A Cross-Cultural Approach to Cognitive State Attribution based on Inter-turn Speech Pauses

Theresa Matzinger*^{1,2}, Michael Pleyer³, Elizabeth Qing Zhang⁴, and Przemysław Żywiczyński³

*Corresponding Author: theresa.matzinger@univie.ac.at

Non-corresponding Authors:

pleyer@umk.pl; zqelizabeth@gmail.com; przemyslaw.zywiczynski@umk.pl

¹Department of English and American Studies, University of Vienna, Spitalgasse 2, 1090 Vienna, Austria

²Cognitive Science Hub, University of Vienna, Kolingasse 14-16, 1090 Vienna, Austria ³Center for Language Evolution Studies, Nicolaus Copernicus University in Toruń, ul. Fosa Staromiejska 3, 87-100 Toruń, Poland

⁴School of Linguistic Sciences and Arts, Jiangsu Normal University, Shanghai Road 101, Tongshan New District, 221116, Xuzhou, Jiangsu, China

Biographical notes:

Theresa Matzinger: Theresa Matzinger is a post-doctoral researcher at the Department of English and American Studies at the University of Vienna. Her research interests lie

at the intersection of linguistics, biology, psychology and cognitive sciences. Her current work focuses on the relationship between communication and cooperation and sheds light on how language can foster social behavior. She also explores the aesthetic qualities of languages and linguistic patterns, offering insights into their role in the evolutionary emergence of language.

Michael Pleyer: Michael Pleyer is an Assistant Professor at the Centre for Language Evolution Studies at Nicolaus Copernicus University in Toruń, Poland. His research focuses on the cognitive, social and interactional foundations of language, especially in the domains of language acquisition and language evolution. He is author (with Stefan Hartmann) of *Cognitive Linguistics and Language Evolution* (2024).

Elizabeth Qing Zhang: Elizabeth Qing Zhang is a lecturer at the School of Linguistic Sciences and Arts at Jiangsu Normal University, China. Her research focuses on the interaction between language and cognition in language evolution, especially the role of vocal learning in language and music evolution. As a native Chinese speaker, she recently has been paying more attention to how Chinese can contribute to language evolution research.

Przemysław Żywiczyński: Przemysław Żywiczyński is a full professor of linguistics at Centre for Language Evolution Studies, Nicolaus Copernicus University in Toruń, Poland. His research combines theoretical reflection on evolutionary theory with empirical studies of improvised gestural communication.

Abstract

This study explores how inter-turn speech pauses influence the perception of cognitive states such as knowledge, confidence, and willingness to grant requests in conversational settings. Longer pauses are typically associated with lower competence and willingness, but Matzinger et al. (2023) discovered that this attribution varies when non-native speakers are involved. They found that listeners were more tolerant of long pauses from non-native than from native speakers when assessing their willingness to grant requests. This may result from the fact that listeners may attribute long pauses to the additional cognitive load non-native speakers face when processing and responding in a second language. This tolerance towards long pauses by non-native speakers did not extend to judgments about non-native speakers' knowledge and confidence - potentially because knowledge questions are less socially engaging than requests. Here, we replicated and extended Matzinger et al.'s (2023) experiment, which focussed on speakers of Polish, to a cross-cultural context with speakers of Chinese. Our results confirmed that non-native accent mediates perceptions of willingness, but not knowledge or confidence. These findings suggest that inter-turn speech pauses play a nuanced role in cognitive state attribution of native and non-native speakers and that cultural factors minimally influence these perceptions. This may indicate that the mechanisms involved are rooted in evolutionarily fundamental aspects of human social communication and cognition.

Keywords: speech pauses, non-native accents, knowledge, confidence, willingness, cognitive state attribution, cross-cultural comparisons

1. Introduction

In many interactions, it is important to assess others' cooperativeness, and there are many different resources that one can draw upon to judge who will or will not be a faithful cooperation partner. One of the most important of those resources is language. However, it is not just the semantic content of potential interaction partners' utterances that can be used as proxies for their cooperativeness but also many other interactional characteristics that are part of the evolved 'human interaction engine' (Levinson 2006). For example, turn-taking and turn-timing have been shown to be particularly important for judging others' cooperative potential from a pragmatic point of view (Levinson 2006).

Especially inter-turn speech pauses, i.e. the time it takes an interlocutor to take the floor, or to respond in a conversation, can serve to express and assess the cognitive states of interaction partners. Such cognitive states include intentions, knowledge, confidence, willingness to grant requests, honesty, comfort with the conversation topic, or psychological and linguistic difficulties - all of which can also be used as proxies for others' cooperative potential (Fox Tree 2002, Harumi & King 2020, Matzinger et al. 2023). In general, Conversation Analysis – which looks at conversations based on adjacency pairs, such as question-answer, request-request granting or offer-offer acceptance (Sacks and Schegloff 1973) – suggests that inter-turn pauses are longer before dispreferred answers, e.g. refusing a request or giving a wrong answer, than before preferred ones, e.g. granting a request or giving a correct answer (Wootton 1981, Levinson 1983). A line of quantitative research supports this prediction. For example, Stivers et al. (2009) and Stivers (2010) showed that in polar questions, confirmations, i.e., preferred answers, are delivered on average between 100 and 500 ms faster than disconfirmations, i.e., dispreferred answers. The difference in timing

between preferred and dispreferred answers reached statistical significance in seven of the ten studied languages and showed similarities across languages and cultures (Stivers et al. 2009). Kendrick and Torreira (2015), studying adjacency pairs in requests, offers, invitations, proposals and suggestions in a corpus of telephone conversations, found that the proportion of dispreferred answers was significantly higher after a long pause (more than 700 ms). Because of those differences in pause duration between preferred and dispreferred answers, interlocutors may use those pauses to infer whether the answers will be preferred or dispreferred.

The interpretation of inter-turn pauses was the focus of a study by Roberts et al. (2011). They used a cross-linguistic perception design to investigate universal patterns in interpreting silence. They tested English, Italian and Japanese participants' rating of speakers' willingness to comply with requests or agree with assessments. To do so, they manipulated telephone conversations in English, Italian and Japanese to obtain three inter-turn silence lengths (0 ms, 600 ms and 1200 ms) and then asked native speaker participants to provide ratings of the speakers' attitudes. Roberts et al. (2011) found that, across languages, as the inter-turn silence became longer, the perceived willingness to comply with requests or agree with assessments became lower. This supports the idea that inter-turn speech pauses are reliable cues to others' cognitive states.

From a cross-cultural and evolutionary perspective, one crucial question is to what degree these cognitive state attributions based on inter-turn pause length are universal. While cross-cultural research on the interpretation of pause length is rare, inter-turn pause production has received more attention. Recent Conversation-Analytic research supports the view that inter-turn pause patterns are similar across languages. For example, psycholinguistic studies showed that most turn reactions come within 500 ms after the end of the preceding turn (Levinson and Torreira 2015). This is impressively

short as it takes around 600 ms to plan for the articulation of a single lexeme (Levelt et al. 1999) and as much as 1500 ms, for the articulation of a simple utterance (Griffin and Bock 2000, Gleitman et al. 2007). These results are in line with the view that the organization of talk follows a universal principle to minimize gaps between conversants' successive contributions and to avoid overlaps between their contributions (Weilhammer and Rabold 2003, Heldner and Edlund 2010, Levinson and Torreira 2015, Roberts et al. 2015). The operation of this general principle combined with the rapidity of taking turns argues against much cultural variation in inter-pause length across cultures.

In contrast, there is a long-standing belief in the literature on ethnography of speaking that there are important differences between cultures in the timing of turn-taking. Sometimes, these differences are expressed in terms of the preference for silence between turns, which anecdotally is characteristic of Nordic languages (Reisman 1974, Lehtonen and Sajavaara 1985), versus the preference for simultaneous speaking (or cooperative overlapping), which was claimed for Antiguans (Reisman 1974) and New York Jews (Tannen 1985). Cross-cultural studies also suggest that Eastern Asians are more tolerant of longer pauses during conversations than Westerners. For example, Shigemitsu (2005) showed that the Chinese and Japanese are comfortable with inter-turn pauses between interlocutors as long as 5 seconds, although the language they were speaking to each other was English instead of their native languages. Another study showed that, compared with Canadians, Chinese used longer pauses in conversations (Crown and Feldstein, 1985). Also, the above-discussed study by Roberts et al. (2011) showed minor culture-specific variations, with the Japanese participants being the most and the Italians the least tolerant towards longer pauses, with the English participants in between. Similarly, the work by Dingemanse and Liesenfeld (2022) on a large number of different languages indicates that the

interpretation of gaps varies with culture but does so minimally and within the thresholds identified by psycholinguistic studies.

Another context where inter-turn pauses may be interpreted differently in different cultures and communities of practice is conversational interaction of L1 and L2 speakers. Here, plenty of research has focused on intra-turn pauses and differences and similarities in L1 vs L2 speakers (e.g. Bosker et al. 2014; De Jong 2016; Trouvain et al. 2016; Matzinger et al. 2020; Raupach 1980; Tavakoli 2011). Much less research has focused on L2 turn-taking behavior in general (Feng, 2022), and inter-turn pausing differences specifically (Peltonen 2017; van Os et al. 2020; Matzinger et al. 2023). L2 speakers with low L2-proficiency, for example, make longer and more frequent inter-turn pauses, are more hesitant to begin a turn, and seem to lack the automaticity to respond to a turn, in contrast to high-proficiency L2 speakers (Peltonen 2017; cf. Feng 2022).

On the level of content of utterances, research in sociolinguistics, raciolinguistics, accentism, and linguistic prejudice has also shown that L2 speakers are generally evaluated more negatively (Lippi-Green 2012; Alim et al. 2016; Roessel et al. 2020). Non-native speech is generally seen as less prestigious and socially attractive (Coupland and Bishop 2007), and non-native speakers are often judged as less employable (Timming 2017; Ramjattan 2022), less competent (Gluszek and Dovidio 2010), and less intelligent (Tsalikis et al. 1991). Lev-Ari and Keysar (2010) found that native speakers rated statements as less true if spoken by a non-native speaker, affecting perceptions of credibility and trustworthiness. While research has shown these evaluations are influenced by racialized stereotypes (Lippi-Green 2012; Alim et al. 2016), the exact features of non-native accents that lead to these judgments, potentially including pause patterns, remain unclear. Overall, these discussions show

that the question whether inter-turn pauses are interpreted similarly across cultures and speaker backgrounds deserves more investigation.

To address questions about inter-turn pauses and nativeness vs. non-nativeness, Matzinger et al. (2023) investigated how differences in pause length interact with attributions of knowledge, confidence and willingness in native and non-native speakers of Polish. In their study, participants listened to short mini-conversations that included either knowledge questions or requests. The questions were asked by Polish native speakers and answers were either given by native speakers of Polish, or Chinese non-native speakers of Polish. The Chinese accent was chosen because it is easily recognized by native speakers of Polish but is not associated with either strong positive or negative sentiments (Wardęga 2017; CBOS 2020). In addition, previous research has shown that native speakers of Polish are influenced by different non-native accents (e.g., Czech, French, Arabic and Russian) in their ratings of the truthfulness of English-language trivia statements (Hanzlíková and Skarnitzl 2017). This made it a likely hypothesis that Polish speakers would show a similar effect when assessing statements in a Chinese non-native accent in Polish.

Matzinger et al. (2023) manipulated the inter-turn pause lengths, so that the answers were either given after a short or a long pause, and predicted that:

- a) raters would generally attribute lower knowledge, confidence and willingness to non-native speakers (based on previous findings such as those of Lev-Ari and Keysar (2010) and Hanzlíková and Skarnitzl (2017)); and
- b) knowledge, confidence and willingness would be rated higher if answers were given after a short pause vs a long pause, as longer pauses are indicative of hesitation and more time needed to think about the answer (Chafe 1980;

Brennan and Williams 1995; Nakane 2007; Roberts et al. 2011; cf. discussion above).

