*TIESPJ, Vol. 12 No. 2, 2021*

# Gaming to Learn: L2 Learners' Conversational Pitch Convergence during Story-Telling

**Chen-Hsiu Grace Kuo**
National Taiwan University
graciakuo@ntu.edu.tw

## Abstract

Conversational convergence is the tendency for individuals to modify their phonetic speech behaviors to converge with the behaviors of their interlocutors. Evidence of conversational convergence in a native language (L1) is robust, yet research regarding its development in the second language (L2) is sparse. In order to examine the conversational convergence in spoken dialogues, we developed a natural speech corpus in which spontaneous conversations of 15 dyads of English learners were collected and analyzed. The conversational convergence in this paper focuses on the convergence of pitch, which includes the maximum, minimum, and average pitch as well as the pitch range. A personality measure (i.e., autism-spectrum quotient score; AQ) and its correlation with the pitch measures were examined. The results reveal that in this casual board game setting, (a) an English learner tends to use higher f0 max, higher f0 min, higher f0 mean, and a narrower f0 range when speaking in their L2, and most of the pitch values (except f0 max) naturally converge; (b) personality does not directly influence language learners' pitch, but it does influence their speaking rate.

***Keywords*:** pitch convergence, story-telling, autistic-spectrum quotient, L2

DOI: 10.6706/TIESPJ.202112_12(2).0003

## 1. Introduction

Conversational convergence is the tendency for individuals to unconsciously alter their verbal and nonverbal behaviors to align with the behaviors of their conversation partner(s). The empirical evidence of conversational convergence in *nonverbal* behaviors includes gestures, postures, and facial expressions (Fuchs & Reichel, 2016; Lakin et al., 2003; Shockley et al., 2003). The evidence of conversational convergence in *verbal* behaviors can be found at different levels. At the lexical and the syntactic levels, speakers tend to converge with their interlocutor's choice of words (Brennan & Clark, 1996; Nenkova et al., 2008), syntactic structure (Reitter et al., 2006), and linguistic style (Gonzales et al., 2009; Neiderhoffer & Pennebaker, 2002). At the phonetic level, speakers' pronunciation (Pardo, 2013) and accent (Kraljic et al., 2008) become more similar over time, and so do the delicate phonetic behaviors, including vocal intensity (Natale, 1975), speaking rate (Giles et al., 1991), and pitch properties (Babel & Bulatov, 2012; Levitan & Hirschberg, 2011). In addition to the above speech behaviors, conversational convergence tends to positively correlate with relationships in scenarios such as dating (Michalsky & Schoormann, 2017), approval-seeking (Danescu-Niculescu-Mizil & Lee, 2011), and tutoring (Thomason et al., 2013), and with social aspects of conversations such as task success (Nenkova et al., 2008; Reitter & Moore, 2014), rapport (Nubold et al., 2021), perceived attractiveness, and likability (Michalsky & Schoormann, 2017; Schweitzer et al., 2017).

Recently, studies on games and gaming (digital and non-digital) have shown that gameplay can be linked to learning because "participants have to understand the rules to proceed and the system to interact or engage with the games" (Steinkuehler et al., 2012). Among the different types of games, board games are useful for eliciting spontaneous speech in language learning in that they require participants to play face-to-face, to follow a set of rules, and to use playing pieces (e.g., cards, dice, tokens) and playing surfaces. In addition, some board games are designed for the purpose of simplifying some complex issues and encouraging participants to do computational thinking in an informal setting.

In this study, we had our participants engage in a board game that requires high cooperation, and they had to present the output (i.e., telling a coherent story) within a limited time. The participants had to join two sessions – one in their native language (Mandarin) and the other in their second language (L2; English). We investigated the participants' conversational convergence at one of the phonetic levels – more specifically, we examined their performance of pitch change (aka "pitch convergence"). Furthermore, we would also like to see whether their personality differences would influence pitch convergence; therefore, a personality assessment, autism-spectrum quotient (AQ), was calculated for each participant, and the correlations between their personality differences and the pitch measures were examined. The findings of this research are expected to provide a preliminary view of speakers' pitch convergence in an informal but cooperative scenario, and the two main research questions are:

(i) Does the participants' pitch in casual speech converge differently when the languages in use are different?

(ii) Do participants' personality differences affect their performance in conversational pitch convergence?

The paper is organized as follows: Section 2 reviews the studies on pitch convergence. Section 3 presents the methodology. Section 4 discusses the results. Section 5 provides a general discussion. Section 6 concludes this paper.

## 2. Theoretical Background

### 2.1 Pitch convergence measures

Pitch is a salient property in phonetic convergence. Pitch is a highly imitable feature, especially in the shadowing task (Goldinger, 1997) and imitation task (Cole & Shattuck-Hufnagel, 2011). Fundamental frequency accommodation is constantly found in human-to-human interaction in both "back mimicry" (i.e., a speaker uses a pitch contour used previously by the interlocutor) and "forward influence" (i.e., a speaker's pitch contour is later found in the interlocutor's speech) (Gregory et al., 1994; Heldner et al., 2010). A recent study by Gálvez et al. (2020) found that pitch is associated with how users perceive the capability of the avatars.

