Exploring the Linguistic Signature of Interpersonal Liking in Second Language Interaction

Pavel Trofimovich,[a] Anamaria Bodea,[a] Kim McDonough,[a] and Masatoshi Sato[b]

[a]Concordia University, Montreal, Canada [b]Universidad Andrés Bello, Santiago, Chile

Trofimovich, P., Bodea, A., McDonough, K., & Sato, M. (2024). Exploring the linguistic signature of interpersonal liking in second language interaction. *System.* https://doi.org/10.1016/j.system.2024.103565

Highlights

- Speakers tend to underestimate their liking by conversation partners.
- English L2 speakers' conversations were coded for fluency and engagement behaviors.
- Speakers also provided perceived and actual ratings of each other's liking.
- Speakers with higher perceived ratings provided more content across shorter turns.
- No linguistic measure predicted speakers' actual liking by conversation partners.

Exploring the Linguistic Signature of Interpersonal Liking in Second Language Interaction

Abstract

People worry about how they are seen by others, but their insights (called metaperceptions) are often too negative. For instance, many speakers believe that their interlocutors like them less than they actually do, and these overly negative metaperceptions inform speakers' actions such as asking for advice or pursuing friendships. Our goal was to understand if low, underconfident metaperceptions are associated with specific interactional behaviors for second language (L2) speakers, as a way of identifying a "linguistic signature" of insecure metaperceivers. We analyzed 10-minute dyadic conversations by 37 L2-speaking university students discussing academic texts. Following the conversation, students provided their metaperceptions (how much they thought their partner liked them) and their actual assessments (how much they liked each other). We coded the conversations for eight measures of utterance fluency (repetitions, repairs, filled pauses, discourse markers) and speaker engagement (lexical content, mean length of turn, backchannels, overlapping speech). Whereas several measures predicted students' metaperceptions, none contributed to their actual assessments. Speakers who felt appreciated by their partner provided more lexical content across shorter conversational turns, whereas those who felt insecure assumed a dominant role speaking in long turns. These findings provide initial insights into how speakers' metaperceptions manifest in their interactional behavior.

Keywords: metaperception; second language; interaction; interpersonal liking; speaking; interactional behavior; linguistic measures

1. Introduction

People often worry about the impressions they make on others, and these judgments are known as metaperceptions (Kenny & DePaulo, 1993). For example, after interviews, applicants may feel like they failed to impress a potential employer or that the interviewer did not like them. These experiences illustrate a common trend for people to overestimate the degree to which they are being negatively evaluated or judged by others (Savitsky et al., 2001). Overestimation of negative perceptions is especially prevalent in public speaking situations, where nervousness levels escalate (Savitsky & Gilovich, 2003). Similarly, during conversation, speakers often form metaperceptions that their interlocutors do not appreciate them, when in fact the contrary is true. Researchers refer to the mismatch between metaperceptions and actual assessments as the liking gap, which captures people's tendency to underestimate how much others like them (Boothby et al., 2018). The liking gap also occurs in second language (L2) conversations, where speakers' similarly underestimate their interlocutors' assessments (Le et al., 2024; Trofimovich et al., 2023; Zheng et al., 2024). However, it is not yet known whether specific interactional behaviors such as repairs or hesitations play a role in speakers' tendency to provide underconfident metaperceptions and whether these behaviors reflect speakers' actual assessments by their interlocutor. To address this issue, we explored the relationship between speakers' interpersonal liking—captured through both metaperceptions and actual assessments—and speakers' linguistic behaviors in L2 conversations.

2. Background Literature

2.1 Metaperceptions in Interaction

According to social psychologists, metaperceptions (i.e., what I believe others think about me) are extension of one's self-perception (i.e., what I think about myself). Although

metaperceptions are informed by people's self-perceptions, they are different constructs because metaperceptions require people not only to develop an accurate appraisal of their own skills but also to entertain a theory of mind to understand the perspective of another person toward themselves (Kenny & DePaulo, 1993). People tend to project their self-views onto others because they assume that what is evident to them will be also evident to their interlocutors (Savitsky & Gilovich, 2003). The systematic mismatch between metaperceptions and what interlocutors truly think typically occurs in the form of underestimation of social qualities such as openness, attractiveness, popularity, and especially interpersonal liking, which is the extent to which we make favorable impressions (Elsaadawy & Carlson, 2022). Put simply, people believe that their interlocutors perceive them more negatively than they do. Boothby et al. (2018) showed that speakers consistently underestimated how much they were liked by their conversation partners. The liking gap occurred after short and long conversations (e.g., 2 vs. 45 minutes) and could persist for up to six months for people who communicate daily. The same tendency for speakers to underestimate their liking occurs in group settings with up to 12 members (Mastroianni et al., 2021) and for children as young as 5 years old, which is when they start creating a self awareness (Wolf et al., 2021). As mentioned previously, speakers' tendency to underestimate their liking has been extended to L2 interactions (Le et al., 2024; Trofimovich et al., 2023; Zheng et al., 2024).

What is especially interesting about metaperceptions is that people appear to act on them, often with negative consequences (Sandstrom & Boothby, 2021). For instance, engineering students who believe that their interlocutors do not like them appear less willing to ask for help, give advice, and collaborate on another project (Mastroianni et al., 2021). Workplace employees who feel disliked by their colleagues might be less motivated to volunteer an opinion, enjoy their

job, or put in extra work effort (Byron & Landis, 2020). Similarly, L2 university students who feel uncertain as to how they are perceived may express less interest in communicating with fellow students inside and outside class settings (Trofimovich et al., 2023). And people who overestimate the perceived awkwardness of their requests tend to avoid asking for help (Dungan et al., 2022). Put simply, metaperceptions can deter people from otherwise beneficial actions such as maintaining relationships or requesting help.

Although researchers have examined metaperceptions and their consequences in many conversational situations, less is known about which speaker variables are associated with metaperceptions. Given that speakers' evaluation of themselves relies on the social context (Kenny & Malloy, 1988), metaperceptions may reflect the nonverbal and verbal behaviors that occur during interaction. In terms of nonverbal behaviors, Albright et al. (2001) found that speakers were influenced by visual cues such as eye contact and gestures when forming metaperceptions. Similarly, Ohtsubo et al. (2009) argued that speakers using facial expressions and body language are less likely to underestimate how they are seen by their interlocutor. However, studies to date have not established if specific verbal behaviors are similarly associated with interlocutor impressions. For example, underconfident metaperceivers might show a linguistic signature, insofar as they produce many hesitations, contribute little conversational content, or avoid collaborative, interlocutor-focused behaviors such as backchanneling (e.g., *uh huh*, *right*). To understand metaperception and minimize its negative consequences, it may therefore be important to establish a linguistic profile of metaperceivers.

2.2 Metaperceptions and Linguistic Behaviors

When it comes to verbal behaviors associated with metaperceptions, social psychologists who investigate communication between speakers in their first language (L1) have identified

several speaker concerns involving their ability to initiate, maintain, and end conversations, including knowing what to say, speaking without awkward or undue pausing, and disclosing too much or too little (Sandstrom & Boothby, 2021). These concerns are persistent and require targeted interventions to mitigate (Sandstrom et al., 2022). They are also speaker-centered, meaning that speakers tend to blame themselves, not their interaction partners, for any conversational shortcomings (Welker et al., 2023), attributing the lowest points in a conversation to their use of dysfluency markers (e.g., pausing) and social cues (e.g., laughter, eye contact). These concerns about the content and delivery of speech seem to feed into speakers' fear of being perceived incompetent, evaluated unfavorably, or blamed if the conversation is not successful or positive (Savitsky et al., 2001).

Communication researchers have documented similar linguistic behaviors with respect to L1 and L2 speakers' experience with public speaking anxiety, a social apprehension commonly reported in relation to real or anticipated public speaking events (Bodie, 2010). Compared to individuals experiencing low levels of anxiety, high-anxious speakers tend to produce less talk and more frequent pauses, repetitions, self-corrections, and slips of the tongue (Gallego et al., 2022; Gregersen, 2005, 2009; Hofmann et al., 1997; Lewin et al., 1996; MacIntyre & Gardner, 1994). These observable behaviors of anxious speakers are intertwined with their metaconcerns about interlocutor responses, including what the listener will think or how the listener will react (Goberman et al., 2011; Grieve et al., 2021).

