possibly construed by listeners as a clue to deception in the making (Andersen, 1999) or conversely, simply thoughtfulness (Burgoon, Buller, & Guerrero, 1995). This research suggests, although it is less cross-culturally robust, that interpretations of inter-silence as "trouble" or thoughtfulness may be grounded in generalized cognitive processes related to memory, retrieval, integration of cues, and so on. Most of this cognitively oriented research has studied silence (and disfluency) in terms of individual rather than collaborative activity, although alternative approaches have been advocated (cf. Clark, 1994; Schober & Bloom, 2004).

What differentiates this study from both anthropological and psychological studies is that it entertains the possibility of a cultural difference in orientation to silence, but examines it at the level of specific collaborative conversational activities (i.e., requests and assessments). Rather than addressing the orientation to silence as *either* culturally specific *or* cognitively generalized, we address perception of silence in terms of both and, more important, in terms of particular discourse environments where specific socially appropriate actions are expected. This is distinct from research on response latency and filled pauses in which study participants are evaluating responses to factual questions (e.g., Brennan & Williams, 1995) or direct interrogatives concerning beliefs (e.g., Fox Tree, 2002).

We base our design on the turn-taking model as empirically derived by Sacks, Schegloff, and Jefferson (1974). From that vantage point, it is clear that silence in conversation is a great deal more than an absence of speech. Indeed, Conversation Analysts[2] have demonstrated, as briefly noted earlier, that when inter-turn silence grows subsequent to a speaker's utterance that sets up conditional relevance, it is indicative of possible trouble at that point in the interaction (Davidson, 1984; Pomerantz, 1984). This trouble is empirically available through speakers' routine practice of producing "subsequent versions" (Davidson, 1984, p. 104) of their turn at talk. The generalization that dispreferred actions can be structured around inter-turn silence (i.e., realized as delays in launching turns at talk) is borne out in qualitative studies of conversation in typologically disparate languages such as Japanese (e.g., Mori, 1999; Tanaka, 2005), Finnish (e.g., Sorjonen, 1996), and Italian (e.g., Monzoni, 2007), to name a few.

---

[2]We capitalize Conversation Analyst to represent those scholars working within a particular tradition that is rooted in phenomenological sociology and ethnomethodology. Heritage (1984) explicates these intellectual roots and connections (see also Maynard & Clayman, 1991).

What is not known, however, is how particular courses of action in the context of particular inter-turn silence lengths are perceived by native speakers of different languages. Such comparative research will help us refine our understanding of the extent to which interactional norms are part of the variable "vocabulary" of language (Chomsky, 1995) or whether tolerances for specific lengths of silence are universal and, therefore, have implications for understanding the social (as well as cognitive) processing of speech. To explore these possibilities, we proceeded with the design and implementation of this study based on the following exploratory questions:

1. Do linguistic/cultural groups differ significantly in their judgments of addressees' willingness to comply with requests or agree with assessments as a function of inter-silence?
2. Do groups differ in their judgments of addressees' willingness to comply with requests or agree with assessments based on the sequential environment (i.e., requests vs. assessments) in the context of inter-turn silence?

The second question brings a unique perspective to the study of inter-turn silence. Because we choose "adjacency pair" as the relevant level of analysis, we move the study of silence away from descriptions of generalized norms and toward a comparative investigation based on judgments concerning the same courses of action (requests and assessments). We are, therefore, better positioned to draw conclusions about the orientation to inter-turn silence at interactionally relevant and comparable moments.

LANGUAGE GROUPS: WHY ITALIAN AND JAPANESE?

Stereotypes abound about the interpersonal styles of many linguistic/cultural groupings. In terms of orientation to inter-turn silence (or propensity for overlapping talk) one relatively unexplored explanation for differing conversational styles is suggested in Schegloff, Ochs, and Thompson (1996). They noted that structures of projectabilty are likely different across typologically disparate languages; and, therefore, the opportunities for overlap would be less or more possible (pp. 28–29). They suggested a possible "confluence of grammar, culture and turn-taking organization," which could "raise the possibility ... of early entry [into a preceding turn] as a common practice ..." (p. 30). We chose, therefore, to study speakers of typologically distinct languages, which also have distinct conversational stereotypes attached to them, to provide an initial snapshot of the degree to which culture/language background is a relevant construct in orientation to inter-turn silence.

Two disparately stereotyped groups, from an Anglo-centric standpoint, are Italian speakers and Japanese speakers: One is characterized as voluble and possibly less tolerant of silence (in that speakers are more prone to overlapping talk), whereas the other is viewed as less voluble and more tolerant of silence. These distinctly opposed stereotypes, for better or worse, have generated interest among scholars, although with mixed results.

Italian literature on silence and its functions in conversation is scarce; the general claim that Italians do not tolerate silence and that they tend to enter conversation in overlap more frequently than speakers of other languages do is often mentioned or implied, but rarely demonstrated. A prefatory chapter in Bazzanella (2002) is promisingly dedicated to silence, but contains few references to empirical studies. The studies of Italian that consider features of the turn-taking system with respect to silence (Monzoni, 2004, 2007; Shultz, Florio, & Erickson, 1982) provide contrasting analyses—one that draws on culture as an explanation (Shultz et al., 1982) and one that draws on participation structure (Monzoni, 2004, 2007). Nonetheless, these empirical studies do indicate that there is some orientation to minimizing inter-turn silence among Italians in multi-party settings, and that this tends to produce overlapping talk.

It is possible to draw quite a different conclusion about Japanese speakers. Based on anthropological and sociolinguistic research, one could infer that the Japanese differ from both Americans and Italians in their perceptions of silence. Sociocultural norms have been described that stress both hierarchy and harmony in interpersonal relationships (Watanabe, 1993). Thus, there may be potential social and psychological preferences for agreement among Japanese speakers along with a more generic structural preference (Pomerantz, 1984) for minimized inter-turn gaps at sequentially relevant moments.[3] Such a confluence could lead to less tolerance for long gaps. Conversely, a strong underlying social orientation to agreement might predict *more* tolerance for long gaps, as there might be a presumed expectation for agreement; thus, delayed affirmative responses would not necessarily elicit negative judgments.

