"Silence Also Speaks: Temporal Effects of Turn-Taking in Interlocutor Attitude Judgment"

Authors: Yu Wenbo, Fan Rong, Pan Wanke, He Yi, Liang Dandan, Liang Dandan

Date: 2025-06-05T00:00:00+00:00

Abstract

Although acoustic verbal symbols constitute the primary vehicle for conveying discourse meaning, silent turn-transition timing is also considered to serve a function in communicating interlocutors' attitudes. The present study conducted three experiments to examine how turn-transition timing influences individuals' judgments of interlocutors' attitudes under acceptance, rejection, and neutral response conditions. The experiments employed an eavesdropping paradigm, in which participants listened to conversational fragments containing request speech acts and rated the interlocutors' attitudes. Leveraging Bayesian inference-based methods for ordinal and binary data analysis, we found that turn-transition timing predicted judgments of interlocutors' attitudes across different response types, manifesting a universal temporal effect; however, the predictive effect was strongest under acceptance responses and weakest under neutral responses. Moreover, under acceptance and rejection response conditions, participants' initial judgments and judgment processes regarding interlocutors' attitudes showed significant differences. Finally, participants did not construe neutral responses as corresponding to negative attitudes, a finding that contradicts pragmatic theoretical perspectives. This study contributes to a comprehensive exploration of the cognitive mechanisms through which interlocutors construct discourse meaning.

Full Text

Silent but Audible — The Timing Effect of Turn-Taking in Interlocutors' Attitude Judgment

Yu Wenbo[1], Fan Rong[1], Pan Wanke[2], He Yi[1,3], Liang Dandan[1,4]

[1]School of Chinese Language and Culture, Nanjing Normal University, Nanjing 210097

[2]School of Psychology, Nanjing Normal University, Nanjing 210097
[3]School of Chinese Language and Culture, Beijing Normal University, Beijing 100875
[4]Interdisciplinary Research Center for Linguistic Science, University of Science and Technology of China, Hefei 230026

Abstract

Although audible speech symbols constitute the primary vehicle for conveying discourse meaning, silent turn-taking time is also believed to play a role in communicating interlocutors' attitudes. This study employed three experiments to examine how turn-taking timing influences individuals' judgments of interlocutors' attitudes under accepting, rejecting, and neutral response conditions. Using an overhearing paradigm, participants listened to conversational fragments containing request speech acts and rated the interlocutors' attitudes. Through Bayesian inference applied to ordinal and binary data analysis methods, we found that turn-taking time predicted attitude judgments across all response types, demonstrating a universal timing effect. However, the predictive effect was strongest under accepting responses and weakest under neutral responses. Furthermore, participants' initial judgments and judgment processes for interlocutors' attitudes differed significantly between accepting and rejecting response conditions. Finally, contrary to pragmatic theory, participants did not interpret neutral responses as conveying negative attitudes. This study contributes to a comprehensive understanding of the cognitive mechanisms through which interlocutors construct discourse meaning.

Keywords: turn-taking, time effect, request speech acts, Bayesian inference, ordinal data

1. Introduction

Language's most important function is communicative interaction and sharing knowledge and feelings [?]. Conversation, as the most common vehicle of language, plays a crucial role in conveying interlocutors' attitudes and establishing communication bridges between two (or more) parties. While audible linguistic symbols are the preferred vehicle for information transmission, silent turn-taking processes also exist in conversation. Previous research on turn-taking duration has primarily employed descriptive and conversation-analytic introspective methods, lacking investigation of cognitive processes. Turn-taking typically lasts 0.2s to 2s, raising the question of whether this duration can convey information or express interlocutors' attitudes just as audible linguistic symbols do. This study investigates this question.

1.1 The Timing Effect of Turn-Taking from a Linguistic Perspective

The most salient feature of conversation is smooth turn-taking [?, ?, ?], where interlocutors alternate occupying the floor and producing speech content within very short transition times. Stivers et al. [?] examined turn-taking times for "yes/no questions" across ten languages, finding they generally occur around 200ms. Liang [?] analyzed turn-taking times in dialogues from mainland Chinese film and television dramas, finding they typically range between 400ms and 600ms. Nevertheless, numerous instances of delayed transitions occur in conversation. In this study, we refer to the cognitive processing effects triggered by delayed transitions as the timing effect of turn-taking.

From a pragmatic perspective, delayed transition is defined as a phenomenon where, according to turn-taking rules [?], after the current speaker selects the next speaker, the selected next speaker fails to take over the floor within the standard turn-taking time, resulting in inter-turn silence [?, ?]. Compared to smooth turn-taking, delayed transition is relatively rare. Unlike explicit audible segments and rapid transitions, prolonged silent segments are ambiguous in meaning and considered a "marked" conversational phenomenon. Longer turn-taking times represent particularly special components in conversational contexts and play an important role in conveying discourse meaning and communicative intentions [?, ?, ?]. Relevance theory posits that discourse understanding is an ostensive-inferential process based on "cognition-inference," and any ostensive communicative act must ensure its maximal relevance [?]. Marked delayed transitions, as significant conversational pauses, serve as interlocutors' "ostensive communicative acts," thereby guiding interlocutors toward cognitive processing and inference. Since rejecting behavior violates social norms and represents a lower-priority action, delayed transition actually serves to avoid the pressure and embarrassment of potential verbal communicative acts [?]. Thus, the timing effect of turn-taking reflects interlocutors' unwilling negative attitudes or lower positive attitudes. Consider example (1)[1]: When the listener (Ai) encounters a delayed transition, they actively engage in cognitive processing, inferring the relatively long turn-taking time (1.3s) as the speaker (Dong)'s concealed rejection or unwilling attitude.

(1) (Ai and Dong are classmates; Ai calls Dong for help borrowing a crucible)
Ai: >Who's in your lab with the door open? I want to use your (.) .h crucible:
(0.4)
Dong: Crucible:?
(0.4)
(1.3)
Dong: Um:::hh. (0.2) This afternoon: I'll come over this afternoon.

1.2 The Timing Effect of Turn-Taking from a Psychological Perspective

[1]Selected from Yu Guodong. (2021). *A Conversational Analysis Study of Speech Acts*. Beijing: Foreign Language Teaching and Research Press.

While pragmatic theories can introspectively discuss the timing effect of turn-taking, these studies primarily focus on how conversational organization and form reflect discourse meaning, mapping discourse meaning onto speech content and form from the speaker's perspective, without examining from a cognitive processing perspective whether listeners can (accurately) recognize the discourse meaning of timing effects. Recent psycholinguistic research has begun investigating how individuals "perceive" different turn-taking times [?, ?, ?, ?], particularly how the duration of turn-taking in request speech acts affects judgments of interlocutors' attitudes—that is, whether delayed transition phenomena increase or decrease requesters' estimations of recipients' attitudinal willingness. Since it is difficult to have participants complete conversational tasks in real situations and estimate speakers' true intentions, most studies employ an overhearing task as the experimental paradigm. The experiment requires participants to listen to conversations containing request behaviors, with different turn-taking times set between the request and response, after which participants provide Likert scale ratings of the speaker's true willingness or acceptance willingness. Roberts et al. [?] designed experimental materials including three levels of turn-taking time (0ms non-delay condition, 600ms weak delay condition, and 1200ms strong delay condition), three conversation topics (taking flyers, viewing furniture, and exercising), and two speech acts (request and evaluation). After both request and evaluation behaviors, speakers provided positive response content ("yeah" or "sure"), as in example (2):

(2) A: Hello?
B: Rachel?
A: Yeah,
B: Hey it's me.
A: Hey how's it goin.
B: Good. I just called Kinkos.
A: uh huh
B: And the call-out flyers are ready. Can you give me a ride over there?
A: Sure!