However, they also hypothesized that

c) the effect of pause length would be significantly less pronounced when judging the knowledge, confidence and willingness of non-native speakers.

They hypothesized that in these cases, raters would interpret a longer pause before granting a request not as being due to a lack of knowledge, confidence or willingness, but instead attribute it to non-native speakers needing more cognitive processing time and their difficulties in formulating a response in a non-native language (Cenoz 2000; Van Os et al. 2020; Guyer et al. 2019; Goupil et al 2021).

Hypotheses a) and b) were confirmed by their experiment. However, regarding c) a differential effect of pause length in native vs. non-native speakers was only found for the perceived willingness to grant a request, and not for knowledge and confidence. That is, both native and non-native speakers who answered after a long pause were judged to be less knowledgeable and confident than those who answered after a short pause. In addition, non-native speakers, regardless of whether they answered after a short or long pause, were generally rated to be less knowledgeable and confident than native speakers answering both after a short or long pause. In contrast, non-native speakers were judged as equally willing to grant a request, irrespective of the length of their inter-turn pauses, and the negative effect of long pauses only emerged for native speakers.

A crucial question is whether these results are cross-culturally valid. Specifically, in the light of Conversation Analysis (see above) but also cultural psychology, it is of interest whether the same interactions between cognitive state attribution, pausing behavior

and accents will emerge in participants from cultures and languages that differ strongly from the ones investigated in the original study. Therefore, the current experiment replicates the study by Matzinger et al. (2023) with participants from different cultural backgrounds. Specifically, we replicate the original study with the languages in the experiment reversed, i.e. the participants and native speakers in the mini-conversations were speakers of Mandarin Chinese and the non-native speakers were native Polish-speaking learners of Mandarin Chinese. It is possible that this cross-cultural design could yield different results from the previous studies given that Poland, a country whose East-Central and Roman Catholic Baltic cultural profile is more proximate to Western countries, differs from cultures in China and other countries that are farther East in terms of cultural and psychological dispositions (Schwartz 2008; Tychmanowicz et al. 2021). Still, given the universal relevance of assessing others' cognitive states and the similarity of pausing patterns and turn-taking across cultures (see discussions above), we predicted for the current study similar results as in the original (i.e. no significant effect of language group when comparing the data from Matzinger et al. 2023 and the data of the current study in a linear mixed effects model). That is, we expected that short pauses would be associated with higher knowledge and confidence than long pauses in both native and non-native speakers. In contrast, short pauses should only be associated with a higher willingness to grant requests for native speakers but not for non-native speakers. In addition, overall, we predicted that native speakers would be rated as more knowledgeable, more confident and more willing than non-native speakers. In contrast, if cross-cultural influences have a bigger effect on the results than expected, we would observe a difference between the two experiments, manifesting in a significant effect of language group in a linear mixed effects model.

2. Methods

The study by Matzinger et al. (2023) and the current study were two rating experiments that just differed in the participant pool (native Polish-speaking vs. native Chinese-speaking participants) and in the languages used in the experiment (Polish vs. Chinese) but were identical in all other aspects. Also, this study includes an analysis comparing the data of the two experiments. Because of that, and for the sake of easier accessibility, we report the methods and results of both experiments here¹.

2.1. Experimental Conditions and Procedure

In each of the two experiments, participants listened to brief mini-conversations between friends (Roberts et al. 2011), where one person asked a question or made a request and another person, who was either a native or non-native speaker, responded to the question or the request. As in Matzinger et al. (2023), we experimentally manipulated the pauses that the respondents made before giving their answers to be either short (200 ms) or long (1200 ms). Thus, the two experiments had four conditions each: conversations in which the respondents were 1. native speakers making short pauses, 2. native speakers making long pauses, 3. non-native speakers making short pauses and 4. non-native speakers making long pauses. After listening to each conversation, participants evaluated the respondents in terms of their (a) knowledge,

_

¹ Also, for the sake of clarity, accessibility and comparability, we based the structure and the text of the methods and results sections on Matzinger et al. (2023).

(b) confidence (for the knowledge questions) and (c) willingness to grant the requests (for the requests).

The mini-conversations were presented in two blocks. The first block was the block with the knowledge questions, and the second block was the block with the requests. In each block, participants listened to 32 mini-conversations. We used a within-subjects design; each participant was tested on all four experimental conditions. This means that, in each block of the 32 mini-conversations, there were 8 mini-conversations with native speakers making short pauses, native speakers making long pauses, non-native speakers making short pauses and non-native speakers making long pauses. Within the two blocks, the individual mini-conversations were presented in a randomized order.

After listening to each knowledge question and answer, participants used a slider to indicate (a) how knowledgeable and (b) how confident they perceived the respondents to be (ranging from 0 = not knowledgeable/confident at all to 100 = very knowledgeable/confident). After listening to each request and answer, participants used the slider to indicate how willing they perceived the respondents to grant the requests (ranging from 0 = not willing at all to 100 = very willing). Thus, in total, each participant listened to 64 mini-conversations and provided 96 ratings (two ratings for each knowledge question and one rating for each request).

Before each block of mini-conversations, there was a short practice phase where participants listened to two mini-conversations with intermediate pause durations (700 ms) so that they became familiarized with the experimental setup. The responses of the practice phase were not included in the analyses.

To explore the motivations behind the participants' choices, after the two blocks of rating tasks, participants answered a short questionnaire. In this questionnaire, they

were asked (a) if they had noticed differences in pause durations, (b) if they thought that differences in pause duration had influenced their ratings, (c) if they thought that differences in accent had influenced their ratings and (d) if they thought that they based their decisions on factors other than pause durations and accents. Participants could leave additional comments for all of those questions.

2.2. Participants and Setting

For the Polish experiment, Matzinger et al. (2023) tested 100 participants (50 female, 46 male, 4 other; mean age: $27.0 \pm SD~8.9$ years), who were all native speakers of Polish. A total of 88 Polish participants were recruited via the crowdsourcing platform Prolific and received £ 2 for their participation. A total of 12 Polish participants were students at Nicolaus Copernicus University in Toruń, who received course credits for their participation. For the Chinese experiment, we tested 100 participants (50 female, 48 male, 2 other; mean age: $33.6 \pm SD~10.2$ years), who were all native speakers of Mandarin Chinese. A total of 89 Chinese participants were recruited via the crowdsourcing platform Prolific and received £ 2 for their participation. A total of 11 Chinese participants were students at Jiangsu Normal University, who received course credits for their participation. Participants had the opportunity to drop out from the experiment at any time while still being fully remunerated (dropping out happened for 14 additional participants for the Polish experiment and 15 additional participants for the Chinese experiment, who were not included in the above-described sample).

For all participants, the experiment was administered via the experiment platform Labvanced (Finger et al. 2017). Participants received a link to the experiment and completed the experiment on their own devices, using the built-in loudspeakers or their headphones.

After the experiment, participants provided demographic data (gender and age), information about their language background and information on their pause and accent perception via a questionnaire. Immediately after testing, there was a short debriefing, and participants were informed where they could obtain additional information about the study. In total, the experiment lasted about 20 minutes per participant. The study protocol of the Polish experiment was approved by the Scientific Research Ethics Committee at the University of Warmia and Mazury in Olsztyn, Poland (reference number: Decision no. 8/2018). The Chinese experiment was approved by the Ethics Committee at the University of Vienna (Reference number: 00993). All participants gave their informed consent in accordance with the Declaration of Helsinki.

2.3. Stimulus material

The stimuli, i.e., the mini-conversations with knowledge questions, requests and answers, were scripted and standardized in terms of their format, difficulty and scope. Also, the set of conversations used for the Polish and the Chinese experiment were largely identical. The knowledge questions were yes/no trivia questions, for which it was very unlikely that participants would know the answers to (e.g., Polish: "Jakie były pierwsze warzywa wyhodowane w kosmosie: ziemniaki czy sałata?", Chinese: "宇宙种

植的第一种蔬菜是什么:土豆还是生菜?", English translation: "What were the first vegetables to be grown in space: potatoes or lettuce?"). Most of the questions were inspired by the little-known trivia statements used in the work by Lev-Ari and Keysar (2010) and Boduch-Grabka and Lev-Ari (2021). As these studies had previously shown that non-native speech influences how participants rate these little-known trivia statements, we designed our trivia statements in a similar manner, i.e., we made it unlikely that raters would know the answers to these trivia questions, given many different areas of knowledge they covered. To further decrease the influence of the participants' potential own prior knowledge about the answers, we only used correct answers in the mini-conversations. The answers to the knowledge guestions all started with the phrase "I think" (Polish: "myśle", Chinese: "我觉得/我认为"), so that the respondents' stances towards the answers were standardized. Moreover, most answers repeated parts of the questions or provided some additional context so that participants had more opportunities to recognize if the answer was given by a native or non-native speaker (e.g., Polish: "Myślę, że ziemniaki były pierwszymi warzywami wyhodowanymi w kosmosie", Chinese: "我认为是土豆", English translation: "I think that potatoes were the first vegetables to be grown in space").

The requests first introduced the context of the conversations and were followed by friendly questions that asked the respondents to perform small tasks for the speakers (e.g., Polish: "Tutaj jest bardzo gorąco. Czy możesz otworzyć okno?", Chinese: "这儿太热了,你能开窗户吗?", English translation: "It's really hot in here. Can you open the window?"; Roberts et al. 2011). To standardize the answers, in all cases, the respondents granted the requests without reservation and started their responses with the phrase "Okay, I can..." (Polish: "Dobrze, mogę...", Chinese: "好啊!我(能)..."). Like the answers to the knowledge questions, the answers to the requests also repeated parts of the questions or added some extra information to give the participants the

chance to recognize the accents better (e.g., Polish: "Dobrze, mogę wpuścić trochę świeżego powietrza.", Chinese: "好啊, 我打开窗户透透气。", English translation: "Okay, I can open the window to get some fresh air."). We ensured that all answers had approximately the same number of words to minimize variation in the duration of the spoken answers. In addition, to make the requests sound more natural and culturally appropriate, we changed the content of some of the requests in the Chinese version. For example, whereas in the Polish version, there is a request to make some coffee, in the Chinese version the request is to boil some hot water, and instead of making pancakes in the Polish version, the respondent is asked to boil some noodles in the Chinese version. The full list of trivia questions, requests and the respective answers can be found in the Appendix. Requests which have been adapted to a Chinese cultural setting are marked with an asterisk and also feature the English translation of the adapted Chinese utterances.

All questions and answers were recorded by university staff or students (age range: 20–35). The knowledge questions and requests were recorded by a male Polish native speaker for the Polish experiment and a male Mandarin Chinese native speaker for the Chinese experiment. For the Polish experiment, the answers to the questions and requests were recorded by four Polish native speakers (=native speakers; 2 female, 2 male) and by four native Chinese learners of Polish who were students of Polish philology at Nicolaus Copernicus University in Toruń, who had studied Polish for at least 5 years (CEFR levels: upper A2/lower B1) and who lived in Poland at the time of the study (=non-native speakers; 2 female, 2 male). All Chinese L2 learners of Polish had a clearly recognizable Chinese accent when speaking Polish. For the Chinese experiment, the answers to the questions and requests were read by four Mandarin Chinese native speakers (=native speakers; 2 female, 2 male) and by four native Polish learners of Mandarin Chinese who had studied Chinese for at least 3-4 years (CEFR

levels: upper A2/lower B1) as part of the curriculum and who lived in Poland at the time of the study (=non-native speakers; 2 female, 2 male).² All Polish L2 learners of Chinese had a clearly recognizable Polish accent when speaking Chinese.