Most previous pitch convergence studies focused on the imitation effect found in speakers' L1 rather than in their L2. For example, tone language speakers tend to adjust their pitch to their interlocutors in their native language (Levitan et al., 2015; Xia et al., 2014), but the Mandarin Chinese speakers in Ip and Cutler's (2020) study did not adopt such a convergence strategy when the stimuli were in a foreign language (i.e., English in their study). In order to examine pitch convergence in the native language and a second language from a holistic perspective, we controlled most of the social factors (e.g., gender, language proficiency, relationship with the interlocutor) and focused on the changes in participants' pitch change with language as one of the manipulating factors.

### 2.2 Personality measure – autistic traits

The implement for measuring the personality difference in this study is the autism-spectrum questionnaire (Baron-Cohen et al., 2001). The scores obtained from this questionnaire are called the AQ scores, and they reflect the autistic traits from five perspectives – social skill, attention to detail, attention switching, imagination, and communication. Cross-linguistic comparisons demonstrated that an individual with a higher AQ score (i.e., more autistic traits) possessed poorer social skills, paid more attention to details, had more difficulty in attention switching, had poorer communication ability, and lacked imagination. Previous phonetic studies have shown that the autistic traits are associated with variations in speech production (Yu 2013), discourse processing, (Nieuwland et al., 2010; Xiang et al., 2013) and prosodic processing (Bishop & Kuo, 2016; Jun & Bishop, 2015). The aforementioned production and perception effects were closely related to the process of language learning, especially for second language acquisition. Therefore, autistic traits are used as a measure for personality difference in this study.

# 3. Methodology

## 3.1 Participants

Thirty (22 female, eight male) undergraduate students (age: 20 – 25) were recruited. All participants were native speakers of Mandarin and had English as their second language (L2). The participants were all English-major students who could carry on meaningful and logical conversations fluently in English; thus, their English proficiency level was between upper-intermediate and advanced or even above. According to their self-reports, none had identified intellectual, language, or speech impairments. Participants paired up by themselves and signed up for the recordings as a dyad. The gender was not controlled, so some dyads were female-male ($n$ = 6), while others were female-female ($n$ = 8) and male-male ($n$ = 1). The female-male groups were also known as "opposite-gender" in the group gender factor, whereas the female-female and male-male groups were labeled "same-gender."

## 3.2 Procedure

Fifteen spontaneous dyadic conversations were collected. Each group participated in two recording sessions – first in Mandarin (L1) and then in English (L2). Each recording lasted 20 minutes. The participants were told that their goal was to tell a complete story in the target language using the tangible materials provided in the board game at each session. The participants first discussed how to rearrange six randomly selected story cards and make a logical and coherent story based on the arrangement for 15 minutes, and they had to take turns telling the story in the last 5 minutes. There was a timer in front of them, and they were told to keep their conversations going until the time was up. In this study, we only examined their performance when they took turns telling the 5-minute story.

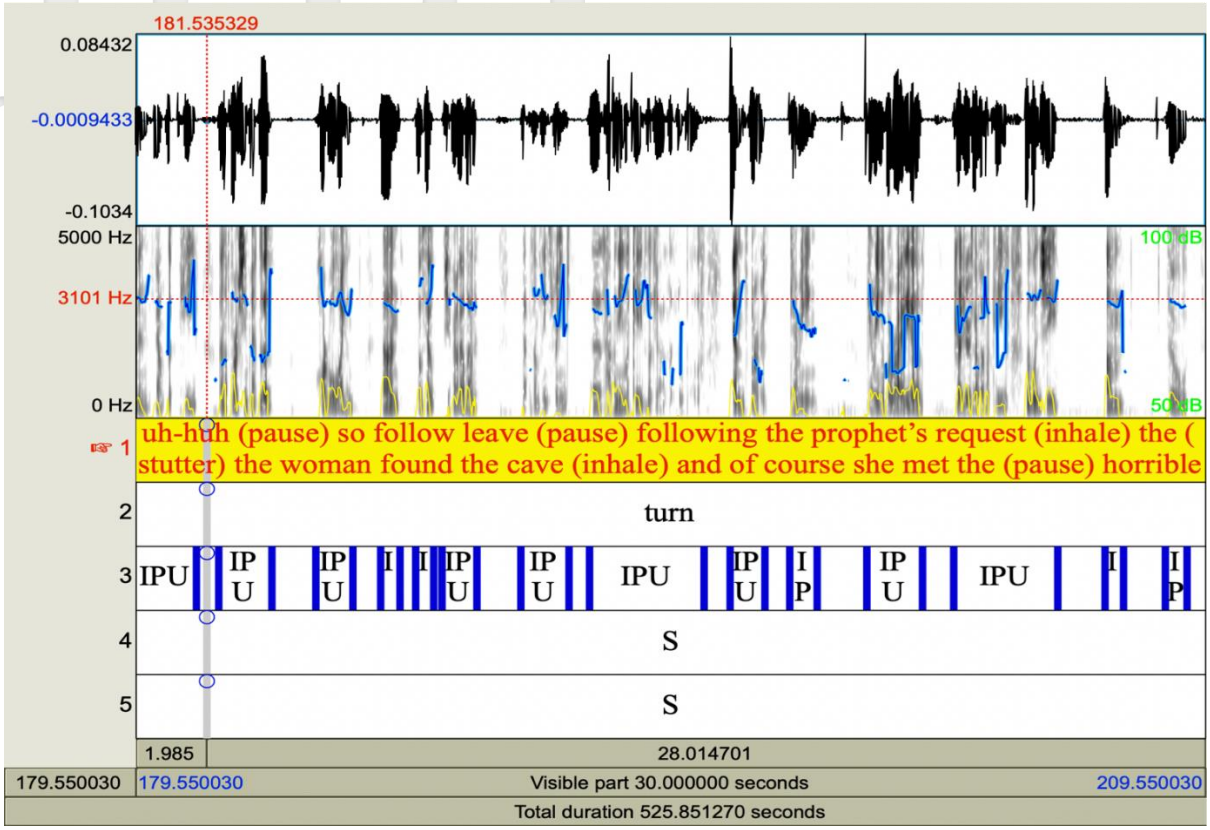
The recordings were made in a quiet room. Each participant had to speak into a unidirectional dynamic microphone (AKG P5) in front of their mouth, and both microphones were connected to a Zoom LiveTrack L-8 digital mixer during the recordings. There were separate audio channels (one channel for each speaker), and the standard settings (48 kHz; 16-bit sampling rate) were employed.