Finally, research on L2 speaker fluency (e.g., Suzuki et al., 2021) and engagement (e.g., Lambert et al., 2017; Qiu & Lo, 2017; see Hiver et al., 2024, for a recent review) has revealed verbal behaviors associated with more versus less optimal interactive performances. In dialogue, fluency often manifests through low frequency of filled pauses and repairs (Tavakoli, 2016), and

dialogic speech is perceived as fluent when interlocutors provide backchannels, use discourse markers as hedges (e.g., *sort of*, *like*), and produce overlapping speech to show collaboration (Brown et al., 2023; Sato, 2014; van Os et al., 2020). In terms of speaker engagement, which broadly captures a person's action with respect to another person or task (Mercer, 2019), engaged speakers provide more conversational content, elaborate their ideas, and show more responsiveness toward their interlocutor through backchanneling (Dao & McDonough, 2018; Lambert et al., 2017; Nakamura et al., 2021; Qiu & Lo, 2017). If speakers feel more or less confident about the impressions they make on others, they might reveal it through fluency and engagement behaviors in interaction.

2.3 The Present Study

To summarize, speakers' tendency to underestimate their interpersonal liking has been documented in many contexts. Although some nonverbal behaviors have been linked to speakers' metaperception (Albright et al., 2001), less is known about the potential relationship between speakers' verbal behaviors and metaperception, particularly in L2 contexts. Considering that people's metaperceptions have important consequences for their future action (Byron & Landis, 2020), uncovering a linguistic profile of metaperception seems especially important as a way of understanding how interpersonal insecurity manifests in dialogue and ultimately affects interlocutors. In our review of research from diverse fields, two sets of measures emerged as relevant to metaperception. The first concerns various aspects of breakdown and repair fluency such as the incidence of pauses, fillers, hedges, repetitions, and repairs (Brown et al., 2023; Goberman et al., 2011; Lewin et al., 1996). The second includes engagement measures capturing the amount of content produced by speakers and their interlocutor-focused behaviors such as backchanneling and overlapping speech (Lambert et al., 2017; van Os et al., 2020; Welker et al.,

2023). We targeted eight measures from these two categories to determine if speakers' metaperceptions are reflected in their verbal behaviors.

To address this goal, we used an existing dataset in which L2 speakers showed metaperceptions that were consistently lower than their actual assessments from conversation partners (Trofimovich et al., 2023). We first coded speaker performances from a 10-minute paired discussion task for the eight linguistic measures, then explored whether these measures predicted speakers' metaperceptions and their actual assessments. Because we relied on existing data in exploratory analyses, we did not generate formal hypotheses. We however expected that lower metaperceptions would be associated with speakers who had more breakdown and repair dysfluency and less interlocutor-focused engagement. Put simply, we anticipated that particularly underconfident metaperceivers would reveal their insecurity through some linguistic measures. In the absence of prior work, we had little insight as to which measures would instead be linked to speakers' actual liking. Our study was guided by the following research question: Are L2 speakers' metaperceptions or their actual assessments of interpersonal liking related to their verbal behaviors in interaction?

3. Method

3.1 Paired Interactions

The paired interactions were drawn from a prior dataset, in which 74 L2 English university students in pairs completed a 10-minute discussion about two academic texts (Trofimovich et al., 2023). Considering that each speaker's performance is influenced by their experience with their conversation partner, which violates the criterion of the independence of observations, we selected one student from each pair (13 female–female, 13 male–male, 11 mixed gender) without consulting transcripts and before coding any linguistic measures. Our

selection was semirandom by which we ensured that the students represented a diverse sample of L1 backgrounds and a balanced distribution of self-reported genders. The selected 37 students (19 males, 18 females) were on average 24 years old (*SD* = 6) and came from 14 L1 backgrounds, the largest being Mandarin (13), Bengali (5), French (5), Arabic (3), and Farsi (2). All were students in an English-medium university in Montréal, a city where French and English are commonly used, located in the French-speaking Canadian province of Québec. The students had studied English previously for a mean of 13 years (*SD* = 5); they used a 100-point scale (where 100 meant "fluent") to provide a mean self-rating of speaking skills of 79 (*SD* = 16), which implied high L2 proficiency.

For the discussion task, which focused on the topic of nature versus nurture, the two interacting students received a brief academic text (around 200 words), each supporting a different side of the debate (see Appendix A). The students read each text individually and then engaged in a 10-minute conversation guided through five prompting questions available to both students (e.g., Which side do you agree with in this debate? Why have scientists been debating this question for centuries?). The conversations were audio recorded using a recorder placed outside the students' direct view, and all students were aware of their interaction being recorded.

After the 10-minute mark, each conversation was stopped; the students then proceeded to individual rooms to complete online questionnaires through the LimeSurvey platform (https://www.limesurvey.org), which took approximately 20–30 minutes. The current analysis is based on the rating scale for interpersonal liking (see Appendix B). They first used a 100-point sliding scale (where 100 meant "strongly agree") to evaluate how much they liked their conversation partner through four statements (e.g., I liked the student, I would like to get to know the student better). They then recorded their metaperceptions, assessing how much their partner

liked them through similar items (e.g., I think the student liked me, I think the student would like to get to know me better). Lastly, the students provided demographic information by completing a brief biographical questionnaire, which included details about their age, gender, languages spoken, education, L2 proficiency, and duration of their residence in Canada.

3.2 Coded Measures

Audio-recordings were transcribed, then verified by a researcher. No conversation was shorter than 10 minutes, so only the first 10 minutes of each conversation were analyzed. The transcripts were coded for eight measures, as described below (with longer examples shown in Table 1). These measures were derived and finalized through iterative discussions followed by exploratory coding of transcripts as part of biweekly meetings in the researchers' lab. In the initial step, we generated a potential list of conceptual themes (i.e., speaker engagement, fluency, speaking anxiety), along with associated markers of speakers' interpersonal or linguistic insecurity (e.g., utterance length, self-corrections, pausing). In the next step, we performed and discussed preliminary coding of several transcripts. In the last step, we finalized the coded categories by applying agreed-upon criteria, namely, targeting only measures that we believed were defensible (e.g., filled pauses can be reliably identified by a coder) and nonredundant (e.g., type frequency and type–token ratio are highly interrelated, so one is sufficient).

In terms of utterance fluency, we coded four measures capturing aspects of repair and breakdown fluency (i.e., repetitions, repairs, filled pauses, and discourse markers). We anticipated that speakers who underestimate how well they are liked by their partner may show distinct dysfluency phenomena, including repetitive talk and high incidence of self-repairs, filled pauses, and lexical fillers such as discourse markers. Put differently, we expected insecure metaperceivers to demonstrate lack of fluency through increased repetitions, self-repairs, and

pausing or to compensate for their lack of fluency through greater use of discourse markers.

1. A measure of repair fluency, repetitions are units of consecutively repeated words (e.g., "*how* um *how*"), word combinations (e.g., "based on your parents *do you do you* think you"), or brief clausal units (e.g., "um *they found out they found out*"), even if each repeated item was separated by a filled pause (Bosker et al., 2013). Each instance of repetition was counted once, irrespective of how many times the speaker repeated a particular unit, although repetitions typically occurred two or three times.
2. Conceptually related to repetitions as a measure of repair fluency, repairs are immediate self-corrections by the speaker (Bosker et al., 2013). If the speaker self-corrected the same item multiple times (e.g., "*app… appreciate… appreciation*"), the repair was counted only once so as not to inflate repair counts.
3. As a measure of breakdown fluency, filled pauses are defined as any pauses filled with nonlexical content such as *uh* or *um* (Bosker et al., 2013). These were different from similar intrusions counted toward the backchannel category (see below) because filled pauses always occurred during a speaker's turn, whereas nonlexical backchannels (e.g., *uh huh*) were quick, isolated responses to a partner. We counted only filled (rather than silent) pauses because they can be reliably identified by a coder; additionally, how to define and interpret a practical threshold for a silent pause (e.g., 0.25 vs. 0.40 vs. 1.00 s) is a matter of controversy (Shea & Leonard, 2019) that we wished to avoid.
4. Discourse markers are vaguely semantic words which are often used as fillers and characterize more natural, idiomatic delivery and greater fluency (Wolk et al., 2021). They were coded broadly across the categories of personal (e.g., *right*, *yeah*, *okay*) and cognitive (e.g., *so*, *I mean*, *sort of*, *you know*, *like*) discourse markers, but excluding

referential and structural ones (e.g., *but*, *because*, *and*, *now*, *how about*) which predominantly help speakers organize discourse rather than overcome speech planning difficulty (Wolk et al., 2021). Although *yeah* and *okay* were also coded as backchannels, they were counted as discourse markers only when embedded in a longer response such as a full speaking turn (e.g., "*Yeah* I really don't know I have no idea how"), and no item was coded toward both discourse markers and backchannels simultaneously.