Experimental results found that participants' ratings under both delay conditions were significantly lower than under the non-delay condition, meaning participants believed request recipients were unwilling to help requesters. This negative timing effect also exists among native speakers of American English, Italian, and Japanese [?]. Further analysis indicated that the significant decrease in acceptance willingness occurred in the 700ms–800ms range [?]. Additionally, some studies suggest that turn-taking time demonstrates interlocutors' attitudes and relationships. Templeton et al. [?] recorded audio of strangers chatting and subsequently asked both conversation participants and strangers to rate the social connection between conversational partners. Results showed that all raters consistently agreed that when turn-taking times were shorter, the social relationship between conversational partners was closer. Other research using perceived sincerity as the dependent variable found that delayed transitions negatively affected perceptions of speakers' sincerity when answering questions across voice, video, and situational skits, and that participants generally believed conversational delays occurred because interlocutors were concealing true thoughts or intentions, serving as the causal bridge between conversational delay and weak sincerity [?].

1.3 Research Questions

Both pragmatics and psycholinguistics have identified the timing effect of turn-taking from theoretical and empirical perspectives, recognizing that this silent segment can convey interlocutors' implicit attitudes and reveal social relationships. However, several unresolved issues remain. First, psychological experiments have designed conversational materials using only accepting responses, without examining results under other response conditions. In fact, request speech acts involve multiple response forms including direct acceptance (normal acceptance), qualified acceptance, direct rejection (flat rejection), and qualified rejection. Although direct rejection is uncommon, it still occupies a certain proportion in conversation[2] [?]. More importantly, the timing effect of turn-taking may have both universal patterns and specific characteristics across different response types. If the timing effect under accepting responses manifests as decreased acceptance willingness intensity, does the timing effect still exist under rejecting and neutral responses? Do its manifestations change? These questions require further comparative investigation. Notably, results under neutral responses (using "zaishuo" in this study) warrant particular examination. Renfu et al. [?] consider "zaishuo" a culturally distinctive refusal expression in Chinese. Since Chinese culture emphasizes "politeness" and "harmony," language often exhibits ambiguous meaning, with "zaishuo" being a typical representative of negative refusal through vague responses that defer issues. Through long-term evolution, the illocutionary force of "zaishuo" as refusal has gradually strengthened, with its semantics becoming increasingly fixed [?]. Therefore, pragmatics considers "zaishuo" to mean surface deferral, actual refusal, manifesting as euphemistic refusal to address face-threatening acts. However, this perspective lacks quantitative research, and whether the timing effect of turn-taking exists under this special response condition also warrants investigation. Finally, regarding methodology, previous studies have primarily used ANOVA based on normal distribution assumptions to analyze Likert rating data [?]. However, rating data are actually ordinal data, and using ANOVA can distort parameter estimation, increase Type I error rates, and reduce statistical power [?, ?]. Additionally, calculating rating means may overlook potential cognitive differences. For example, a mean rating of 4 could result from scores clustering around 4 or from a distribution with many 1s and 7s—two distribution types corresponding to completely different cognitive mechanisms.

[2]In Dutch corpora, direct rejection accounts for approximately 9.2% of responses.

Based on the above analysis, this study examined whether individuals' judgments of interlocutors' true attitudes are influenced by turn-taking duration under three response conditions: acceptance, rejection, and neutrality—that is, the timing effect. Experiment 1 used "haode" (okay) as the accepting response, Experiment 2 used "buhao" (not okay) as the rejecting response, and Experiment 3 used "zaishuo" (let's talk later) as the neutral response. The first two experiments required participants to provide seven-point ratings of the strength of acceptance (or rejection) attitudes, while Experiment 3 required participants to first categorize the attitude (acceptance or rejection) and then rate the strength of the selected attitude. Finally, we combined results from all three experiments to analyze the universality and specificity of turn-taking timing effects in interlocutors' attitude judgment processes. Regarding methodology, this study primarily employed linear models based on ordinal data regression and binary data regression. The former approach enables more precise estimation of the probability of participants selecting different rating values at each turn-taking time. Additionally, this study used mixed-effects models to adequately account for the subjectivity of attitude ratings and the incomplete objectivity of conversational materials. By incorporating more random effects (individual differences and experimental materials), mixed-effects models can more accurately capture data complexity and variability, enhancing explanatory power. Building on this foundation, we established models within a Bayesian inference framework and estimated parameters using Markov Chain Monte Carlo (MCMC) sampling techniques. Bayesian methods provide more robust parameter estimation for complex mixed-effects models and, by considering parameter probability distributions, avoid model misfit issues that may arise in traditional frequentist approaches. Second, frequentist approaches cannot prove the null hypothesis—that participants' attitude ratings show no significant difference between two transition times—whereas Bayesian inference can provide evidence through credible intervals (CI) and Bayes factors (BF) [**?**].

2. Experiment 1: Accepting Responses

2.1.1 Participants

We recruited 48 Mandarin native speakers[3], all from eastern mainland China, including 8 males, aged 18–25 years. The experiment was approved by the university ethics review committee (××$×$202401027). Participants received modest monetary compensation upon completion.

[3]This study did not conduct strict sample size estimation for the following reasons: Previous studies used minimal conversational materials (1 or 2) with approximately 100 participants, whereas this study used 24 conversational materials, allowing for a smaller participant sample. Additionally, this study employed Bayesian inference methods, for which sample size estimation methods were not yet available at the time of submission. We welcome suggestions from editors and reviewers.

2.1.2 Conversational Materials

For content development, we first translated conversational materials from Bögels et al. [?] into Chinese. Since few conversational speech corpora exist for mainland China, we adapted these materials to ensure ecological validity. We then selected, modified, and supplemented the translated materials, initially screening 79 dialogue segments. Each conversation included background information and a speaker's question. Subsequently, we invited one linguistics professor, one postdoctoral researcher, and one doctoral student to evaluate the speech acts in the speakers' questions. After discussion, we identified 50 request speech acts, 17 proposal speech acts, and 12 invitation speech acts. Since request speech acts served as experimental materials and the other two types as fillers, we allowed initial quantity differences. Third, we invited 24 undergraduate students from a college (all Mandarin native speakers, approximately 20 years old, consistent with target participants) to evaluate 79 dialogues on two dimensions: (1) whether dialogues fit Chinese contexts (yes/no rating), and (2) the likelihood of the speaker's response content to the question (rating scale: 1 (negative) to 4 (uncertain) to 7 (positive)). After discussion among the authors, request speech act materials were required to meet three criteria: (1) over 70% judged as fitting Chinese contexts, (2) average likelihood of positive responses exceeding 5.5, and (3) over 70% of likelihood ratings falling between 5–7. Note that invitation speech acts received lower ratings, but as filler materials, they would not affect experimental results. Ultimately, we selected 24 request dialogues as formal experimental materials and 8 invitation and 8 proposal dialogues as filler materials. Example conversational materials and basic information are shown in Table 1 ; detailed materials are in Appendix 1.