## 3.3 Segmentation and labeling

Three trained research assistants manually coded the 5-minute audio files and annotated the utterances by their turns and IPUS using Praat TextGrid (Boersma & Weenink, 2021). An IPU (inter-pausal unit) is a pause-free unit of speech separated from the others by at least 50 ms (Levitan & Hirschberg, 2011). Maximal sequences of IPUs from one speaker's uninterrupted speech make a "speaker turn." There were five tiers in each TextGrid: Tier 1 provided the transcription of the utterances; Tier 2 labeled the "speaker turns"; Tier 3 labeled the IPUs, and the types of speaker turn were labeled in Tier 4 and Tier 5 by the two different assistants. Figure 1 shows an annotation example.

**Figure 1**

*Annotation example*




*Note*. This 30-second fragment shows a speaker telling a story in English. Tier 1 provides the word-by-word transcription. Tier 2 labels where the turns are; in this example, the turn is longer than 30 seconds. Tier 3 labels the IPUs. Tier 4 and tier 5 show the types of turns, which are not discussed in this paper.

### 3.4 Data Analysis

#### *3.4.1 Pitch convergence measures*

The measurements of the pitch feature included the measures of maximum f0, minimum f0, average f0, and f0 range, and they were obtained at the initial position and the final position of each turn. The consideration of position (i.e., turn-initial and turn-final) is to observe if speakers' pitch would be more alike (i.e., pitch convergence) later in a turn.

#### *3.4.2 Personality measure*

All participants completed the autism-spectrum quotient questionnaire (AQ) before the recordings, and each one received an AQ score denoting the speaker's autistic traits. The total score was 50, and a higher AQ score suggested that the speaker possessed poorer social skills, was more attentive to details and patterns, had a stronger focus of attention, had poorer communication ability, and lacked imagination. Spearman's correlation coefficients were used to examine the relationship between the personality measure (i.e., the AQ score) and the pitch measures (i.e., f0 max, f0 min, f0 mean, and f0 range).

# 4. Results

The four pitch measures – f0 max, f0 min, f0 range, and f0 mean – were first obtained from running a Praat script (Boersma & Weenink, 2021), and the extreme values were manually checked and fixed by three trained research assistants. Then, to make proper comparisons of the pitch values from men and women, the maximum f0, minimum f0, and average f0 values were normalized by gender using z-scores, and the f0 ranges were normalized as semitones. The measures were later entered into a factorial ANOVA with language (2 levels: Mandarin, English) and position (2 levels: turn-initial, turn-final) as the within-subject factors and group gender (2 levels: same-gender, opposite-gender) as the between-subjects factor. If participants did not treat native language and second language differently, no difference would be found in "language." If participants did not have pitch convergence, no difference would be found in "position." If participants used the same pitch to speak regardless of the partner's gender, no difference would be found in "group gender."