Speakers were instructed to read the answers in a flat but still natural tone, in order to minimize and standardize the influence of intonation on how knowledgeable, confident or willing the respondents were perceived. We recorded the stimuli with a ZOOM Handy Recorder (H4n, ZOOM Corporation, Japan; sampling rate: 44.1 kHz, bit depth: 16 bits, format: way) in a guiet office.

All recordings were edited using the software package Praat (version 6.2.19, Boersma and Weenink 2022); first, all individual questions, requests and answers were cut so that there were no silent intervals preceding or following them. Moreover, the amplitude of the individual questions, requests and answers was scaled so that the recordings of all speakers had the same loudness. Then, questions, requests and answers were concatenated, and silent pauses of exactly 200 ms (short pauses) or 1200 ms (long pauses) were inserted between the questions/requests and answers (see the extended discussion about pause durations in our introduction and in Matzinger et al. 2023).

In total, for each of the two experiments, we created 4 different sets of all 64 mini-conversations, in which we semi-randomly varied which speaker answered which question and which pause duration was used after which question. Therefore, in the overall stimulus collection, each question was answered by two different native speakers (one with a short pause and one with a long pause) and two different non-native speakers (one with a short pause and one with a long pause). This ensured that the individual questions were not linked to specific speakers, accents or pause durations. For each experiment, participants were semi-randomly assigned to one of

-

² They participated in a BA programme in Far Eastern studies at Nicolaus Copernicus University in Toruń, Poland.

the four sets of mini-conversations, so that half of the participants heard a specific question being answered by a native speaker and the other half of the participants heard the same question being answered by a non-native speaker. Similarly, one half of the participants heard a specific question being answered after a short pause, and the other half of the participants heard the same question being answered after a long pause.

2.4. Analyses

To test whether the ratings of perceived knowledge, perceived confidence and perceived willingness were influenced by *pause duration* and *nativeness*, i.e., whether different pause durations led to different ratings when the pauses were made by native and non-native speakers, and whether this differed across *language groups*, i.e. Polish and Chinese participants, we applied three Linear Mixed Models (Baayen 2008; Field et al. 2012), in which *perceived knowledge*, *perceived confidence* and *perceived willingness*, respectively, were the response variables.

Initially, in all of those models, pause duration, nativeness, language group and their interactions were included as fixed effects (full triple interaction model: pause duration * nativeness * language group). In the model testing perceived knowledge, there was no significant triple interaction effect of pause duration, nativeness and language group (see the results section), which is why, for perceived knowledge, we excluded the triple interaction from the model and computed a single interaction model that only included an interaction between pause duration and nativeness, and included language group

as a separate predictor (single interaction model: pause duration * nativeness + language group). Also, in this model, the interaction effect was non-significant, which is why our final model for perceived knowledge was a model that only included pause duration, nativeness and language group, but no interactions, as fixed effects (reduced model: pause duration + nativeness + language group).

In all of our models, we entered random intercepts of the participants and speakers into the model. To avoid inflated type I error rates, each model included a random slope (Barr et al. 2013; Schielzeth and Forstmeier 2009) of nativeness within the participants. Models that also included a random slope of pause duration within the participants did not converge. In the final model testing perceived knowledge and in the model testing perceived willingness, model optimization was performed using the "bobyqa" algorithm (Powell 2009).

The sample size for each model was 6400 data points (100 participants tested on 32 utterances each, for two language groups each). We used the short pause duration, the native language condition and the Polish language group as reference levels in the models.

The models were fitted in R (version 4.3.0; R Development Core Team 2023), using the function *Imer* of the R-package "Ime4" (version 1.1-33; Bates et al. 2015). Additionally, we used the R-package "ImerTest" (version 3.1-3; Kuznetsova et al. 2017) to obtain p-values for the effects of the individual predictors.

We used likelihood ratio tests to test the overall significance of the full models as compared with respective null models comprising only the random effects (R function anova; Dobson 2002).

For each linear mixed effects model, we visually inspected a qqplot as well as the residuals plotted against fitted values. Those inspections checked whether the assumptions of normally distributed and homogeneous residuals were fulfilled and indicated no obvious deviations from normality or homoscedasticity.

For calculating effect sizes that serve as indicators of the goodness-of-fit of our models, we used the function r.squaredGLMM of the R package "MuMIn" (version 1.47.5; Bartón 2018) to obtain the marginal and conditional R² for each model (Johnson 2014, Nakagawa and Schielzeth 2013, Nakagawa et al. 2017). The marginal R² (R²m) indicates the variance explained by the fixed effects, whereas the conditional R² (R²c) indicates the variance explained by the fixed and random effects.

All hypotheses, study protocols and analyses for the two individual experiments, including the R files with our models, were preregistered (Polish experiment: https://osf.io/z83yx/, Preregistration date: 29 September 2022; Chinese experiment: https://osf.io/8w2gj/, Preregistration date: 30 June 2023). The models combining the data from the Polish and Chinese experiments, have not been preregistered. However, they differ from the preregistered models only in the inclusion of an additional predictor of language group and its interaction with the other predictors. This addition became possible by combining the datasets from both experiments. The R files for this combined analysis can be found on osf (https://osf.io/8w2gj/).

3. Results

3.1. Rating data

3.1.1. Perceived Knowledge

Regarding the perceived knowledge of speakers, the comparison of the full triple interaction model and the null model showed an effect of either pause duration, nativeness, language group or their interaction on how knowledgeable speakers were perceived (likelihood ratio test: χ 2 = 50.40, df = 7, p < 0.001; effect size for the full model: R2m = 0.05, R2c = 0.26). Specifically, we found a significant main effect of pause duration on perceived knowledge, indicating that speakers who made shorter pauses were perceived as more knowledgeable than speakers who made longer pauses (Table 1, Reduced model; Figure 1A left & right panel; Figure 1B left & right panel). Furthermore, there was a significant effect of nativeness on perceived knowledge, with native speakers being perceived as more knowledgeable than non-native speakers (Table 1, Reduced model; Figure 1A left & right panel; Figure 1B left & right panel). However, there was no significant effect of language group on perceived knowledge (Table 1, Reduced model; Figure 1A left & right panel; Figure 1B left & right panel), indicating that Polish and Chinese participants gave similar ratings. Also, there was no significant triple interaction effect of pause duration, nativeness and language group (Table 1, Full model including a triple interaction effect), and no significant single interaction effect of pause duration and nativeness (Table 1, Model including a single interaction effect). This shows that the raters did not interpret the

duration of pauses differently depending on whether the pauses were made by native or non-native speakers, and that also in this regard, Polish and Chinese participants behaved similarly. The negative effect of non-native accents on perceived knowledge was approximately four times as big as the negative effect of long pause durations (Table 1, Reduced model).

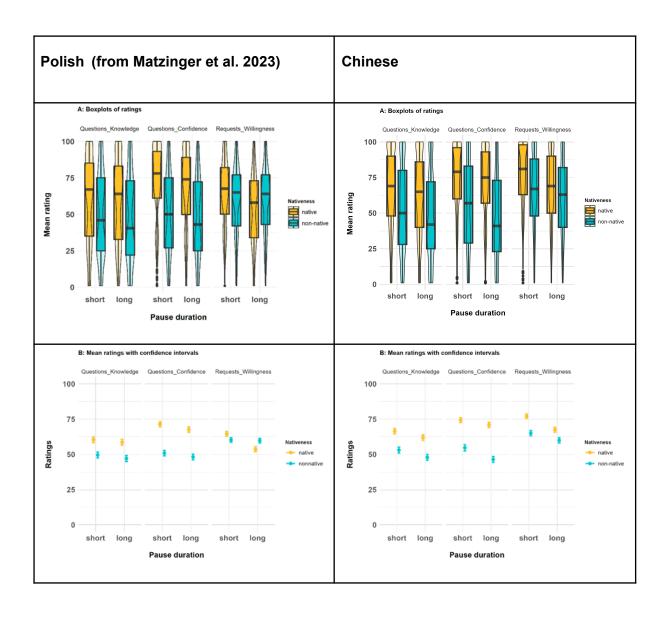


Figure 1. Ratings of the perceived knowledge and confidence of speakers answering knowledge questions, and of the perceived willingness of speakers to comply with requests. Answers were given either by native or non-native speakers and were preceded by either a short (200 ms) or long (1200 ms) pause. Ratings range from 0 all) (not knowledgeable/confident/willing at to 100 (very knowledgeable/confident/willing). Left column: data from the Polish experiment, right column: data from the Chinese experiment. (A): Boxes depict medians and quartiles, whiskers depict minimum and maximum values, and black dots depict outliers. Violin shapes around the boxes depict the distribution of the ratings. The width of the violin shapes at a given y coordinate corresponds to the number of ratings in this region. Note that, when evaluating the speakers' knowledge, confidence and willingness, participants used the full range of the rating scale. (B): Mean and 95% confidence intervals of participants' responses. Non-overlapping confidence intervals indicate significant differences between the groups.

Table 1. Results of the linear mixed models exploring the effects of pause duration, nativeness and language group on perceived knowledge. The table reports estimated model coefficients (Estimate), standard errors (SE) and p-values (p). Because of the non-significant interaction effects in the full triple interaction model and the single interaction model, to determine the effects of pause duration, nativeness and language group alone, we also report the results of a reduced model excluding those interaction effects.

Full model including a triple interaction effect	Estimate	SE	p
Intercept	60.38	2.40	
PauseDuration_Long	-1.75	1.31	0.182
Nativeness_Nonnative	-10.83	2.98	0.001
LanguageGroup_Chinese	5.40	3.39	0.121
PauseDuration_Long:Nativeness_Nonnative	-0.67	1.86	0.720
PauseDuration_Long:LanguageGroup_Chinese	-1.90	1.86	0.306
Nativeness_Nonnative:LanguageGroup_Chines e	-1.90	4.22	0.657
PauseDuration_Long:Nativeness_Nonnative: LanguageGroup_Chinese	-0.61	2.62	0.816
Model including a single interaction effect			
Intercept	61.41	2.09	

PauseDuration_Long	-2.70	0.91	0.004
Nativeness_Nonnative	-11.78	2.07	<0.001
LanguageGroup_Chinese	3.34	2.53	0.192
PauseDuration_Long:Nativeness_Nonnative	-0.97	1.31	0.460
Reduced model excluding the interaction effects			
Intercept	61.66	2.07	
PauseDuration_Long	-3.19	0.66	<0.001
Nativeness_Nonnative	-12.27	1.96	<0.001
LanguageGroup_Chinese	3.34	2.53	0.193

3.1.2. Perceived Confidence

The comparison of the full triple interaction model and the null model that investigated perceived confidence showed an effect of either pause duration, nativeness, language group or their interaction on how confident speakers were perceived (likelihood ratio test: χ 2 = 87.73, df = 7, p < 0.001; effect size for the full model: R2m = 0.12, R2c = 0.32). More specifically, pause duration had a significant effect on perceived confidence, with speakers who made shorter pauses being perceived as more confident than speakers who made longer pauses (Table 2; Figure 1A left & right panel; Figure 1B left & right panel). In addition, nativeness had a significant effect on perceived knowledge, i.e., native speakers were perceived as more confident than non-native speakers (Table 2; Figure 1A left & right panel; Figure 1B left & right panel). The effect of nativeness on perceived confidence was approximately five times bigger than the effect of pause duration (Table 2). There was no significant main effect of language group (Table 2). However, there was a significant triple interaction effect of pause duration, nativeness and language group, indicating that how pause duration is modulated by nativeness differs across language groups (Table 2; Figure 1A left & right panel). Figure 1B (left & right panel) indicates that the difference between long and short pauses made by non-native speakers is bigger in the Chinese experiment than in the Polish experiment. This shows that Chinese raters perceive non-native speakers making long pauses as less confident than Polish raters.