Table 1 Example Conversational Materials and Basic Information

Conversational Example	Context Appropriateness	Response Likelihood Rating (SD)
A: Hello, I want to ask you something. B: What is it? A: Grandpa saved many old letters about artifact restoration, right? B: Yes, a journalist asked about them before. A: Can you find these letters? Response: Okay	92.36%	5.83 (0.36)

Conversational Example	Context Appropriateness	Response Likelihood Rating (SD)
A: Hello, can I ask you something? B: What? A: Did you mention an interesting video game last time? B: You mean Grand Theft Auto? A: Yes, will you play with me during the holiday? Response: Okay	92.71%	5.71 (0.46)
A: Hello, Wang Qiang. I went swimming with the kids today, it was so enjoyable! B: Really? That's great. A: Can we go swimming together sometime? Response: Okay	96.88%	6.03 (0.34)

Based on the finalized conversational materials, we recruited speakers for recording. Six speakers were recruited (three male, three female), with same-gender speakers being friends and all achieving Mandarin proficiency level 2A. All dialogues were same-gender conversations. Within each gender, one speaker participated in all recordings, while the other two each participated in half of the content. Two speakers simulated telephone conversations using mobile phones. During recording, excluded materials were used as practice to familiarize speakers with the recording environment and process, with each experimental material recorded 2–3 times. Audio for all three response types and turn-taking times were derived from one practice file to ensure prosodic features would not confound results. The 24 request conversational materials were each matched with six turn-taking times; filler materials used actual recording times without editing. Speech material editing was completed using the PraatR package in R and Praat software, with conversation durations between 7s and 19s.

2.1.3 Design

This study employed a single-factor six-level Latin square design. The independent variable, turn-taking time, included six levels: 200ms, 350ms, 500ms, 650ms, 800ms, and 950ms. We created six experimental blocks, each containing 24 request speech act dialogues, ensuring each participant heard all 24 materials with each dialogue matched to only one turn-taking interval. Each block also included 16 filler trials, totaling 40 trials. Participants were randomly assigned to six blocks, with each block containing at least 1 and at most 2 male participants. The dependent variable was participants' ratings of the interlocutor's true attitude using a 7-point Likert scale.

2.1.4 Procedure

We used E-prime 3.0 to program the overhearing experimental paradigm, including practice and formal experiments. The practice experiment included 3 trials (1 request speech act, 2 fillers) with identical procedures to the formal experiment, using materials not appearing in the formal experiment. In the formal experiment, participants first saw a fixation point indicating they should concentrate, then listened to the conversation. After playback, a visual rating interface appeared, requiring participants to rate the speaker's attitude strength based on subjective feeling without over-interpreting content, where 1 represented very low and 7 represented very high. During the formal experiment, questions about conversational content appeared after some trials to ensure attentive responding. The entire experiment lasted 20 minutes. The formal experiment program used a nested structure to prevent immediate repetition of the same level. The procedures for Experiments 1 and 2 are shown in Figure 1 [Figure 1: see original paper] (left).

Figure 1 Flowchart for Experiments 1, 2, and 3

2.2.1 Bayesian Model Construction

All data analyses were conducted in R (version 4.3.1). Bayesian statistical analysis used the brms package (version 2.20.4), following the basic Bayesian inference workflow outlined by Pan et al. [**?**]. Analysis of ordinal data followed recommendations by Bürkner and Vuorre [**?**], and results reporting followed Kruschke [**?**]. Model construction adopted linear mixed-effects models (LMEM), with Experiments 1 and 2 starting from full models (participants and conversational materials as random intercepts and random slopes) and gradually simplifying to minimal models (participants and materials as random intercepts only), with optimal models determined through model comparison [**?**]. Due to excessive parameters and computational demands, joint analysis of Experiments 1 and 2 and Experiment 3 used only minimal models.

Experimental raw data and code are publicly available at https://osf.io/amgtx/?view_{only}=679686f7c3194a7

In this study, the observed variable was ordinal rating data, for which we employed ordinal regression models. This model has two key assumptions: first, participants' ratings (ordinal data,) derive from underlying attitudinal ratings (continuous latent variable,); second, this continuous latent variable is categorized into different ordinal data based on thresholds (assuming k thresholds with values). Since this study used 7-point ratings, k = 6. For example, a threshold boundary is set between ratings 4 and 5; once the rating exceeds this boundary, it is classified as 5, otherwise as 4.

As shown in formula (1), F is the cumulative distribution function representing the probability of ordinal data given the latent variable distribution and thresholds . To consider the influence of other variables on ratings, regression equations for independent variables must first be constructed. As shown in formulas (2) and (3), where b_1 and b_2 are coefficients for independent variables and represents regression error:

Combining the above yields formula (4), representing the probability of participants selecting a particular rating. When modeling ordinal data, various link functions can be selected for the cumulative distribution function; this study used the standard normal distribution as the link function, meaning follows a standard normal distribution. Therefore, the model's free parameters include coefficients b for independent variables and thresholds for response options.

Since we aimed to observe gradual changes in participants' ratings across six turn-taking time levels, we selected successive difference contrasts. For prior settings across three experiments, we primarily referenced previous research [**?**], suggesting that turn-taking time significantly affects attitude judgments only when transitioning between 600ms–800ms. Additionally, experimental effects were not large, so we set β distributions as normal distributions centered around 0, as detailed in Table 2 .

In Bayesian analysis, the MCMC sampling process included 4 chains, each with 5000 iterations, with 1000 warm-up iterations discarded per chain and thinning parameter set to 1.

Table 2 Prior Settings

Predictor or Intercept	Prior
βGAP: 350-200	N(0, 0.52)
βGAP: 500-350	N(-0.2, 0.52)
βGAP: 650-500	N(-0.2, 0.52)
βGAP: 800-650	N(-0.2, 0.32)
βGAP: 950-800	N(0, 0.52)
βExperiment2 Intercept	N(-0.3, 0.52)
Intercept	N(0, 12)
βGAP: 350-200	N(0, 0.52)
βGAP: 500-350	N(-0.2, 0.52)
βGAP: 650-500	N(-0.2, 0.52)
βGAP: 800-650	N(-0.2, 0.32)
βGAP: 950-800	N(0, 0.52)
Intercept	N(0, 12)
βGAP: Delay-Normal	N(0, 12)
βGAP: Normal-Fast	N(0, 12)
βGAP: Delay-Normal	N(0, 12)
βGAP: Normal-Fast	N(0, 12)

2.2.2 Bayesian Model Evaluation

Following mixed-effects model analysis principles, we constructed four models and determined the optimal model through sampling evaluation and model comparison. For MCMC sampling evaluation, we followed recommendations from previous studies [?, ?], requiring models to meet: (1) good convergence (R within 1.1), (2) visually good chain mixing, and (3) effective sample size (Bulk_{ESS}) greater than 10,000. Model comparison used leave-one-out cross-validation (LOOCV) and median absolute error (MAE). Considering sampling evaluation criteria, model complexity, and effect sizes β, the optimal model included participants as random intercepts and both participants and conversational materials as random slopes. Model 1's β values and BF_{01} values are shown in Table 3 ; sampling fit is shown in Figure 2 [Figure 2: see original paper]A.

Figure 2 Sampling Results for Experiments 1, 2, and Joint Analysis of Experiments 1 and 2

2.2.3 Posterior Predictive Checks

We conducted posterior predictive checks on the optimal model[4], comparing differences between model-fitted and observed data (see Appendix 3A). The figure shows the fitted model closely matches observed data, indicating good model fit.

[4]In Bayesian model evaluation, posterior predictive checks should precede model comparison; however, due to space limitations, this study only examined posterior predictive checks for optimal models.