In sum, from the few studies of Italian and Japanese that focus on silence, it is not clear whether norms for orientation to silence are operating at the level of participation structure or at a broader cultural level or whether, in fact, there is any stable cultural norm. This study is certainly not the last word on cultural orientation to silence, but it does provide a stable footing of reasonably

---

[3]Although there is an important distinction in the literature between social and structural understandings of preference, these are not mutually exclusive conceptions. Indeed, they may be related under a broader understanding of affiliation or social solidarity. (On the distinctions between, and relations among, these 2 conceptions of preference, see Heritage, 1984; Lerner, 1996; Levinson, 1983.)

controlled, consequential discourse environments (requests and assessments) on which to test and build further investigation.

METHOD

Overview

We devised a series of telephone conversations that were based on actual telephone calls between college-aged friends. These dialogues were then performed by American, Italian, and Japanese university students who were native speakers. Digital recordings were made for subsequent sound editing. We inserted inter-turn silences between requests or assessments and the affirmative responses that followed them. Undergraduates at study sites in their home countries were recruited as study participants to listen to the dialogues and provide their judgments (as ordinal ratings) of addressees' willingness, within the recorded dialogues, to comply with requests or agree with assessments.

A central challenge in designing this study was to control, as much as possible, for the vocal qualities of the American, Italian, and Japanese speakers who performed the dialogues. To fully account for any vocal differences between the speakers, we would have had to enlist trilingual speakers to perform all of the dialogues, but even this "matched-guise" approach (Lambert, Hodgson, Gardner, & Fillenbaum, 1960) has been called into question (Bradac, Cargile, & Halett, 2001). That technique only works under the assumption that the multilingual speaker is equally skilled in each language or language variety, when in fact the speaker may present idiosyncratic differences of fluency in the various guises that could still affect hearer judgments (Bradac et al., 2001, pp. 139–141). Thus, we proceeded by controlling for production features *within* each language group, as detailed later, but we did not control paralinguistic qualities to the extent of finding one speaker who could perform in all three languages.

Study Participants

Native-speaking American English ($n = 70$), Italian ($n = 72$), and Japanese ($n = 80$) undergraduate students, all living in their home countries at the time of data collection, were recruited and compensated according to an approved institutional review board protocol. The sample populations did not differ reliably in age (American: $M = 19.97$, $SD = 2.57$; Italian: $M = 21.96$, $SD = 2.60$; Japanese: $M = 20.14$, $SD = 1.26$), $F(1, 217) = 2.35$, ns; but they differed in the representation of the genders. For the American English group, 60% were women ($n = 42$); for the Italian group, 74% were women ($n = 53$); and for the Japanese group, 84% were women ($n = 67$), $\chi^2(1, N = 222) = 54.50$,

$p < .01$. However, as discussed later, gender did not affect results across or within language groups.

Materials

To fully replicate the design of the earlier study of American English (Roberts et al., 2006), the same nine dialogues were used as the foundation for the Japanese and Italian study materials. The dialogues are based on the transcription of the opening of an actual telephone conversation between two female American college-aged friends (Roberts & Robinson, 2004). For this study, one native speaker for each language group translated the dialogues from English into the target languages. These were then back-translated into English by a second native speaker from the target language groups. This was to check that the dialogues were roughly equivalent in the different languages, yet still idiomatic and comprehensible for their local contexts. Each conversation culminated with a request or an assessment by the caller; the call recipient responded with an affirmative one-word response (further details follow).

Of the nine dialogues, six were manipulated for purposes of the study and three remained as they had been produced by the actors. These three dialogues ("lures") were interspersed among the six target stimuli to mask the manipulations. Although the stimuli with long gaps clearly had a different flavor than the un-manipulated dialogues, study participants had a variety of guesses concerning the purpose of the study: in response to our verbally posed debriefing question, "What do you think this study was about?," study participants mostly said "tone of voice" or "emotion in conversation," but some (roughly $\frac{1}{4}$) were able to pinpoint that it was about "pauses" in speech. This elicitation of study aims was done as a precursor to disclosing the actual aims of the study and was not designed as part of the data collection. Thus, although our conclusion about study participants' awareness of study aims is impressionistic, we can say that it was not generally obvious to participants that the study had targeted gaps in conversation.

Construction and recording of dialogues.   Each conversation, whether target or lure, included a mundane request or assessment made by the caller. These were about everyday topics concerning school flyers, going to the gym, or picking up a new computer. In response to the requests and assessments surrounding these three themes, the call recipient responded in the affirmative.

The target stimuli controlled for theme, but differed on speech act. For example, if the theme of the call concerned flyers for a school function, it ended with the caller either formulating a request in terms of that topic (e.g., getting a ride to pick up the flyers) or offering an opinion on the topic at hand (e.g., reporting that the flyers look good). The call recipient answered in

the *affirmative* for both sequence types: In English, the token "sure" followed requests, and "yeah" displayed agreement with the assessments. The roughly equivalent tokens used for Italian were *certo* for "sure" and *si* for "yeah." For Japanese, *iyo* was used for "sure" and *so danay* for "yeah." Appendix A contains examples in English and Appendix B contains the targeted translated sequences along with notes on similarity of syntactic projectablity.

All of the dialogues (target and lure) were enacted by native-speaking students in the same age range as the target study population. To control for gender and voice quality, the target stimuli in each language were performed by the same two women, who maintained the same caller-call recipient identity. Familiarity in all of the conversations was maintained by features such as omission of the caller's name, use of an informal register, and/or truncating the greeting sequence (Hopper, 1992; Schegloff, 1979). The person portraying the call recipient in the recording was directed to make responses to requests and assessments sound agreeable, but not unusually enthusiastic. The student actor was instructed to maintain a normal register and to respond affirmatively, with no sense of hesitation (i.e., to sound willing and agreeable in an everyday sort of way).