In terms of engagement, a speaker can show engagement with an interlocutor by contributing conversational content, elaborating ideas, and expressing approval or disapproval through backchanneling or nodding (Dao & McDonough, 2018; Lambert et al., 2017; Nakamura et al., 2020; Qiu & Lo, 2017). We captured some of these interactional behaviors through four measures (i.e., lexical content, mean length of turn, backchannels, and overlapping speech). We expected that speakers who feel particularly apprehensive about how they are perceived might contribute little content (e.g., speaking in short conversational turns) or might avoid collaborative, interlocutor-focused actions such as backchanneling and overlapping speech.

5. Lexical content is a broad measure of conversational output (Nakamura et al., 2020; Qiu & Lo, 2017), defined as the total number of content word types (nouns, adjectives, verbs, adverbs) produced by the speaker and counted through a lexical profiler (Cobb, 2023).
6. As a broad index of the time invested in speaking (Lambert et al., 2017; Nakamura et al., 2020), mean length of turn is the total number of words produced by the speaker, after the removal of nonlexical content (i.e., repair attempts, fillers such as *uh* and *um*, and backchannels) and discourse markers, and divided by the total number of the speaker's turns. Word counts were computed through a lexical profiler; speaker turns were derived from transcripts.

7. A measure of social reciprocity in conversation (Storch, 2002), backchannels are quick, generic verbal or nonverbal responses (e.g., *yeah*, *okay*, *uh huh*) by a speaker while listening to their partner, often used to signal willingness to listen and provide feedback that facilitates conversation (Bjørge, 2010; Lambert et al., 2017). To be coded, backchannels had to occur either alone or with another single, short-phrase response; backchannels were not counted when they were paired with longer responses (e.g., "*Yeah* I really don't know I have no idea how") or when they involved substantive responses requiring extended context for interpretation (e.g., "Oh come on, are you serious?").
8. Overlapping speech, conceptualized as an index of collaboration and reciprocity (van Os et al., 2020), was coded as any overlapping content between two speakers (Sacks et al., 1974). Overlapping speech was coded regardless of its purpose (e.g., supply affirmation, provide correction) and irrespective of its length (e.g., a single word, longer phrase) and source (e.g., whether or not a turn was stolen), considering that overlaps characterize a speaker's investment to scaffold and sustain interaction (Lerner, 1989). Although overlapping speech involves both positive and negative connotations (e.g., as an instance of collaboration and support vs. a case of interruption and speaking over), we were unable to straightforwardly determine—from transcripts alone—the potential impact of each coded instance on a given speaker; therefore, all cases of overlapping speech were categorized together regardless of the affective valence they may have conveyed.

Table 1. *Examples of Coded Categories*

Coded measure	Example
Repetitions	P4: So the text I read about what about uh the was the experience where they brought 100 um when **they they** are um 100 identical twins participating **in the in the** experience

Repairs	P45: So um they **will be more agr** um **they will agree** to that um like agree to that discussion that um they find that cartoon more humorous
Filled pauses	P26: Ok so… **uh** so after this in the result what they have so in the result they found that **um** the… most of the twins have strong influence of genetics on their personality.
Discourse markers	P33: %I think% the answer **like** for this **like** what I think is curiosity, because the reason why we are in space today is because we are curious what is out there
Backchannels	P8: For some people so that's why this is important to you this um is important to you
	P7: **Yeah**
Overlapping speech	P62: Like nature and **%nurture%**
	P61: **%yeah and%** I feel like language acquisition is somehow similar to the intelligence like how it related how

Note. Words placed between the % symbol represent overlapping speech.

All transcripts were first coded by one researcher. Another trained coder then used the same coding scheme to code four transcripts (10%) for all measures except those computed through a lexical profiler (total word count, lexical content). Inter-rater reliability (Cronbach's alpha) was high across all measures: repetitions (1.00), repairs (.97), filled pauses (.97), discourse markers (1.00), mean length of turn (1.00), backchannels (.94), and overlapping speech (.99). Because all conversations (and associated transcripts) corresponded to 10 minutes of total interaction time per speaker, all coded variables were directly comparable without any data transformation, meaning that each count reflected the same baseline of conversation length.

4. Data Analysis

The students' ratings of interpersonal liking revealed high consistency (Cronbach's alpha)

for metaperceptions (.93), which targeted how much the student believed that their partner liked them, and for actual assessments (.91), which focused on how much the partner liked the student. Therefore, composite values were created by averaging across the four relevant assessments per student. The measure of perceived liking was thus the composite metaperception score provided by each student; the measure of actual liking was the composite assessment score provided by the student's partner. As in the initial study (Trofimovich et al., 2023), to describe the students' ratings of interpersonal liking, we used the lme4 package (version 1.1-31; Bates et al., 2015) in R (version 4.2.2; R Core Team, 2023) to compute a linear mixed model with rating type (perceived vs. actual), followed by speaker and partner gender (and their interactions) as fixed effects. To capture speaker-specific variance, we also included by-speaker intercepts as a random effect and controlled for several student background variables (five personality traits, plus speakers' age, length of residence in Canada, and their self-assessed L2 speaking ability), given that interpersonal liking might depend on the students' age, L2 experience and proficiency, and specific personality traits (see Trofimovich et al., 2023).

Before the main analysis, all measures were checked for assumptions of normality through visual inspection (Q–Q plots) and tests (Shapiro-Wilk). Whereas the measure of perceived liking revealed a normal distribution, the measure of actual liking and most coded variables violated normality, demonstrating a negative skew (e.g., actual liking) or a positive skew (e.g., repetitions, repairs, overlapping speech, mean length of turn), which was not unexpected because the students were high-proficiency L2 speakers. We also checked for intercorrelations between the eight coded measures; these analyses (reported in Appendix C) revealed no relationships that surpassed the benchmark of |.70| (Field, 2018), with the highest association reaching .64 (between discourse markers and lexical content).

For the main analysis, to explore relationships between the eight linguistic measures and the students' perceived and actual liking, we computed generalized additive models using the mcgv package (version 1.9-1; Wood, 2011). An extension of linear regression, it is an ideal analysis which accommodates a range of data, including nonnormal distributions, and does not assume linear relationships between the response variable and its predictors (Wood, 2011). For instance, overlapping speech might be associated positively with perceived liking until a certain threshold, after which the relationship might diminish or change in directionality, demonstrating a curvilinear function. We ran two models, one for each response variable (perceived liking, actual liking), using the eight coded measures as predictors (called smooth terms). We treated speakers' gender and self-assessed L2 speaking ability as fixed-effect covariates and by-speaker intercepts as a random effect. Because generalized additive modeling yields complex output, the recommended way of interpreting data is through visualization, for which we used the itsadug package (version 2.4.1; van Rij et al., 2022).