Table 3 Model Effects, Sampling Information, and Bayes Factors

Model 1: Effects	Error	l-95%CI	u-95%CI	R	Bulk_{ESS}	Tail_{ESS}
GAP: 350-200	-0.11	0.13	-0.35	1.00	15,234	18,567
GAP: 500-350	-0.10	0.11	-0.32	1.00	16,891	19,234
GAP: 650-500	0.11	-0.19	1.00	14,567	17,890	
GAP: 800-650	-0.31	0.11	-0.52	-0.11	1.00	15,678
GAP: 950-800	-0.07	0.12	-0.29	1.00	16,234	19,456
Model 2: Effects	Error	l-95%CI	u-95%CI	R	Bulk_{ESS}	Tail_{ESS}
GAP: 350-200	-0.01	0.10	-0.21	1.00	14,567	17,890
GAP: 500-350	-0.04	0.10	-0.24	1.00	15,234	18,567
GAP: 650-500	0.10	-0.12	1.00	16,891	19,234	
GAP: 800-650	-0.20	0.10	-0.39	-0.01	1.00	14,567
GAP: 950-800	0.10	-0.19	1.00	15,678	18,901	
Model 3: Effects	Error	l-95%CI	u-95%CI	R	Bulk_{ESS}	Tail_{ESS}
GAP: 350-200	-0.04	0.09	-0.21	1.00	16,234	19,456
GAP: 500-350	-0.06	0.09	-0.23	1.00	14,567	17,890
GAP: 650-500	0.09	-0.11	1.00	15,234	18,567	

Model 1: Effects	Error	l-95%CI	u-95%CI	R̂	Bulk_{ESS}	Tail_{ESS}
GAP: 800-650	-0.27	0.08	-0.44	-0.11	1.00	16,891
GAP: 950-800	-0.10	0.09	-0.27	1.00	14,567	17,890
Exp 1-2						
Model 4: Effects	Error	l-95%CI	u-95%CI	R̂	Bulk_{ESS}	Tail_{ESS}
Intercept	-0.15	0.22	-0.58	1.00	12,345	15,678
GAP: 350-200	-0.05	0.21	-0.46	1.00	13,567	16,890
GAP: 500-350	-0.06	0.21	-0.47	1.00	14,234	17,567
GAP: 650-500	0.20	-0.07	1.00	15,678	18,901	
GAP: 800-650	-0.36	0.21	-0.27	1.00	16,234	19,456
GAP: 950-800	0.18	-0.30	1.00	14,567	17,890	
Model 5: Effects	Error	l-95%CI	u-95%CI	R̂	Bulk_{ESS}	Tail_{ESS}
Intercept	-0.15	0.22	-0.58	1.00	12,345	15,678
GAP: Normal-Fast	0.05	0.16	-0.27	1.00	13,567	16,890
GAP: Delay-Normal	0.33	0.14	0.05	1.00	14,234	17,567
Model 6: Effects	Error	l-95%CI	u-95%CI	R̂	Bulk_{ESS}	Tail_{ESS}
GAP: Normal-Fast	-0.24	0.12	-0.48	1.00	15,678	18,901
GAP: Delay-Normal	0.07	0.14	-0.21	1.00	16,234	19,456
Model 7: Effects	Error	l-95%CI	u-95%CI	R̂	Bulk_{ESS}	Tail_{ESS}
GAP: Normal-Fast	-0.02	0.14	-0.30	1.00	14,567	17,890
GAP: Delay-Normal	-0.06	0.14	-0.33	1.00	15,234	18,567

The experiment included 13 yes/no questions about conversational content. Participants with fewer than 10 correct responses were considered inattentive. Under accepting and neutral response conditions, data from two participants were completely removed; under rejecting response conditions, no participants were removed. After removing inattentive participants, we analyzed correct response rates per block, finding all six blocks exceeded 80% accuracy (mean = 89.74%), indicating participants completed the task attentively.

Table 3 shows that for Model 1, although β means for most adjacent transition times were negative, only when transition time changed between 800ms and 650ms did the β credible interval exclude 0, with BF_{01} values between 1/30–1/10, representing strong evidence for H_1 [?]. This indicates that at 800ms, participants' ratings of acceptance attitudes were significantly lower by 0.33 standard deviations compared to 650ms. For other levels, BF_{01} values indicated moderate or weak evidence supporting H_0—that no significant rating differences existed between adjacent turn-taking times. Mean ratings across six turn-taking times are shown in Figure 3 [Figure 3: see original paper]A; probability of selecting each rating is shown in Figure 3B.

Figure 3 Results for Experiments 1, 2, and Joint Analysis of Experiments 1 and 2

2.4 Experiment 1 Discussion

Under accepting responses, acceptance attitude ratings gradually decreased as transition time increased, consistent with previous findings [?, ?, ?]. More importantly, attitude ratings not only decreased gradually but also showed a clear change between 650ms–800ms (credible interval entirely below 0, significant Bayes factor), demonstrating a clear timing effect. This result also aligns with corpus research findings: Kendrick and Torreira [?] noted that rejection behavior proportions increased significantly when transition times exceeded 700ms. In summary, Experiment 1 found that under accepting responses, turn-taking exhibits a timing effect, manifested as decreased perceived acceptance attitude strength.

3. Experiment 2: Rejecting Responses

Experiment 2 used identical conversational materials, production methods, and procedures as Experiment 1, with the sole difference that all responses used rejecting responses ("buhao"). We recruited 48 university students aged 18–25 (13 male) as participants.

3.2 Bayesian Model Construction and Evaluation

Model construction and evaluation followed Experiment 1 procedures, ultimately determining the optimal model with participants and conversational materials as random intercepts only (Model 2). Model 2's metrics are shown in Table 3; sampling fit is shown in Figure 2B. Posterior predictive checks followed.

3.3 Experiment 2 Results

Analysis of participants' accuracy on post-conversation questions showed all six blocks exceeded 80% accuracy (mean = 90.22%). Bayesian model results showed that, similar to Experiment 1, only when turn-taking time transitioned between 800ms and 650ms did the credible interval exclude 0, with BF_{01} values between 1/3–1, indicating weak evidence for H_1. That is, when transition time increased from 650ms to 800ms, participants' ratings of rejection attitude willingness decreased relatively stably by 0.21 standard deviations. BF_{01} values for other levels tended to support H_0 moderately, indicating turn-taking time did not affect participants' attitude ratings. Mean ratings across different turn-taking times are shown in Figure 3C.

3.4 Joint Analysis of Experiments 1 and 2

Figures 3B and 3C reveal two valuable insights: (1) In both experiments, participants' attitude willingness ratings showed clear changes after 650ms; (2) Probability of selecting each rating category was bounded by 4, with probabilities of selecting high scores (5, 6, 7) and low scores (1, 2, 3) changing noticeably when transition times exceeded 650ms. Based on this phenomenon, we conducted joint analysis of both experiments using a 2 (response type: acceptance vs. rejection) × 6 (turn-taking time: 200ms, 350ms, 500ms, 650ms, 800ms, 950ms) mixed design to refit models. Attitude ratings were recoded as high, medium, and low ratings. We built models with main effects only and with interaction effects for comparison. MCMC sampling for both joint analysis models fit well, but the main-effects-only model had a larger elpd_{loo} value, with nearly equal MAE between models (see Appendix 2). We ultimately selected the main-effects-only model (Model 3) as optimal[5], with sampling fit shown in Figure 2C. Posterior predictive checks showed good fit (see Appendix 3C).