Clearly, one cannot control for all intervening variables when using human voices; some voices just sound friendlier or more sincere or more needy and those qualities can influence judgments. However, to perfectly match the acoustic characteristics through machine produced language would seriously undercut face validity and, as mentioned earlier, a matched-guise approach was neither possible nor completely reliable. Our best approach was to competently direct actors in their native languages to produce the kind of friendly, familiar demeanor we were seeking. We did not undertake a full study, either before or after choosing our actors, of whether their voices were considered "friendly" within their own cultural contexts. Such a study might yield some reassurance that each of the voices was, indeed, friendly, whatever that would mean in each culture, but whether they were equally so across the groups is simply impossible to calibrate or control. Readers concerned with issues of voice quality are invited to request digital files of the experimental stimuli by contacting Felicia Roberts.

Agreement token manipulations. To control for possible confounding from the acoustic qualities of the actors' slightly different agreement token pronunciations (during recording) we identified, following procedures detailed in Roberts et al. (2006), two median agreement tokens (1 "yeah" and 1 "sure" and their equivalents in the other languages) from all of the agreement tokens produced by each actor. The "sure" token and the "yeah" token, which fell in the middle in terms of duration, pitch range, and pitch change, were chosen as the default agreement tokens and edited into the appropriate dialogues. In other words, within each language condition the agreement tokens were identical in each dialogue so that only the inter-turn silence between the request or

assessment and the response differed. Thus, the acoustic quality of the responses was fully controlled.

Inter-turn silence manipulations.   Once the median agreement tokens were edited into the dialogues, silences were inserted between the first pair part (the request or assessment) and the agreement token ("sure" or "yeah"). Three lengths of inter-turn silence were used to test the interaction between silence and sequence type: 0 ms (no lag time), 600 ms, and 1,200 ms. These lengths were chosen to provide a baseline simulation of no gap between the request/assessment and the affirmative response, and then equal increments leading up to Jefferson's (1989) proposed limit of "approximately 1 s" as a standard maximum of silence in conversation. Silence was taken from other dead space in the dialogue (i.e., it was not a machine-produced silence) to best maintain the natural acoustic environment. These natural silences were spliced to the end of phonation on the request/assessment utterance as visually apparent from the sound wave.

Design

In a 3 × 2 × 3 mixed model with repeated measures, *silence length* (0 ms gap, 600 ms, and 1,200 ms) and *sequence type* (request and assessment) were within-group factors; *language group* (English, Italian, or Japanese) was the between-groups factor. These independent variables were manipulated in the context of constructed dialogues that simulated telephone conversations between two college-aged female friends. The dependent variable was rater perceptions of an addressee's willingness to comply with requests or agree with assessments as elicited from a written question following each dialogue. Judgments were captured using an ordered scale ranging from 1 (*not willing* or *not in agreement*) to 6 (*very willing* and *very much in agreement*). Raters were encouraged to use the whole scale, and some study participants circled adjacent values (i.e., they could provide a rating of 2.5 by circling the 2 and the 3 on the scale).

Three different orders of presentation were used. These were counterbalanced based on the inter-turn silence manipulations of the target stimuli. Admittedly, this was not a full counterbalancing of sequence types (2 levels) and inter-turn silence length (3 levels), which would have produced six presentation orders. Given our recruitment and administration procedure (in classrooms of consenting instructors; see the Procedure section) and no evidence from an item analysis that ratings between the six counterbalanced orders were significantly different for English (see Roberts et al., 2006, Experiment 1), we chose to take this more economical path. Thus, across the three presentation orders, each inter-turn silence appeared in each position an equal number of times, but not every Speech Act × Gap combination was heard in every possible position.

The dialogues intended to mask the purpose of the study (lures) always appeared in the same place in the audio presentation. Each presentation order began with a masking dialogue to get study participants oriented to the task. The other masking dialogues appeared fourth and seventh in the series of nine conversations.

Each of the three audio presentation orders consisted of nine conversations: Three ended with the assessment/response pairing ("I think it/they look[s] pretty good," followed by "Yeah"), which was separated by one of the three inter-turn silence lengths (0 ms, 600 ms, or 1,200 ms); three conversations ended with the request/response pairing ("Can you give me a ride over there," followed by "Sure"), separated by the three inter-turn silence lengths; and three non-target conversations ended with an assessment or request (also paired with the affirmative response tokens "sure" and "yeah"), but they were un-manipulated. The natural inter-turn silence, as originally produced by the actors, was left intact for the non-target stimuli. Thus, nine stimuli were heard by each study participant, and each Silence Length × Speech Act type manipulation was heard once.

Procedure

Study participants were recruited by the study authors from classrooms where instructors had given permission for the last 10 min of the class to be used for student participation in "a study about communication." Students who were willing remained in the classroom and were given consent forms to review and sign. Thus, from these classrooms, independent samples of convenience were drawn. Different classrooms were used for each administration such that no student ever participated twice.

For each sample group, students were given the same recorded instructions: that they would be hearing "several telephone conversations among a group of friends" and that each friend "was just relaxing at home." The recording also said that after each conversation there would be a question to answer about the conversation. Including the consent process and debriefing, the process took about 15 min. About 6 min were required for the actual administration of the experimental protocol. Study participants (and the student actors who recorded the dialogues) were compensated for their time.

RESULTS

The fact that the study participants were homogenous in terms of age (see the Participants section) and that exploratory analysis revealed that their gender had no significant effect on their judgments of the silences—$F(1, 218) = 0.896$,

ns—analyses reported here do not include gender or age as covariates.[4] In the remainder of this section, therefore, we report results from an omnibus analysis of variance (ANOVA), including calculation and interpretation of effect size for relevant variables. Key interactions and major findings are then explored in subsections. For ease of presentation, the inter-turn silence length variable is referred to as "gap," and the sequence type (request or assessment) is referred to as "speech act."

The ANOVA indicated statistically significant main effects for the three independent variables: gap, $F(2, 438) = 529.76$, $p < .01$; language group, $F(2, 219) = 51.82$, $p < .01$; and speech act, $F(1, 219) = 50.32$, $p < .01$. The interaction of gap and language group was statistically significant, $F(4, 438) = 10.57$, $p < .01$; as were the interactions of speech act and language group, $F(2, 219) = 44.54$, $p < .01$; and gap and speech act, $F(2, 438) = 3.21$, $p < .01$. There was no three-way interaction ($p = .381$).