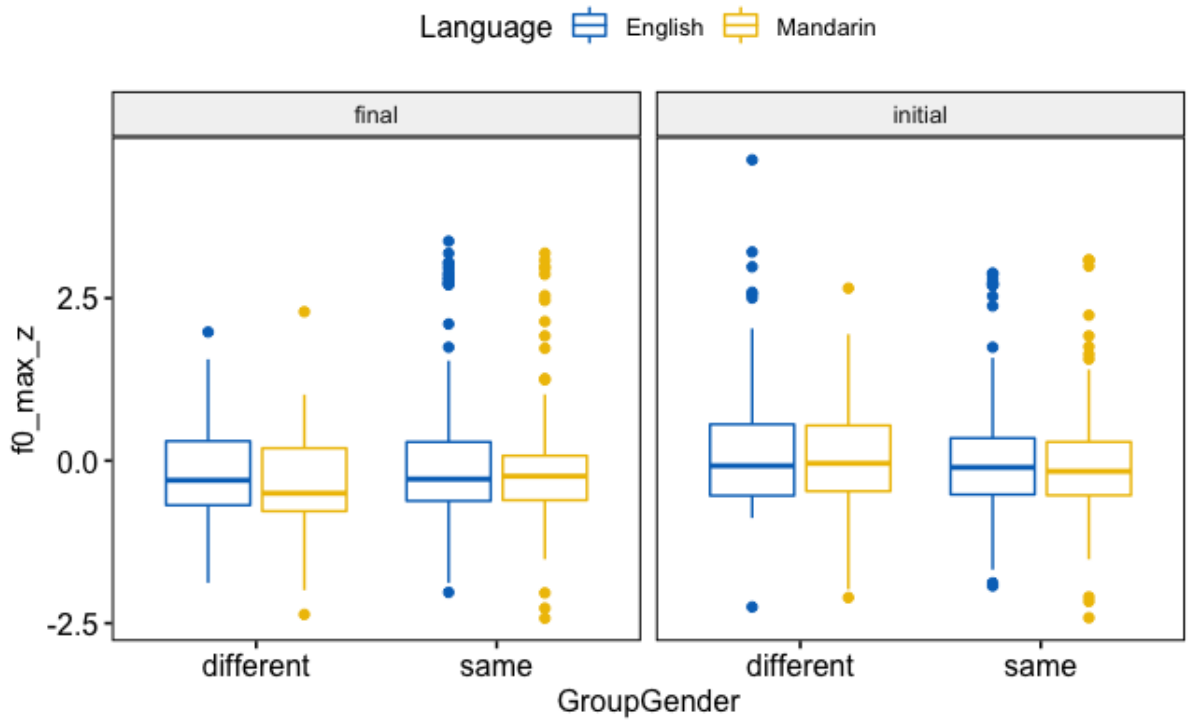
## 4.1 Pitch measures

### *4.1.1 Maximum pitch*

For f0 max, there was a main effect of **language** ($F$ (1, 762) = 3.19, $p < 0.05$) and a statistically significant two-way interaction of **group gender** and **position** ($F$ (1, 762) = 5.97, $p < 0.05$). More specifically, participants tended to use the higher maximum pitch when using a nonnative language. Furthermore, the pairwise comparisons (aka TukeyHSD) showed that the significant interaction mainly came from the f0 max difference at the turn-initial and turn-final position when the interlocutor was of the opposite gender. In other words, speakers demonstrated high pitch at a turn-initial position and low pitch at a turn-final position. However, this f0 max convergence was only observed when a participant was speaking to an opposite-gender partner. Figure 2 demonstrates f0 max displayed by **group gender**, **language**, and **position.**

**Figure 2**

*Maximum pitch (f0 max in z-score) grouped by* ***group gender,*** *colored by* ***language,*** *and faceted by* ***position***

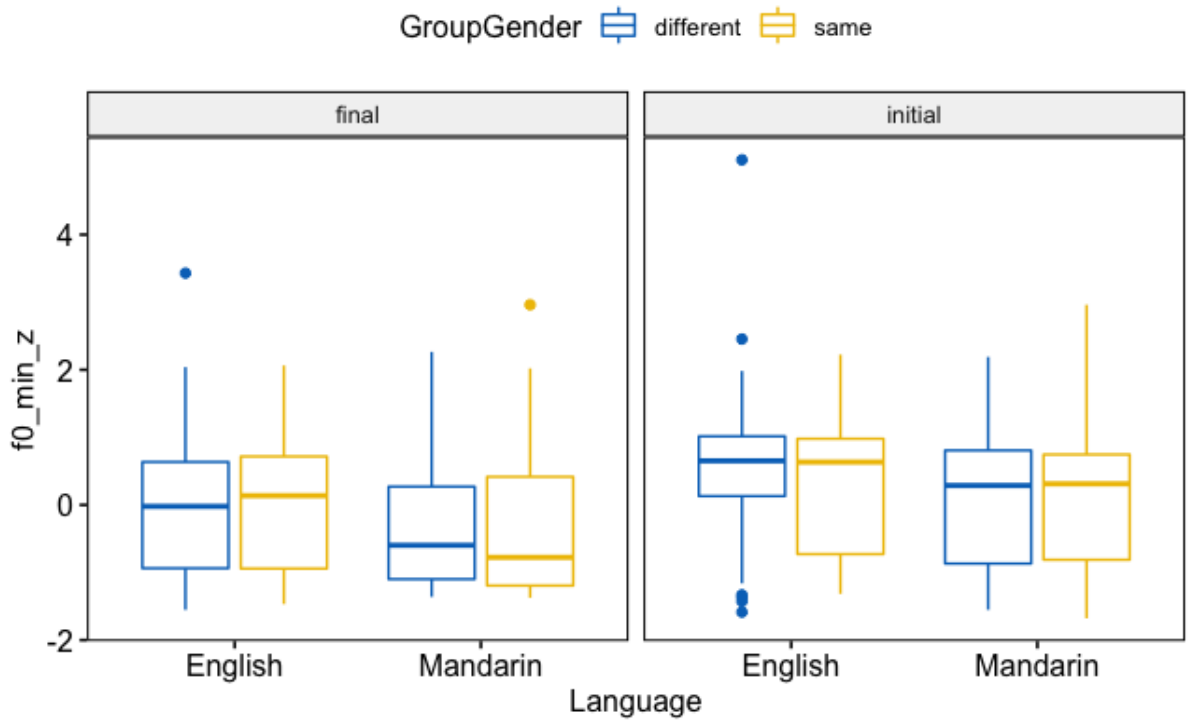



### *4.1.2 Minimum pitch*

For f0 min, main effects were found in language ($F$ (1, 762) = 18.41, $p < 0.05$) and position ($F$ (1, 762) = 34.37, $p < 0.05$), but not in group gender nor any interaction. Speakers tended to use higher f0 min in their second language than in their native language. Most importantly, they demonstrated pitch convergence in f0 min. Figure 3 is the visualization of f0 min displayed by group gender, language, and position.