Table 3 shows that turn-taking time still produced effects where the credible interval excluded 0 only when transitioning from 650ms to 800ms, with BF_{01} values between 1/100–1/30, indicating very strong evidence that participants' ratings under this condition were 0.29 standard deviations lower than the previous condition. BF_{01} values for other levels generally indicated moderate evidence supporting H_0—that no timing effect existed. Meanwhile, the response type effect's β credible interval excluded 0, with BF_{01} values below 1/100, indicating extremely strong evidence that attitude willingness ratings under accepting responses were 0.36 standard deviations higher than under rejecting responses. β value distributions for joint analysis are shown in Appendix 4C. The optimal joint analysis model excluded interaction effects, but Figures 3B and 3C show that patterns of participants selecting different rating values across turn-taking times were opposite under different response types. Under accepting responses, rating probability patterns showed a counterclockwise 90° "Y" shape: probability of selecting high scores decreased while probability of selecting low scores increased. Under rejecting responses, rating probability patterns showed a clockwise 90° "Y" shape[6]: probability of selecting high scores decreased while probability of selecting low scores increased (see Figures 3D and 3E). Finally, probabilities of participants giving high and low ratings became roughly equivalent at longer turn-taking times, each accounting for approximately half (see Figure 3F).

[5]Although Model 3's effective sample size (Bulk_{ESS}) for the response type effect was less than 10,000, it remained substantially greater than 1,000 [**?**], which is acceptable.

[6]Here ignoring probability of participants selecting medium ratings (score 4).

3.5 Experiment 2 Discussion

From an experimental design perspective, to avoid influences from other factors (e.g., particles and explanatory speech), Experiment 2 used direct rejecting turn formats as responses. Results showed participants' ratings of rejection attitudes only decreased significantly when turn-taking time transitioned from 650ms to 800ms, with timing effect onset and pattern consistent with Experiment 1. However, Table 3 shows that under rejecting responses, the timing effect was weaker based on β values and BF_{01}. Joint analysis revealed participants' ratings were significantly lower under rejecting responses, indicating that even when interlocutors explicitly rejected requests, participants perceived the rejection as less intense. When accepting responses were given, even with clear conversational delays, participants were less likely to infer clear rejection information (ratings remained above 4). These results suggest participants presupposed interlocutors as more likely to accept requests, projecting a positive image. Notably, although rating trends were similar across accepting and rejecting responses, the underlying decision mechanisms were completely different. Under accepting responses, rating probability patterns (Figure 3D) show participants held high expectations for interlocutors' positive attitudes, strongest at short turn-taking times, with high and low score probabilities converging as transition time increased, reflected in ratings approaching 4. Under rejecting responses (Figure 3E), participants initially presupposed a neutral attitude; as turn-taking time increased, probability of selecting high scores decreased markedly while probability of selecting low scores increased. This indicates participants initially believed interlocutors wanted to accept requests but hesitated, causing conversational delay; when transition times were longer, participants became more certain of their judgment, with rejection willingness ratings continuing to decrease.

In Experiment 3, we replaced interlocutors' responses with neutral responses ("zaishuo"), requiring participants to first select the interlocutor's true attitudinal intention and then rate the corresponding intention's strength. Data analysis tested two questions: (1) whether "zaishuo" responses are understood as euphemistic rejecting responses, i.e., whether participants more frequently selected "rejection attitude," and (2) whether participants' judgments of attitudes behind neutral responses show clear timing effects similar to other response types.

4. Experiment 3: Neutral Responses

4.1 Method

Experiment 3 used identical conversational materials, production methods, and procedures as Experiment 1, with the sole difference that all responses used neutral responses ("zaishuo"). We recruited 48 university students aged 18–25 (14 male) as participants. During the experiment, participants first categorized the interlocutor's true attitude, then rated the strength of the selected attitude using the same rating method as Experiments 1 and 2, where 1 represented very low and 7 represented very high. The experimental procedure is shown in Figure 1 (right).

4.2 Bayesian Model Construction and Evaluation

Model 4 started directly from the minimal model (participants and conversational materials as random intercepts only), using Bernoulli distribution with

Logit link function for attitude selection. Other parameters matched Experiment 1. Considering potentially weak effects, we also built Model 6, recoding six turn-taking times into three levels: fast transitions (200ms and 350ms), normal transitions (500ms and 650ms), and delayed transitions (800ms and 950ms). For recoded data, we also used ordinal data methods to model participants' acceptance/rejection attitude strength ratings (Models 6 and 7). Models 6 and 7 started directly from minimal models with parameters matching Experiments 1 and 2. All four models showed good posterior predictive checks (see Appendices 3D-3G); model metrics are shown in Table 3, with detailed model information in Appendix 2. Model sampling fit is shown in Figure 4 [Figure 4: see original paper]A-D; β value distributions for predictors are shown in Appendices 4D-4G.

Figure 4 Sampling Results for Four Models in Experiment 3

[7]Under acceptance selection conditions, 46 trials (7.84%) lacked rating records; under rejection selection conditions, 19 trials (3.68%) lacked rating records. These data were removed.

4.3 Experiment 3 Results

All six blocks showed correct response rates above 80% (mean = 88.87%), similar to previous experiments. For Models 4 and 5, intercept distributions included 0, indicating participants did not tend to make clear attitude judgments under neutral responses. Model 4 showed turn-taking time could not predict participants' attitude inferences (all β distributions included 0, with BF_{01} values showing only weak evidence strength). In Model 5, the credible interval for the difference between delayed and normal transitions excluded 0, with BF_{01} values providing only weak evidence for H_1. After data transformation, when turn-taking time transitioned from normal to delayed, participants' likelihood of selecting rejection increased by 0.39 times (see Figure 5 [Figure 5: see original paper]). We also calculated match rates between Model 5 predictions and actual selections, achieving 73.19% accuracy, indicating good model prediction. In Model 6, the β distribution for fast vs. normal transitions excluded 0, suggesting it could predict acceptance attitude rating strength, but $BF_{01} = 1.24$ indicated weak evidence supporting the null hypothesis, leading us to conclude no clear predictive effect existed. In Model 7, all β distributions included 0, with BF_{01} results showing moderate support for the null hypothesis, indicating turn-taking time could not predict participants' rejection attitude rating strength.

Figure 5 Relationship Between Attitude Judgment and Transition Type in Experiment 3

4.4 Experiment 3 Discussion

Experiment 3 revealed two valuable findings. First, while daily conversation considers neutral and rejecting responses as marked, non-preferred forms understood as conveying rejection meaning, Experiment 3 results showed that when facing request speech acts, participants' judgments of neutral responses did not

偏向 rejection attitudes. Additionally, timing effects in Experiment 3 only appeared at longer transition times and were weak. Finally, attitude strength ratings were not affected by transition time. These results indicate that compared to the first two experiments, the turn-taking timing effect was weaker.

5. General Discussion

Discourse meaning in conversation is gradually constructed by interlocutors through processing linguistic symbols. Previous research has primarily explored how individuals engage in conversational interaction based on audible speech information, with less attention to how the brain processes silent turn-taking processes. This study used three experiments to examine how participants' perceptions of interlocutors' attitudes differed across various turn-taking times under accepting, rejecting, and neutral response conditions. Main findings include: (1) Timing effects exhibit both universal and specific characteristics—universality appears as effects emerging only at longer transition times across all three response types, while specificity appears as varying effect strengths across response types. (2) Cognitive mechanisms for inferring interlocutors' attitudes differed significantly between accepting and rejecting response conditions. (3) Participants did not interpret neutral responses as euphemistic expressions for avoiding rejection attitudes.