Table 2. Results of the linear mixed models exploring the effects of pause duration, nativeness and language group on perceived confidence. The table reports estimated model coefficients (Estimate), standard errors (SE) and p-values (p).

Full model including a triple interaction effect	Estimate	SE	р
Intercept	71.46	2.88	
PauseDuration_Long	-3.82	1.24	0.002
Nativeness_Nonnative	-20.54	3.95	<0.001
LanguageGroup_Chinese	2.37	4.07	0.568
PauseDuration_Long:Nativeness_Nonnative	1.10	1.75	0.531
PauseDuration_Long:LanguageGroup_Chinese	0.93	1.75	0.600
Nativeness_Nonnative:LanguageGroup_Chines e	1.48	5.59	0.794
PauseDuration_Long:Nativeness_Nonnative: LanguageGroup_Chinese	-6.03	2.48	0.015

Because the knowledge and confidence ratings were based on the same mini-conversations, we computed Spearman's rank correlation coefficient to assess the relationship between perceived knowledge and perceived confidence. There was a

medium positive correlation between the two variables for both the Polish (r = 0.56, p < 0.001) and the Chinese experiment (r = 0.67, p < 0.001; Cohen 1992).

3.1.3. Perceived Willingness to Grant Requests

A comparison of the full triple interaction model and the null model testing the perceived willingness of speakers to grant requests revealed that there was an effect of either pause duration, nativeness, language group or their interaction on how willing speakers were perceived (likelihood ratio test: χ 2 = 239.97, df = 7, p < 0.001; effect size for the full model: R2m = 0.06, R2c = 0.41). More specifically, pause duration had a significant main effect on perceived willingness, with speakers who made shorter pauses being perceived as more willing than speakers who made longer pauses (Table 3; Figure 1A left & right panel; Figure 1B left & right panel). However, there was also a significant interaction effect of pause duration and nativeness (Table 3), indicating that the perception of pause duration was mediated by the speakers' accents: the overall negative effect of long pauses was less pronounced when speakers had a non-native accent (Figure 1A left & right panel; Figure 1B left & right panel). In addition to that, our model revealed a significant triple interaction effect of pause duration, nativeness and language group (Table 3). This shows that how much the influence of pause duration on perceived willingness was mediated by nativeness differed across language groups: for the Polish raters, the mediating effect of nativeness was larger than for the Chinese group of raters. Polish raters gave similar willingness ratings for non-native speakers, irrespective of whether they made short or long pauses before their answers. In contrast, Chinese raters perceived non-native speakers making long pauses as less

willing than those making short pauses but this difference between long and short pauses was much smaller for the non-native speakers than for the native speakers (Figure 1B left & right panel). Thus, the mediating effect of nativeness on the perception of pause duration was present in both the Polish and the Chinese raters but it was less pronounced in the Chinese raters. Finally, the model also revealed a significant main effect of language group (Table 3), indicating that, overall, the Chinese raters provided higher willingness ratings than the Polish raters (Figure 1A left & right panel; Figure 1B left & right panel). In contrast to perceived knowledge and confidence, we did not find a significant main effect of nativeness, i.e. the willingness of native and non-native speakers was, overall, rated similarly (Table 3). This result was mostly driven by the strong mediating effect of nativeness in the Polish raters.

Table 3. Results of the linear mixed models exploring the effects of pause duration, nativeness and language group on perceived willingness. The table reports estimated model coefficients (Estimate), standard errors (SE) and p-values (p).

Full model including a triple interaction effect	Estimate	SE	p
Intercept	64.47	3.47	
PauseDuration_Long	-10.75	0.99	<0.001

Nativeness_Nonnative	-4.19	4.67	0.385
LanguageGroup_Chinese	12.42	4.91	0.022
PauseDuration_Long:Nativeness_Nonnative	10.23	1.40	<0.001
PauseDuration_Long:LanguageGroup_Chinese	1.24	1.40	0.377
Nativeness_Nonnative:LanguageGroup_Chines e	-7.14	6.60	0.298
PauseDuration_Long:Nativeness_Nonnative: LanguageGroup_Chinese	-5.62	1.98	0.005

3.2. Questionnaire data

Here, we exploratorily report the answers of the post-study questionnaire, which help to understand why participants rated the mini-conversations in the way they did. Note that, in the questionnaire, the participants were explicitly asked to comment on our experimental manipulations, which is why it is impossible to disentangle if the participants were really influenced by the manipulations during the experiment or if they

just believed they had been influenced by them in retrospect when their attention was drawn to the manipulations in the questionnaire.

Our post-study questionnaire showed that the factors influencing the ratings of the Polish and the Chinese participants were largely similar. About two thirds of the Polish participants (65.0%) and almost three quarters of the Chinese participants (72.0%; Table 4) had noticed the differences in pause durations in the mini-conversations. Regarding their ratings, approximately half of the participants in both the Polish and the Chinese group reported that the duration of the pauses influenced their ratings of knowledge and confidence (53.0% in both groups; Table 4), whereas a little less than half of the participants in each group reported that the duration of the pauses influenced their ratings of willingness (Polish: 44.0%, Chinese 46.0%; Table 4), A little more than half of the Polish participants (53.0%) and a little less than half of the Chinese participants (43.0; Table 4) reported that the accent of the speakers influenced their ratings. It is interesting that, overall, the influence of accent on the ratings was bigger than the influence of pause duration (Figure 1B), although the number of participants who reported that pause duration and accent influenced their ratings were comparable. This might indicate that, when being influenced by accent, participants gave more extreme ratings than when being influenced by pause durations.

Table 4. Results of the descriptive analyses of the questionnaire. Chi-squared tests revealed no significant differences between the percentages of participants reported for the Polish and the Chinese experiment.

	i	
	Percentage of Polish	Percentage of
	Participants	Chinese Participants
Noticed differences in pause duration	65.0	72.0
Reported that pause duration	53.0	53.0
influenced their knowledge and		
confidence ratings		
Confidence ratings		
Reported that pause duration	44.0	46.0
influenced their willingness ratings		
Reported that the speakers' accent	53.0	43.0
influenced their overall ratings		
Reported that other factors influenced	50.0	53.0
their ratings		
then ratings		

Around half of the participants (Polish: 50.0%, Chinese: 53.0%; Table 4) reported that factors other than pause durations and accent influenced their ratings. Out of the 50 Polish participants who reported that other factors influenced their ratings, 46 provided further comments. Approximately half (25) of those comments mentioned talking speed as an additional influencing factor, with faster speech indicating higher knowledge, confidence and willingness. In the Polish samples, on average, native speakers had a faster talking speed than that of non-native speakers, with native speakers needing, on

average, $2.28 \pm \text{SD}\ 0.87$ s and the non-native speakers needing, on average, $3.68 \pm \text{SD}\ 1.84$ s per answer. Out of the 53 Chinese participants who reported that other factors influenced their ratings, all provided further comments. 45 of them mentioned talking speed as an additional influencing factor, with faster speech indicating higher knowledge, confidence and willingness. In our Chinese samples, on average, native speakers had a faster talking speed than that of non-native speakers, with native speakers needing, on average, $1.73 \pm \text{SD}\ 0.51$ s and the non-native speakers needing, on average, $2.86 \pm \text{SD}\ 1.08$ s per answer.

Those differences in talking speed between native and non-native speakers were unavoidable because, naturally, most L2 learners speak more slowly than native speakers. Because of this difference in speech tempo between native and non-native speakers, it is possible that the observed effects of accent can, to a large extent, be explained by speech tempo, and that speech tempo, which is closely related to perceptions of fluency, might have been more relevant for the ratings than other characteristics of non-native speech, such as deviations in the pronunciation of phonemes or intonation (cf. Bosker et al. 2014).

4. Discussion

In this study, we extended Matzinger et al.'s (2023) study on inter-turn pause length and differences in cognitive state attribution to a cross-cultural setting, and compared whether inter-turn pause length in native and non-native speech was perceived differently by Polish and Chinese participants. Due to the cross-cultural and cross-linguistic relevance of cognitive state attribution in communication partners, we had predicted that inter-turn pause perception would be similar across cultures, even though Polish and Chinese cultures differ in many respects (Schwartz 2008). We found

that there are some cultural differences in the precise results but that, overall, confirming our predictions, the duration of inter-turn speech pauses is used similarly across Polish and Chinese participants to make inferences about the knowledge, confidence and willingness of conversation partners.

More precisely, both Polish and Chinese raters perceived native and non-native speakers that made long pauses before their answers as less knowledgeable than speakers that made short pauses. Regarding confidence, results were similar: both Polish and Chinese raters perceived native and non-native speakers that made long pauses before their answers as less confident than speakers that made short pauses. For the Chinese raters, the long pauses in non-native speech had a slightly bigger negative effect than for the Polish raters. When it comes to perceived willingness to grant requests, we found that across the two language groups, there was a mediating effect of non-native accent, indicating that in both groups, long pauses in non-native speakers were not perceived as negatively as in native speakers. Here, the mediating effect was slightly bigger for the Polish participants, where there was hardly any difference between short and long pauses, than for the Chinese ones, who still showed a slight preference for short pauses. Also, throughout all conditions, the Chinese raters gave slightly higher ratings for willingness than the Polish ones. Overall, those results support the universalistic assumptions of Conversation Analysis about the interpretation of inter-pause silence (see Introduction) and corroborate the findings of previous research that pause length is a cue that we use to evaluate cognitive states of interlocutors - in our case, their knowledge, confidence and willingness (cf. Roberts et al., 2015).

In addition, we found that both rater groups consistently rated speakers with non-native accents as being less knowledgeable and confident. This replicates previous studies

and extends their findings to Polish-accented Chinese. There are several factors that likely contributed to this perception. First, non-native speakers tend to have slower speech rates, which can affect how fluent, and thus how knowledgeable and confident, they are perceived to be (Van Os et al. 2020; Suzuki et al. 2021). Also, native listeners may experience 'processing difficulties' with non-native accents and may misattribute those challenges to low knowledge and confidence on the side of the speakers (Lev-Ari and Keysar 2010). Furthermore, other factors such as the challenge of being recorded speaking a non-native language in an unfamiliar setting might have led to genuine reductions in the non-native speakers' confidence. This may have been reflected in, for example, prosodic features of the non-native speakers and may have been picked up by the raters (for a more in-depth discussion, see Matzinger et al. 2023).

Besides our main result that cognitive state attribution based on pauses and accents happens similarly across cultures, it is also interesting to see that people reacted differently to non-native speaker pauses before knowledge questions and requests. The reason why raters were more tolerant towards long pauses in non-native speakers in requests than in knowledge questions might be that requests are more socially engaging and reflect cooperation and trust, while knowledge questions focus more on competence and ability (Matzinger et al. 2023). This might also be the reason why the minor cultural differences that we found in our study, emerged in the responses to requests. In accordance with literature on politeness, requests have a large potential to disrupt the interpersonal balance between conversants, i.e. directly target the requester's needs to be supported by others (Goffman 1963; Brown and Levinson 1987). The Western conception of politeness (Eelen 2001) is based on one's individualistic wants to be appreciated and unimpeded by others. One the other hand, the Chinese *limao* (roughly translated as politeness) refers to a set of obligations an individual has towards society: "[politeness is] a sanctioned belief that an individual's

behavior ought to live up to the expectations of respectfulness, modesty, attitudinal warmth and refinement" (Gu 1990: 245). Given the interactional status of requests and the societal weight of *limao*, the Chinese raters may have been less lenient in interpreting long pauses even if they were produced by non-natives.