**Figure 3**

*Minimum pitch (f0 min in z-score) grouped by group gender, colored by language, and faceted by position.*

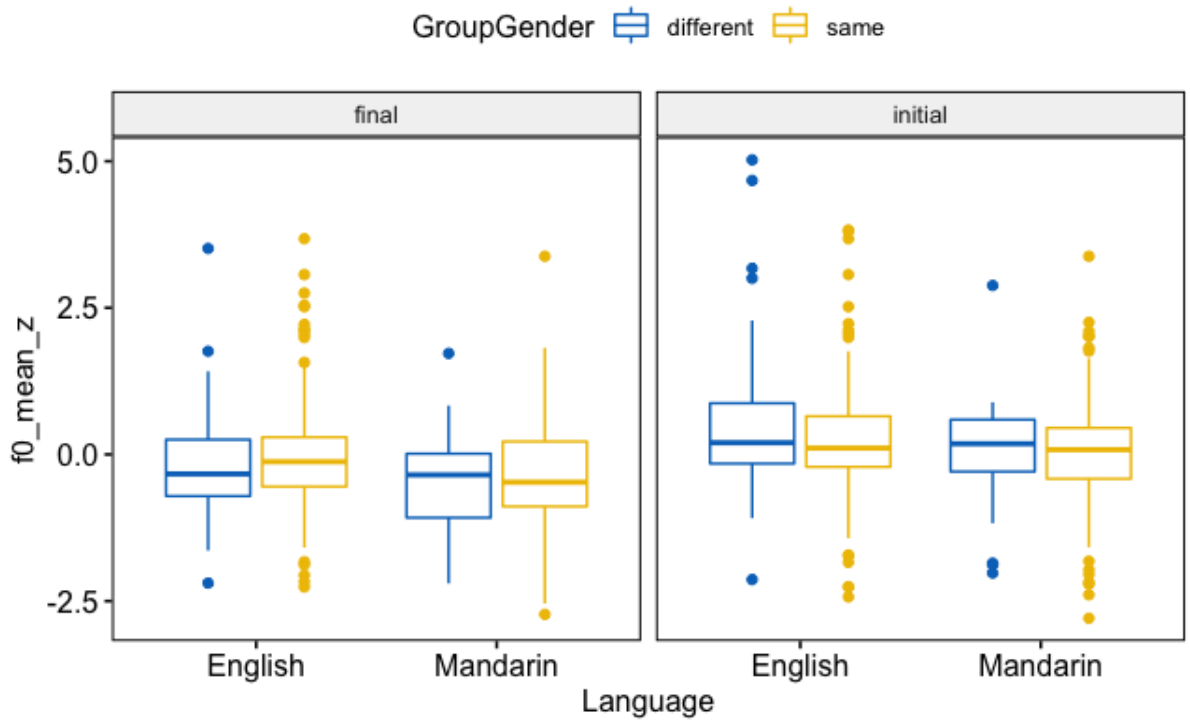



### *4.1.3 Average pitch*

For f0 mean, significant main effects were found in language ($F$ (1, 762) = 19.51, $p < 0.05$) and position ($F$ (1, 762) = 25.04, $p < 0.05$), but not in group gender. The f0 mean in English was significantly higher than that in Mandarin; the f0 mean at the turn-initial position is significantly higher than that at the turn-final position. Figure 4 shows the visualization of f0 mean by group gender, language, and position.

**Figure 4**

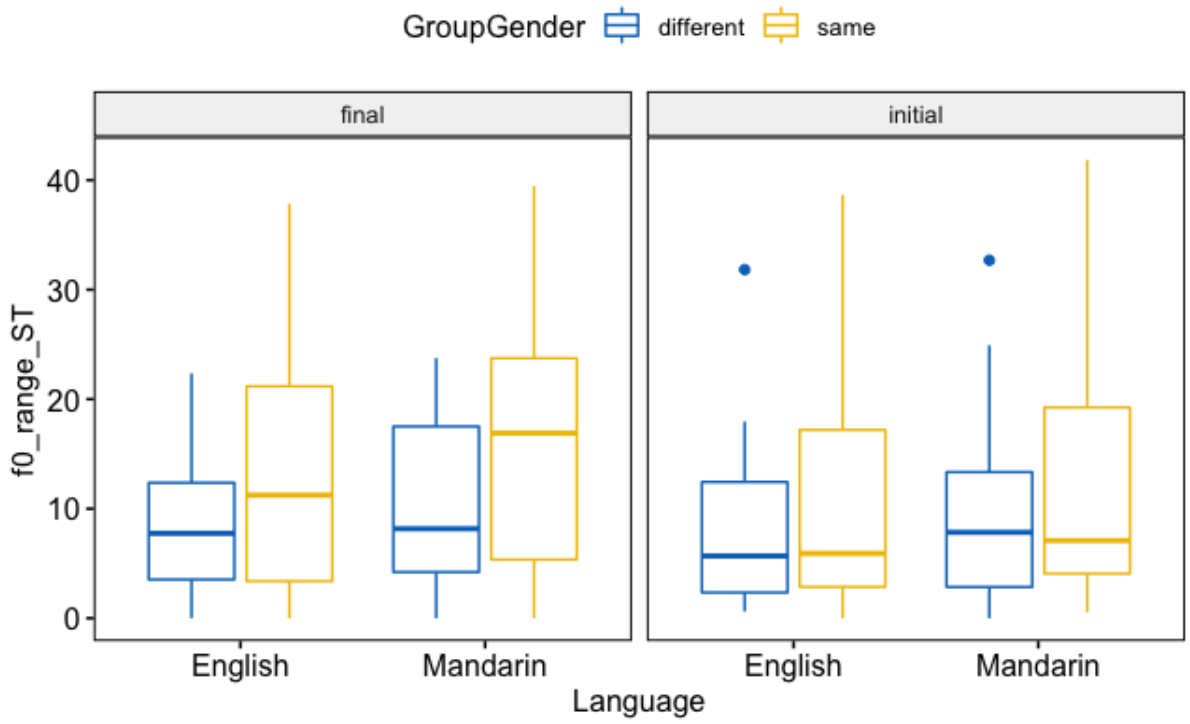
*Average pitch (f0 mean in z-score) grouped by group gender, colored by language, and faceted by position.*