5. Results

5.1 Preliminary Analyses

Before exploring linguistic correlates of interpersonal liking, we first carried out descriptive analyses of the students' ratings and the coded measures derived from their conversations. In terms of liking, the students' metaperceptions (perceived liking: $M = 63.96$, $SD = 17.92$, range = 27–100) were lower than their partners' actual assessments (actual liking: $M = 76.25$, $SD = 21.44$, range = 20–100). We found a significant effect of rating type (perceived vs. actual), $Estimate = –12.29$, $SE = 4.47$, $t = –2.75$, $p = .008$, 95% CI [–21.21, –3.38], where the students underestimated how much their partner liked them by approximately 12 points on a 100-point scale. Neither the inclusion of speaker or listener gender, $\chi^2(1) < 2.97$, $p > .085$, nor their

interaction with rating type, $\chi^2(2) < 5.02$, $p > .081$, improved fit, suggesting that the liking gap was similar for the female and male students regardless of their partner's gender. Finally, control covariates (five personality traits, speakers' age, length of residence, and self-assessed L2 speaking ability) did not improve model fit, $\chi^2(8) = 12.21$, $p = .142$ (see Appendix D). Thus, regardless of their own or partner's gender, the students' metaperceptions were lower than their actual assessments (see Figure 1).[1]

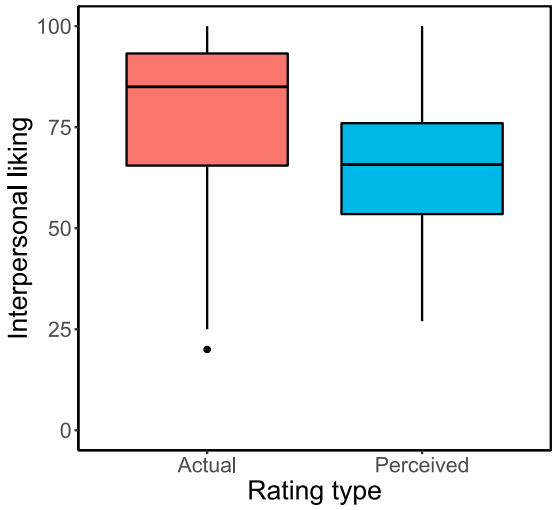

Figure 1. Boxplot for perceived liking (metaperception) and actual liking.

In terms of the linguistic behaviors (see Table 2), the students produced an average of 628.35 words (range = 150–1,119) in a 10-minute conversation, providing a range of lexical content ($M = 216.16$) across conversational turns that varied in length ($M = 17.72$). Discourse markers occurred the most frequently ($M = 56.30$), followed by filled pauses ($M = 21.03$) and backchannels ($M = 16.68$), all with a wide range of occurrences across the students. Repetitions

[1]As indicated by an anonymous reviewer, the magnitude of the liking gap might reflect both a speaker's underconfident perceived score (perceived liking) and an overly generous actual assessment by their conversation partner (actual liking). For this reason, we did not use the liking gap as the outcome measure and instead modeled speakers' perceived and actual liking separately (see Boothby et al., 2018, for in-depth discussion of the liking gap and its potential sources).

(M = 7.24), repairs (M = 6.32), and overlapping speech (M = 4.32) occurred less frequently, displaying a more restricted range.

Table 2. *Descriptive Statistics for Coded Measures*

Linguistic measure	*M*	*SD*	Range
Repetitions	7.24	6.17	0–25
Repairs	6.32	3.95	1–17
Filled pauses	21.03	14.30	2–62
Discourse markers	56.30	26.30	2–98
Lexical content	216.16	49.25	81–326
Mean length of turn	17.72	9.15	7–54
Backchannels	16.68	10.33	0–46
Overlapping speech	4.32	3.86	0–17

5.2. Predicting Perceived Liking

Our main analysis examined if the students' perceived ratings were associated with their verbal behaviors. We fit a generalized additive model using the eight coded measures as predictors of perceived ratings, treating speakers' gender and self-assessed L2 speaking ability as fixed-effect covariates. As shown in Table 3 (see Appendix E for model diagnostics), speakers' gender and L2 speaking ability made no significant contributions to the model. However, two predictors were significantly associated with perceived liking. Lexical content predicted perceived liking positively (p = .036); because effective degrees of freedom (edf) equalled 1.00 for this predictor, its effect was linear. Mean length of turn predicted perceived liking negatively (p = .046); because the edf value approached 2.00, this effect could be described through a polynomial quadratic function. The model was a good fit (Wood, 2011), shown by the amount of

variance explained (adj. R^2 = .82) and its degree of generalization (deviance explained = 97%).

Table 3. Model output for predictors of perceived liking

Coefficients	Estimate	*SE*	*z*	*p*
(Intercept)	3.95	0.33	11.83	< .001
Speaker gender	0.14	0.10	1.39	.163
Self-rated L2 speaking	0.01	0.01	0.31	.756
Smooth terms	edf	ref.df	χ^2	*p*
Repetitions	1.00	1.00	1.10	.294
Repairs	1.00	1.00	0.01	.915
Filled pauses	1.00	1.00	0.04	.841
Discourse markers	1.00	1.00	0.19	.667
Lexical content	1.00	1.00	4.42	.036
Mean length of turn	1.92	2.05	8.71	.046
Backchannels	1.00	1.00	0.23	.632
Overlapping speech	2.75	2.90	0.44	.937
Speaker (intercept)	17.59	26.00	202.22	< .001

As shown in Figure 2, the students producing less lexical content (top panel) across longer turns (middle panel) tended to have lower perceived ratings than those providing more content across shorter turns. The contour plot (bottom panel), which models the effect of both predictors simultaneously, shows that that optimal region of greater perceived liking (with ratings above 70 on a 100-point scale, plotted as gradients of warmer colors) involves a student contributing high lexical content (around 300 unique content words over 10 minutes) while speaking in turns which vary between 10 and 30 words. The region of low perceived liking (with

ratings below 70, depicted as gradients of cooler shades) involves a student speaking in long turns (30–50 words) irrespective of the lexical content provided. To sum up, the students who felt more appreciated by their partner provided more content across shorter turns, but those who felt less liked by their partner spoke in longer turns.