Effect size is one way of estimating the contribution of a particular factor to an observed effect. Although the ANOVA indicated statistically significant main and two-way interaction effects for all three independent variables, the effect sizes, calculated using partial eta-squared (η_p^2), for Gap × Language ($\eta_p^2 = .08$) and Gap × Act ($\eta_p^2 = .01$), were small. Although interpretation of effect sizes is made cautiously, subject to several limitations (Cohen, 1988), it does appear that gap length ($\eta_p^2 = .70$) and language group ($\eta_p^2 = .32$) contribute large percentages (70% and 32%, respectively) of the overall variance. Speech act accounts for 19% ($\eta_p^2 = .19$) and Sequence × Language Group is 29% ($\eta_p^2 = .29$). Although these values are not additive (i.e., partial eta-squared does not provide estimates that sum to 100%), they nonetheless provide an indication of the contribution of each factor (or interaction) as if it were the only variable (Young, 1993). The salient result from the effect size calculations is that inter-turn silence length (gap) has, by far, the strongest effect on study participants' ratings of an addressee's willingness to comply with requests and agree with assessments.

American, Italian, and Japanese Raters All Judged Longer Silences More Negatively, but Differed Sufficiently to Suggest Disparate Cultural Orientations to the Inter-Turn Silences

As Figure 1 illustrates, regardless of language background, all raters judged speakers to be less willing to comply with requests or agree with assessments the longer the speaker paused before agreeing: linear trend, $F(1, 219) = 794.08$, $p < .01$. This indicates that, across all three language groups, there is a decreasing sense of speaker agreement or compliance with increased silence.

---

[4]We tested this using an analysis of covariance model with all levels of silence length and speech act as within-subjects variables, language group as a between-subject variable, and gender and age as covariates.
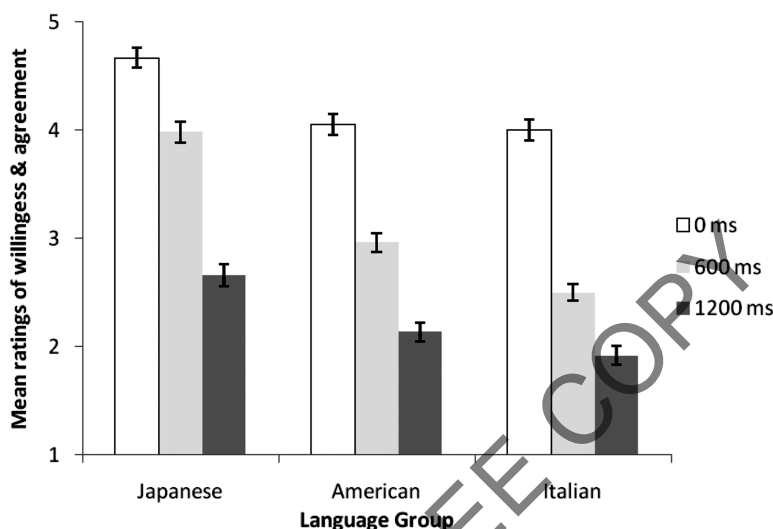
FIGURE 1   Effect of inter-turn silence length on judgments of speakers' agreement and compliance (error bars represent standard error of the mean).

The effect of language group is also salient in Figure 1. Overall, the Japanese study participants rated the speakers as more agreeable than either the Italian or the American participants. Between-groups differences were statistically reliable in *post hoc* comparisons (Bonferroni-corrected) across all gap conditions when comparing the Japanese and American groups ($M_{diff} = 0.721$, $p < .01$) and the Japanese and Italian groups ($M_{diff} = 0.935$, $p < .01$). The only statistically reliable mean difference between the American and Italian raters was in the 600 ms gap condition ($M_{diff} = 0.46$, $p < .01$).

The distinctly different ratings between the Japanese and the other groups suggests that what constitutes "agreeable" in the context of conversational gaps is calibrated slightly differently for them, and that they may be entering the scale at a different point. The statistically reliable differences between all three groups in the 600 ms gap condition suggests that a roughly $\frac{1}{2}$ s gap may be enough to distinguish cultural differences in orientation to inter-turn silence.

Italian Raters Judged the Smaller Inter-Turn Gaps as More Problematic, Whereas the Japanese Judged the Longer Inter-Turn Gaps as More Problematic

To further explore the interaction of gap length and language group, we examined raters' judgments of the inter-turn silences within their language group and then compared the groups descriptively. Most notable is the comparison between the

Japanese and the Italian raters. As visualized in Figure 1, the magnitude of the drop in ratings by the Italians from 0 ms to 600 ms is double that of the Japanese, whereas the drop in ratings for the Japanese is double that of the Italians when comparing the 600 ms and 1,200 ms gaps. This pattern of difference ratings points to the possibility that the different language groups do not simply differ in where they enter the scale of "agreeableness," but may also have different thresholds at which they find silence problematic.

Americans, Italians, and Japanese Respond Slightly Differently to the Sequence Types (Requests vs. Assessments)

A unique contribution of our approach to the valence of inter-turn silence was to examine the effects of sequence type ("speech act") on raters' judgments. Because most prior research on gaps in conversation is situated in the environment of question-asking (factual/trivia questions), we chose to manipulate the stakes of the face threat by using more naturalistic, collaborative, and consequential social actions. To further refine our understanding of the significant main effect of speech act and the interaction of speech act and language group, we explored differences and similarities within each condition.

Overall, raters judged the request sequences as sounding more agreeable (i.e., the ratings were generally higher in this condition; see Table 1), but only the Japanese rated requests substantially higher than they did assessments ($M_{\text{diff}} = 1.04$, $p < .01$). For all of the groups, this is likely due to the relative strength of the agreement carried in the different tokens ("sure" vs. "yeah," as we discuss further later).