### *4.1.4 Pitch range*

For f0 range, the three factors – **group gender** ($F$ (1, 762) = 17.36, $p < 0.05$), **language** ($F$ (1, 762) = 6.64, $p < 0.05$), and **position** ($F$ (1, 762) = 16.97, $p < 0.05$) – all exhibited significant differences. Speakers tended to use a wider pitch range when interacting with a speaker of the same gender when they were using their native language and when the utterance made was at the end of a speaker turn. Figure 5 is the visualization of f0 range displayed by group gender, language, and position.

**Figure 5**
*Pitch range (f0 range in semitone) grouped by **group gender**, colored by **language**, and faceted by **position**.*




### 4.2 Autistic traits

The personality difference is measured by the total score of each participant's autistic traits (AQ score). Previous studies reported that males tended to have higher AQ scores than females; hard science students scored higher than students in humanities and social science (Baron-Cohen et al., 2001; Stewart & Austin, 2009). Given that our participants were all English-major students, the only factor under examination here was gender (2 levels: male and female). One-way ANOVA was performed to determine whether there was a main effect of gender on autistic traits. The *p*-value is larger than 0.05, which means there was no significant difference in the AQ scores between the male and female participants in this study.

Furthermore, in order to see if autistic traits could predict the patterns of individuals' f0 measures, correlational analyses between the AQ scores and the four pitch measures (i.e., f0 max, f0 min, f0 mean, and f0 range) were implemented in R. The AQ scores and the four pitch measures were tested for their normality with the Shapiro-Wilk test. The results showed that the values of these measures were significantly different from the normal distribution ($p < 0.05$). Therefore, the non-parametric correlation method Spearman correlation coefficient was used. The correlation results between the AQ scores and the phonetic measures from the two recording sessions are presented in Table 1.

**Table 1**

*Spearman's Rho between AQ scores and the four pitch measures*

| **Mandarin** | | **English** | |
|---|---|---|---|
| Speaking Rate | $r$ = 0.2* | Speaking Rate | $r$ = -0.28* |
| f0 max | $r$ = 0.1 | f0 max | $r$ = -0.035 |
| f0 min | $r$ = 0.21* | f0 min | $r$ = 0.077 |
| f0 range | $r$ = -0.085 | f0 range | $r$ = -0.1 |
| f0 mean | $r$ = 0.14 | f0 mean | $r$ = -0.012 |

*Note*. The phonetics measures from the two recording sessions are separated. The significant ones are labeled with an asterisk.

As a comparison, the result of an additional phonetic measure "speaking rate" is reported in the table. The speaking rates were calculated as the number of syllables per second. A two-way mixed ANOVA analysis with language (2 levels: Mandarin and English) as the within-subjects factor and group gender (2 levels: same-gender and opposite-gender) as the between-subjects factors was conducted on speaking rate, and the results revealed a statistically significant main effect of language ($F$ (1, 28) = 170.57, $p < 0.05$), but not of group gender.

The results of the correlational analysis in Table 1 showed that AQ scores mainly correlated with speaking rates (though in different directions in the two recording sessions). And among all the pitch measures, only f0 min had a positive correlation with AQ score when the language in use was Mandarin. In other words, individuals with more autistic traits tended to use a faster speaking rate when using their L1 but a slower speaking rate when using their L2. In addition, individuals with more autistic traits tended to use higher f0 min when using their L1.

## 5. General discussion

In this study, we examined conversational pitch convergence of Mandarin speakers in a casual board game setting where each individual was asked to make up a story with a partner based on the materials given using Mandarin and English. Table 2 below provides a brief summary of the statistical results on the four pitch measures. An obvious "language" effect is found across all measures, and the pitch convergence is found in f0 min, f0 mean, and f0 range. The effect of the interlocutor's gender is only found in the f0 range.

**Table 2**
*Significant main effects of the three factors on the four pitch measures*

| | **Language** | **Position** | **Group Gender** |
|---|---|---|---|
| f0 max | English > Mandarin | | |
| f0 min | English > Mandarin | turn-initial > turn-final | |
| f0 mean | English > Mandarin | turn-initial > turn-final | |
| f0 range | Mandarin > English | turn-final > turn-initial | same > opposite |

The four pitch measures and a personality measure (i.e., AQ score) were used to answer the two research questions:

(i) Does the participants' pitch in casual speech converge differently when the languages in use are different?
(ii) Do participants' personality differences affect their performance in conversational pitch convergence?