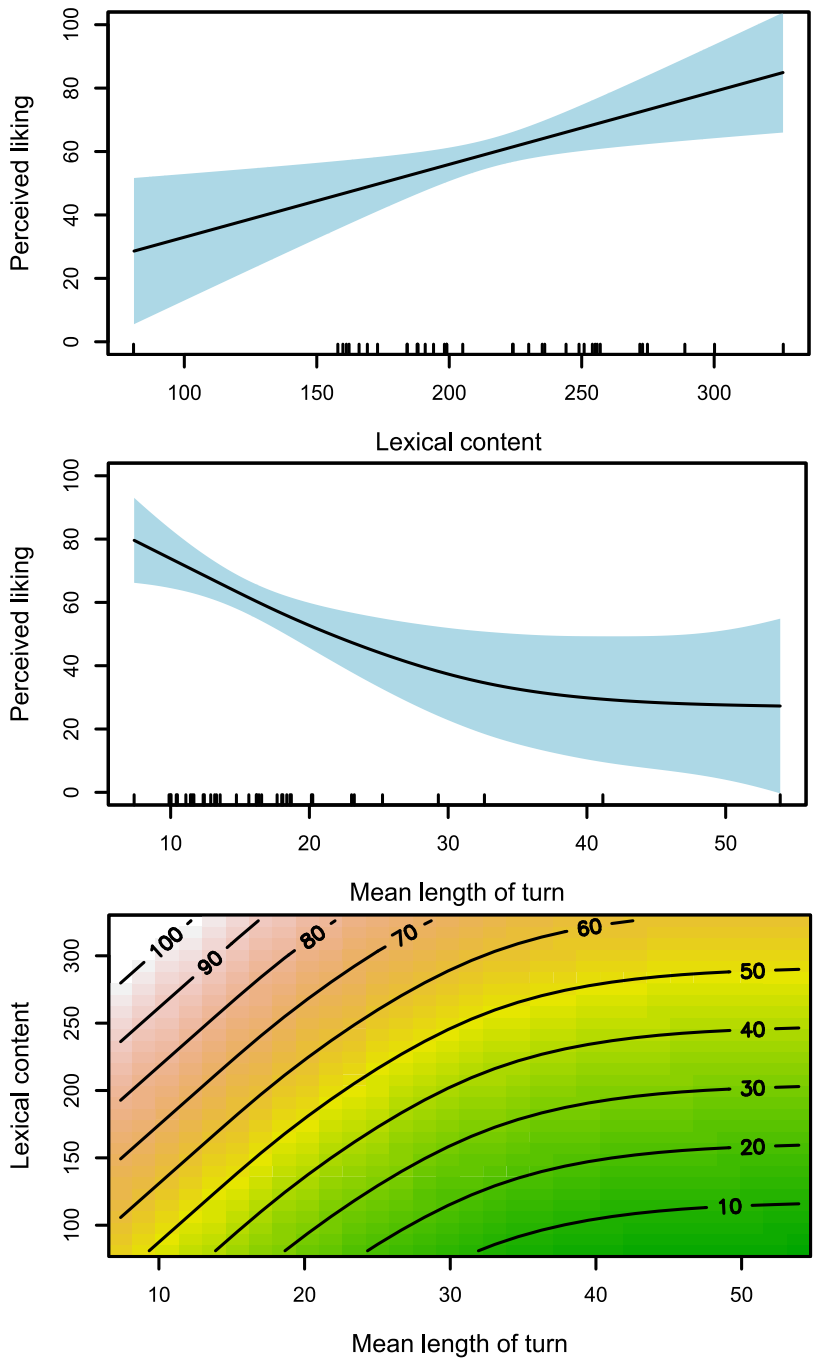

Figure 2. Relationship between perceived liking (metaperception) and lexical content (top panel) and mean length of turn (middle panel), and their joint effect on perceived liking (bottom panel). Shading corresponds to 95% confidence intervals. Warmer colors designate greater perceived liking depicted in 10-point increments through contour lines.

These findings left us with a question: Why would the two measures meant to capture the same dimension of speaker interactional behavior diverge in their relationship with metaperceptions? As shown in Figure 2 (bottom panel), whereas both lexical content and mean length of turn predicted metaperceptions at the higher end of the scale, only mean length of turn explained metaperceptions at the lower scale end. We addressed this issue in an unplanned, exploratory post hoc analysis. Based on our review of transcripts, we reasoned that unlike lexical content, which captured a speaker's conversational output, mean length of turn may have reflected a speaker's orientation toward interaction, where those who assumed a dominant role produced longer turns than those acting collaboratively. To explore this possibility, we used the framework of dyadic interaction (Storch, 2002) to designate each speaker as passive (i.e., a speaker contributes little content beyond backchannels as simple acknowledgements), dominant (i.e., a speaker contributes content but largely ignores their partner's contributions), or collaborative (i.e., a speaker both contributes content and engages with the partner such as through follow-up questions and feedback). In essence, passive speakers contributed little talk; dominant speakers approached the conversation through largely monologic performances while excluding their partner's speech; and collaborative speakers co-constructed meaning with their partner through tightly intertwined turn-taking (see Appendix F for extended examples and additional figures).

As shown in Table 4, only four of the 37 speakers were passive; the remaining were dominant (18) or collaborative (15), which allowed for exploratory comparisons. According to Mann-Whitney tests (Cohen, 1988), the dominant and collaborative speakers did not differ in the lexical content they contributed, $Z = –0.53$, $p = .300$, $r = .09$; however, compared to the collaborative speakers, the dominant speakers produced longer turns, $Z = –4.05$, $p < .001$, $r = .71$

(large effect), and also provided lower metaperceptions (perceived liking), $Z = -1.84$, $p = .033$, $r = .32$ (medium effect). To sum up, the students' dominant versus collaborative interaction pattern had no bearing on the amount of lexical content they produced. However, the two patterns were associated with different turn-taking behaviors and metaperceptions, where the dominant speakers held the floor longer and also felt less appreciated by their interaction partner than the collaborative speakers.

Table 4. *Descriptive Statistics for Post Hoc Analysis of Speakers' Interaction Patterns*

	Passive (n =4)		Dominant (n = 18)		Collaborative (n = 15)	
Measure	*M*	*SD*	*M*	*SD*	*M*	*SD*
Lexical content	165.00	71.56	227.33	49.34	216.40	36.91
Mean length of turn	11.57	1.02	23.23	10.35	12.75	3.42
Perceived liking	70.13	11.51	55.74	19.36	69.42	17.91

5.3 Predicting Actual Liking

We fit a similar generalized additive model using the same eight measures as predictors of the students' actual ratings provided by their partners. As shown in Table 5 (see Appendix G for model diagnostics), speakers' gender and self-rated L2 speaking ability made no significant contributions to the model. Though the model parameters implied a good fit to the data (adjusted $R^2 = .91$, deviance explained = 99%), most of the variance was attributed to speaker-specific, idiosyncratic performance, as captured through highly significant by-speaker random intercepts. Among the eight linguistic measures, none emerged as a significant predictor of actual liking.

Table 5. Model output for predictors of actual liking

Coefficients	Estimate	*SE*	*z*	*p*
(Intercept)	4.14	0.59	7.03	< .001
Speaker gender	–0.20	0.17	–1.14	.256
Self-rated L2 speaking	0.00	0.01	0.44	.659
Smooth terms	edf	ref.df	χ^2	*p*
Repetitions	1.00	1.00	0.63	.428
Repairs	1.00	1.00	0.59	.442
Filled pauses	1.00	1.00	2.45	.118
Discourse markers	1.00	1.00	1.65	.199
Lexical content	1.00	1.00	0.27	.600
Mean length of turn	1.00	1.00	0.11	.739
Backchannels	1.00	1.00	0.56	.453
Overlapping speech	1.00	1.00	0.09	.763
Speaker (intercept)	24.26	26.00	2097.36	< .001

6. Discussion

In this exploratory study, we examined the relationship between L2 speakers' perceived and actual assessments of interpersonal liking and their linguistic behaviors. This work was inspired by previous social-psychological research which shows that speakers underestimate how much they are liked by their conversation partners and that these judgments often have negative consequences (Boothby et al., 2018; Mastroianni et al., 2021). Our goal was to uncover a linguistic profile of metaperception as a way of understanding how interpersonal insecurity manifests in dialogue. Two of the eight coded measures (lexical content, mean length of turn)

predicted the speakers' metaperceptions; however, none of the coded measures were associated with actual liking as assessed by the speakers' conversation partners. These findings provide initial insights into how speakers' metaperceptions manifest in their linguistic behavior.

6.1 Linguistic Correlates of Metaperception

Two measures which we derived from the broader literature on speaker engagement and categorized under speaker's interactional behavior—lexical content and mean length of turn—contributed to explaining metaperceptions, where speakers who produced less unique lexical content and spoke in longer conversational turns tended to show lower metaperceptions. To illustrate, in our dataset, the students with particularly low metaperceptions—that is, beliefs that they were not particularly appreciated or liked by their conversation partner (around 30 on a 100-point scale)—provided about 188 unique content words in a 10-minute conversation across speaking turns of around 20–50 words. In contrast, the students with high metaperceptions—that is, those who felt relatively certain that they were appreciated or liked by their partner (above 80)—provided around 236 unique content words across conversational turns that were 7–25 words in length. These relationships remained significant after accounting for speakers' gender and self-reported L2 speaking ability and modeling random variation in their performance.

These findings align broadly with prior work on speaking anxiety, where anxious speakers tend to produce less talk or speak in shorter utterances while also demonstrating various dysfluency phenomena (Gallego et al., 2022; Lewin et al., 1996). These results also echo previous social-psychological research, which shows that speakers, especially those experiencing communication apprehension, worry about contributing genuine, new content to a conversation (Sandstrom & Boothby, 2021). Thus, the linguistic profile of an insecure metaperceiver is similar but not identical to that of an anxious public speaker. Just as anxious speakers, insecure

metaperceivers contribute less content but additionally speak in longer conversational turns.