As such, the results, and explanations for the differences in the cognitive and interactional function of inter-turn pauses also converge with developments in cross-cultural psychology, which have shown that there is a breadth of cultural differences in motivation, emotion, cognition and conceptualisation. Previous research has found a number of differences in cognitive, motivational and attentional tendencies between Western and Chinese populations in areas such as social cognition (e.g. Morris & Peng 1994), preferences (e.g. Corriveau et al. 2017), independent vs interdependent orientation (e.g. Kitayama & Salvador 2024), visual attention (e.g. Ji et al. 2000), holistic vs analytic processing (e.g. Koo et al. 2018), values (e.g. Ji et al. 2001), and others (cf. Carstensen et al. 2024; Nisbett 2003). These results in turn call "for a thorough reconceptualization of human psychology as situated and realized through a complex of various cultural practices and meanings" (Kitayama & Salvador 2024). These results are also highly relevant for intercultural communication, as research has shown that people from different cultures and communities of practice might engage in different evaluations and cognitive processing of their interaction partners' communicative behavior. Indeed there might be misunderstandings and different evaluations based on different interpretations and cultural interpretations of communicative behavior, including inter-turn speech pauses (Harumi & King 2020; Tannen 1985; Nakane 2012, Verouden & van der Sanden 2018; Yamada 1997). While our study showed that the general tendencies observed for Polish raters also held for Chinese ones, it would be worthwhile to explore the minor cross-cultural cognitive and linguistic differences in the interpretation of pauses more fully in future investigations.

In terms of intercultural contact and communication, it is also worth noting that a big proportion of the Chinese-speaking raters, who were recruited via Prolific, currently live in places where Chinese is not the main community language - this is in contrast to the Polish raters, most of whom lived in Poland at the time of the study. The difference between Chinese-speaking raters who live in China and those who live outside China could subtly change the results of evaluating the credibility and willingness of the speakers. In terms of language, studies have shown that foreign language learning affects speech processing and production of the first language (van Hell & Dijkstra, 2002; Linck, Kroll & Sunderman, 2009; Chang, 2012). Immersion in a second language inhibits the phonological system of the native language (Levy et al., 2007). In terms of culture, it is also possible that living in a country with a different culture explicitly influences the perception of their native culture (Liu, Rigoulot & Pell, 2016). It might also influence their evaluation of non-native speakers of languages in general, as our raters likely lived in settings characterized by greater degrees of intercultural communication and contact than would be the case for raters living in China.

Overall, we have shown that inter-turn speech pauses have an important influence on attributing cognitive states such as knowledge, confidence and willingness to interlocutors and that this effect can be fruitfully studied for different languages, accents, and cultures. Future studies should extend the range of languages and cultures being studied, to further probe the overall universality of this effect as well as the cultural differences of the effect on a more fine-grained level.

From an evolutionary perspective, the similarities in cognitive state attribution across different languages and cultures suggest that the mechanisms involved may stem from early stages in language and cognitive evolution. This is supported by the fact that social timing more generally is crucial for interaction in many species (Verga et al.

2023), and turn-taking is found in many mammals, birds, insects and anurans (Pika et al. 2018; Abreu & Pika 2022). Interestingly, recent research has shown that chimpanzee communicative interactions displayed a similar average pause length between turns as those found in human language. The underlying rules for social timing and turn-taking in interaction might therefore be based on "shared ancestral mechanisms or convergent strategies that enhance coordinated interactions or manage competition for communicative 'space'" (Badihi et al. 2024). Investigating how different inter-turn pause durations contribute to cognitive state attribution and meaning-making in non-human animals would offer valuable insights into these mechanisms.

5. Funding information

This research was supported by the Polish National Science Centre (NCN) under grant agreement UMO-2023/49/B/HS2/0222, and by project No. 2021/43/P/HS2/02729 co-funded by the Polish National Science Centre (NCN) and the European Union Framework Programme for Research and Innovation Horizon 2020 under the Marie Skłodowska-Curie grant agreement No. 945339.

6. Acknowledgements

We thank Klaudia Karkowska for helping with formatting and proofreading the manuscript and the references.

7. References

- Abreu, F. & Pika, S. (2022). Turn-taking Skills in Mammals: A Systematic Review Into Development and Acquisition. *Frontiers in Ecology and Evolution*, 10.
- Alim, H. S., Rickford, J.R. & Ball, A.F. (Eds.) (2016). *Raciolinguistics: How Language Shapes Our Ideas about Race*. Oxford University Press.
- Atkinson, J. M. & Drew, P. (1979). Order in Court: The Organization of Verbal Interaction in Juridical Settings. Macmillan.
- Baayen, R. H. (2008). Analyzing Linguistic Data. Cambridge University Press.
- Badihi, G., Graham, K. E., Grund, C., Safryghin, A., Soldati, A., Donnellan, E., Hashimoto, C., Mine, J. G., Piel, A. K., Stewart, F., Slocombe, K. E., Wilke, C., Townsend, S. W., Zuberbühler, K., Zulberti, C., & Hobaiter, C. (2024).
 Chimpanzee gestural exchanges share temporal structure with human language. *Current Biology*, 34(14), R673–R674.
 https://doi.org/10.1016/j.cub.2024.06.009
- 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, 255–78.
- Bartón, K. (2018). *MuMIn: Multi-Model Inference. R Package Version 1.42.1*. Available online: https://cran.r-project.org/package=MuMIn (accessed on 2 January 2023).
- Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using Lme4. *Journal of Statistical Software*, 67, 1–48.
- Boduch-Grabka, K. & Lev-Ari, S. (2021). Exposing Individuals to Foreign Accent Increases their Trust in What Nonnative Speakers Say. *Cognitive Science*, 45, e13064.

- Bosker, H. R., Quené, H., Sanders, T. & De Jong, N. H. (2014). The perception of fluency in native and nonnative speech. *Language Learning*, 64, 579–614.
- 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 and Language*, 34, 383–98.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*.

 Cambridge University Press.
- Carstensen, A., Cao, A., Tan, A. W, Liu, D., Liu, Y., Bui, M. K, et al. (2024). Cognitive diversity in context: US-China differences in children's reasoning, visual attention, and social cognition. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 46.
- CBOS (Centrum Badania Opinii Publicznej). (2020). *Tożsamość narodowa i postrzeganie praw mniejszości narodowych i etnicznych*. Warszawa: Fundacja Centrum Badania Opinii Społecznej.
- Cenoz, J. (2000). Pauses and hesitation phenomena in second language production.

 ITL: Review of Applied Linguistics, 127, 53–69.
- Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, 40(2), 249–268. https://doi.org/10.1016/j.wocn.2011.10.007
- Chafe, W. L. (1980). Some reasons for hesitating. In H.W. Dechert & M. Raupach (Eds.) *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler* (pp. 169-180). De Gruyter Mouton.
- Cohen, J. (1992). A Power Primer. Psychological Bulletin, 112, 155–59.
- Corriveau, K. H., DiYanni, C. J., Clegg, J. M., Min, G., Chin, J., & Nasrini, J. (2017).

 Cultural differences in the imitation and transmission of inefficient actions.

 Journal of Experimental Child Psychology, 161, 1–18.

 https://doi.org/10.1016/j.jecp.2017.03.002

- Crown, C. L., & Feldstein, S. (1985). Psychological correlates of silence and sound in conversational interaction. In D. Tannen & M. Saville-Troike (Eds.), Perspectives. on silence (pp. 31-54). Ablex Publishing.
- Coupland, N. & Bishopm H. (2007). Ideologised values for British accents 1. *Journal of Sociolinguistics*, 11, 74–93.
- Davis, N. (2003). God's Playground A History of Poland: Volume II: 1795 to the Present. Oxford University Press.
- De Jong, N. H. (2016). Predicting pauses in L1 and L2 speech: The effects of utterance boundaries and word frequency. *International Review of Applied Linguistics in Language Teaching*, 54, 113–32.
- Dingemanse, M. & Liesenfeld, A. (2022). From text to talk: Harnessing conversational corpora for humane and diversity-aware language technology. In *Proceedings* of the 60th Annual Meeting of the Association for Computational Linguistics 1 (Long Papers) (pp. 5614–33). Association for Computational Linguistics,
- Dobson, A. J. (2002). *An Introduction to Generalized Linear Models*. Chapman and Hall.
- Eelen, G. (2001). A critique of politeness theory. Routledge.
- Feng, R. (2022). Cognitive Factors Influencing Utterance Fluency in L2 Dialogues:

 Monadic and Non-monadic Perspectives. *Frontiers in Psychology*, 13, 926367.
- Field, A, Miles, J. & Field, Z. (2012). *Discovering Statistics Using R*. Los Angeles: SAGE.
- Finger, H., Goeke, C., Diekamp, D. Standvoß, K. & König, P. (2017). *LabVanced: A unified JavaScript framework for online studies. In International Conference on Computational Social Science* (Cologne) (pp. 1–3). University of Osnabrück,
- Fox Tree, J. E. (2002). Interpreting pauses and ums at turn exchanges. *Discourse Processes*, 34, 37–55.

- Gleitman, L. R., January, D., Nappa, R. & Trueswell, J. C.. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language*, 57, 544–69.
- Gluszek, A., & Dovidio, J.F.. (2010). Speaking with a nonnative accent: Perceptions of bias, communication difficulties, and belonging in the United States. *Journal of Language and Social Psychology*, 29, 224–34.
- Goffman, E. (1963). Embarrassment and Social Organization. In N. J. Smelser & W. T. Smelser (Eds.), *Personality and social systems* (pp. 541–548). John Wiley & Sons, Inc..
- Goupil, L., Ponsot, E., Richardson, D., Reyes, G., & Aucouturier, J.-J. (2021). Listeners' perceptions of the certainty and honesty of a speaker are associated with a common prosodic signature. *Nature Communications*, 12, 1–17.
- Griffin, Z. M. & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, 11, 274–79.
- Gu, Y. (1990). Politeness phenomena in modern Chinese. *Journal of Pragmatics*, 14(2), 237-257.
- Guyer, J. J., Fabrigar, L. R. & Vaughan-Johnston, T. I.. (2019). Speech rate, intonation, and pitch: Investigating the bias and cue effects of vocal confidence on persuasion. *Personality and Social Psychology Bulletin*, 45,389–405.
- Hanzlíková, D., & Skarnitzl, R. (2017). Credibility of native and non-native speakers of English revisited: Do non-native listeners feel the same? *Research in Language*, 15, 285–98.
- Harumi, S., & King, J.. (2020). East Asian Perspectives on Silence in English Language Education: An Introduction. In J. King & S. Harumi (Eds.) *East Asian Perspectives on Silence in English Language Education* (pp. 1-16). Multilingual Matters

- Heldner, M. & Edlund, J. (2010). Pauses, gaps and overlaps in conversations. *Journal of Phonetics*, 38, 555–68.
- Ji, L. J., Peng, K., & Nisbett, R. E. (2000). Culture, control, and perception of relationships in the environment. *Journal of Personality and Social Psychology*, 78(5), 943–955.
- Ji, L. J., Nisbett, R. E., & Su, Y. (2001). Culture, change, and prediction. *Psychological Science*, 12(6), 450-456.
- Johnson, P. C. D. (2014) Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. *Methods in Ecology and Evolution* 5, 44–946.
- Kendrick, K. H., & Torreira, F. (2015). The timing and construction of preference: A quantitative study. *Discourse Processe*, 52, 255–89.
- Kitayama, S., & Salvador, C. E. (2024). Cultural psychology: Beyond east and west.

 Annual Review of Psychology, 75(1), 495-526.
- Koo, M., Choi, J. A., & Choi, I. (2018). Analytic versus holistic cognition: Constructs and measurement. In J. Spencer-Rodgers & K. Peng (Eds.), *The psychological* and cultural foundations of East Asian cognition: Contradiction, change, and holism (pp. 105–134). Oxford University Press.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). ImerTest Package:

 Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82, 1–26.
- Lehtonen, J. & Sajavaara, K. (1985). The Silent Finn. In D.Tannen & M. Saville-Troike (Eds.) *Perspectives on Silence* (pp 193–201). Ablex Publishing.
- Lev-Ari, S. & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, 46, 1093–96.
- Levelt, W. M. (1999). Producing spoken language: A blueprint of the speaker. In C. M. Brown & P. Hagoort (Eds.) *The Neurocognition of Language* (pp. 83–122).