TABLE 1
Mean Ratings for Each Language Group at Each Inter-Turn Silence Length
and for Each Sequence Type

Inter-Turn Silence Length	Language	n	Assessments		Requests	
			M	SD	M	SD
0 ms	Japanese	80	4.08	1.61	5.26	0.70
	English	70	3.96	1.05	4.14	0.98
	Italian	72	3.92	1.15	4.21	0.96
600 ms	Japanese	80	3.56	1.07	4.40	1.20
	English	70	2.89	0.99	3.04	0.73
	Italian	72	2.63	0.83	2.38	0.99
1,200 ms	Japanese	80	2.10	0.93	3.23	1.22
	English	70	2.10	0.75	2.17	0.85
	Italian	72	2.01	0.79	1.82	0.90

Although the comparison of raters' judgments of the speech act types may be obscured by the semantics of the different response tokens, the result for the request condition *between* the three language groups is clear: Participants' ratings are significantly different, $F(2, 221) = 102.731$, $p < .01$—a finding that was reliable in all post hoc pairwise comparisons. This may reflect different underlying cultural orientations to compliance with requests in general, but this is an initial speculation and remains to be systematically explored in further studies.

In sum, what seems to be driving the interaction effect of sequence type and language group is the fact that (a) the Japanese participants rated requests and assessments quite differently; and (b) between the three language groups, primarily requests were rated significantly differently from one group to the next.

DISCUSSION

This study was designed to examine the effect of inter-turn silence and sequence type (speech act) on native speakers' judgments of an addressee's willingness to comply with requests or agree with assessments. We tested for these effects across identical "friendly telephone conversations," which were idiomatically translated from American English for Italian and Japanese study participants. The limitations of this exploratory experimental approach are acknowledged and addressed, but we begin with the strengths of our design, a discussion of the main findings, and their relevance for building research that could help synthesize the contributions of cognitive and social approaches to discourse studies.

The basis for the experimental design was the empirically derived turn-taking model as described in Sacks et al. (1974). Empirical research based on that model indicates both a preference for progressivity in talk (Stivers & Robinson, 2006) and that gaps arising in the context of conditional relevance (utterances designed for, and expecting, responses of some sort) will prove problematic for speakers (Davidson, 1984; Pomerantz, 1984). What the prior research could not help us to tease apart is the extent to which this perception of "trouble" would be truly universal (all speakers responding the same to the same silence lengths) or relative (cultural groups displaying different norms), or whether there was some intersection of these two dimensions. Our aim, therefore, was to maximally control for a variety of acoustic, thematic, and speech act factors so that we could explore potential relations among language/cultural group, sequence type, and inter-turn silence.

Among the American, Italian, and Japanese undergraduate students in our sample, length of the inter-turn silent gap matters. Across language groups, the longer the inter-turn silence length, the lower the ratings of willingness to comply

with requests or agree with assessments. This supports emerging evidence of a widely shared social orientation to silence as problematic (Stivers et al., 2009).

A socially oriented theory might posit that a preference for progressivity in conversation is what underlies this "distaste" for the growing silence. A more individualistic (cognitive) theory would claim that elapsed time reflects searching and monitoring activity as participants spin out attributions in their minds. We underscore the argument presented elsewhere (e.g., Clark, 1996; Levinson, 2006) that these are complementary dimensions of analysis and should not be either/or explanations. These levels must be integrated in future research as discourse scholars become more familiar with and find ways to explore and test the notion of a socially embedded interaction engine (Levinson, 2006). This study provides an initial step in that direction.

Despite the negative judgments of longer silences across language groups, results also indicate that the American, Italian, and Japanese raters differed in their tolerance for the different gaps. By *tolerance*, we simply mean that ratings differ between language groups in a way that reflects slightly different underlying expectations for agreement (among peers, in this study) and for the rapidity with which agreement is offered. The Japanese average ratings were consistently highest on ratings of willingness and agreement, the Italians tended to give the lowest ratings, and the American judgments were generally in between (although tending closer to the Italian group's ratings).

An alternative explanation for the different ratings among groups would be based on an assumption that the different linguistic structures of the languages differentially stall a projectable possible completion point (Schegloff et al., 1996). Briefly, projectability refers to the hearer's ability to anticipate that a particular type of utterance is unfolding. Because the target stimuli in all of the languages were similar on this dimension (see Appendix B), we were confident that responses to the inter-turn silences were not based on differential syntactic processing of the utterance underway. We suggest, therefore, that projectability being equal in these dialogues it is the inter-turn silence itself that is driving the disparity in judgments, not hesitation due to completion of syntactic processing.

Thus, any group distinction in orientation to silence is not about differential need for syntactic processing based on formal features of the language, but about cultural expectations for the speed with which agreement is offered. If there is a baseline sense that agreement is expected (e.g., in the Japanese context), it makes sense that ratings of willingness and agreement among the Japanese are less attenuated, as we found, from 0 ms to 600 ms. Conversely, if there is less expectation of agreement, then as soon as some silence is heard, judgments are affected and any additional increments, although still negatively valenced, are perhaps less marked (e.g., in the Italian context, as we found). This pattern merits further investigation using additional, and incremental, lengths of silence. In this

way, it might be possible to get a sense of a dose response and to more finely tune the overall finding of culturally differentiated response patterns within a universally obvious sense of decreasing agreement.

The fact that addressees' responses to requests were rated as reflecting greater willingness than responses to assessments is best explained by the fact that their response to the requests ("sure") is a stronger form of agreement than the response used for the assessments ("yeah"; on the full character of weak agreements, see Davidson, 1984, and Pomerantz, 1984). Thus, although the response tokens were prosodically similar, they were semantically distinct. We cannot conclusively state that it was the sequential environment itself (the conditional relevance set up by the request or the assessment) that was driving the perceptions. It could well be that the weaker form of agreement, which is broader in scope, is also more ambiguous and, therefore, did not elicit equally strong positive responses.

It is also possible that the Japanese study participants' significantly higher ratings of the "sure" condition (*iyo* in Japanese) is simply reflecting an even stronger distinction between the tokens *iyo* and *so danay* than for the English "sure" and "yeah" or Italian *certo* and *si*. It would be possible to retest the findings about sequence type using parallel terms for English and Italian (e.g., "yeah" and "si" for both requests and assessments), but it would be harder to match them prosodically. For Japanese, other tokens such as "eh" might be appropriate for gauging response to sequence type without risk of confounding from the semantics of the agreement tokens. This is clearly an area for further exploration within each language.