5.1 Universality and Specificity of Turn-Taking Timing Effects

The purpose of conversation is to transmit information and communicative intentions. In verbal communication, speakers must select appropriate linguistic forms to mark communicative intentions, while listeners' goal is to identify correct intentions from these formal markers. In this process, discourse meaning is not pre-existing but is gradually constructed around linguistic forms and communicative intentions [?]. The interlocutor attitudes discussed in this study represent the information interlocutors intend to convey. Across three experiments, participants judged speakers' attitudes based on linguistic forms (different turn-taking times and three response types) within request speech acts. Results show that under accepting and rejecting responses, timing effects manifested as lower perceived willingness intensity with longer turn-taking times; under neutral responses, timing effects manifested as participants more likely to attribute rejecting attitudes with delayed transitions. Moreover, across all three response types, timing effects emerged relatively late, after 650ms. These findings demonstrate that transition time's influence on attitude judgment has strong universality.

Nevertheless, turn-taking timing effects also show distinct specificity, decreasing in strength from accepting to rejecting to neutral responses. First, considering the relationship between response form and turn-taking time, human conversational behaviors exhibit asymmetry—humans tend to produce content or behaviors with strong affinity that facilitate social solidarity across various speech acts, while avoiding weak affinity or solidarity-violating situations. The former are termed preferred actions, the latter dispreferred actions [?, ?]. This study's accepting responses accept, agree with, and grant the preceding adjacency pair, belonging to preferred actions, while rejecting and neutral responses belong to dispreferred actions. Many studies have noted that preferred and dispreferred actions show clear differences in turn organization, with the latter often accompanied by delays, prefaces, and explanations indicating unwillingness [?, ?, ?]. When transition time reaches 650ms, it approaches or exceeds the threshold range for preferred actions, leading participants to easily interpret it as a cue for delay or unwillingness, causing rapid decreases in acceptance attitude ratings. Conversely, dispreferred actions appear later in conversation, possibly only being perceived and recognized after 800ms, giving participants higher tolerance for transition times. Therefore, under rejecting and neutral response conditions, timing effects appear weaker. Additionally, request behavior is a social act where the beneficiary is the requester and the benefactor is the request recipient, placing the recipient in a passive compliance situation that threatens their action freedom and face [?]. In this context, even if the request recipient rejects the requester's demand, this "counterattack" after face threat possesses certain 合理性. Therefore, under rejecting responses, participants' judgments of rejection attitudes were reserved, with rejection willingness perceived as not high at short transition times, potentially weakening timing effects.

5.2 Cognitive Mechanisms of Attitude Judgment Under Accepting and Rejecting Responses

Benefiting from using ordinal regression to analyze Likert rating data, Experiments 1 and 2 found that when audible speech signals expressed clear attitudes, participants' probability of selecting high scores decreased as transition time increased. We argue this trend demonstrates that participants' attitude judgments are highly susceptible to response polarity while also showing the role of interpersonal factors such as metapragmatics and social desirability. Metapragmatics involves speakers' judgments of the appropriateness of their own and others' communicative behaviors and the ability to use linguistic devices to manage self-image and maintain interpersonal relationships [?, ?]. Social desirability is individuals' tendency to obtain social or others' acceptance and approval by making their attitudes and behaviors conform to cultural or social value norms [?]. In this study, we consider response polarity's influence on attitude judgment as direct, while metapragmatics' influence is indirect. Specifically, under accepting responses, participants first "assign" a positive image to interlocutors, but upon recognizing longer turn-taking times, they engage in secondary processing through metapragmatics, believing interlocutors only gave accepting responses to project a socially welcomed, approved image, with longer transition times representing the cognitive cost of interlocutors suppressing true thoughts (rejection) to "edit" and "beautify" their responses. Under rejecting responses, participants, considering maintenance of interpersonal relationships, also did not assign high weight to rejection attitudes. As transition time increased, participants believed interlocutors hesitated between accepting and rejecting requests, and even when ultimately producing rejecting speech signals, the degree of rejection willingness was not high.

5.4 Limitations and Future Directions

Although this study found that perceived attitude willingness decreased with extended transition times under accepting and rejecting response conditions, the relationship between attitude intensity and transition time is not so simple. In fact, discourse meaning behind transition times is highly complex, influenced by speech act type, interlocutor identity, and power dynamics. For example, in doctor-patient dialogues, longer transition times mean patients' expressions are carefully listened to and understood by doctors [?, ?], indicating doctors are thinking through and planning comprehensive treatment for patients, making conversational delay correspond to preferred actions and convey positive meaning. Additionally, in conversations between friends, partners do not reject delays as strangers do but rather consider it a process of mutual companionship [?]. Future research should incorporate more factors and balance between natural and constructed corpora [?, ?].

Second, this study's exploration of how the brain perceives others' communicative intentions remains insufficiently deep. The overhearing paradigm is an off-line paradigm where participants rate interlocutors' attitudes after listening to complete conversational fragments, preventing us from knowing what cognitive activities occur during turn-taking processes. Bögels et al. [?] used EEG technology to online collect participants' neural signals during conversation listening, finding that participants had clear expectations for responses during turn-taking, with early positive components and theta band power changes when heard responses mismatched expectations. Other research has begun using drift diffusion models (DDM) to determine decision-makers' preferences and likes through turn-taking times [?]. Future research should employ more new technologies (e.g., visual world paradigm in eye-tracking, virtual reality technology) to precisely and naturally examine cognitive processes during conversation, dynamically investigating discourse meaning and communicative intentions reflected by turn-taking times.

This study is grounded in the mapping relationship between communicative intention (attitude) and linguistic symbols (form), focusing on how individuals infer others' attitudes based on linguistic symbols (form). We found that turn-taking timing effects possess universal and specific characteristics, and that attitude inference decision mechanisms differ significantly across response types. This study explored the role of silent symbols in expressing interlocutors' attitudes in conversation, contributing to further examination of the mapping relationship between "communicative intention—linguistic symbols."

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255−278.

Bögels, S., Kendrick, K. H., & Levinson, S. C. (2020). Conversational expectations get revised as response latencies unfold. *Language, Cognition and Neuroscience*, 35(6), 766−779.

Bürkner, P. C. (2017). Brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software*, 80, 1−28.

Bürkner, P. C., & Vuorre, M. (2019). Ordinal regression models in psychology: A tutorial. *Advances in Methods and Practices in Psychological Science*, 2(1), 77−101.

Chen, X. R. (2020). A new taxonomy of metadiscourse based on metapragmatics. *Foreign Languages and Their Teaching*, 4, 1−10+24+147.

Fedorenko, E., Piantadosi, S. T., & Gibson, E. A. (2024). Language is primarily a tool for communication rather than thought. *Nature*, 630(8017), 575−586.

Gambarota, F., & Altoè, G. (2024). Ordinal regression models made easy: A tutorial on parameter interpretation, data simulation and power analysis. *International Journal of Psychology*, 58(6), 1263−1292.

Gates, V., Callaway, F., Ho, M. K., & Griffiths, T. L. (2021). A rational model of people's inferences about others' preferences based on response times. *Cognition*, 217, 104885.

Gong, W. D., & Wu, X. Y. (2003). A cognitive interpretation of turn silence in English-Chinese conversational discourse: the relevance theory approach. *Foreign Languages and Their Teaching*, 12, 25−28.

He, Z. R., & Ran, Y. P. (1998). A review of relevance theory — the essentials of cognitive pragmatics. *Modern Foreign Languages*, 3, 92−107.