 Oxford University Press.

- Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.
- Levinson, S. (2006). Cognition at the Heart of Human Interaction. *Discourse Studies*, 8 (1): 85–93. doi:10.1177/1461445606059557.
- Levinson, S. C. (2016). Turn-taking in Human Communication–Origins and Implications for Language Processing. *Trends in Cognitive Sciences* 20: 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.
- Levy, B. J., McVeigh, N. D., Marful, A., & Anderson, M. C. (2007). Inhibiting Your Native

 Language. *Psychological Science*, 18(1), 29–34.

 https://doi.org/10.1111/j.1467-9280.2007.01844.x
- Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing Access to the Native

 Language While Immersed in a Second Language: Evidence for the Role of
 Inhibition in Second-Language Learning. *Psychological Science*, 20(12),
 1507–1515. https://doi.org/10.1111/j.1467-9280.2009.02480.x
- Lippi-Green, R. (2012). English with an Accent: Language, Ideology, and Discrimination in the United States. Routledge.
- Liu, P., Rigoulot, S., & Pell, M. D. (2017). Cultural immersion alters emotion perception:

 Neurophysiological evidence from Chinese immigrants to Canada. *Social Neuroscience*, 12(6), 685–700.

 https://doi.org/10.1080/17470919.2016.1231713
- Lundholm-Fors, K. (2015). Production and Perception of Pauses in Speech. [Doctoral dissertation, University of Gothenburg], Gothenburg University Publications Electronic Archive. http://hdl.handle.net/2077/39346
- Matzinger, T., Ritt, N. & Fitch, W. T. (2020). Non-native speaker pause patterns closely correspond to those of native speakers at different speech rates. *PLoS ONE*, 15: e0230710.

- Matzinger, T., Pleyer, M., & Żywiczyński, P. (2023). Pause Length and Differences in Cognitive State Attribution in Native and Non-Native Speakers. *Languages*, 8(1), 26.
- Morris, M. W., & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology*, 67(6), 949–971. https://doi.org/10.1037/0022-3514.67.6.949
- Nakagawa, S. & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4, 133–42.
- Nakagawa, S., Johnson, P. C. D. & Schielzeth, H. (2017). The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, 14, 20170213.
- Nakane, I. (2007). Silence in Intercultural Communication. John Benjamins.
- Nakane, I. (2012). Silence. In C. B. Paulston, S. F. Kiesling, & E. S. Rangel (Eds.), *The Handbook of Intercultural Discourse and Communication* (pp. 158-179).

 Blackwell.
- Nisbett, R. (2003). The Geography of Thought: How Asians and Westerners Think

 Differently... and why. Simon and Schuster.
- Peltonen, P. (2017). Temporal fluency and problem-solving in interaction: An exploratory study of fluency resources in L2 dialogue. *System*, 70, 1–13.
- Pika, S., Wilkinson, R., Kendrick, K. H. & Vernes, S. C. (2018). Taking turns: bridging the gap between human and animal communication. Proceedings of the Royal Society B, 285, 28520180598.
- Powell, M. J. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. Cambridge NA Report NA2009/06, University of Cambridge, Cambridge, 26, 26-46.

- R Development Core Team. (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Ramjattan, V. A. (2022). Accenting racism in labour migration. *Annual Review of Applied Linguistics*, 42, 87–92.
- Raupach, M. (1980). Temporal variables in first and second language speech and perception of fluency. In H. W. Dechert & M. Raupach (Eds.) *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler (pp.* 263–70). De Gruyter Mouton.
- Reisman, J. M., & Yamokoski, T. (1974). Psychotherapy and Friendship: An Analysis of the Communications of Friends. *Journal of Counseling Psychology*, 21 (4), 269–73. doi:10.1037/h0036621.
- 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, EL471–EL477.
- 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, 331–54.
- Roberts, S. G., Torreira, F., and Levinson, S. C. (2015). The effects of processing and sequence organization on the timing of turn taking: a corpus study. *Frontiers in Psychology,* 6, 509. doi: 10.3389/fpsyg.2015.00509
- Roessel, J., Schoel, C. & Stahlberg, D. (2020). Modern notions of accent-ism: Findings, conceptualizations, and implications for interventions and research on nonnative accents. *Journal of Language and Social Psychology*, 39, 87–111.
- Sacks, H. (1992). Lectures on Conversations. Oxford: Blackwell, vols. I and II.
- Sacks, H. & Schegloff, E. (1973). Opening up Closings. Semiotica, 7, 289–327.
- Schwartz, S. (2008). Cultural value orientations: Nature & implications of national differences. Psychology. *Journal of Higher School of Economics*, 5(2), 37-67.

- Schielzeth, H. & Forstmeier, W. (2009). Conclusions beyond support: Overconfident estimates in Mixed Models. *Behavioral Ecology*, 20, 416–20.
- Shigemitsu, Y. (2005). Different interpretations of pauses in natural conversation--Japanese, Chinese and Americans. 東京工芸大学工学部紀要. 人文·社会編, 28(2), 8-14.
- Stivers, T. (2010). An overview of the question–response system in American English conversation. *Journal of Pragmatics*, 42, 2772–81.
- Stivers, T, Enfield, N. J., Brown, P., Englert, E., Hayashi, M., Heinemann, T., Hoymann, G. Rossano, F., de Ruiter, J. P., Yoon, K.-E., et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106, 10587–92.
- Suzuki, S., Kormos, J. & Uchihara, T. (2021). The relationship between utterance and perceived fluency: A meta-analysis of correlational studies. *The Modern Language Journal*, 105, 435–63.
- Tannen, D. (1985). Silence: Anything but. In D. Tannen and M. Saville-Troike (Eds.)

 *Perspectives on Silence (pp. 93–11). Ablex.
- Tavakoli, P. (2011). Pausing patterns: Differences between L2 learners and native speakers. *ELT Journal*, 65, 71–79.
- Timming, A. R. (2017). The effect of foreign accent on employability: A study of the aural dimensions of aesthetic labour in customer-facing and non-customer-facing jobs. Work, *Employment and Society*, 31, 409–28.
- Trouvain, J., Fauth, C. & Möbius, B. (2016). Breath and Non-breath Pauses in Fluent and Disfluent Phases of German and French L1 and L2 Read Speech. Speech Prosody (SP8), 31, 31–35.
- Tsalikis, J., DeShields O. W. Jr., & LaTour, M. S. (1991). The role of accent on the credibility and effectiveness of the salesperson. *Journal of Personal Selling and Sales Management*, 11, 31–41.

- van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, 9(4), 780–789. https://doi.org/10.3758/BF03196335
- Van Os, M., De Jong, N. H., & Bosker, H. R. (2020). Fluency in dialogue: Turn-taking behavior shapes perceived fluency in native and nonnative speech. *Language Learning*, 70, 1183–217.
- Verga, L., Kotz, S. A., & Ravignani, A. (2023). The Evolution of Social Timing. *Physics of Life Reviews*, 46, 131–51. doi:10.1016/j.plrev.2023.06.006.
- Verouden, N.W. and van der Sanden, M.C.A. (2018) Silence in intercultural collaboration:
 - A Sino-Dutch research centre. Advances in Applied Sociology, 8, 125–151.
- Wardęga, J. (2017). Chińczycy w Polsce–niewidoczna mniejszość. *Studia Migracyjne-Przegla d Polonijny*, 2, 121–41.
- Weilhammer, K. & Rabold, S. (2003). Durational aspects in turn taking. *Proceedings of the international conference of phonetic sciences*, 2145-2148.
- Wootton, A. J. (1981). The Management of Grantings and Rejections by Parents in Request Sequences. *Semiotica*, 37, 59–90.
- Yamada, H. (1997). Different Games, Different Rules: Why American and Japanese

 Misunderstand Each Other. Oxford University Press.
- Żywiczyński, P., Gontier, N. & Wacewicz, S. (2017). The Evolution of (Proto-)Language:

 Focus on Mechanisms. *Language Sciences*, 63 (September), 1–11.

 doi:10.1016/j.langsci.2017.06.004.

8. Appendix

8.1. Knowledge questions and answers

Table A1. Knowledge questions and answers that were used as stimuli for testing perceived knowledge and confidence of the respondents. The mini-conversations are presented in English translations and the original Polish and Chinese versions that were used in the study.

1 Q: What was invented before: the cigarette lighter or matches?

A: I think that the cigarette lighter was invented before.

P: Co wymyślono wcześniej: zapalniczkę czy zapałki?

O: Myślę, że zapalniczka została wynaleziona wcześniej.

问: 打火机还是火柴是先被发明的?

答:我认为打火机。

Q: What's faster: a sneeze traveling out of your mouth or a leopard running at full speed?

A: I think a sneeze is faster than a leopard.

P: Co jest szybsze: kichnięcie przez usta czy lampart biegnący z pełną prędkością?

O: Myślę, że kichnięcie jest szybsze niż lampart.

问:喷嚏打出口的速度和豹子全速跑的速度哪个更快?

答:我认为喷嚏打出口的速度更快。

3	Q: For how much money was the most expensive coin in the world sold? For 1.5
	million dollars or for 7 million dollars.
	A: I think it was sold for 7 million dollars.
	P: Za ile pieniędzy sprzedano najdroższą monetę na świecie? Za półtora miliona
	dolarów czy za siedem milionów dolarów?
	O: Myślę, że najdroższą monetę sprzedano za siedem milionów dolarów.
	问:世界上成交最贵的硬币多少钱?一百五十万还是七百万?
	答:我认为是七百万那个。
4	Q: Which of the two animals can't stick out their tongues: crocodiles or dolphins?
	A: I think crocodiles can't stick out their tongues.
	P: Które z tych dwóch zwierząt nie potrafi wystawić języka: krokodyl czy delfin?
	O: Myślę, że krokodyl nie potrafi wystawić języka.
	│ │问:哪种动物的舌头不能伸出来:鳄鱼还是海豚?
	 答∶我觉得是海豚。
5	Q: Which animal can't jump unless its tail is touching the ground: a kangaroo or a
	cat?
	A: I think it's a kangaroo.
	P: Które z tych dwóch zwierząt potrafi skoczyć jedynie, kiedy jego ogon dotyka ziemi:
	kangur czy kot?
	O: Myślę, że kangur.
	│ │问:哪种动物只能需要尾巴接触地面才能跳起来:袋鼠还是猫?
	│ 答 : 我觉得是袋鼠。
6	Q: For how long can polar bears swim without taking a rest: for about 30 km or 100
	km?

A: I think polar bears can swim for 100 km. P: Jak daleko niedźwiedzie polarne mogą pływać bez odpoczynku: około trzydzieści kilometrów czy sto kilometrów? O: Myślę, że niedźwiedzie polarne potrafią przepłynąć 100 km bez odpoczynku. 问:北极熊如果不休息最长可以游多远:大约30公里还是100公里? 答: 我觉得北极熊能游大约100公里。 Q: How high can a flea jump: 120 times or 350 times its body length? A: I think a flea can jump 350 times its body length. P: Jak wysoko może skoczyć pchła: na stu dwudziesto-krotność lub trzystu pięćdziesięciu-krotność długości swojego ciała? O: Myślę, że pchła może skoczyć na trzystu pięćdziesięciu-krotność długości swojego ciała. 问: 跳蚤能跳多高: 身长的120倍还是350倍? 答: 我觉得跳蚤能跳它身长350倍高。 Q: Which animals have striped skin: zebras or tigers? A: I think tigers have striped skin. P: Które z tych dwóch zwierząt ma prążkowaną skórę: zebry czy tygrysy? O: Myślę, że tygrysy mają prążkowaną skórę. 问:哪种动物自身的皮有条纹:斑马还是老虎? 答:我觉得是老虎。 Q: What is bigger: an ostrich's eye or its brain?