Of the two predictors, mean length of turn was especially important to describing insecure metaperceivers irrespective of the amount of lexical content they produced. While lexical content measured the quantity of conversational output, mean length of turn—as implied by our post hoc analysis—captured a broader collaborative–dominant continuum of interaction (Storch, 2002). Collaborative speakers engaged in partner-sensitive conversation management, where the speaker did not dominate the floor and the speaking turns were fairly short (Ducasse & Brown, 2009). This turn-taking behavior, which is typical of the negotiation for meaning that facilitates comprehension and conversational flow (Foster & Ohta, 2005; McCarthy, 2010), seemed to characterize relatively secure metaperceivers. However, dominant speakers may have approached the interaction through monologic performances, guided by prompting questions, and this behavior was common among insecure metaperceivers. Thus, as far as metaperceptions are concerned, interpersonal insecurity manifested through a dominant orientation toward interaction, where the speaker held the floor speaking in long conversational turns.

We have thus far avoided discussing any directional or causal relationships between speakers' metaperceptions and linguistic behaviors. In fact, establishing directionality may not be possible or desirable, given that metaperception–behavior links could be reciprocal and self-reinforcing. For instance, speakers who are generally concerned about the impressions they make on others may have deliberately chosen a dominant role, as a way of asserting themselves in the conversation. At the same time, low metaperceptions may have also arisen because of speakers' frustration with how they handled the dialogue, reflecting concerns about their partner's opinion. Above all, metaperception–behavior links are most likely intertwined with speaker positioning toward conversational goals. In this study, the discussion was set up as an information-exchange

task which was high on informational goals such as sharing content but low on relational outcomes such as developing personal rapport (Yeomans et al., 2022). Insecure metaperceivers may have prioritized the task's informational component assuming a dominant role in the conversation; secure metaperceivers may have instead approached it in social terms engaging in collaborative dialogue. Thus, whether and how metaperceptions relate to speakers' interaction behavior and their assumed conversational goals are worthwhile targets of future work.

6.2 Other Linguistic Markers of Metaperceptions?

In this dataset, measures of speaker interactional behavior patterned with speakers' metaperceptions while utterance fluency showed no association with metaperceptions. Considering that people observe their own behaviors to determine how they are perceived by others (Kenny & DePaulo, 1993), it may be that some linguistic behaviors might be more observable and therefore perceptible to speakers than other behaviors, in light of their relevance to a given task. From this perspective, in a discussion task that requires speakers to exchange information, metaperceptions might indeed reflect such interactional behaviors as contributing conversational content and assuming a certain conversational role (with consequences for turn-taking). In contrast, in a task that prioritizes language accuracy or fluency, metaperceptions might instead be interlinked with various (dys)fluency phenomena in speakers' speech. In essence, the linguistic signature of insecure metaperceivers might differ by task, reflecting the behaviors that become more or less relevant to the speaker's success.

If this interpretation is correct, then such interactional behaviors as overlapping speech and backchanneling might emerge as predictors of more versus less secure metaperceivers in communicative situations with a higher relational load, for example, as speakers seek advice, apologize, or break bad news. These behaviors would be not only central to these interactive

contexts but also salient to a speaker, because overlapping speech signals the speaker's investment to scaffold and sustain interaction (Lerner, 1989; van Os et al., 2020) while backchannels mark the speaker as willing to listen and provide feedback (Bjørge, 2010; Lambert et al., 2017). Other linguistic behaviors, including various utterance fluency phenomena, might instead be useful at distinguishing more versus less secure metaperceivers in situations that are high on both the informational and relational continua such as giving an oral presentation, brainstorming ideas, or persuading an interlocutor. In these contexts, self-repetitions could signal a speaker's communicative investment (Rydland & Aukrust, 2005) whereas filled pauses and discourse markers could create the impression of an idiomatic, natural delivery (Clark & Fox Tree, 2002; Wolk et al., 2021). Thus, understanding how speakers show interpersonal insecurity through verbal behavior is likely possible only in relation to specific interactive goals.

6.3 Linguistic Predictors of Actual Liking

It is unsurprising that the students' actual assessments of likeability were unrelated to any measure derived from their speech. After all, it may be unrealistic to expect that L2-speaking students would draw heavily on their partner's linguistic behavior to consider them a likeable person or a potential friend. According to research on social attractiveness (Weiss et al., 2021), in non-romantic scenarios, speaker likeability is associated with such descriptors as high proficiency, enunciation clarity, high articulation rate, and relaxed delivery characterized by lower pitch (Rosenberg & Hirschberg, 2021). Considering that our students were high-level L2 speakers from 14 L1 backgrounds interacting in gender-balanced pairs, any linguistic dimensions of likeability may have been difficult to detect in this highly heterogenous sample. Alternatively, our measures failed to capture other consequential attributes of likeability such as a particularly strong accent, slow articulation rate, or especially high-pitch delivery. It therefore remains for

further work to determine linguistic correlates of social attractiveness for L2 speakers. For now, it is striking that while the dominant and collaborative speakers differed in their metaperceptions (see Table 4), they elicited similar actual ratings (77 vs. 80), implying that a speaker's turn-taking behavior during interaction made little difference in how likeable they were seen by their partner.

7. Implications

Our findings offer tentative implications for L2 speakers and language instructors. Even though some measures of speakers' linguistic behavior predicted their metaperceptions (i.e., beliefs about how likeable or appreciated they appeared to their conversation partner), none in fact mattered for their actual liking, which might be welcome news for many students, including those who struggle with speaking anxiety (Bodie, 2010). Instructors could therefore remind students that some linguistic behaviors are not as conspicuous or consequential to others as students might believe. Instructors could similarly engage students in discussion of verbal behaviors which might primarily feed into students' insecurity versus those which might be appreciated by their interlocutor (e.g., speaking comprehensibly, smiling, providing verbal and nonverbal feedback). Instructors could also capitalize on activities targeting various metaperception biases, including speakers' tendency to underestimate their liking (Bootby et al., 2018) and overestimate the extent to which their nervousness is visible to others (Savitsky & Gilovich, 2003). These activities might focus on providing unambiguous and explicit feedback, encouraging peer evaluations, and discussing dominant versus collaborative communication styles. Finally, instructors could also tell students that they are often more likable than they believe and encourage them to reflect on their speaking and communication skills. For instance, after students complete a discussion activity, they could analyze how they provided and interpreted various signals (e.g., laughter, backchanneling, verbal affirmation) and consider

whether these signals contributed to the success of their interaction.

8. Limitations

This study is not without limitations. To begin with, in the absence of prior work on speaker insecurity, our chosen linguistic measures generally reflected our intuition rather than principled decision-making. Therefore, we may have inadvertently not included relevant linguistic markers of interpersonal insecurity relevant to metaperceptions. For instance, unfilled pauses as a marker of breakdown fluency might have been more consequential to speakers' metaperceptions than our measure of filled pauses, especially in dialogic performances (Tavakoli, 2016). Without video recordings, we also could not examine speakers' nonverbal behaviours and assess their comparative role in metaperceptions. Speakers' nonverbal cues such as displays of emotion, facial expressions, eye gaze, gestures, and body movement may have provided signals as to how they were perceived by each other, with consequences for their own linguistic and nonlinguistic behavior. A worthwhile focus of future research, therefore, would include establishing a holistic, embodied picture of how speakers react to and feel about each other in interaction and how these perceptions manifest in their verbal and nonverbal behavior.

In terms of the sample, the students were on average 24 years old. In this demographic, some aspects of discourse such as the incidence of dysfluency might be more or less prevalent than in other groups (Bortfeld et al., 2001), with consequences for how observable these behaviors are to speakers. Similarly, the students were proficient in L2 English, with a mean of 13 years of prior language study, and nearly all reported having successfully passed internationally recognized standardized tests. Our findings, therefore, remain specific to young L2 speakers studying at an English-medium university located in a complex sociolinguistic environment, where English is widely spoken around the university campus but is considered a

second language outside this context.