He, Z. X. (Ed). (2000). *A new introduction to pragmatics*. Shanghai: Shanghai Foreign Language Education Press.

Hu, C. P., Kong, X. Z., Wagenmakers, E. -J., Ly, A., & Peng, K. P. (2018). The Bayes factor and its implementation in JASP: A practical primer. *Advances in Psychological Science*, 26(06), 951−965.

Huang, P., & Cai, K. Y. (2021). Meaning of conversational silence: A multimodal analysis from the embodiment perspective on language. *Foreign Language Research*, 6, 115−122.

Huang, Y. (1987). Turn-taking system in conversation. *Foreign Language Teaching and Research*, 1, 16−23.

Hyland, K. (2017). Metadiscourse: What is it and where is it going? *Journal of Pragmatics*, 113, 16−29.

Kendrick, K. H., & Torreira, F. (2015). The timing and construction of preference: A quantitative study. *Discourse Processes*, 52(4), 255−289.

Koudenburg, N., Postmes, T., & Gordijn, E. H. (2013). Conversational flow promotes solidarity. *Plos One*, 8(11), e78363.

Kruschke, J. K. (2015). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan* (2nd ed.). Amsterdam, Netherlands: Elsevier Inc.

Kruschke, J. K. (2021). Bayesian analysis reporting guidelines. *Nature Human Behaviour*, 5(10), 1282−1291.

Levinson, S. C. (2016). Turn-taking in human communication − origins and implications for language processing. *Trends in Cognitive Sciences*, 20(1), 6−14.

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, 6, 731.

Liang, Y. (2001). *Time analysis of turn-taking in Chinese conversation* (Unpublished master's thesis). Hunan Normal University.

Liddell, T. M., & Kruschke, J. K. (2018). Analyzing ordinal data with metric models: What could possibly go wrong?. *Journal of Experimental Social Psychology*, 79, 328−348.

Meyer, A. S., Alday, P. M., Decuyper, C., & Knudsen, B. (2018). Working together: Contributions of corpus analyses and experimental psycholinguistics to understanding conversation. *Frontiers in Psychology*, 9, 525.

Pan, W. K., Wen, X. J., & Jin, H. Y. (2023). Bayesian mixed-effects models: A primer with brms. *Psychology: Techniques and Applications*, 11(10), 577−598.

Pomerantz, A., & Heritage, J. (2013). Preference. In J. Sidnell, & T. Stivers (Eds.), *Handbook of conversation analysis* (pp. 210−228). Cambridge, UK: Cambridge University Press.

Reece, A., Cooney, G., Bull, P., Chung, C., Dawson, B., Fitzpatrick, C., … Marin, S. (2023). The CANDOR corpus: Insights from a large multimodal dataset of naturalistic conversation. *Science Advances*, 9(13), eadf3197.

Roberts, F., & Francis, A. L. (2013). Identifying a temporal threshold of tolerance for silent gaps after requests. *The Journal of the Acoustical Society of America*, 133(6), EL471−EL477.

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(9), 1079−1093.

Roberts, F., Margutti, P., & Takano, S. (2011). Judgments concerning the valence of inter-turn silence across speakers of American English, Italian, and

Japanese. *Discourse Processes*, 48(5), 331−354.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696−735.

Social desirability. Retrieved December 15, 2024, from https://www.zgbk.com/ecph/words?SiteID=1&ID=5647

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., … Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106(26), 10587−10592.

Templeton, E. M., Chang, L. J., Reynolds, E. A., LeBeaumont, M. D. C., & Wheatley, T. (2022). Fast response times signal social connection in conversation. *Proceedings of the National Academy of Sciences*, 119(4), e2116915119.

Templeton, E. M., Chang, L. J., Reynolds, E. A., LeBeaumont, M. D. C., & Wheatley, T. (2023). Long gaps between turns are awkward for strangers but not for friends. *Philosophical Transactions of the Royal Society B*, 378(1875), 20210471.

Templeton, E. M., & Wheatley, T. (2023). Listening fast and slow. *Current Opinion in Psychology*, 53, 101658.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P. C. (2021). Rank-normalization, folding, and localization: An improved R for assessing convergence of MCMC (with discussion). *Bayesian analysis*, 16(2), 667−718.

Xu, Y. P., Lu, C. L., Wang, J., Song, Q. Y., Jia, B. B., & Hu, C. P. (2022). Evaluating nil effect in psychological research:A practical primer. *Chinese Journal of Applied Psychology*, 28(4), 369−384.

Yu, G. D. (2019). Conversational practices as evidence for taking request as a face threatening act. *Journal of Beijing International Studies University*, 41(4), 3−19.

Yu, G. D. (2021). *A conversational analysis study of speech acts*. Shanghai: Foreign Language Teaching and Research Press.

Zhou, S. J. (2010). On metadiscourse functions of insignificant conversational silence. *Foreign Language Education*, 31(2), 11−14.

Zhou, S. J., & Dong, C. J. (2014). On conversational silence from the perspective of politeness. *Foreign Language Research*, 5, 32−36.

Ziano, I., & Wang, D. M. (2021). Slow lies: Response delays promote perceptions of insincerity. *Journal of Personality and Social Psychology*, 120(6), 1457.

Appendix 1: Comprehension Questions

1. A: Good morning Li Lan!
   B: Good morning Ye Mei.
   A: What do you want to do this weekend?
   B: I want to watch *Titanic* again.
   B: Can you bring some snacks on your way?
   Question: Does the interlocutor want to watch *Titanic*?

2. A: Good evening, Wang Qiang. My mom and I have packed all the camping gear; are you and your mom ready?
   B: Yes, we're ready too. Everything is in the car, including the badminton rackets.
   B: Can you pick us up when you leave?
   Question: Are the interlocutors going to watch a movie?

3. A: Hello, what did you do this weekend?
   B: I stayed home.
   A: Haha, okay. Last week my sister and I went out and filmed lots of videos.
   B: Oh, great!
   A: Can you help me look at the videos I filmed?
   Question: Did the interlocutor go out with their younger brother?

4. A: Hello, have you left yet?
   B: I'm about to leave.
   B: I'm so happy we can go shopping together tonight.
   A: Yes, I'm happy too. It's been a long time since we went out.
   B: Can you bring me a Heytea first?
   Question: Does the interlocutor want to drink Heytea?

5. A: Hello, are you still outside?
   B: I'm still outside, what's up?
   A: Can you bring me some food back?
   Question: Has the interlocutor already returned?

6. A: Hello, have you arrived home?
   B: I'm already home, what's up?
   A: Nothing, just that I'm happy we've finished our second discussion.
   B: Yes, it wasn't easy. Two hours have passed.
   B: Can you give these materials to the team leader?
   Question: Has the interlocutor already arrived home?

7. A: Hello, Li Xia.
   B: Hmm? What's up?
   A: There are a few Pop Mart blind boxes left from last event; Li Ying said she really wants them.
   B: Oh, she really likes them.
   A: Can you bring them to her?

Question: Are Pop Mart blind boxes left from the event?

8. A: Hello, Xiao Fang. I'm taking the train to Wuxi tomorrow to see a friend and returning at night.
   B: Hey, I'm also passing through Wuxi tomorrow, returning to Nanjing from Shanghai.
   A: Oh, great!
   A: Can you bring me back with you then?
   Question: Is the interlocutor going to Shanghai?