7

8

A: I think it's the ostrich's eye.

P: Co jest większe: oko strusia czy jego mózg?

O: Myślę, że oko strusia jest większe.

问:鸵鸟的眼睛和大脑哪个更大?

答:我觉得鸵鸟的眼睛更大。

10 Q: How many eyelids do camels have to protect them from the sand: one or three?

A: I think they have 3 eyelids.

P: lle powiek mają wielbłądy, zeby chronic sie przed piaskiem: jedną czy trzy?

O: Myślę, że mają trzy powieki.

问: 骆驼有几层眼皮用来保护它们的眼睛不进沙子: 一层还是三层?

答:我认为有三层。

11 Q: Which planet has no seasons because it does not tilt as it goes around the Sun:

Venus or Mars?

A: I think Venus doesn't have seasons.

P. Na której z tych dwóch planet nie ma pór roku, ponieważ nie przechyla się ona krążąc wokół Słońca: na Wenus czy na Marsie?

O: Myślę, że na Wenus nie ma pór roku.

问:哪个星球因为它绕太阳公转没有斜角而没有四季:金星还是火星?

答:我觉得是金星。

12 | Q: Which animal can go longer without water: a camel or a giraffe?

A: I think a giraffe can go longer.

P: Które z tych dwóch zwierząt może dłużej wytrzymać bez wody: wielbłąd czy żyrafa?

O: Myślę, że żyrafa może dłużej wytrzymać.

问:哪种动物不喝水能走更久:骆驼还是长颈鹿。

答:我觉得长颈鹿能走更久。

13 Q: Which animal's heart can pump more blood in one minute: a giraffe's or an elephant's? A: I think it's the heart of the giraffe. P: Serce, którego z tych zwierząt może przepompować więcej krwi w ciągu jednej minuty: żyrafy czy słonia? O: Myślę, że serce żyrafy. 问:哪种动物的心脏每分钟可以供血更多:长颈鹿还是大象? 答:我觉得是长颈鹿。 14 Q: Which animal has the hardest teeth: a sea snail or a beaver? A: I think it's the teeth of the sea snail. Q: Jakie zwierze ma twardsze zęby: ślimak morski czy bóbr? A: Myślę, że ślimak morski. 问:哪种动物牙齿最坚硬:海螺还是海狸? 答:我觉得是海螺。 Q: How many percent of the total mass of the solar system does the sun have? - 93.5 15 percent or 99.8 percent? A: I think the sun has 99.8 percent of the mass. Q: lle procent całej masy Układu Słonecznego to masa Słońca? - dziewięćdziesiąt trzy i pięć dziesiątych procenta czy dziewięćdziesiąt dziewięć i osiem dziesiątych procenta? A: Myślę, że masa Słońca to dziewięćdziesiąt dziewięć i osiem dziesiątych procenta. 问:太阳的质量占整个太阳系的百分之多少?93.5%还是99.8%? 答:我觉得是99.8%。 Q: Does the sun shrink or expand five feet every hour?

A: I think the Sun shrinks five feet every hour.

P: Czy słońce kurczy się czy się rozszerza o półtora metra co godzinę?

O: Myślę, że Słońce kurczy się o półtora metra co godzinę.

问:太阳每小时大约缩小还是膨胀1.5米?

答:我觉得是缩小大约1.5米。

17 | Q: Which are the only animals that are born with horns: giraffes or chameleons?

A: I think it's giraffes.

P: Które z tych dwóch zwierząt rodzi się z rogami: żyrafa czy kameleon?

O: Myślę, że żyrafa.

问:哪种是唯一一种天生有犄角的动物:长颈鹿还是变色龙?

答:我觉得是长颈鹿。

18 Q: The two hemispheres of a dolphin's brain work independently in alternating "shifts". After which time do they change: after 30 minutes or after 8 hours?

A: I think it's after 8 hours.

P: Dwie półkule mózgu delfina pracują niezależnie w naprzemiennych "rundach". Po jakim czasie następują te rundy: po trzydziestu minutach czy po ośmiu godzinach?

O: Myślę, że po ośmiu godzinach.

问:海豚的左右脑独立工作且换班:那么它们多久会换班一次呢?30分钟还是8小时?

答:我觉得是8小时。

19 Q: When was the can opener invented: half a year or 48 years after the invention of the can?

A: I think it was after 48 years.

P: Kiedy wynaleziono otwieracz do puszek: pół roku czy czterdzieści osiem lat po wynalezieniu puszki?

O: Myślę, że po czterdziestu ośmiu latach.

问:开罐器是什么时候发明的:在罐头发明之后的半年还是48年?

答: 我觉得是在罐头发明后48年。

20 | Q: How long was the longest wedding veil: longer than 21 or than 63 football fields?

A: I think the longest wedding veil was longer than 63 football fields.

P: Jak długi był najdłuższy welon ślubny: dłuższy niż dwadzieścia jeden boisk piłkarskich czy sześćdziesiąt trzy boiska piłkarskie?

O: Myślę, że najdłuższy welon był dłuższy niż sześćdziesiąt trzy boiska piłkarskie.

问:最长的婚纱有多长:长于21个足球场还是63个足球场?

答: 我觉得比63个足球场长。

21 Q: How many different viruses does the ocean contain: almost 200,000 or almost one million?

A: I think the oceans contain almost 200,000 different kinds of viruses.

P: Ile różnych rodzajów wirusów znajduje się w oceanach: prawie dwieście tysięcy czy prawie milion?

O: Myślę, że oceany zawierają prawie dwieście tysięcy różnych rodzajów wirusów.

问:海洋存在多少不同的病毒:将近20万种还是将近100万种?

答:我觉得将近20万种。

22 | Q: How many different dog breeds are there: more than 200 or more than 600?

A: I think that there are around 200 different breeds of dogs.

P: lle jest ras psów: ponad dwieście czy ponad sześćset?

O: Myślę, że istnieje około dwustu ras psów.

问:一共有多少种狗:200多种还是600多种?

答:我觉得一共有200多种。

23	Q: Where was the first public library in the world: in Vienna or in Peterborough?
	A: I think it was in Peterborough.
	P: Gdzie otwarto pierwszą bibliotekę publiczną na świecie: w Wiedniu czy w
	Peterborough?
	O: Myślę, że w Peterborough.
	问:世界上第一所图书馆在哪里:维也纳还是彼得伯勒?
	答:我觉得是彼得伯勒。
24	Q: How many babies are globally born every second: around 5 or around 15?
	A: I think it's around 5 babies.
	P: lle dzieci rodzi się na świecie co sekundę: około pięciu czy około piętnastu?
	O: Myślę, że około pięciu dzieci.
	问:全球每秒有多少新生儿:大约5个还是15个?
	答:我觉得大约5个。
25	Q: Which of the two planets spins in the opposite direction of the earth: Saturn or
	Uranus?
	A: I think it's Uranus.
	P: Która z tych dwóch planet obraca się w przeciwnym kierunku do Ziemi: Saturn czy
	Uran?
	O: Myślę, że Uran.
	问:哪个星球和地球的自转方向相反:土星还是天王星?
	答:我觉得是天王星。
26	Q: How many grapes go into one bottle of wine: around 500 or around 1000?
	A: I think it's around 500.

P: Ile winogron potrzeba, żeby uzyskać jedną butelkę wina: około pięciuset czy około tysiąca? O: Myślę, że około pięciuset. 问:一瓶葡萄酒有多少葡萄酿成:大约500个还是1000个? 答:我觉得大约500个。 27 Q: Where is the original designer behind the game "Tetris" from: from Norway or from the Soviet Union? A: I think the original designer behind the game "Tetris" is from Norway. P: Skąd pochodzi twórca gry "Tetris": z Norwegii czy ze Związku Radzieckiego? O: Myślę, że twórca gry "Tetris" pochodzi z Norwegii. 问:俄罗斯方块是原始设计来源于哪个国家:挪威还是苏联? 答:我觉得是挪威。 28 Q: How many feathers are on an eagle: approximately 7,000 or 19,000 feathers? A: I think that there are approximately 7,000 feathers on an eagle. P: lle piór ma orzeł: około siedmiu tysięcy czy dziewiętnastu tysięcy? O: Myślę, że orzeł ma około siedmiu tysięcy piór. 问:一只鹰有多少根羽毛:接近7千根或者1万9千根? 答:我觉得接近7千根。 Q: What kind of light are most birds' eyes most sensitive to: visible light or ultraviolet 29 light? A: I think it's ultraviolet light. P: Na jaki rodzaj światła najbardziej wrażliwe są oczy większości ptaków: na światło widzialne czy światło ultrafioletowe? O: Myślę, że na światło ultrafioletowe.

问: 鸟对哪种光最敏感: 可见光还是紫外线? 答:我觉得是紫外线。 Q: The skin of which sex of mature blue sharks is more than twice as thick as the 30 skin of the opposit sex: males or females? A: I think it's the skin of females. P: Skóra której płci dojrzałych rekinów błękitnych jest ponad dwukrotnie grubsza od skóry płci przeciwnej: samców czy samic? O: Myślę, że skóra samic. 问:成年大青鲨的皮肤是雌性比雄性厚两倍多还是雄性比雌性厚两倍多? 答:我认为是雌性。 31 Q: How long can large crocodiles survive without food: for over a week or for over a year? A: I think large crocodiles can live without food for a year. P: Jak długo duże krokodyle mogą przetrwać bez jedzenia: ponad tydzień czy ponad rok? O: Myślę, że duże krokodyle mogą przetrwać bez jedzenia przez rok. 问:大型鳄鱼多久不进食还可以活下来:一周还是一年? 答:我觉得是一年。 Q: What were the first vegetables to be grown in space: potatoes or lettuce? 32 A: I think was potatoes. were the first vegetables to be grown in space. P: Jakie były pierwsze warzywa wyhodowane w kosmosie: ziemniaki czy sałata? A: Myślę, że ziemniaki były pierwszymi warzywami wyhodowanymi w kosmosie. 问:宇宙种植的第一种蔬菜是什么:土豆还是生菜? 答:我觉得是土豆。

8.2. Requests

Table A2. Requests and answers that were used as stimuli for testing perceived willingness of the respondents to grant the requests. The mini-conversations are presented in English translations and the original Polish and Chinese versions that were used in the study. Requests which have been adapted from a Polish to a Chinese cultural setting are marked with an asterisk and also feature the English translation of the adapted Chinese utterances.

1* R: I start getting really tired. Could you make me black coffee?

A: Okay, I can make you black coffee.

P: Zaczynam czuć zmęczenie. Czy możesz mi zrobić czarną kawę?

O: Dobrze, mogę zrobić Ci czarną kawę.

R: I start getting really thirsty. Could you boil some water for me?

A: Okay, I can boil some water.

问: 我好渴啊, 你能帮我烧点热水吗?

答:好啊,我给你烧

2 R: I need a rest. Can you move your bags so that I can sit here?

A: Okay, I can move my bags over there.

P: Muszę odsapnąć. Czy możesz wziąć te siatki, żebym mógł tutaj usiąść?

O. Dobrze, mogę zabrać stąd swoje siatki.

问:我需要坐这里休息一下,你能把你放在这里的包拿到别处吗?