Whether the sequence types were tapping a semantic difference (based on response token) or a pragmatic difference (based on assessment vs. request condition) or even a difference in face threat, it is important to note that the requests not only received consistently higher ratings across all groups and gap lengths, but they also were more distinctive between groups. Because requests can be considered face-threatening acts (i.e., in Brown & Levinson's, 1987, terms, a threat to negative face or the desire for autonomy), study participants may be reflecting cultural attitudes about not infringing on others' autonomy. As a more obvious threat, this may have been a more salient concern than that represented in the threat posed by an assessment (i.e., agreement with another's opinion).

In sum, the somewhat murky results for speech act open up the possibility for further clarification of the relative weight of the threat posed by the utterance versus the semantics of the response. Because all responses were affirmative, we were able to isolate the effect of the inter-turn silence, which was clearly the most powerful factor coloring study participants' judgments; however, the question does remain open as to the role of the response token itself in shading response to the sequence types.

Given our finding that ratings are more similar between all language groups at 1,200 ms than at 600 ms, there is reason to consider Jefferson's (1989) proposal of "approximately 1 s" (she recognized a span from 0.9 s–1.2 s) as a possible maximum for silence in conversation. However, because our silences were machine timed (and Jefferson was using an analogue procedure, measuring silence in relation to the surrounding talk), it may be that in absolute terms we overshot the mark. In as much as the Italian and American ratings were similar, and the Japanese ratings remained significantly different from both of those groups, we do not yet see strong support for the proposal.

Challenges for Cross-Linguistic Study of Discourse

There are numerous weaknesses in any experimental study of human social interaction. One cannot control for all of the natural, in-the-moment inventiveness and complexity of real conversation, particularly in a cross-linguistic study. One can raise questions about the representativeness of the study population (undergraduate students), voice comparability of the recorded speakers, recognize possible differences in cultural preference/aversion to the types of requests and assessments being made, and consider whether true translation of base dialogues from English to the other languages is even possible. All of these caveats are implicit in our discussion of findings and as limits to their generalizability.

Indeed, one could argue that the robust result concerning different cultural responses to silence could simply be an artifact of the method; although the dialogues were parallel in terms of semantic content, the quality of the voices of the different speakers in the different languages may have had some effect. Unfortunately, there is no body of acoustics literature comparing issues of "tone" (e.g., friendliness of voice) in these languages, nor sufficient research on what such acoustic correlates might be for each language. Thus, researchers must privilege their native speaker auditory perception over acoustic measurements until there are, if ever, reliable cross-cultural measures of the paralinguistic correlates of speaker affect.

As noted in the Method section overview, only a matched-guise approach using trilingual actors might have addressed such concerns about paralinguistic comparability, but even that technique is not perfect (Bradac et al., 2001). Multilingual speakers may not have full command over their affective orientation to the languages they speak and may, therefore, have differences of fluency or vocal quality in the various guises they are asked to perform, which then can affect hearer judgments.

The most relevant and potentially feasible comparison to make, given the aims of this study, would be to assess speaking rates because a silence might sound long or short depending on the speed of the surrounding talk. However, speaking rates are calculated on different units of analysis for the languages in this study:

Japanese is a mora-timed language, English is stress-timed language, and Italian is generally considered a syllable-timed language; but such classifications are not uncontroversial (see Fletcher, 2010).[5]

From our vantage point as scholars of language and social interaction, the major weakness of this study is the reliance on contrived conversations and assessment of external perceptions (listener judgments), rather than internally motivated analyses of natural interaction. We recognize that our approach is not a substitute for the internal proof procedures of qualitative analysis (Heritage, 1995) and that we tread dangerously by treating "sequence type" (request and assessment) as an abstract category (i.e., a speech act), rather than as a particular and real course of action (see Schegloff, 1995, pp. xxviii–xxix).

However, within the confines of the experimental method, we have worked to ground the design in findings from studies of talk-in-interaction and based the materials on actual telephone calls. Our aim was to investigate a measurable and manipulable aspect of the operation of the turn-taking machinery (gaps) within specific courses of action; we were not attempting to provide an account or analysis of the course of action itself. We forced, in fact, a particular hearing of the silences—one in which agreement was at stake rather than, for example, friendliness or thoughtfulness or comfort or comprehension. We make no claim that disagreement is the only interpretation that study participants might have conjured from the silence.

CONCLUSION

There is a striking overall similarity of judgments in response to inter-turn silence across American, Italian, and Japanese raters, yet there are also compelling fine-grained differences between the groups. This lends support to furthering the investigation of the complex interrelation among cognition, social practices, and discourse processes (see, *inter alia*, Gee, 1992; Lave & Wegner, 1991; Levinson & Enfield, 2006; Ochs, 1996; Rogoff, 1990; for discussions of usage-based approaches to cross-linguistic studies of language universals, see Sidnell, 2007, and Tomasello, 2003).

---

[5]Because Italian and English are, theoretically, comparable on syllable units, we calculated speech rates, *post hoc*, for the English and Italian speakers in our stimuli. They were remarkably similar: 319 syllables per minute and 336 syllables per minute, respectively (using a measure of speaking rate, not articulation rate.) Although we cannot compare this to a Japanese rate of speech measure, we can say that the same semantic content was produced by that speaker in roughly the same amount of time as the Italian rate of speech (7.35 s and 6.22 s, respectively.) Ultimately, however, this analysis is vacuous because the Americans only used 4.8 s to produce the same stretch of talk, yet their rate of speech is the same as the Italians. Thus, from various vantage points, speech rate is a problematic comparison.

Clearly, something generalized about human perceptual processing generates an observable-reportable phenomenon of silence as indicative of trouble in conversation. In turn, the valence of that silence (as good, bad, long, or short) is, as has been argued, culturally conditioned. This type of Pan-human, ontological basis for social cues has been strongly argued by researchers of non-vocal communication, as in (albeit controversial) proposals on the universality of several facial displays of emotion (Ekman, Sorenson, & Friesen, 1969), the deployment of which is nonetheless governed by cultural display rules (e.g., Ekman & Friesen, 1975; Saarni, 1993).