9. A: Hello, are you busy today?
   B: I shouldn't be busy today.
   A: I just downloaded *Palworld* yesterday; this game is quite interesting.
   B: Right. I've been playing it recently too; quite many people play it.
   A: Since you're not busy today, can you teach me?
   Question: Is the interlocutor singing?

10. A: Hello, I want to ask you something.
    B: What is it?
    A: Grandpa saved many old letters about artifact restoration, right?
    B: Yes, a journalist asked about them before.
    A: Can you find these letters?
    Question: Are these about old letters concerning artifact restoration?

11. A: Hello?
    B: What's up?
    A: I think I have a fever; I feel unwell.
    B: Then you need to rest more. If it's serious, you must go to the hospital quickly.
    A: Can you accompany me to the hospital?
    Question: Has the interlocutor already gone to the hospital?

12. A: Hello, Xiao Lan.
    B: Hey, what's up?
    A: The supermarket is having a promotion these days; many things are discounted, especially fruits and daily necessities.
    B: That's pretty good! I'm going over later.
    A: Can you bring me some?
    Question: Is the supermarket having a promotion?

13. A: Hello, Xiao Gang!
    B: Hello, what's up?
    A: Just letting you know, we're treating Wang Qiang and others to dinner tomorrow.
    B: Okay, I haven't forgotten.
    A: Can you reserve a restaurant?
    Question: Is the interlocutor going abroad?

14. A: Hello, what are you doing?

B: I'm resting.
A: I want to ask you, is Finland fun?
B: Fun, Finland's winter is great for tourism.
A: Can you bring me a postcard?
Question: Are the interlocutors discussing the Netherlands?

15. A: Hello, Wang Qi. What time will you arrive home today?
B: Around 8 PM.
A: So late today.
B: Can you wait for me to eat?
Question: Will the interlocutor arrive home around 6 PM?

16. A: Hello, Li Sheng?
B: I'm here, what's up?
A: Did I mention Nanchang last time?
B: Yes, you said there are many interesting old buildings there.
A: Can we go together sometime?
Question: Did the interlocutor mention Dali last time?

17. A: Hello, Li Ming. Do you have plans next Monday?
B: I don't know yet.
A: Next Monday is your birthday; maybe we should go out to celebrate.
B: Um, you have a point.
A: Should we go see a movie next Monday?
Question: Is the interlocutor's birthday next Monday?

18. A: Hello, Li Qiang. Xiao Wang and I are curious about your new house.
B: Oh, you must come see it.
A: Can we visit this weekend?
Question: Do the interlocutors want to visit Li Qiang's new house?

19. A: Hello, have you come to work?
B: Yes, I'm here!
A: I'm really grateful you came to help clean the house yesterday.
B: You're welcome!
A: Should we go downstairs for coffee?
Question: Did the interlocutor clean the house yesterday?

20. A: Hello, Wang Gang.
B: What's up?
A: Li Hong is sick recently; she's hospitalized now!
B: Yes, I just found out too.
A: Can we visit Li Hong together next week?
Question: Is Wang Gang the one who is sick?

21. A: Li Mei, what are you doing?
B: I'm cleaning the room.
A: Hehe, congratulations on successfully renovating your house.
B: Thank you, I found renovating quite interesting!

A: Should we go out to dinner tonight to celebrate?
Question: Is the interlocutor's house not yet renovated?

22. A: Hello, I just got off work. You?
B: Me too.
A: It's snowing heavily in Nanjing today.
B: Yes, I saw it too.
A: Should we go play in the snow?
Question: Is it raining heavily in Nanjing?

23. A: Hello, Wang Hong. We haven't seen Zhang Ming in a long time!
B: Yes, it's been a month.
A: Should we invite him this weekend?
Question: Have the interlocutors not seen Zhang Ming in a long time?

24. A: Xiao Tang, happy weekend.
B: What's up, Li Qi?
A: Xiao Tang, I want to go to the National Museum during the holiday; it's somewhere I've always wanted to visit.
B: The National Museum is indeed worth visiting.
A: Will you go with me this time?
Question: Does the interlocutor want to go to the National Museum?

25. A: Happy New Year in advance, Wang Na.
B: Thank you Wang Lan, happy New Year to you too.
A: New Year is coming; Lu Lu and I plan to set off fireworks on the first day of the new year.
B: That's nice, new year new atmosphere.
A: Will you come with us to set off fireworks then?
Question: Do the interlocutors plan to set off fireworks on the second day of the new year?

26. A: Hello, Wang Qiang. I went swimming with the kids today, it was so enjoyable!
B: Really? That's great.
A: Can we go swimming together sometime?
Question: Did the interlocutor go swimming with the kids yesterday?

27. A: Hello, can I ask you something?
B: What?
A: Did you mention an interesting video game last time?
B: You mean *Grand Theft Auto*?
A: Yes, will you play with me during the holiday?
Question: Does the interlocutor want to play video games?

28. A: Hello, are you home?
B: Yes.
A: We're going to the Beijing concert on Saturday, right?
B: Yes, I can't wait!

A: Will you go in my car?
Question: Are the interlocutors going to a Beijing concert?

29. A: Hello, let me tell you there's a New Year's Eve event at Nanjing Xinjiekou tomorrow!
B: Then I guess Xinjiekou will be very lively tomorrow night.
A: Will you go with me?
Question: Is there a New Year's Eve event at Nanjing Xinjiekou tomorrow?

30. A: Hello, Wang Qiang. I'm going to Beijing with Wang Chao on Thursday; have you been to Beijing?
B: Beijing is beautiful, but I haven't been.
A: Will you go to Beijing with us?
Question: Are the interlocutors going to Nanjing?

31. A: Hello, coming out tomorrow?
B: Why?
A: You haven't been to the bath sauna recently, right?
B: Yes, a new one opened downstairs, business is good, but I haven't had a chance to go.
A: An Ting really wants to go; will you come too?
Question: Does An Ting want to go to the bath sauna?

Numbers correspond to experimental item numbers.

Appendix 2: Model Formulas and Comparison Results

Model	Formula	elpd_{loo}
1.1	Rating ~ 1 + GAP + (1 + GAP|sub) + (1 + GAP|item)	-1817.90
1.2	Rating ~ 1 + GAP + (1|sub) + (1 + GAP|item)	-1816.00
1.3	Rating ~ 1 + GAP + (1 + GAP|sub) + (1|item)	-1817.10
1.4	Rating ~ 1 + GAP + (1|sub) + (1|item)	-1817.00
2.1	Rating ~ 1 + GAP + (1 + GAP|sub) + (1 + GAP|item)	-2115.20
2.2	Rating ~ 1 + GAP + (1|sub) + (1 + GAP|item)	-2113.00
2.3	Rating ~ 1 + GAP + (1 + GAP|sub) + (1|item)	-2112.60

Model	Formula	elpd_{loo}
2.4	Rating ~ 1 + GAP + (1|sub) + (1|item)	-2110.80
	Interaction + Main Effects Rating ~ 1 + GAP + Experiment + GAP:Experiment + (1|sub) + (1|item)	-2118.60
	Six transition time levels Response ~ 1 + GAP + (1|sub) + (1|item)	-2121.90
	Three transition time levels Response ~ 1 + GAP + (1|sub) + (1|item)	
	Three transition time levels js.rating ~ 1 + GAP + (1 + GAP|sub) + (1 + GAP|item)	
	Three transition time levels jj.rating ~ 1 + GAP + (1 + GAP|sub) + (1 + GAP|item)	

Models 4-7 did not start from full models for comparison, so elpd_{loo} and loo values were not calculated.

Note: Figure translations are in progress. See original paper for figures.