答:好啊,我把包拿走。

3	R: I can't find my way out of the building. Can you walk me to the exit?
	A: Okay, I can show you where the exit is.
	P: Nie wiem, jak wyjść z budynku. Czy możesz odprowadzić mnie do wyjścia?
	O: Dobrze, mogę Cię odprowadzić do wyjścia.
	问:我找不到出口了,你能带我走出去吗?
	答:好啊,我可以告诉你出口在哪。
4	R: I have some more questions that I'd like to ask her. Can you give me her phone
4	number?
	number?
	A: Okay, I can look for her number and give it to you.
	P: Jest jeszcze kilka rzeczy, o które chciałbym ją spytać. Czy możesz mi dać jej
	numer?
	O: Dobrze, mogę poszukać jej numeru i Ci go dać.
	问:有些问题想问她一下, 你能告诉我她的电话号码吗?
	答:好啊,我找一下给你。
5	R: It's really hot in here. Can you open the window?
	A: Okay, I can open the window to get some fresh air.
	P: Tutaj jest bardzo gorąco. Czy możesz otworzyć okno?
	O: Dobrze, mogę wpuścić trochę świeżego powietrza.
	问:这儿太热了, 你能开窗户吗?
	答:好啊,我打开窗户透透气。
6	R: I'm really hungry. Can you make me pancakes?
	A: Okay, I can make pancakes for you.
	P: Jestem naprawdę głodny. Czy możesz zrobić mi naleśniki?
	O: Dobrze, mogę zrobić Ci naleśniki.

R: I'm really hungry. Can you make me some noodles? A: Okay, I can make some noodles for you. 问: 我好饿啊, 你能帮我做点吃的吗? 答:好啊,我给你下点面条。 7 R: I don't have enough money to buy our tickets. Can you pay me in advance? A: Okay, I can withdraw the money to pay you in advance. P: Nie starczy mi pieniędzy, żeby zapłacić za nasze bilety. Czy możesz mi zapłacić z góry? O: Dobrze, mogę wypłacić pieniądze i zapłacić Ci z góry. 问:我钱不够买咱们的票了,你能把你那部分钱先给我吗? 答:好啊,我把我那部分先给你。 R: It's Dad's birthday next week. Can you get something for him? A: Okay, I can get some flowers for him. P: Urodziny taty są w przyszłym tygodniu. Czy możesz coś kupić na prezent dla niego? O: Dobrze, mogę kupić mu kwiaty. 问:下周就是咱爸的生日了,你能帮我给他买点啥吗? 答:好啊. 我给你买束鲜花。 9 R: This is so much work. Can you come over and help me with these papers? A: Okay, I can stop by and help. P: Mam naprawdę dużo do zrobienia. Czy możesz wpaść i pomóc z tymi papierami? O: Dobrze, mogę wpaść i Ci pomóc. 问:这工作量好大啊,你能过来帮我处理这些文件吗?

	答:好啊,我这就过去帮你。
10	R: I would like to drop him off at his parents' house. Can you check the address for
	me?
	A: Okay, I can check and jot it down for you.
	P: Chciałbym go podrzucić do domu jego rodziców. Czy możesz sprawdzić jaki to
	adres?
	O: Dobrze, mogę sprawdzić i Ci go zapisać.
	问:我想把他送到他父母家, 你能帮我查一下他父母家的地址吗?
	答:好啊,我查完写下来给你。
11	R: I haven't had lunch yet. Can you get me something to eat?
	A: Okay, I can get you something from the cafeteria.
	P: Nie jadłem jeszcze lunchu. Czy możesz kupić coś do jedzenia?
	O: Dobrze, mogę przynieść Ci coś z baru.
	│ │问:我还没吃午饭, 你能帮我买点吃的吗?
	答:好啊,我给你去食堂买点。
12	R: The oven is already preheated. Can you help me peel the potatoes?
*	A: Okay, I can help you peel them.
	P: Piekarnik jest rozgrzany. Czy możesz mi pomóc obrać ziemniaki?
	O: Dobrze, mogę pomóc Ci je obrać.
	R: The stewed beef is almost ready. Can you help me peel the potatoes?
	A: Okay, I can help you peel them.
	问:牛肉马上炖好了,你能帮我削下土豆吗?

13 R: This is so complicated. Can you help me fill in this tax report? A: Okay, I can have a look and help you. P: To jest naprawdę skomplikowane. Czy możesz mi pomóc wypełnić to zeznanie podatkowe? O: Dobrze, moge na nie spojrzeć i ci pomóc. 问:这也太复杂了吧,你能帮我填下纳税报表吗? 答:好啊,我看看帮你填一下。 14 R: This is so much work. Can you help me with a few more things? A: Okay, I can help you with some of your tasks. P: Mam naprawdę dużo pracy. Czy możesz pomóc mi z kilkoma rzeczami? O: Dobrze, mogę Ci trochę pomóc. 问:这工作量好大啊,你能再帮我完成一些吗? 答:好啊. 我能帮你完成一些。 R: We ran out of sugar. Can you get some so that we can make the cake? A: Okay, I can go to the supermarket to get it. P: Nie ma cukru. Czy możesz go kupić, żebyśmy mogli zrobić ciasto? O: Dobrze, mogę pójść do sklepu po cukier. 问:我们的糖用完啦,我们得做蛋糕,你能去买点吗? 答:好啊. 我去超市买一袋。 R: We need one more person to be there early. Can you come at ten? A: Okay, I can already come at ten. P: Jeszcze jedna osoba musi pojawić się tam wcześnie. Czy możesz przyjść na dziesiątą? O: Dobrze, mogę przyjść już na dziesiątą.

问:我们还需要一个人早到,你能10点来吗? 答:好啊,我十点能到。 R: I'm not sure who will be there at the meeting tomorrow. Can you ask people and 17 find out? A: Okay, I can send out an email. P: Nie wiem kto będzie jutro na spotkaniu. Czy możesz dowiedzieć się kto się wybiera? O: Dobrze, mogę wysłać maila do ludzi. 问:我不确定明天开会谁会来,你能帮忙问问吗? 答:好啊,我给他们发邮件问问。 18 R: I just called the copy shop and the flyers are ready. Can you give me a ride over there? A: Okay, I can give you a ride. P: Właśnie rozmawiałem z punktem ksero i ulotki są już gotowe. Czy możesz mnie tam podwieźć? O: Dobrze, mogę Cię podwieźć. 问:我刚才给打印店打电话,那边说传单已经印好了,你能载我过去取吗? 答:好啊. 我能载你过去。 R: I wanna go to my yoga class but I cannot find my yoga mat. Can I borrow yours? 19 A: Okay, you can borrow my yoga mat. P: Ide na joge i nie moge znaleźć maty. Czy możesz pożyczyć Twoją?

O: Dobrze, mogę pożyczyć ci moją matę.

答:好啊. 我能借你。

问:我要去上瑜伽课,可我找不到垫子了,能把你的借我吗?

20	R: My new computer is ready for pick-up. Can you come with me and help me carry
	it?
	A: Okay, I can come with you.
	P: Mój nowy komputer można już odebrać. Czy możesz pójść ze mną i pomóc mi go
	przenieść?
	O. Dobrze, mogę Ci pomóc z komputerem.
	问:我的新电脑到店里了,可以去取了,你能陪我一起去帮我拎一下吗?
	答:好啊,我能陪你一起去。
21	R: I forgot to bring cash. Can you lend me some money?
	A: Okay, I can lend you 100 zloty.
	P: Zapomniałem pieniędzy. Czy możesz mi trochę pożyczyć?
	O: Dobrze, mogę pożyczyć Ci 100 złotych.
	│ │问:我忘记带现金了, 你能借我点吗?
	 答:好啊, 我能借你100块。
22	R: Our bathroom is so dirty and we will have guests tomorrow. Can you clean it?
	A: Okay, I can clean it before they come.
	P: Nasza łazienka jest naprawdę brudna, a jutro mamy gości. Czy możesz ją
	posprzątać?
	O: Dobrze, mogę posprzątać łazienkę.
	问:我们的卫生间好脏啊,明天来客人,你能打扫一下吗?
	答:好啊,他们来之前我能打扫好。
23	R: I don't know how to assemble this shelf. Can you help me do it?
	A: Okay, I can help you with the shelf.
	P: Nie wiem jak złożyć tę półkę. Czy możesz mi pomóc?

		O: Dobrze, mogę pomóc ci z tą półką.
		问:我不知道这个架子怎么组装,你能我帮我装一下吗?
		答:好啊,我帮你装一下。
ŀ	24	R: I don't have any tools to fix it. Can you lend me your hammer?
		A: Okay, you can have my hammer.
		P: Nie mam żadnych narzędzi, żeby to naprawić. Czy możesz pożyczyć mi młotek?
		O: Dobrze, mogę pożyczyć Ci młotek.
		问:我没有修理的工具,你能借我个锤子吗?
		答:好啊,我把我的借给你。
ľ	25	R: Mum needs some help in the garden. Can you mow the lawn?
	*	A: Okay, I can cut the grass.
		P: Trzeba pomóc mamie w ogrodzie. Czy możesz skosić trawnik?
		O: Dobrze, mogę skosić trawę.
		R: Mum needs some help in the kitchen. Can you help her cut the vegetables?
		A: Okay, I can cut the vegetables.
		问: 妈在厨房需要帮忙, 你能去帮她切切菜吗?
		答:好啊,我这就去。
	26	R: The washing machine is done. Can you unload it?
		A: Okay, I can hang the clothes.
		P: Pralka skończyła. Czy możesz ją wyładować?
		O: Dobrze, mogę rozwiesić pranie.
		问:洗衣机的衣服洗好啦, 你能把衣服拿出来吗?
		答:好啊,我把衣服晾了。
-1		

27 R: We will make some important decisions. Can you take notes during the meeting? A: Okay, I can take notes of the most important things. P: Musimy podjąć kilka ważnych decyzji. Czy możesz robić notatki podczas spotkania. O: Dobrze, mogę notować najważniejsze rzeczy. 问:我们会做出些比较重要的决定,你能做会议记录吗? 答:好啊,我把重要的部分记下来。 28 R: I don't know if our guests will find their way. Can you await them at the gate? A: Okay, I will wait for them and show them their way. P: Nie wiem czy nasi goście tutaj trafią. Czy możesz poczekać na nich przy bramie? O: Dobrze, poczekam na nich i pokaże im drogę. 问:我不确定要来的客人能找到我们的位置,你能去门口迎接一下他们吗? 答:好啊. 我去门口接一下他们。 R: One of the light bulbs is broken. Can you change it? A: Okay, I can change the light bulb. P: Jedna żarówka się przepaliła. Czy możesz ja zmienić? O: Dobrze, mogę zmienić tę żarówkę. 问:有个灯泡不亮了,你能换一下吗? 答:好啊,我去换一下。 30 R: I still need to finish some things. Can you walk the dog alone today? A: Okay, I can go for a walk with him. P: Mam jeszcze kilka rzeczy do zrobienia. Czy możesz dziś wyprowadzić psa? O: Dobrze, mogę z nim pójść na spacer. 问:我还有事情没忙完,你今天能自己去遛狗吗?

	答:好啊,我自己去。
31	R: I won't have time to go to the travel agency with you. Can you book the flight for
	both of us?
	A: Okay, I can get our tickets.
	P: Nie będę miał czasu pójść z Tobą do biura podróży. Czy możesz zarezerwować lot
	dla nas obojga?
	O: Dobrze, mogę zarezerwować dla nas bilety.
	│ │问:我今天没有时间和你一起去旅行社了, 你能把咱俩的机票都订了吗?
	│ │答∶好啊﹐ 我把咱俩的机票都订了。 │
32	R: My wifi hotspot doesn't seem to work. Can you check the departure time for me?
	A: Okay, I can check it for you.
	P: Moje wifi nie działa. Czy możesz sprawdzić, kiedy jest odjazd?
	O: Dobrze, mogę to dla Ciebie sprawdzić.
	│ │问:我无线网的热点好像不行了, 你能帮我查一下几点出发吗?
	│ │答∶好啊, 我帮你查一下。 │