Tolerance for silence is, of course, a dynamic factor that will significantly vary depending on the types of communicative activities underway and the identities being enacted in and through the deployment of silence. Although we assume that norms are flexible and deployable-avoidable on an "as needed basis" as resources for accomplishing and enacting identity and solidarity, we also maintain that there is room for considering the more universal or generalized cognitive underpinnings of these practices. This study allows us to move beyond descriptive comparisons between languages to a position where the "adaptations and inflections" of generic features of talk-in-interaction (Sidnell, 2007, p. 230) can be specifically examined. This involves breaking down the artificial boundary between social and cognitive. Although it may be somewhat tidier to study them separately, that bifurcation is difficult to maintain as we move toward more comprehensive understandings of human communication in terms of discourse processes and processing.

REFERENCES

Andersen, P. A. (1999). *Nonverbal communication: Forms and functions*. Mountain View, CA: Mayfield.

Basso, K. (1970). To give up on words: Silence in Western Apache culture. *Southwestern Journal of Anthropology, 26*, 213–230.

Bazzanella, C. (2002). Le voci del silenzio [Voices of silence]. In C. Bazzanella (Ed.), *Sul dialogo. Contesti e forme di interazione verbale* (On dialogue. Contexts and structures of verbal interaction) (pp. 35–44). Milano, Italy: Edizioni Angelo Guerini e Associati SpA.

Bradac, J. J., Cargile, A. C., & Halett, J. (2001). Language attitudes: Retrospect, conspect, and prospect. In H. Giles & P. Robinson (Eds.), *The new handbook of language and social psychology* (pp. 137–155). Chichester, UK: Wiley.

Brennan, S. E., & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory & Language, 34*, 383–398.

Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge, UK: Cambridge University Press.

Burgoon, J. K., Buller, D. B., & Guerrero, L. K. (1995). Interpersonal deception: IX. Effects of social skill and nonverbal communication on deception success and detection accuracy. *Journal of Language and Social Psychology, 14*, 289–311.

Chomsky, N. (1995). *The minimalist program.* Boston, MA: MIT Press.

Clancy, P. M. (1986). The acquisition of communicative style in Japanese. In B. B. Schieffelin & E. E. Ochs (Eds.), *Language socialization across cultures* (pp. 213–250). Cambridge, UK: Cambridge University Press.

Clark, H. (1994). Managing problems in speaking. *Speech Communication, 15,* 243–250.

Clark, H. (1996). *Using language.* Cambridge, UK: Cambridge University Press.

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

Davidson, J. (1984). Subsequent versions of invitations, offers, requests, and proposals dealing with potential or actual rejection. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 102–128). Cambridge, UK: Cambridge University Press.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A field guide to recognizing emotions from facial clues.* Englewood Cliffs, NJ: Prentice Hall.

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotions. *Science, 164,* 86–88.

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (pp. 523–602). Malden, MA: Blackwell.

Fox Tree, J. E. (2002). Interpreting pauses and *ums* at turn exchanges. *Discourse Processes, 34,* 37–55.

Garfinkel, H. (1967). *Studies in rthnomethodology.* Englewood Cliffs, NJ: Prentice Hall.

Gee, J. P. (1992). *The social mind.* London, UK: Bergin & Garvey.

Glucksberg, S., & McCloskey, M. (1981). Decisions about ignorance: Knowing that you don't know. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 311–325.

Goodwin, C. (2002). Conversational frameworks for the accomplishment of meaning in aphasia. In C. Goodwin (Ed.), *Conversation and brain damage* (pp. 90–116). New York, NY: Oxford University Press.

Heritage, J. (1984). *Garfinkel and ethnomethodology.* Cambridge, UK: Polity Press.

Heritage, J. (1995). Conversation analysis: Methodological aspects. In U. M. Quasthoff (Ed.), *Aspects of oral communication* (pp. 391–418). Berlin, Germany: de Gruyter.

Hopper, R. (1992). *Telephone conversation.* Bloomington, IN: Indiana University Press.

Jaworski, A. (Ed.). (1997). *Silence: Interdisciplinary perspectives.* Berlin, Germany: de Gruyter.

Jefferson, G. (1989). Preliminary notes on a possible metric which provides for a "standard maximum" silence of approximately one second in conversation. In D. Roger & P. Bull (Eds.), *Conversation: An interdisciplinary perspective* (pp. 166–196). Clevedon, UK: Multilingual Matters.

Lambert, W., Hodgson, R., Gardner, R., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology, 60,* 44–51.

Lave, J., & Wegner, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge, UK: Cambridge University Press.

Lehtonen, J., & Sajavaara, K. (1984). The silent Finn. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (pp. 193–201). Norwood, NJ: Ablex.

Lerner, G. (1996). Finding "face" in the preference structures of talk-in-interaction. *Social Psychology Quarterly, 59,* 303–321.

Levinson, S. C. (1983). *Pragmatics.* Cambridge, UK: Cambridge University Press.

Levinson, S. C. (2006). On the human "interaction engine." In S. C. Levinson & N. J. Enfield (Eds.), *Roots of human sociality* (pp. 39–69). Oxford, UK: Berg.

Levinson, S. C., & Enfield, N. J. (Eds.). (2006). *Roots of human sociality.* Oxford, UK: Berg.

Maynard, D. W., & Clayman, S. E. (1991). The diversity of ethnomethodology. *Annual Review of Sociology, 17,* 385–418.

Monzoni, C. M. (2004). Do Italians "prefer" disagreeing? Some interactional features of disputational talk in Italian multi-party family interaction. In K. Aijmer (Ed.), *Understanding and*

misunderstanding in dialogue. Selected papers from the 8th IADA conference, Göteborg 2001 (pp. 119–130). Tübingen, Germany: Max Niemeyer Verlag.

Monzoni, C. M. (2007). *Disputational talk in Italian multi-person family interaction.* Unpublished manuscript, University of York, UK.

Mori, J. (1999). *Negotiating agreement and disagreement in Japanese: Connective expressions and turn-construction.* Philadelphia, PA: John Benjamins.

Nakane, I. (2006). Silence and politeness in intercultural communication in university seminars. *Journal of Pragmatics, 38*, 1811–1835.