As discussed previously, our findings can only be understood with respect to the discussion task, and they reflect a speaker-specific orientation toward it (i.e., passive, dominant, collaborative). As pointed out by an anonymous reviewer, speakers may have also differed in their familiarity and comfort with the task (i.e., debate over nature vs. nurture), so their knowledge of the task's content and their ability to engage in an argumentative discussion with a partner may have influenced their metaperceptions of each other. Speakers' personality may have also played a role in their performance and its relationship with metaperception. While we accounted for students' personality in our measure of liking, modeling personality through generalized additive modeling proved problematic, as models became too complex to converge. Similarly, our student sample included individuals representing diverse L1s and places of origin, which we hoped would minimize the influence of any given ethnolinguistic background on speakers' metaperceptions. Nevertheless, specific cultural norms such as the value placed on being persuasive versus self-deprecating might have influenced speakers' perceptions of how they are seen by their interlocutor, with consequences for their linguistic behaviors. It would also be important to move beyond the linguistic domain to document how interpersonal insecurity—as captured through interlocutors' metaperceptions of each other—might manifest in their nonverbal behaviors such as the visual cues they provide to each other (e.g., avoidance of eye contact, leg bouncing, rigid posture, lack of gesture). Equally revealing would be to document various socioaffective correlates of speakers' interpersonal insecurity, including displays of emotionality, markers of stress and anxiety, or changes in mood, attitudes, and behaviors, all of which might color interlocutors' metaperception and ultimately their interactive experience.

Finally, a major limitation pertains to our decision to maintain the independence of

observations in speakers' behaviors separately from their partner's behaviors. Interlocutors not only tend to converge in their linguistic behaviors, but their metaperceptions also depend on the partner's expressiveness, where the partner's verbal and nonverbal feedback determines how accurately speakers can judge the impressions they make on others (Hall et al., 2016). To illustrate, for the 18 speakers we identified as dominant, their partner showed a similar dominant communication strategy in 15 cases (83%), and all 15 collaborative speakers had a collaborative partner (100%). Clearly, interpersonal (in)security in interaction might be "contagious," inasmuch as one speaker's linguistic behaviors—and accompanying metaperceptions—might be reciprocated by their partner. Understanding these complex interactional dynamics is an intriguing direction for future research.

9. Conclusion

To understand how interpersonal insecurity manifests in L2 interaction, we examined potential associations between speakers' linguistic behaviors and their perceived and actual judgments of liking in an academic discussion task. Among eight coded linguistic measures, two measures (lexical content, mean length of turn) predicted speakers' metaperceptions (i.e., beliefs about how likeable they appeared to their interlocutor) but none predicted their actual assessments. Speakers who felt more appreciated by their partner tended to provide more content across short conversational turns, whereas those who felt less liked by their partner assumed a dominant role in the conversation speaking in long turns. These findings offer several implications for L2 speakers and their instructors. L2 speakers might benefit from understanding that they are often more likable than they believe and that some of their linguistic behaviors may contribute to speakers' metaperceptions but may be inconsequential to how they are actually perceived. In turn, instructors could encourage L2 speakers to reflect on their speaking and

communication skills, discussing interactive behaviors that feed into speakers' insecurity versus those that are appreciated by interlocutors.

References

Albright, L., Forest, C., & Reiseter, K. (2001). Acting, behaving, and the selfless basis of metaperception. Journal of Personality and Social Psychology, 81, 910–921. https://doi.org/10.1037/0022-3514.81.5.910

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67, 1–48. https://doi.org/10.18637/jss.v067.i01

Bjørge, A. K. (2010). Conflict or cooperation: The use of backchannelling in ELF negotiations. English for Specific Purposes, 29, 191–203. https://doi.org/10.1016/j.esp.2009.04.002

Bodie, G. D. (2010). A racing heart, rattling knees, and ruminative thoughts: Defining, explaining, and treating public speaking anxiety. Communication Education, 59, 70–105. https://doi.org/10.1080/03634520903443849

Boothby, E. J., Cooney, G., Sandstrom, G. M., & Clark, M. S. (2018). The liking gap in conversations: Do people like us more than we think? Psychological Science, 29, 1742–1756. https://doi.org/10.1177/0956797618783714

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. Language and Speech, 44, 123–147. https://doi.org/10.1177/00238309010440020101

Bosker, H. R., Pinget, A. F., Quené, H., Sanders, T., & de Jong, N. H. (2013). What makes speech sound fluent? The contribution of pauses, speed and repairs. Language Testing, 30, 159–175. https://doi.org/10.1177/0265532212455394

Brown, D., Fernández, J., & Huensch, A. (2023). The role of pragmatic markers in perceptions of L2 fluency in dialogue, System, 119, Article 103157. https://doi.org/10.1016/j.system.2023.103157

Byron, K., & Landis, B. (2020). Relational misperceptions in the workplace: New frontiers and challenges. Organization Science, 31, 223–242. https://doi.org/10.1287/orsc.2019.1285

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. Cognition, 84, 73–111. https://doi.org/10.1016/S0010-0277(02)00017-3

Cobb, T. (2023). VocabProfilers [computer program]. Accessed 17 May 2023 at https://www.lextutor.ca/vp

Cohen J. (1988). Statistical power analysis for the behavioral sciences. Routledge.

Dao, P., & McDonough, K. (2018). Effect of proficiency on Vietnamese EFL learners' engagement in peer interaction. International Journal of Educational Research, 88, 60–72. https://doi.org/10.1016/j.ijer.2018.01.008

Ducasse, A. M., & Brown, A. (2009). Assessing paired orals: Raters' orientation to interaction. Language Testing, 26, 423–443. https://doi.org/10.1177/0265532209104669

Dungan, J. A., Munguia Gomez, D. M., & Epley, N. (2022). Too reluctant to reach out: Receiving social support is more positive than expressers expect. Psychological Science, 33, 1300–1312. https://doi.org/10.1177/09567976221082942

Elsaadawy, N., & Carlson, E. N. (2022). Do you make a better or worse impression than you think? Journal of Personality and Social Psychology, 123, 1407–1420. https://doi.org/10.1037/pspp0000434

Field, A. (2018). Discovering statistics using IBM SPSS Statistics. Sage.

Foster, P., & Ohta, A. S. (2005). Negotiation for meaning and peer assistance in second language

classrooms. Applied Linguistics, 26, 402–430. https://doi.org/10.1093/applin/ami014

Gallego, A., McHugh, L., Penttonen, M., & Lappalainen, R. (2022). Measuring public speaking anxiety: Self-report, behavioral, and physiological. Behavior Modification, 46, 782–798. https://doi.org/10.1177/0145445521994308

Goberman, A. M., Hughes, S., & Haydock, T. (2011). Acoustic characteristics of public speaking: Anxiety and practice effects. Speech Communication, 53, 867–876. https://doi.org/10.1016/j.specom.2011.02.005

Gregersen, T. (2005). Nonverbal cues: Clues to the detection of foreign language anxiety. Foreign Language Annals, 38, 388–400. https://doi.org/10.1111/j.1944-9720.2005.tb02225.x

Gregersen, T. (2009). Recognizing visual and auditory cues in the detection of foreign-language anxiety. TESL Canada Journal, 26, 46–64. https://doi.org/10.18806/tesl.v26i2.414

Grieve, R., Woodley, J., Hunt, S. E., & McKay, A. (2021). Student fears of oral presentations and public speaking in higher education: A qualitative survey. Journal of Further and Higher Education, 45, 1281–1293. https://doi.org/10.1080/0309877X.2021.1948509

Hall, J. A., Schmid Mast, M., & West, T. (Eds.) (2016). The social psychology of perceiving others accurately. Cambridge University Press.