The answer to the first question is yes. All the pitch measures, including f0 max, f0 min, f0 range, and f0 mean had **language** as the factor that caused significant differences. In other words, when speakers were in the Mandarin recording session, they spoke with a lower pitch (maximum, minimum, and average) and wider pitch range. The pitch range was calculated as the difference between the f0 max and f0 min, and this suggests that the wider pitch range in Mandarin should result from an extremely low minimum pitch. An explanation is that speakers felt more comfortable using their native language so that they tended to be more articulate and allowed their vocal folds to vibrate more slowly.

Among all the phonetic measures, the f0 range was the one that showed a significant difference in all the listed factors (i.e., **language, group gender**, and **position**). This finding suggested that speakers not only had a strong tendency for pitch convergence but also tended to manipulate the f0 range for various purposes, such as using a different language, and speaking to someone of the opposite gender. Therefore, a language instructor could encourage language learners to manipulate their speech by lowering their pitch (especially the minimum pitch to widen their pitch range) when learning a new language in conversation with a partner.

The answer to the second question is no; the correlations between the AQ scores and the pitch measures were weak. It appears that "speaking rate" was the measure significantly associated with autistic traits in both the Mandarin and English recording sessions, and "f0 min" in the Mandarin session was the only pitch measure that had a positive correlation with autistic traits. As shown in Table 1 in section 4.2, there is a tendency for speakers with a timid personality (i.e., those with a higher AQ score; aka introverts) to speak faster than extroverts when using their native language (*rho* = 0.2). Nevertheless, introverts tended to speak more slowly than extroverts when using a second language (*rho* = -0.28). This finding suggests that a speaker's personality has a direct influence to their speaking rate, but not on their pitch. A similar finding was reported in Xiao et al. (2015) where they found that therapists' empathy

toward patients correlated with their speaking rate but not with the other phonetic measures. Therefore, a language instructor might consider measuring the learners' levels of autistic traits prior to the instruction and encourage those with higher AQ scores (aka the introverts) to enhance their fluency or confidence level when speaking a second language.

## 6. Conclusion

This study used informal board gameplay to capture conversational pitch convergence in a storytelling context. The findings demonstrated that speakers showed pitch convergence in f0 min, f0 mean, and f0 range when using different languages, and their personalities did not directly affect their pitch performance but played an essential role in their speaking rates. The finding is expected to contribute to the study of L2 acquisition and human-machine interaction by emphasizing the importance and the interaction of pitch, language, speaking rate, and personality.

## References

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics, 40*(1), 177-189.

Babel, M., & Bulatov, D. (2012). The role of fundamental frequency in phonetic accommodation. *Language and Speech, 55*, 231-248.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders, 31*, 5-17.

Bayeck, R. Y. (2020). Examining board gameplay and learning: A multidisciplinary review of recent research. *Simulation & Gaming, 51*(4), 411- 431.

Bell, L., Gustafson, I., & Heldner, M. (2003). Prosodic adaptation in human-computer interaction. *Proceedings of the 15th ICPhS*, 2453-2456.

Bishop, J. & Kuo, G. (2016). Do "autistic-like" personality traits predict prosody perception? Presentation at *Laboratory Phonology 15 Satellite Workshop*.

Boersma, P., & Weenink, D. (2021). Praat: doing phonetics by computer [Computer program]. Version 6.2.04, retrieved 18 December 2021 from http://www.praat.org/

Borrie, S. A., & Liss, J. M. (2014). Rhythm as a coordinating device: Entrainment with disordered speech. *Journal of Speech, Language, and Hearing Research, 57*(3), 815-824.

Brennan, S., & Clark, H. (1996). Conceptual pacts and lexical choice in conversation, *Journal of Experimental Psychology, learning, Memory and Cognition, 22*(6), 1482-1493.

Cole, J., & Shattuck-Hufnagel, S. (2011). The phonology and phonetics of perceived prosody: What do listeners imitate? *INTERSPEECH 2011*, 969-972.

Danescu-Niculescu-Mizil, C., & Lee, L. (2011). Chameleons in imagined conversations: A new approach to understanding coordination of linguistic style in dialogs. *ACL HLT 2011, 76*.

Fuchs, S., & Reichel, U. (2016). On the relation between pointing gestures and speech production in German counting out rhymes: Evidence from motion capture data and speech acoustics. *Proceedings of the Phonetics and Phonology Conference (P&P),* 2016.

Gálvez, R. H., Gravano, A., Beňuš, Š., Levitan, R., Trnka, M., & Hirschberg, J. (2020). An empirical study of the effect of acoustic-prosodic entrainment on the perceived trustworthiness of conversational avatars. *Speech Communication, 24,* 46-67.

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. *Contexts of Accommodation*, 1-68.

Goldinger, S. (1997). Perception and production in an episodic lexicon. *Talker Variability in Speech Processing,* 33-66.

Gonzales, A. L., Hancock, J. T., & Pennebaker, J. W. (2013). Language style matching as a predictor of social dynamics in small groups, *Communication Research*, *37*(1), 3-19.

Gregory, S. W., Dagan, K., & Webster, S. (1997). Evaluating the relation of vocal accommodation in conversation partners' fundamental frequencies to perceptions of communication quality. *Journal of Nonverbal Behavior, 21,* 23-43.