Nelson, T. (1993). *Metacognition: Core readings.* Englewood Cliffs, NJ: Prentice Hall.

Nwoye, G. (1985). Eloquent silence among the Igbo of Nigeria. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (pp. 185–191). Norwood, NJ: Ablex.

Ochs, E. (1996). Linguistic resources for socializing humanity. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 407–437). Cambridge, UK: Cambridge University Press.

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

Roberts, F., Francis, A. L., & Morgan, M. (2006). The interaction of inter-turn silence with prosodic cues in listener perceptions of "trouble" in conversation. *Speech Communication, 48*, 1079–1093.

Roberts, F., & Margutti, P. (2008, March). *The interaction of inter-turn silence, sequence type, and prosody in American and Italian perceptions of trouble in interaction.* Paper presented at the 30th annual convention of the American Association of Applied Linguistics, Washington, DC.

Roberts, F., & Robinson, J. D. (2004). Interobserver agreement on first-stage conversation analytic transcription. *Human Communication Research, 30*, 376–410.

Roberts, F., & Takano, S. (2007, May). *The interaction of inter-turn silence and prosody in American and Japanese perceptions of trouble in interaction.* Paper presented at the 57th annual convention of the International Communication Association, San Francisco, CA.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context.* New York, NY: Oxford University Press.

Saarni, C. (1993). Socialization of emotion. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 435–446). New York, NY: Guilford.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*, 696–735.

Schegloff, E. A. (1972). Sequencing in conversational openings, In J. J. Gumperz & D. H. Hymes (Eds.), *Directions in sociolinguistics* (pp. 346–380). New York, NY: Holt, Rinehart, & Winston.

Schegloff, E. A. (1979). Identification and recognition in telephone conversation openings, In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 23–78). New York, NY: Irvington.

Schegloff, E. A. (1992). Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology, 97*, 1295–1345.

Schegloff, E. A. (1995). Introduction. In G. Jefferson (Ed.), *Lectures on conversation* (Vol. 1, pp. ix–lxii). Cambridge, UK: Blackwell.

Schegloff, E. A., Ochs, E., & Thompson, S. A. (1996). Introduction. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 1–51). Cambridge, UK: Cambridge University Press.

Schober, M. F., & Bloom, J. E. (2004). Discourse cues that respondents have misunderstood survey questions. *Discourse Processes, 38*, 287–308.

Scollon, R., & Scollon, S. B. K. (1979). *Linguistic convergence: An ethnography of speaking at Fort Chipewyan, Alberta.* New York, NY: Academic.

Shultz, J. J., Florio, S., & Erickson, F. (1982). Where's the floor? Aspects of the cultural organization of social relationships in communication at home and in school. In P. Gilmore & A. Glatthorn

(Eds.), *Children in and out of school. Language and ethnography Series #2* (pp. 88–123). Washington, DC: Center for Applied Linguistics.

Sidnell, J. (2007). Comparative studies in conversation analysis. *Annual Review of Anthropology, 36*, 229–244.

Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory & Language, 32*, 25–38.

Sorjonen, M. (1996). On repeats and responses in Finnish conversations. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 277–327). Cambridge, UK: Cambridge University Press.

Stivers, T., Enfielf, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences, 106*, 10587–10592.

Stivers, T., & Robinson, J. D. (2006). A preference for progressivity in interaction. *Language in Society, 35*, 367–392.

Tanaka, H. (2005). Grammar and the "timing" of social action: Word order and preference organization in Japanese. *Language in Society, 34*, 389–430.

Tannen, D., & Saville-Troike, M. (1985). *Perspectives on silence*. Norwood, NJ: Ablex.

Tomasello, M. (2003). Introduction: Some surprises for psychologists. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (pp. 1–14). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Tryggvason, M. (2006). Communicative behavior in family conversation: Comparison of amount of talk in Finnish, Swedish Finnish, and Swedish families. *Journal of Pragmatics, 38*, 1795–1810.

Watanabe, S. (1993). Cultural differences in framing: American and Japanese group discussions. In D. Tannen (Ed.), *Framing in discourse* (pp. 176–209). Oxford, UK: Oxford University Press.

Young, M. A. (1993). Supplementing tests of statistical significance: Variation accounted for. *Journal of Speech & Hearing Research, 36*, 644–656.

APPENDIX A
Examples of a Request and an Assessment in English

1. (Telephone rings)
    A: Hello?
    B: Rachel?
    A: Yeah,
    B: Hey it's me.
    A: Hey:: How's it goin.
    B: Good. I: just called the copy shop,
    A: uh huh
    B: An:d the flyers are ready. Can you give me a ride over there?
    A: Sure.

Question: How willing is Rachel to give her friend a ride?

Scale: 1   2   3   4   5   6

2. (Telephone rings)
   A: Hello?
   B: Rachel?
   A: Yeah,
   B: Hey it's me.
   A: Hey.:: How's it goin.
   B: Good. I just saw the flyers. They look pretty good.
   A: Yeah.

Question: How much does Rachel agree with her friend about the flyers?

Scale: 1   2   3   4   5   6

APPENDIX B
Examples of English, Italian, and Japanese Glosses of Targeted Request and Assessment Formulations

1. Request
   "I just called the copy shop and the flyers are ready …"
   English: Can you give me a ride over there?
   Italian: Ti va di darmi un passaggio in macchina?
   Japanese: Kuruma ni nosete tte kurenai?
2. Assessment
   "I just saw the flyers …"
   English: They look pretty good.
   Italian: Mi sembrano belli.
   Japanese: Yoku dekiteru yo ne.

In all three languages, the listener could likely guess that a request is coming prior to the end of the turn. In English, "Can you give me …" projects a yes/no request of some sort early in the turn, as do the formulations "ti va di" (glossable as, "Is it okay/convenient for you") in Italian and after "nosete" in Japanese (glossable as get/give a ride). Likewise, for the assessments, in Italian after "mi sembrano …" (glossable as "Seems to me"), it could be clear that an assessment is being formulated; and in Japanese, perhaps right after "Yoku" ("Well"), but certainly after "dekiteru" ("done").

From this, we conclude that projectability, from both grammatical and pragmatic standpoints, is equivalent in the three languages.