Hiver, P., Al-Hoorie, A. H., Vitta, J. P., & Wu, J. (2024). Engagement in language learning: A systematic review of 20 years of research methods and definitions. Language Teaching Research, 28, 201–230. https://doi.org/10.1177/13621688211001289

Hofmann, S. G., Gerlach, A. L., Wender, A., & Roth, W. T. (1997). Speech disturbances and gaze behavior during public speaking in subtypes of social phobia. Journal of Anxiety Disorders, 11, 573–585. https://doi.org/10.1016/S0887-6185(97)00040-6

Kenny, D. A. & Malloy, T. E. (1988). Partner effects in social interaction. Journal of Nonverbal Behavior, 12, 34–57. https://doi.org/10.1007/BF00987351

Kenny, D. A., & DePaulo, B. M. (1993). Do people know how others view them? An empirical and theoretical account. Psychological Bulletin, 114, 145–161. https://doi.org/10.1037/0033-2909.114.1.145

Lambert, C., Philp, J., & Nakamura, S. (2017). Learner-generated content and engagement in second language task performance. Language Teaching Research, 21, 665–680. https://doi.org/10.1177/1362168816683559

Le, T.-N. N., Trofimovich, P., McDonough, K., & Sato, M. (2024). Interpersonal liking, cultural belonging, and heritage language: Exploring the role of metaperception in interaction between heritage speakers of Vietnamese. Frontiers in Communication, 9, Article 1468943. https://doi.org/10.3389/fcomm.2024.1468943

Lerner, G. H. (1989). Notes on overlap management in conversation: The case of delayed completion. Western Journal of Speech Communication, 53, 167–177. https://doi.org/10.1080/10570318909374298

Lewin, M. R., McNeil, D. W., & Lipson, J. M. (1996). Enduring without avoiding: Pauses and verbal dysfluencies in public speaking fear. Journal of Psychopathology and Behavioral Assessment, 18, 387–402. https://doi.org/10.1007/BF02229142

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. Language Learning, 44, 283–305. https://doi.org/10.1111/j.1467-1770.1994.tb01103.x

Mastroianni, A. M., Cooney, G., Boothby, E. J., & Reece, A. G. (2021). The liking gap in groups and teams. Organizational Behavior and Human Decision Processes, 162, 109–122.

https://doi.org/10.1016/j.obhdp.2020.10.013

McCarthy, M. (2010). Spoken fluency revisited. English Profile Journal, 1, 1–15. https://doi.org/10.1017/S2041536210000012

Mercer, S. (2019). Language learner engagement: Setting the scene. In X. Gao, C. Davison, & C. Leung (Eds.), International handbook of English language teaching (pp. 1–19). Springer. https://doi.org/10.1007/978-3-030-02899-2_40

Nakamura, S., Phung, L., & Reinders, H. (2021). The effect of learner choice on L2 task engagement. Studies in Second Language Acquisition, 43, 428–441. https://doi.org/10.1017/S027226312000042X

Ohtsubo, Y., Takezama, M, & Fukuno, M. (2009). Mutual liking and meta-perception accuracy. European Journal of Social Psychology, 39, 707–718. https://doi.org/10.1002/ejsp.568

Qiu, X., & Lo, Y. Y. (2017). Content familiarity, task repetition and Chinese EFL learners' engagement in second language use. Language Teaching Research, 21, 681–698. https://doi.org/10.1177/1362168816684368

R Core Team. (2023). R: A language and environment for statistical computing. https://www.R-project.org

Rosenberg, A., & Hirschberg, J. (2021). Prosodic aspects of the attractive voice. In B. Weiss et al. (Eds.), Voice attractiveness: Studies on sexy, likable, and charismatic speakers (pp. 17–40). Springer. https://doi.org/10.1007/978-981-15-6627-1_2

Rydland, V. & Aukrust, V. G. (2005). Lexical repetition in second language learners' peer play interaction. Language Learning, 55, 229–274. https://doi.org/10.1111/j.0023-8333.2005.00304.x

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization

of turn-taking for conversation. Language, 50, 696–735. https://doi.org/10.2307/412243

Sandstrom, G. M., & Boothby, E. J. (2021). Why do people avoid talking to strangers? A mini meta-analysis of predicted fears and actual experiences talking to a stranger. Self and Identity, 20, 47–71. https://doi.org/10.1080/15298868.2020.1816568

Sandstrom, G. M., Boothby, E. J., & Cooney, G. (2022). Talking to strangers: A week-long intervention reduces psychological barriers to social connection. Journal of Experimental Social Psychology, 102, Article 104356. https://doi.org/10.1016/j.jesp.2022.104356

Sato, M. (2014). Exploring the construct of interactional oral fluency: Second language acquisition and language testing approaches. System, 45, 79–91. https://doi.org/10.1016/j.system.2014.05.004

Savitsky, K., & Gilovich, T. (2003). The illusion of transparency and the alleviation of speech anxiety. Journal of Experimental Psychology, 39, 618–625. https://doi.org/10.1016/S0022-1031(03)00056-8

Savitsky, K., Epley, N., & Gilovich, T. (2001). Do others judge us as harshly as we think? Overestimating the impact of our failures, shortcomings, and mishaps. Journal of Personality and Social Psychology, 81, 44–56. https://doi.org/10.1037/0022-3514.81.1.44

Shea, C., & Leonard, K. (2019). Evaluating measures of pausing for second language fluency research. Canadian Modern Language Review, 75, 216–235. https://doi.org/10.3138/cmlr.2018-0258

Storch, N. (2002). Patterns of interaction in ESL pair work. Language Learning, 5, 119–158. https://doi.org/10.1111/1467-9922.00179

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,

435–463. https://doi.org/10.1111/modl.12706

Tavakoli, P. (2016). Fluency in monologic and dialogic task performance: Challenges in defining and measuring L2 fluency. International Review of Applied Linguistics in Language Teaching, 54, 133–150. https://doi.org/10.1515/iral-2016-9994

Trofimovich, P., Lindberg, R., Bodea, A., Le, T.-N. N., Zheng, C., & McDonough, K. (2023). I don't think you like me: Examining metaperceptions of interpersonal liking in second language academic interaction. Languages, 8, Article 200. https://doi.org/10.3390/languages8030200

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, 1183–1217. https://doi.org/10.1111/lang.12416

van Rij, J., Wieling, M., Baayen, R., & van Rijn, H. (2022). itsadug: Interpreting time series and autocorrelated data using GAMMs. R package version 2.4.1. https://doi.org/10.32614/CRAN.package.itsadug

Weiss, B., Trouvain, J., Barkat-Defradas, M., & Ohala, J. J. (Eds.) (2021). Voice attractiveness: Studies on sexy, likable, and charismatic speakers. Springer.

Welker, C., Walker, J., Boothby, E., & Gilovich, T. (2023). Pessimistic assessments of ability in informal conversation. Journal of Applied Social Psychology, 53, 555–569. https://doi.org/10.1111/jasp.12957

Wolf, W., Nafe, A., & Tomasello, M. (2021). The development of the liking gap: Children older than 5 years think that partners evaluate them less positively than they evaluate their partners. Psychological Science, 32, 789–798. https://doi.org/10.1177/0956797620980754

Wolk, C., Götz, S., & Jäschke, K. (2021). Possibilities and drawbacks of using an online application for semi-automatic corpus analysis to investigate discourse markers and alternative fluency variables. Corpus Pragmatics, 5, 7–36. https://doi.org/10.1007/s41701-019-00072-x

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society: Series B, 73, 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x

Yeomans, M., Schweitzer, M. E., & Brooks, A. W. (2022). The conversational circumplex: Identifying, prioritizing, and pursuing informational and relational motives in conversation. Current Opinion in Psychology, 44, 293–302. https://doi.org/10.1016/j.copsyc.2021.10.001

Zheng, C., Trofimovich, P., Lindberg, R., McDonough, K., & Sato, M. (2024). Do they like me? Exploring the role of metaperception in L1–L2 speaker interaction. Studies in Second Language Acquisition. Published online 11 November 2024. https://doi.org/10.1017/S0272263124000469

Competing interests: The authors report that there are no competing interests to declare.

Data availability statement: All data and materials for this study are publicly available via the Open Science Framework: https://osf.io/rchfa.

Funding statement: This research was supported by Social Sciences and Humanities Research Council of Canada (SSHRC) grants awarded to the first author (435-2024-1488) and the third and the first author (435-2019-0754).

Acknowledgments: We thank our research group members Noga Broitman, Thao-Nguyen Nina Le, Yoo Lae Kim, Rachael Lindberg, and Chaoqun Zheng for their valuable insights.

CRediT author statement: Pavel Trofimovich (conceptualization, methodology, validation, formal analysis, writing–original draft, writing–review & editing, visualization, supervision, funding acquisition). Anamaria Bodea (conceptualization, methodology, investigation, data curation, validation, writing–original draft). Kim McDonough (methodology, writing–review & editing, supervision, funding acquisition). Masatoshi Sato (methodology, supervision). All authors have read and agreed to the published version of the manuscript.