Heldner, M., Edlund, J., & Hirschberg, J. B. (2010). Pitch similarity in the vicinity of backchannels. In *Proceedings of INTERSPEECH 2010* (pp. 3054–3057). Baixas, France: International Speech Communication Association.

Ip, M. H. K., & Cutler, A. (2020). Universals of listening: Equivalent prosodic entrainment in tone and non-tone languages. *Cognition*, *202*, 104311.

Jun, S., & Bishop, J. (2015). Priming implicit prosody: prosodic boundaries and individual differences. *Language and Speech 58*(4), 459-473.

Kraljic, T., Brennan, S., & Samuel, A. (2008). Accommodating variation: Dialects, idiolects, and speech processing, *Cognition*, *107*(1), 54-81.

Lakin, J. L., Jefferies, V. E., Cheng, C. M., & Chartrand, T. L. (2003). The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior, 27*(3), 145–162.

Levitan, R., Beňuš, Š., Gravano, A., & Hirschberg, J. (2015). Acoustic-prosodic entrainment in Slovak, Spanish, English and Chinese: A cross-linguistic comparison. *SIGDIAL*, 325-334.

Levitan, R., & Hirschberg, J. (2011). Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions. *Proceedings of INTERSPEECH 2011*, 3081-3084.

Lubold, N., Walker, E., & Pon-Barry, H. (2021). Effects of adapting to user pitch on rapport perception, behavior, and state with a social robotic learning. *User Modeling and User-Adapted Interaction, 31,* 35-73.

Michalsky, J., & Schoormann, H. (2017). Pitch convergence as an effect of perceived attractiveness and likeability. *Proceedings of INTERSPEECH 2017*, 2253-2256.

Natale, M. (1975). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology, 32*(5), 790-804.

Nenkova, A., Gravano, A., & Hirschberg, J. (2008). High frequency word entrainment in spoken dialogue. *Proceedings of the 46th Annual meeting of the Association for*

*Computational Linguistics on Human language Technologies (ACL/HTL 2008)*, 169-172.

Niederhoffer, K. G., & Pennebaker, J. W. (2002). Linguistic Style Matching in Social Interaction. *Journal of Language and Social Psychology, 21*(4), 337-360.

Nieuwland, M., S., Ditman, T., & Kuperberg, G. R. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and language 63*, 324-346.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, *19*(4), 2390-2393.

Pardo, J. S. (2013). Measuring phonetic convergence in speech production. *Frontiers in Psychology, 4*, 559.

Reitter, D., Moore, J., & Keller, F. (2006). Priming of syntactic rule sin task-oriented dialogue and spontaneous conversation, In *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, 685-690.

Reitter, D., & Moore, J. (2014). Alignment and task success in spoken dialogue. *Journal of Memory and Language, 86*, 29-46.

Schweitzer, K., Walsh, M., & Schweitzer, A. (2017). To see or not to see: Interlocutor visibility and likeability influence convergence in intonation. *Proceedings of INTERSPEECH 2017*, 919-923.

Shockley, K., Santana, M., & Fowler, C. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception & Performance, 2*, 326–332.

Staun Casasanto, L., Jasmin, K., & Casasanto, D. (2010). Virtually accommodating: speech rate accommodation to a virtual interlocutor. *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, 127-132.

Steinkuehler, C. A., Squire, K., & Barab, S. (Eds.). (2012). *Games, learning, and society: Learning and meaning in the digital age*. Cambridge University Press.

Stewart, M. E., & Austin, E. J. (2009). The structure of the Autism-Spectrum Quotient (AQ): Evidence form a student sample in Scotland. *Personality and Individual Differences, 47*, 224-228.

Thomason, J., Nguyen, H. V., & Litman, D. (2013). Prosodic Entrainment and Turing Dialogue Success. *Artificial Intelligence in Education (AIED) 2013*, 750-753.

Wynn, C. J., Borrie, S. A., & Sellers, T. P. (2018). Speech rate entrainment in children and adults with and without autism spectrum disorder. *American Journal of Speech-Language Pathology, 27*(3), 965-974.

Xia, Z., Levitan, R., & Hirschberg, J. (2014). Prosodic Entrainment in Mandarin and English: A Cross-Linguistic Comparison. *Proceedings of the International Conference on Speech Prosody* (pp. 65–69).

Xiang, M., Grove, J., & Giannakidou, A. (2013). Dependency-dependent interference: NPI interference, agreement attraction, and global pragmatic inferences. *Frontiers in Psychology 4*, 708.

Xiao, B., Imel, Z. E., Atkins, D. C., Georgiou, P. G., & Narayanan, S. S. (2015). Analyzing speech rate entrainment and its relation to therapist empathy in drug addiction counseling. *Proceedings of INTERSPEECH 2015*, 2489-2493.

Yu, A., Abrego-Collier, C., and Sonderegger, M. (2013). Phonetic imitation from an individual difference perspective: subjective attitude, personality and "autistic traits". *Plos One, 8*.

**About the author:**

**Chen-Hsiu Grace Kuo** is an assistant professor at the Department of Foreign Languages and Literatures, National Taiwan University, Taiwan. She received her Ph.D. in Linguistics from the University of California, Los Angeles, US. Her research interests include prosody acoustics and perception, L2 acquisition, and socio-phonetics.